The Industrial Internet of Things: Managing and Assessing Trustworthiness for IIoT in Practice

An Industrial Internet Consortium White Paper

Version 1.0

2019-07-29

CONTENTS

FIGURES

TABLES

Trustworthiness is the degree to which the system performs as expected in the face of environmental disturbances, loss of performance quality and accuracy, human errors, system faults and attacks. Assurance of trustworthiness is the degree of confidence one has in this expectation. A system must be assured as being trustworthy for a business or organization to have confidence in it. Depending on the context of the system and the possible consequences of failures, the effort spent on achieving a specific level of confidence will vary.

Confidence is essential to business, including confidence that the consequences of decisions and processes are acceptable, and that business information is handled properly. It is based on the confidence the business has in the appropriate behaviors of the individuals, organizations and the processes they use. While industrial systems offered many checkpoints and interfaces for human control in the past, connecting them to the internet, into the industrial internet of things (IIoT), requires that the confidence must now extend to the technologies, components and system.

With increasing complexity and integration of such systems into systems of systems, the risks and consequences need a more consistent and rigorous approach to defining what is required for a specific system to be trustworthy and what evidence is necessary to provide confidence in the system's ability to perform as intended for its context. Confidence comes from the assurance that several aspects of the system are under control: security of its data and of its equipment, safety for people and the community, protection of assets, privacy protection of data, reliability of operations and subsystems, and resilience of the system.

We hope to raise awareness in industry of the importance of trustworthiness, context, and assurance, how to measure, analyze and assess it, as well as how to manage and govern it. We build on the trustworthiness material outlined in the IIC Security Framework[1], the IIC Security Maturity Model[2], the IIC Security Maturity Model (SMM) Practitioner's Guide[3], and the concepts explored in the IIC Journal of Innovation trustworthiness issue[4]. We support the concepts with an example, the Factory Operations Visibility and Intelligence (FOVI) system.

1 TRUSTWORTHINESS

Trustworthiness is based on several characteristics and the interactions and tradeoffs among them. The Industrial Internet Consortium (IIC) defines in its Vocabulary[5] document these characteristics as follows:

[1] See [IIC-IISF2016]
[2] See [IIC-SMMD2019]
[3] See [IIC-SMMP2019]
[4] See [IIC-JOI20182]
[5] See [IIC-IIV2018]

Safety is the condition of the system operating without causing unacceptable risk of physical injury or damage to the health of people, either directly, or indirectly as a result of damage to property or to the environment (adapted from ISO/IEC Guide 55:1999, modified for consistency).

Security is the property of being protected from unintended or unauthorized access, change or destruction. Security concerns equipment, systems and information, ensuring availability, integrity and confidentiality of information.

Privacy is the right of individuals to control or influence what information related to them may be collected and stored and by whom and to whom that information may be disclosed (original from ISO TS 17574:2009).

Resilience is the ability of a system or component to maintain an acceptable level of service in the face of disruption. This includes the ability to recover lost capacity in a timely manner (using a more or less automated procedure), or to reassign workloads and functions.

Reliability is the ability of a system or component to perform its required functions under stated conditions for a specified period of time (original from ISO/IEC 27040:2015). This includes expected levels of performance, QoS, functional availability and accuracy.

The characteristics of security, safety, reliability, resilience, and privacy have been identified as collectively defining the trustworthiness of a system.

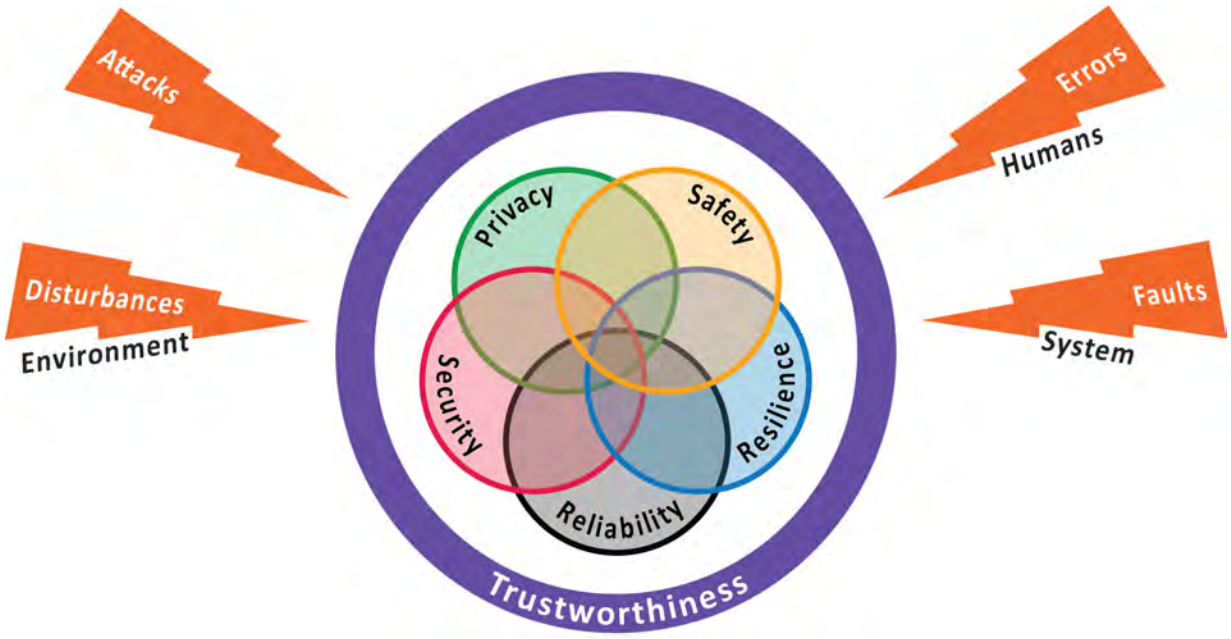


Figure 1-1: Threats and Integrated Trustworthiness Characteristics

Although these characteristics have been addressed in the past, their combination into the single concept of trustworthiness accounts for their interactions and interdependencies (see Figure 1-1). This is due to the integration of the digital and the physical world, the increased level of automation, the extent to which people and processes depend on systems to perform as expected and increased digitization and the volume of data generated.

Managing trustworthiness means understanding the trustworthiness characteristics (safety, security, privacy, reliability and resilience) in the context of a given IIoT system, defining objectives and metrics, determining what evidence is required to confirm the characteristics are providing their needed contribution to operations, assessing trade-offs between these properties, and their corresponding effect on business and operations. For some systems, the context and consequences of failures will require more assessment and evidence to gain the requisite level of confidence about the system's ability to meet the its trustworthiness characteristics versus other systems where the consequences of failure aren't as dire.

1.1 RECONCILING TRUSTWORTHINESS CHARACTERISTICS

Trustworthiness in industrial IoT applications covers IT infrastructure and data as well as physical resources (personnel, equipment), such as manufacturing or construction. Operational technology (or OT, also known in manufacturing as "shop floor") is known to have its own constraints and requirements), must be reconciled with those for information technology (or IT, also known as the "office floor").

In most industrial sites, the first priority is safety: do not kill anyone, do not put public safety at risk, and do not cause environmental harm. The second priority is usually reliability of the physical process (rather than of the control system): keep clean water running in the distributed system, keep gasoline coming out of the refinery and keep the lights on by providing power. A reliable and available control system is a means to the end to a reliable physical process.

Security in information technology ("cybersecurity") is not the same as security in operational technology. Operators of control systems traditionally care about who has physical access to controls—who is turning the dials and throwing the switches—so that the physical equipment is operated safely and reliably. In this context, security, especially security that impacts safety and reliability, is more to do with controlling access to human-machine interfaces than data security.

Several characteristics of trustworthiness, in particular reliability, safety and resilience will be driven primarily by OT concerns. Other trustworthiness characteristics such as security and reliability, apply to both OT and IT but with different emphasis (see Figure 1-2). Privacy is an IT concern only about the information that is collected, stored and managed, but addressing privacy often causes changes that can affect the other characteristics.

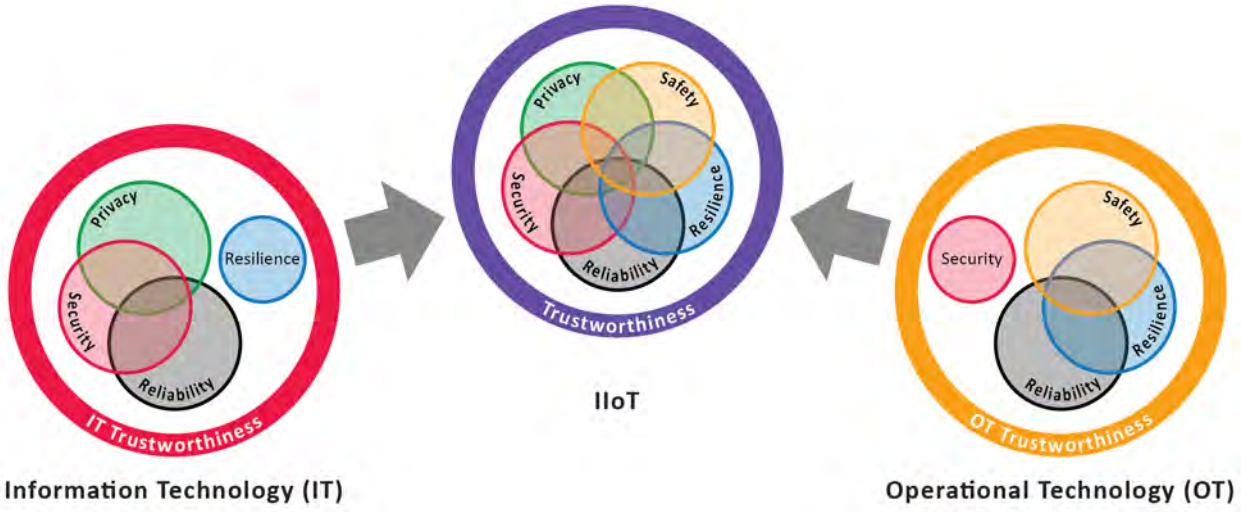


Figure 1-2: Bringing IT and OT Trustworthiness concerns together in the IIoT

IT protection profiles and maturity models focus on security, privacy and reliability, while shop floor profiles depend on the vertical in question, but start with safety and extend to resilience, for example preventing damage of the system itself. Reliability is a concern reflected in both IT and OT affecting trustworthiness of the system and business value.

The meanings and operational targets related to the trustworthiness characteristics need to be tailored for the specific system, the technologies it uses, the hazards and threats it faces, and consequences of things going wrong. Terms have different meanings. OT reliability is about the reliability of equipment; in IT, it is about services and systems. Are we talking of information security? Operational security? Security of a service? Of a process? Similarly, reliability and resilience will have different operational targets in IT and OT as well as for components of the system. Appropriate questions are necessary: Are we talking of the reliability of an assembly machine? Of an IT service? Trustworthiness may be measured using different metrics across IT, OT and subsystems and the how and why of what is done needs to be understood along with what level of rigor and type of evidence is needed to have confidence in the system's ability to perform in a trustworthy manner.

Trustworthiness characteristics interdependencies must also be understood. How will the reliability of an IT service for itinerary routing affect the safety of truck drivers? How will the reliability of this routing IT service depend on the security of the cloud platform on which it is running? Such dependencies must be captured and understood to manage trustworthiness.

There are dependencies between managing the security of a component through software updates and ensuring its stability by avoiding frequent changes and relying on extensive validation. Avoiding software updates so as to not interrupt the reliability and resilience of the manufacturing process conflicts with the need to update software to address known security and other flaws in the current system.

1.2 TRUSTWORTHINESS BUSINESS DRIVERS AND CONTEXT

There are three major drivers for trustworthiness in an industrial context:

Compliance requirements. These come with laws, regulations (industry-specific, governmental or regional) and organization-wide policies. Compliance is necessary to business in most jurisdictions, for product approval, and has marketing value (across regions, international vs. national, multi-jurisdictional,[1] OSHA, GDPR, industry-specific etc.)

Risk avoidance and *risk mitigation*. The evaluation and the handling of risk is helped by frameworks and guidelines, such as Cyber-Security Capability Maturity Model (C2M2) or the CERT Resilience Management Model. Threat modeling helps capture and evaluate risks such as STRIDE for security. The IIC Security Maturity Model brings together business, technology and process concerns, for example.

Performance predictability and *quality*. Operational efficiency and system-specific objectives drive this aspect of trustworthiness. When collaborating with partners, this is subject to an agreement. An example is cloud service (functional) availability as described and contracted upon in service-level agreements (SLAs).

These drivers are each associated with specific trustworthiness characteristics. For example, laws and regulations are direct drivers for safety, security and privacy. Risk analysis has been developed for security, resilience, and privacy. Performance predictability and quality is typically under reliability.

But the mapping is not clear-cut due to dependencies. For instance, GDPR privacy (initially a compliance concern) affects risk management. Business-performance predictability and quality directly drive reliability requirements, but business performance is also affected by other characteristics such as security.

Accordingly, many trustworthiness requirements are initially driven by some concern within one of the above categories, yet such requirements may indirectly motivate other trustworthiness requirements as support. For example, if the reliability of a service is necessary for performance predictability, one way to support this reliability is to enhance the security of the service to avoid delays or shutdowns caused by a denial of service (DOS) attack.

Failure to meet the minimum and recommended trustworthiness requirements in IIoT systems can lead to significant direct consequences such as accidents, equipment damage, personal injury, data breaches, delays and operational interruptions and lack of compliance (for some regional markets or in some industry sectors). These consequences may incur additional indirect costs such as revenue loss, fines, litigation costs, higher insurance costs, reputational damage and opportunity loss.

[1] A system may be deployed across multiple jurisdictions where the regulatory and legal compliance requirements may be different. The IT system must therefore have to have an appropriate level of jurisdictional-awareness.

While implementing trustworthiness measures has a direct financial cost their impact on operations is more difficult to assess: What are the metrics? What is the monitoring process?

These questions cannot be answered in general, as every IIoT system is different. So is the understanding of how trustworthiness characteristics interact with each other within a specific IIoT system, for example:

- Will privacy inhibit security? Will privacy instead support some aspects of security?
- How does IT security support the reliability of a service?
- How much does the safety of personnel depend on the reliability of services and components, or on the resilience of a subsystem?

It may seem counterintuitive that anything less than 'the best achievable' levels of trustworthiness could be appropriate. But an increased level of trustworthiness may result in increased costs and more cumbersome processes, so tradeoffs must be made.

Addressing specific trustworthiness characteristics to reduce risk may cause 'per user' costs to increase and cause some processes to become more cumbersome; however, risk is reduced so that the overall value of a business through the lifecycle of a system increases.

However, greater levels of trustworthiness can generate financial-related benefits in many ways, many intended to avoid costs related to failure, including:

- reduced levels of compensation payments for failed delivery of services or other failures,
- avoided fines to regulatory bodies for non-compliance with regulations, or any individual trust-related events,
- increased levels of sales, and revenues per sale, due to stronger brand image,
- reduced costs of business insurance,
- reduced risk and cost of legal defense,
- reduced costs of funding for non-determined risks and
- increased shareholder-value.

Trustworthiness objectives need to be understood and established for subsystems and the overall system. While it is important to establish global safety objectives based on regulatory requirements and company-wide policies, acting on these requires defining targets at a finer level of granularity, such as on the factory floor or for transport. The risks and therefore the metrics for worker safety will differ for the factory floor and for truck drivers, or for IT workers.

Table 1-1 highlights some examples where excessive emphasis (and investment) in certain areas of trustworthiness may affect other business considerations.

Trustworthiness characteristic	Possible consequences of excessive emphasis	Illustrative example
Security	Increased costs, loss of agility and reduced usability.	Severely limiting user's access to solution configuration parameters can result in increased support costs, loss of system agility for the user and reduced usability.
Safety	Reduced solution flexibility, more complex processes and reduced productivity.	Extensive scenario testing and certification may delay the enhancement or upgrading of an IIoT solution. Reducing expected radiation dosages for nuclear power plant workers to <0.01% of prevailing industry best practice would likely hamper many aspects of the operations of the power plant.
Reliability	Excessive capital and maintenance costs; reduced usability.	An excessive focus on reliability for a factory floor robot may result in a simplified on-board software application and heavier and more expensive hardware components. In a field environment, more reliable equipment may be heavier and so less portable or harder to carry.
Resilience	Excessive capital and maintenance costs, reduced flexibility and functionality.	Implementing a non-critical data logging service over several data centers to provide instant fail-over capabilities with <125ms latency would be an over-specification. A system designed to support geographic redundancy will often be harder to upgrade and enhance than an equivalent system deployed in a single location. Some functionality may even conflict with the need for geographic redundancy due to latency considerations.
Privacy	Unnecessarily cumbersome processes.	It would be excessive, cumbersome, and costly to apply full GDPR-compliant controls to registering user accounts and resolving technical issues that in no way reveal personal information simply as a precautionary measure.

Table 1-1: Example consequences of excessive emphasis on Trustworthiness characteristic

Once a trustworthiness analysis has been conducted, the business impact understood, the objectives established, the assessment procedures and metrics defined, there remain questions of governance and management. Who will be responsible for delivering and enforcing trustworthiness and its various characteristics throughout the system lifecycle? What evidence is being collected to gain confidence in the planned approach's ability to fulfill those required trustworthiness characteristics, and what is the process to do so?

We provide business and operation managers a deeper understanding of these questions, and an approach toward a methodology and tools to address them. We invite the IIoT practitioner to think of IIoT systems in a new way—in terms of their *properties* such as trustworthiness characteristics, not just in terms of their *functions*, which has commonly been the focus in industrial systems. This perspective is essential in managing and controlling IIoT systems.

1.3 TRUSTWORTHINESS ANALYSIS: SOME TOOLS AND PROCESS OUTLINE

1.3.1 DEFINITIONS

First, some definitions are necessary for discussing how to analyze and manage trustworthiness:

Trustworthiness criterion: an indicator for the trustworthiness of a component or subsystem in the context of an IIoT system, for a particular trustworthiness characteristic such as *safety* or *reliability*. For example, the safety of a particular machine in terms of stress injuries, or the reliability of a service in terms of its availability.

Trustworthiness interpretation (for a particular IIoT system): the set of all trustworthiness criteria established in the context of that system, for all its parts and subsystems.

Trustworthiness metric: a way to measure and gather evidence about a single trustworthiness criterion or a combination of these. A metric may be quantitative or qualitative (i.e. producing an ordinal value or score such as "fair", "poor" or "good"). An example of quantitative metric is the total number of days manufacturing personnel have been incapacitated over a year because of a safety deficiency for a machine. Depending on the consequence of that aspect of trustworthiness not meeting its objective, the frequency and preciseness of the measure may be different.

Trustworthiness assessment vector: the aggregation of trustworthiness metrics defined for a trustworthiness interpretation of a system, or of a subset of these considered of interest for controlling the system.

Trustworthiness target: an objective given to a trustworthiness metric.

Key Performance Indicators (KPIs): metrics that apply to the operations of a business, covering all aspects from plant operations and logistics, to its financial and strategic objectives. Such metrics are not trustworthiness metrics but may be affected by trustworthiness implementation considerations.

Change factor: any intentional change in the system that is designed to affect the trustworthiness Interpretation of the system or to improve metric results. This change may include process, policy, configuration, operational, and architecture changes.

Note: change factors considered here are intentional and designed to affect the trustworthiness criteria of the system and their metrics, as opposed to accidental or driven by external factors.

Trustworthiness implementation: A combination of change factors intended to meet trustworthiness targets established for the trustworthiness interpretation.

1.3.2 TRUSTWORTHINESS VECTORS

A convenient way to represent the trustworthiness interpretation of a system and its metrics is the *trustworthiness assessment vector* (TA-vector). The trustworthiness assessment vector aggregates metrics—and their values—for all the criteria of the trustworthiness interpretation as illustrated by Figure 1-3 below. Each element of the vector holds the value of a metric that is used to measure one of the trustworthiness criteria (note that there may be more than one metric for each criterion). Depending on the consequences of a characteristic deviating from the desired value an organization typically adjusts the types of measurements made and puts more effort towards gathering appropriate evidence and insights about those where the most harmful consequences could emerge. Trustworthiness criteria are identified in the diagram as Saf-C1, Sec-C2, or Rel-C1.

A related vector—the *trustworthiness target* vector (TT-vector)—defines the trustworthiness target (such as acceptable value range) for each metric in the trustworthiness assessment vector (TA-vector). The figure below illustrates an example of (abstracted) trustworthiness vectors for a system where resilience and privacy are not relevant—such as in the simplified manufacturing use case described in the next section (FOVI system).

These two vectors (TA-vector and TT-vector) are used to assess and track over time the trustworthiness of the entire system and its parts and would also reflect the level of confidence the organization has in the underlying data/evidence represented by the vector. They are also tools for evaluating the impact of trustworthiness on business and operations, and for adjusting this impact while satisfying trustworthiness targets.

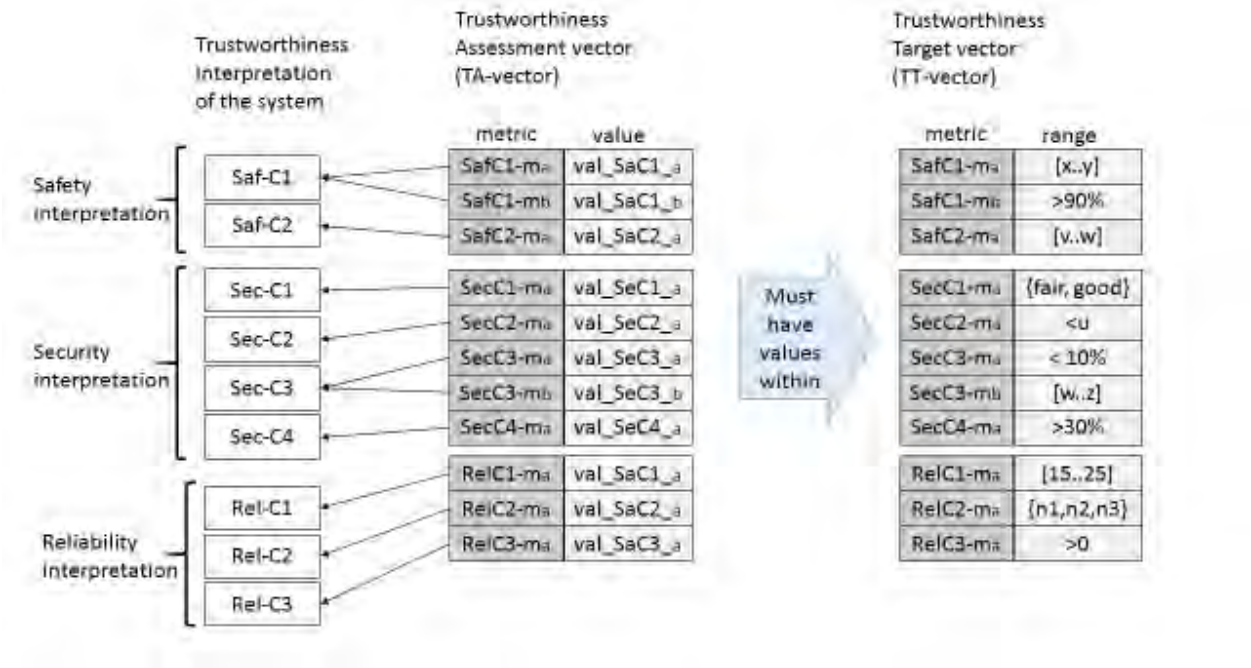


Figure 1-3: Trustworthiness assessment and target vectors

Another way to represent the trustworthiness vectors is to use a spatial representation as shown in Figure 1-4. Below is a simplified spatial representation for the trustworthiness of a system. All the metrics related to each particular characteristic (say, all the reliability metrics) of the trustworthiness interpretation of the system are summarized into a single value that locate the trustworthiness status of the system along a spatial axis (here, the reliability axis).

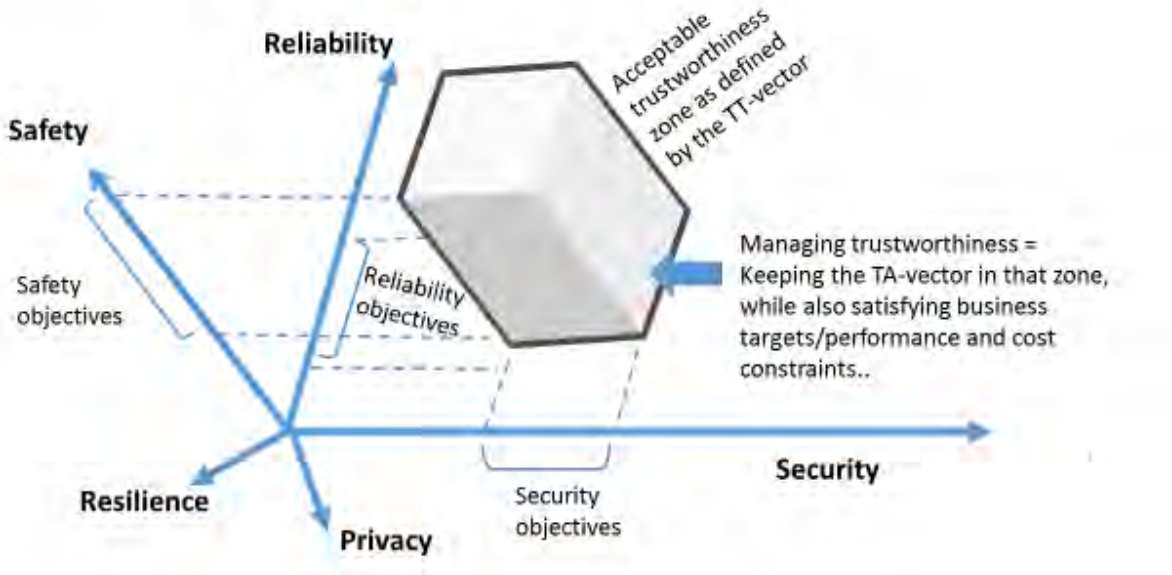


Figure 1-4: A spatial representation for managing trustworthiness

In both Figure 1-3 and Figure 1-4, the representation of trustworthiness targets (or objectives) is simplified and conflates two notions: (a) the range of *acceptable* values for a metric, with (b) the range of *desirable* values for that metric. In reality, and as discussed elsewhere in this document, both notions and quantities should be distinguished for a refined management of trustworthiness.

1.3.3 An Outline of a Trustworthiness Assessment and Management Process

The process for assessing and controlling trustworthiness is summarized below:

Phase 1: conduct a trustworthiness analysis of the system.

1. Define a *trustworthiness interpretation* of the system to include system-specific criteria for each of the trustworthiness characteristics that are relevant to this system and its components and determine the quality of supporting evidence required given the consequences of a characteristic being off from the desired level.
2. *Trustworthiness assessment vector*: Define specific metrics for each criterion of the trustworthiness interpretation. This leads to defining the TA- vector for the system. In this phase, acceptable ranges or thresholds for each metric can initially be defined. This leads to defining the trustworthiness target vector (TT-vector).
3. *Change factors*: Identify possible factors to influence the values of the trustworthiness vector or to adjust the level of confidence needed for the vector components.
4. *Business metrics*: Identify specific metrics and KPIs for business and operations performance to assess the impact of trustworthiness.
5. *Dependencies and impact*: Identify dependencies between the trustworthiness criteria of the system, including those due to trustworthiness factors. List the expected impact on business and operations. Supporting evidence about these will need be captured for confirmation or invalidation with a rigor appropriate to the harmful consequences of deviation from expected.

Phase 2: Evaluation and control:

1. Implement the metrics and monitoring tools for the trustworthiness interpretation.
2. Measure the system under operation and generate values for the trustworthiness assessment vector.
3. Adjust the factors to control the trustworthiness assessment vector as desired toward reaching acceptability values (defined by the trustworthiness target vector), while monitoring their costs and impact on business operations and performance criteria.
4. Refine priorities: In case there is conflict between (1) the targets for the trustworthiness assessment vector, (2) the cost of achieving acceptable trustworthiness with an appropriate level of confidence, and (3) the metrics and KPIs targets for business and operations performance, re-establish priorities and redefine either (1) or (3) as allowed by the business context.

Iterate between these two phases is needed until the desired outcome is reached for the trustworthiness assessment vector while satisfying the business and operations performance targets.

2 A TRUSTWORTHINESS USE CASE IN MANUFACTURING

2.1 THE FACTORY OPERATIONS VISIBILITY AND INTELLIGENCE (FOVI) SYSTEM

Fujitsu's FOVI factory system is a relatively simple IIoT system in manufacturing plants, with the goal of bringing more visibility of operations to plant managers in near-real time – a goal seen as valuable in many factories. This system has been developed as one of the testbeds in the IIC[1] as a brownfield system where IoT capability is added to existing equipment and processes.

The goal is to reduce human errors, bring more predictability to assembly and delivery, and optimize the production while ensuring sufficient safety, security and reliability (privacy and resilience are not a significant concern in this case).

In this case, many variants of a same product, here a network appliance, need to be manufactured on the same line. This is as a "high-mix" product assembly line, which poses various challenges including interruptions due to machine configuration changes or supply, human interventions and stress or optimization of diverse production lots. Changing the configuration of the machines entails delays, human errors and other inefficiencies and risks.

Figure 2-1 illustrates the system components. This architecture is at the core of one of the IIC testbeds developed and tested by Fujitsu, Inc.

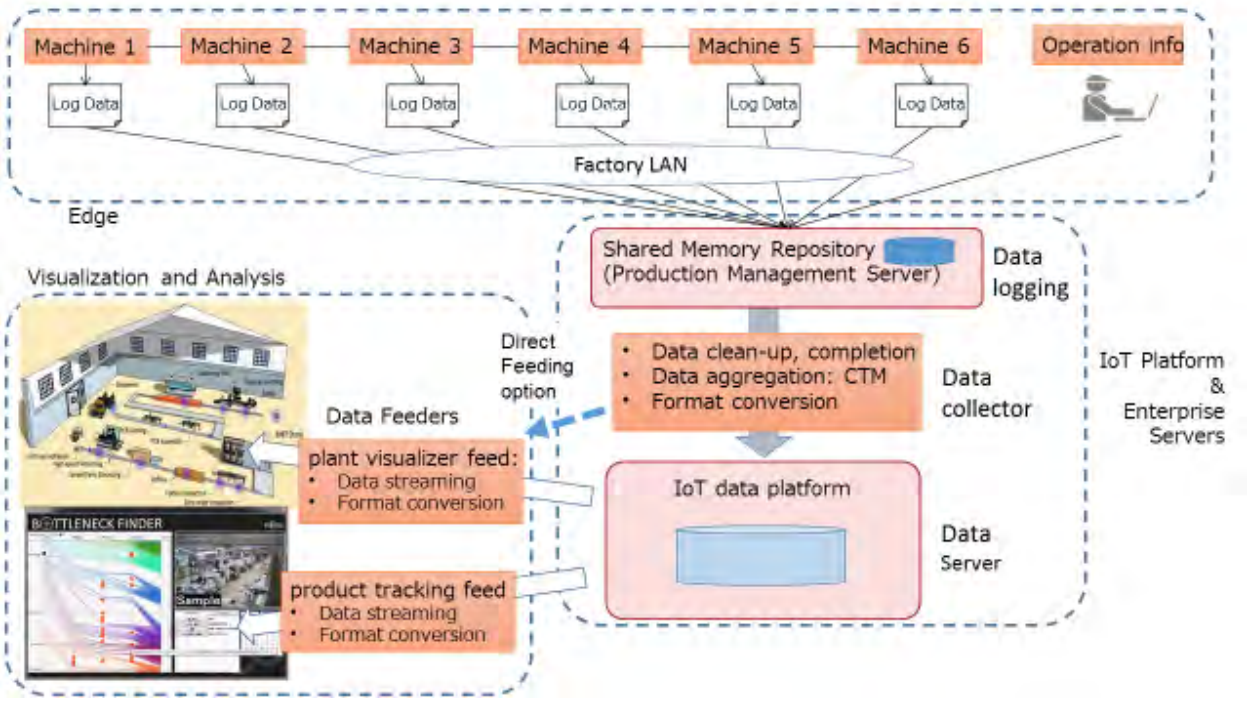


Figure 2-1: Manufacturing Operations Visualization System

[1] See [IIC-FOVI]

2.2 ILLUSTRATING SOME DEFINITIONS

We illustrate some of the definitions given in section 1.3 over the FOVI architecture.

Trustworthiness criterion: each trustworthiness characteristic (reliability, security, etc.) has a different meaning for various parts of a system. A trustworthiness criterion is a precise way to define what a trustworthiness characteristic means within the context of a particular IIoT system or part of it. Consider two subsystems of FOVI:

- an assembly line; where a reliability criterion for it is: (1) Capacity of assembly machines to be operational and available as needed.
- a cloud storage server; where a reliability criterion for it is: (2) Cloud storage service availability.

Trustworthiness criteria are often tied to a subsystem or component, although they may apply to an entire system. For example, one may talk of the safety of a particular machine, or of an entire manufacturing plant.

Trustworthiness metric: every trustworthiness criterion should have one or more associated metrics. For example:

- *rel-m1* is a metric for the reliability criterion: "*Capacity of machines on an assembly line to be operational and available as needed*". A metric for this criterion is the delay caused by unexpected machine downtime or malfunction, as percentage of the total operation time to process a production lot.
- *rel-m2* is a metric for the reliability criterion: "*Cloud storage service availability*". A common metric for this criterion is the service uptime percentage.

Trustworthiness target: a trustworthiness target translates into a threshold or a range for a quantitative metric, or into some acceptable values for a qualitative metric. For example:

- A target for *rel-m1* may be to keep this metric value under 10%, over a one-month period.
- A common target for a cloud service availability metric (*rel-m2*) is to reach at least 99%.

2.3 A TRUSTWORTHINESS ANALYSIS OF THE FOVI SYSTEM

2.3.1 DEFINING A TRUSTWORTHINESS INTERPRETATION OF THE IIOT SYSTEM

We now develop a simplified trustworthiness analysis of the FOVI system. Each relevant trustworthiness characteristic (reliability, safety etc.) is first interpreted in the particular context of this system, resulting in a set of trustworthiness criteria for the system: a *trustworthiness interpretation*. Objectives are associated with each trustworthiness criterion, which must translate into targets for specific metrics. Some example of *change factors* that influence these criteria, are identified. These change factors—if proven effective—also give clues on the cost of achieving trustworthiness targets, provide hints on their business impact and insights on how trustworthiness criteria depend on each other.

For an easier reading, each trustworthiness criterion is also listed along with its related metric(s) and identified change factor(s):

Safety: for this IIoT solution and its context, we identify three safety criteria, along with change factors either known or suspected to act positively on these criteria (measurements and metrics will confirm or invalidate these):

- *Saf-C1: Reduce accidents*. (metric: number of accidents/quarter)
  Change factors: reduce itineraries and trips on premises, reduce time pressure for human moves on the floor-plan or across campus.

- *Saf-C2: Avoid human stress injuries using specific equipment*. (metric: number of injuries/quarter)
  Change factors: detect risk-prone situations or conditions and reduce the frequency of machine manipulations on factory floor by personnel. For example: when an assembly chain operates for a high-mix production lot, provide enough time to human personnel to operate machine configuration changes.

Note: change factors have their own metrics (not mentioned here) to evaluate their effectiveness in contributing to a trustworthiness criterion, as well as for detecting their impact on others.

Reliability: for this IIoT solution, we identify five reliability criteria and some of their immediate Change factors:

- Rel-C1: Visualization consoles operational and available. (metric: downtime events / month)
  Change factors: ensure proper monitoring and maintenance, timely detection of failures, quick notification to and intervention from IT expert on premises through effective and available human machine interfaces.
- Rel-C2: Data platform operational and available 24/7 (that supplies data-stream to the visual systems on the floor plan). (metric: uptime percentage per month)
  Change factors: ensure proper IT oversight and ownership, keep track of downtimes and minimize their duration as objective.
- Rel-C3: Cloud infrastructure Service availability (3rd party, data archiving for future analytics). (metric: uptime percentage per month)
  Change factors: proper SLA terms, monitoring of service uptime percentage metric.
- Rel-C4: Assembly machines operational and available as needed. (metrics: downtime events per month, reduced functionality total operation time per month)
  Change factors: predictive or preventive maintenance, detection of particular conditions that are conducive to failures (such as too high speed for an assembly chain, patterns of brief errors before a critical outage, or incorrect machine configuration - human error).
- Rel-C5: Reliability of the visualization function (consistently provide timely and accurate visibility of operations to operation managers on the factory floor so they can make

correct decisions—both for real-time and replay) (metric: number of times the average latency of visualized signal is larger than five seconds per day).

Change factors: timely detection of data-stream backlog and of local network outage, ensuring that sensors are operational and signals can be validated (e.g. redundant). Periodic verification of consistency between the data log on edge gateway and the log on the data platform.

- Rel-C6: Reliability of the analytics function (metrics: average latency of an alert from its input events per month).

  Change factors: timely upload of data to the cloud data service, checking quality and relevance.

Security: several aspects and objectives of security are identified for this IIoT solution. There are both IT and OT security aspects. IT security criteria across the system are illustrated in Figure 2-2.

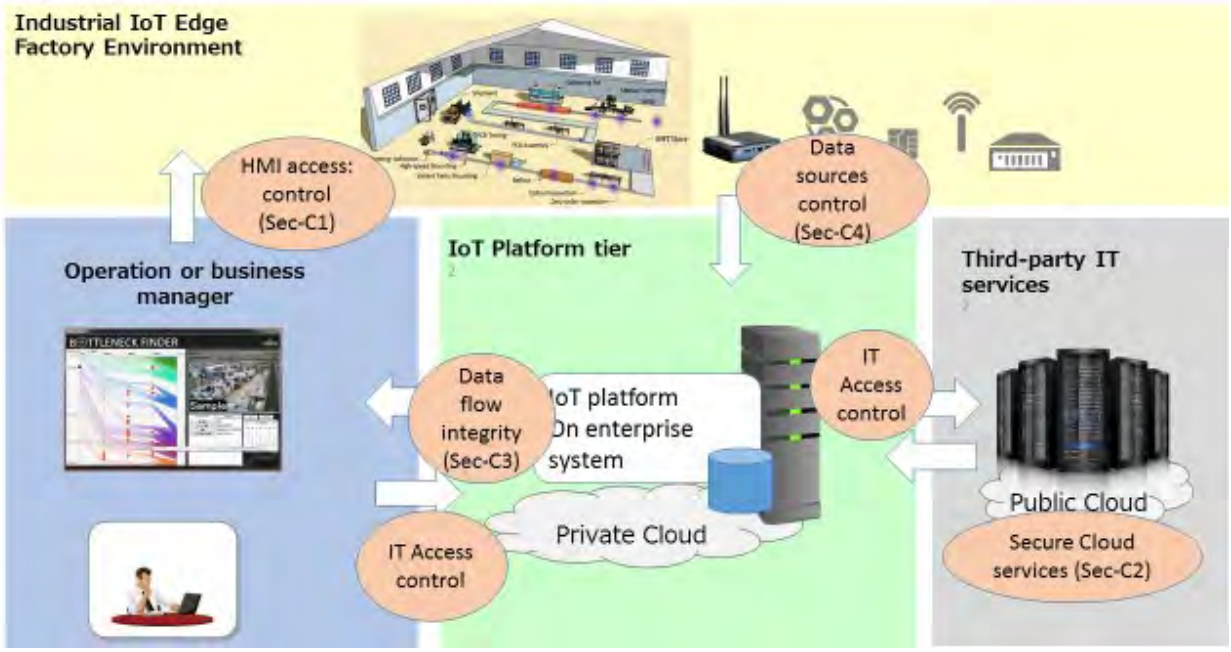


Figure 2-2: Security Overview of Manufacturing platform

Some of the main security criteria to be assessed are:

- Sec-C1: *Access control to management consoles and other human machine interfaces (HMI)*. (metrics: number of unauthorized access including illicit reuse of security badges/month).

  Change factors: proper authentication of personnel, role-based access levels, access event logging, proper monitoring of the location of mobile consoles (tablets) on floor plan.
- Sec-C2: *Secure Cloud environment for services, for logged data*. (metrics: usual security metrics for such providers, evaluation of the security by an external security consulting company and comparing with industrial standard (benchmarking.))

Change factors: VPN and secure access for user and management operations, prevention and detection of DoS/DDoS attacks on provider side.

- Sec-C3: *Integrity of data flows from assembly machines to public cloud*. (metrics: volume of data loss from cache overflow and delays)
  Change factors: SSL encryption, IPsec (SSLVPN).
- Sec-C4*: Control of authorized data sources*. (metrics: number of delayed invalidations of unauthorized device registrations as a proportion of all unauthorized device registrations, volume of unauthorized data ingestions)
  Change factors: ensuring only registered devices can upload to data platform, verification that registered devices are uploading data only when expected to.

2.3.2 IDENTIFYING TRUSTWORTHINESS INTERACTIONS AND IMPACT ON BUSINESS

When implementing the above, the effect of trustworthiness criteria and their change factors on business and operational performance must also be evaluated. For this, common indicators for operations and business (KPIs and other system-specific metrics) must be selected, that are likely to be affected (positively or negatively) by trustworthiness targets and factors. For example:

- (PI-1) unplanned delays and idle time (e.g. due to manual errors on equipment, or configuration changes),
- (PI-2) maintenance costs,
- (PI-3) assembly chain productivity and
- (PI-4) device on/off boarding and management overhead.

Examples of impact are:

- *Reliability* (Rel-C4) (*Assembly machines operational and available as needed)* impacts positively: (PI-1) unplanned delays and idle time, (PI-3) assembly chain productivity.
- *Security* (Sec-C4) (*Control of authorized data sources*) impacts negatively: (PI-4) device on/off boarding and management overhead.
- *Safety* (Saf-C2) (*Avoid stress injuries on specific equipment)* impacts negatively: (PI-3) assembly chain productivity (by slowing down a high-mix assembly chain), but impacts positively: (PI-1) unplanned delays and idle time (by reducing manual errors and wrong machine configurations.)

In addition to affecting business and operations, trustworthiness criteria also affect each other. An interdependency analysis will assess how trustworthiness criteria affect each other, in a positive or negative way.

Example where a security criterion reinforces a reliability criterion:

- Security (*Sec-C1*) (*Access control to management consoles, HMI*) reinforces Reliability (*Rel-C4*) (*assembly machines operational and available as needed),* by ensuring that only qualified personnel are manipulating them, thus avoiding misuse and incorrect configurations.

Example where a change factor initially designed to reinforce a safety criterion (Saf-C2), also indirectly affects operational and business performance in both positive and negative ways:

- Saf-C2 Change factor (slowing down of assembly line) negatively impacts productivity (operational performance) and negatively affect revenues (business performance)
- Safety (Saf-C2) in turn positively impacts predictability of production time and deliveries (operational performance) and positively reduces health care expenses & insurance rates (business performance)

The interdependencies (negative and positive) that are identified between criteria, along with their operational and business impact can be illustrated graphically. The diagram below illustrates trustworthiness criteria analysis for this particular FOVI system as well as their impact on the operations and the business bottom-line. Figure 2-3 shows an example scenario where a change factor of assembly line is slowed down is used in order to reduce personnel stress injuries:

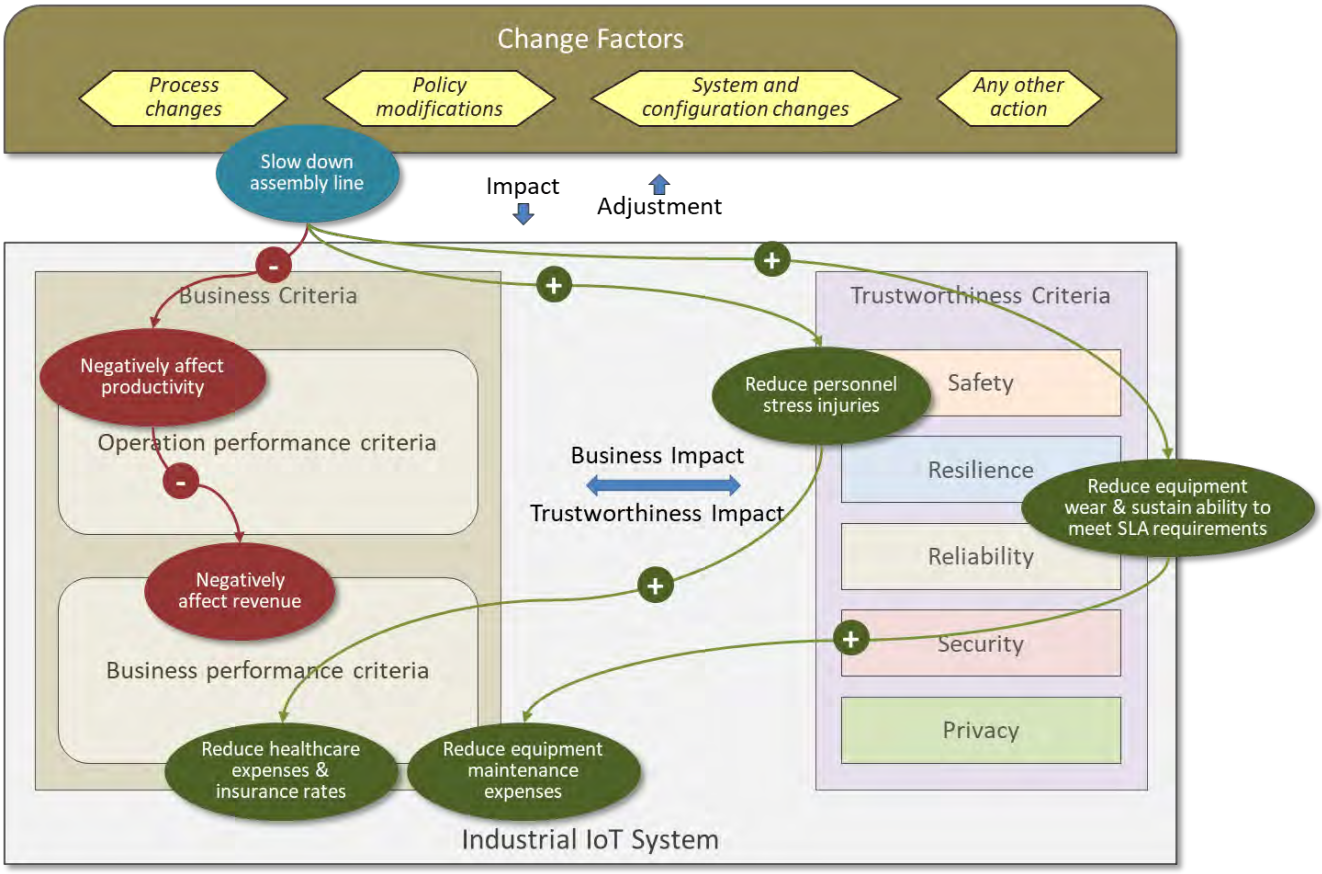


Figure 2-3: Example Trustworthiness Analysis

In this example, slowing down the assembly line will also negatively affect the operation and business performance criteria of the IoT system (see previous impact example.) Perhaps a different change factor can be used to improve safety without negatively affecting performance?

A more precise assessment of trustworthiness, criteria dependencies and impact will rely on metrics the definition of which is beyond the scope of this paper.

As a summary, the above trustworthiness analysis involved the following steps:

- Define the trustworthiness interpretation for the system, and identify some change factors.
- Define metrics and initial targets for the trustworthiness criteria within that interpretation.

Identify dependencies between trustworthiness criteria, and assess the impact these criteria and their change factors will have on business and operational performance (to be monitored once implemented.)

3 MANAGING TRUSTWORTHINESS OBJECTIVES

3.1 AN INTEGRATED APPROACH: DEFINING OBJECTIVES AND THEIR ASSESSMENT

While trustworthiness interpretations vary from one system to the other, some variations are regional-specific or industry-specific. Also, the importance of the different trustworthiness characteristics will vary according to industry vertical and regional interpretations. For example, there are different expectations of privacy in Europe from in the United States. Business strategy and individual targets will depend on regulations, laws and the industry. Moreover, the level of confidence in the delivery of the different trustworthiness characteristics will vary with the consequences of breaching the trustworthiness objectives given a specific IIoT systems context. If lives or the environment will be harmed there are often higher degrees of rigor and completeness required so that the likelihood of undesirable outcomes can be assessed more accurately, and better mitigated.

Managers want a synthetic view of trustworthiness in their organization, and of the expected levels for each characteristic. The following “spider diagram” in Figure 3-1 is an abstract representation of the trustworthiness vector and its targets, that rely on aggregated metrics for each of the five major trustworthiness characteristics.

The diagram collapses each one of the trustworthiness characteristics and all criteria and metrics, into an aggregated metric along each of the five dimensions. Although this simplification ignores the details as reported by the trustworthiness assessment and target vectors, it gives a good summary of the overall trustworthiness status of a system or organization compared with the minimum baseline requirements and relative to the desired state. It typically reports three levels of trustworthiness:

- Current State: Actual state as it exists now
- Minimum State: Non-negotiable minimum level as mandated by legal, regulatory and other types of requirements
- Target State: Level of trustworthiness to achieve based on corporate vision, desired return on investment (ROI) and risk considerations, etc.

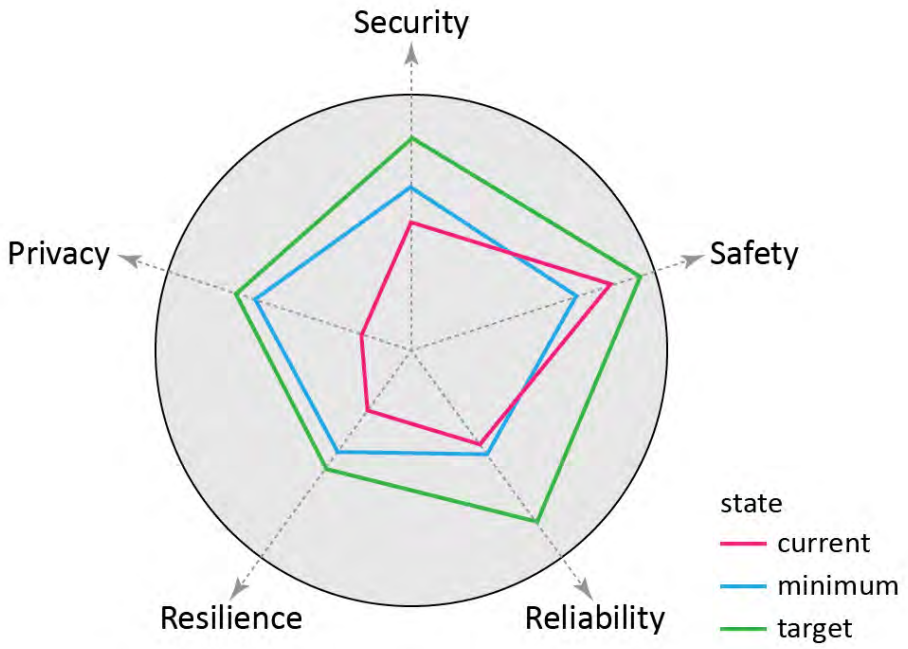


Figure 3-1: Trustworthiness Spider Diagram

A trustworthiness spider diagram can be defined for an entire system so as to give an overall perspective of where one stands with respect to objectives and requirements. It can then be broken down in more detailed spider diagrams for various subsystems, or for various markets (where the minimum state and target state may vary).

In Figure 3-1, the current state for safety already meets the Minimum State required. The diagram also shows that the Current States of the other trustworthiness characteristics (on security, privacy, resilience and reliability) are below the required Minimum State levels, which means that work is required on them to meet the Minimum State requirements.

Each of the five characteristics of the trustworthiness of the IoT solution has its own "journey" to change from the current state to the target state as represented in Figure 3-2:

- Each of these "journeys" will have Current State → Minimum State and Minimum State → Target State segments
- Each segment will require its own technical and financial justification

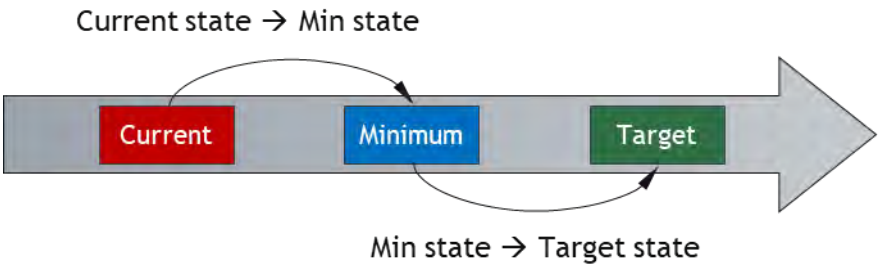


Figure 3-2: Trustworthiness States

3.2 MEASURING TRUSTWORTHINESS

Measuring the current state of trustworthiness characteristics is necessary to assess and manage the implementation of trustworthiness, since the feedback is needed to make appropriate changes (change factors applying to the system controls, processes and architecture).

Consider a cloud storage service used to collect, contextualize and archive edge data. Such a service is often provided externally, such as through a contract with a provider as expressed by a service level agreement (SLA). The SLAs of cloud-service providers involve service measurement metrics. One of the most common metrics is service availability, often defined as the percentage of time the service is up over a billing cycle (service-uptime percentage). Service availability falls under the reliability characteristic of trustworthiness (as this is *functional* availability, not *data* availability often considered as part of security). Thus, service uptime percentage is one way to measure the reliability of a service.

Metrics are more than just performance and assessment indicators. When assigned a target—for example, 98% service availability, or zero injuries, or 10 seconds or less to diagnose a DNS attack—a metric becomes an operational management tool.

There could be other metrics for the reliability of this storage service. For example, metrics that capture the ability of the service to handle large volumes of data without performance degradation, or also measuring various types of service interruptions and their properties (average duration, variance, triggering factors), or yet the responsiveness of its administrators in case of problems. A basket of metrics such as represented in a trustworthiness assessment vector (TA-vector) will give a well-rounded assessment of the trustworthiness interpretation of an IIoT system, its services or components.

There are different reasons for using metrics. For example, metrics on our above storage service reliability will help to:

- clarify the expected service level with the service providers, as defined by an SLA,
- evaluate how well the service performs and provide feedback to providers. It also supports assessment of penalties in case of failure to fulfill SLA or SLO targets, and ability to compare providers and
- understand the nature of shortcomings and failures of a component so that these can be mitigated by the system, or negotiated with the provider. For example, in a system where edge devices periodically invoke the cloud storage service directly every minute while having the capacity to handle a backlog of only up to ten minutes of data stream, it is important to prevent data storage service downtimes of more than 10 min. An adequate metric measuring the duration of downtimes—not just the uptime average—will be the basis for negotiating adequate service quality to minimize data loss.

Now, our cloud storage service is just a component of the edge data collection subsystem of an IIoT system—the part of the system that collects and channels edge data up to a processing platform. The data collection subsystem has other reliability metrics for its components:

- ability of gateways to handle and aggregate bursts of edge data with minimal loss,

- device robustness under stressful conditions and
- network latency and ability to handle heavy volumes.

Metrics associated with these components will provide a good grasp on their behaviors. This is in turn necessary to control the trustworthiness of entire subsystems, here assessed by metrics such as (staying here in the reliability characteristic):

- variability of end-to-end data latency from source to storage. Keeping such variability within limits will clearly depend on many factors: device caching and configuration settings, network latency, storage uptime and
- elapsed time between detection of stress conditions and dynamic scalability operations to restore overall performance expectations.

Thus, trustworthiness metrics depend on each other and so will their targets.

Finally, business operations and performance have their own metrics. Assessing the cost or benefits of trustworthiness measures will ultimately show in these business metrics—but a precise assessment of the correlation between trustworthiness characteristics and business, will require correlating metric results in both areas. For example: Did the reduction in data loss contribute to better, more timely decisions and actions? And how did these better decisions contribute to the overall bottom-line? Or: How did a more reliable monitoring of an assembly chain contribute to more accurate product shipment time or prediction, and how did this in turn contribute to avoid delay penalties or to improve transport logistics?

Metrics will be key to managing the trustworthiness of a system, including an evaluation of how the trustworthiness of various components contribute to it. Such components also include third-party providers, in which case metrics are a basis for the contracting aspect. Finally, metrics will be necessary to assess the costs and benefits of trustworthiness with respect to operations and business bottom-line.

3.3 Middle-Out Management Strategy for Trustworthiness

We have described a top down perspective on trustworthiness. Managing and controlling a trustworthy system may require combination of bottom-up and top-down elements—a middle-out approach as illustrated in Figure 3-3:

Because of the cross-functional nature of the work and the need to coordinate both top-down and bottom-up activities it is useful to create a cross-functional team to manage trustworthiness activities in the organization, a *trustworthiness steering committee*. This committee can steer the activities related to trustworthiness, the *trustworthiness program*. The steering committee will need to consider both bottom-up and top-down approaches:

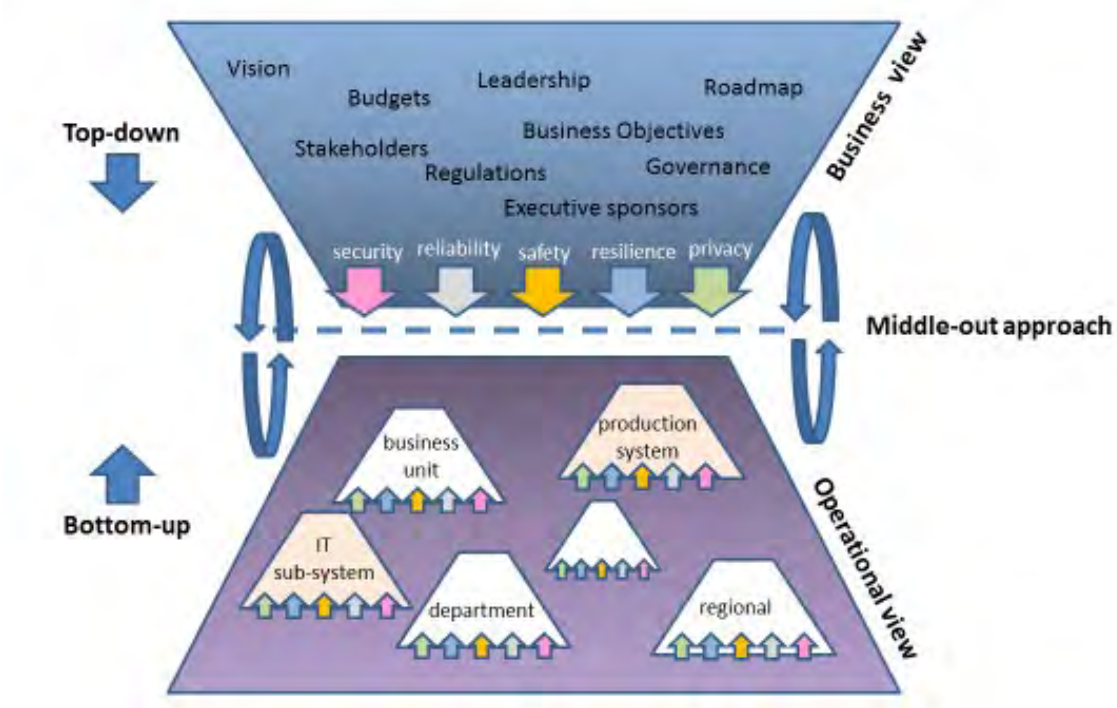


Figure 3-3: Trustworthiness Approaches

Bottom-up:

- The *stakeholders*: Operational, production and regional managers, with their partners and customers.
- The *drivers*: There are both operational "local" drivers, and "top-down" drivers. Operational drivers: safety and continuity of local operations, risk mitigation with respect to local production and operational objectives, reliability of equipment and services involved in local operations, resilience of operational systems with respect to known risks.
- The *challenges*: Fragmented objectives and governance across departments, business units. Harmonizing and integrating the separate trustworthiness objectives and measures so that they align with and contribute to corporate and regulatory drivers. Understanding the interdependencies of various trustworthiness objectives.
- *How to implement*: empower operational personnel and managers to establish trustworthiness objectives and metrics for local operations (unit, regional, department). Establish a trustworthiness council across business units to address fragmentation and interdependency challenges. The different stakeholders responsible for the various dimensions) of trustworthiness define the current states and identify minimum states of their respective domains. They also identify the requirements to move from the current to the minimum states, including technical roadmaps, budget requirements and resource requirements.

Top-down:

- The *stakeholders*: corporate executives, business managers. Trustworthiness should be assigned a corporate sponsor who can mandate and track its realization within the organization (including timelines).
- The *drivers*: regulatory compliance, global market requirements, corporate-wide policies, industry standard and practices.
- The *challenges*: Trustworthiness crosses group and departmental boundaries within the organization. Organization-wide objectives must be translated in a consistent way across departments or units.
- *How to implement*: Assign trustworthiness to a corporate sponsor who can mandate its realization within the organization (including timelines). This sponsor may, for strategic or competitive reasons, mandate targets for trustworthiness that can exceed the Minimum State level. A steering committee comprising representatives from these groups and departments is tasked with the responsibility of transforming and translating the trustworthiness mandate into specific requirements for each group.

3.4 ESTABLISHING GOVERNANCE STRUCTURE AND POLICIES

Trustworthiness poses a governance challenge besides its measurement. The different characteristics of trustworthiness have typically been managed by different communities. Each community uses their own vocabulary, techniques, assessment practices, standards and regulations. For example, the safety community is generally distinct from the security community. They need to be brought together through a shared understanding of how trustworthiness characteristics are managed affects business value.

Trustworthiness governance is distributed in two ways:

- The security aspect, the resilience aspect, safety aspect, etc. even if governed organization-wide are likely under different management and responsibilities—e.g. a chief security officer, a chief safety officer.
- Each business unit, region, or department has its own requirements and local objectives for trustworthiness.

Because trustworthiness is a cross-functional and cross-departmental responsibility we must create an organization-wide trustworthiness program managed by a steering committee. This committee must include, as a minimum, senior members of the groups that represent the different characteristics of trustworthiness. The steering committee (and the program itself) must have a leader who is mandated and empowered by the corporate sponsor to steer the program and achieve its objectives.

An effective project management method to chart the responsibilities of the various groups involved in the IIoT trustworthiness program is the RACI Matrix (R=responsible, A=accountable, C=consulted, I=informed). In this matrix, the rows list the individual tasks and functions (varying degrees of granularity), and the columns show the individual parties involved in each of the tasks listed. Each cell of the matrix is filled with the actual role of the party for that task:

- *Responsible*: Parties who do the work to complete the specific task. Each row must have at least one party assigned an R.
- *Accountable*: The single party who is ultimately answerable for the correct and thorough completion of the task, and the one who delegates the work to those responsible.
- *Consulted*: The parties whose opinions are sought (example SMEs) for the execution of the tasks.
- *Informed*: Parties who are kept up-to-date on progress and/or completion of the task.

A variant of RACI called RASCI and introduces a variant to the R role: the S or supporting role.

When using RACI in the context of trustworthiness governance, each characteristic is associated with a set of tasks. The Table 3-1 is a high-level example of RACI model for trustworthiness, where each row aggregates all tasks supportive of a particular trustworthiness characteristic:

Tasks	Corporate	Finance	OT	IT	Security	Operation	Business	Legal
Security	C	C	R	R	RA	R	I	C
Safety	C	C	R	I	C	RA	C	I
Reliability	I	C	RA	I	C	R	I	I
Resilience	I	C	RA	I	C	R	C	I
Privacy	I	C	R	R	R	C	C	RA

Table 3-1: RACI Matrix

Notes about the above RACI Matrix:

- The tasks listed, the parties identified, and the roles assigned to them must be expanded to reflect the unique requirements and characteristics of the specific use case and market.
- The table shows that each task has one and only one party accountable, one or multiple parties responsible, and any number of parties consulted or informed.
- The key party for ensuring the expected level of trustworthiness is the accountable party. This party varies depending on how critical its role is for each trustworthiness characteristic. It should be invested with proper authority to supervise the assessment and implementation of trustworthiness.

3.5 THE CONTINUING IOT TRUSTWORTHINESS JOURNEY

Concerns about establishing confidence that an IIoT system meets the trustworthiness requirements must be addressed throughout the lifecycle of the system. This means that industrial IoT trustworthiness is not a project with a finite start and a finite end. It is a journey that must be powered by an established program.

Figure 3-4 shows an example of a trustworthiness journey. It shows how the actual level of trustworthiness (red line) must be increased rapidly to reach the level of required trustworthiness (blue line) and then maintained above that minimum level.

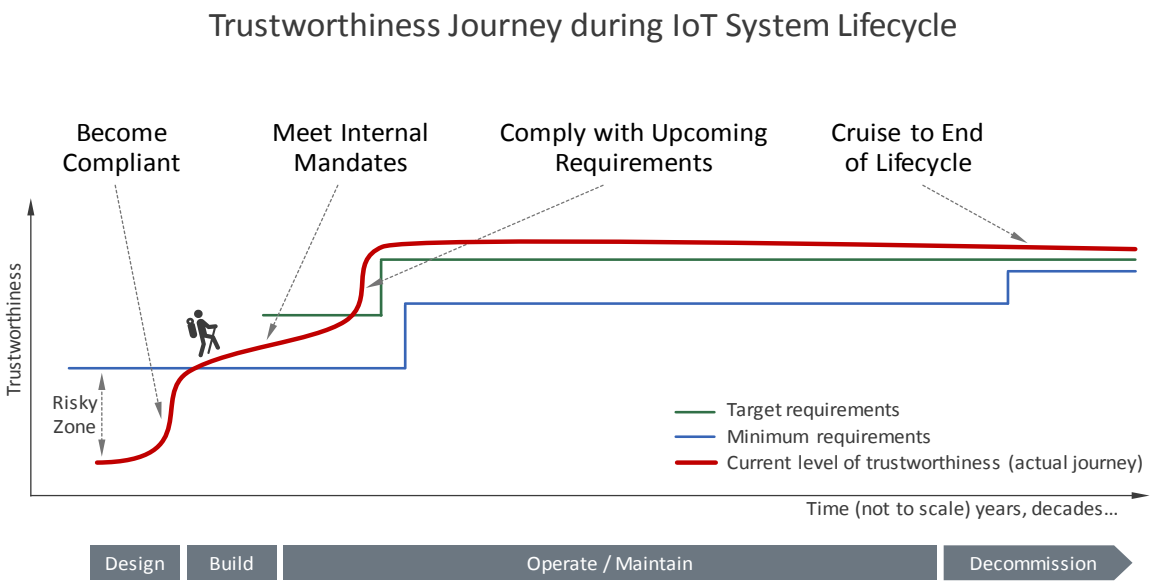


Figure 3-4: Trustworthiness Over Time

For the trustworthiness Journey to be managed to be responsive to the needs of the organization, it must be governed via a trustworthiness program: a framework for organizing, directing, implementing and maintaining trustworthiness of a system throughout its lifecycle.

The program must have a formalized governance structure with the stakeholders and their roles clearly identified. It must also have a corporate sponsor who sets the trustworthiness mandate and objectives.

New information must be taken into account. When inputs change, the trustworthiness analyses must be updated. Trustworthiness management practices need to include a method for notifying downstream consumers of data when the trustworthiness associated with a data feed has changed, so that downstream users can appropriately re-work their trustworthiness solutions. For instance, if the original source for a weather information data feed changes, then users of that data feed should re-assess their trustworthiness positions given the new SLAs (and trustworthiness) associated with the new data feed.

4 Best Practices for Managing Trustworthiness

A best practice approach comprises four phases:

- baselining, which includes the gathering of basic information to input into the process,
- analysis, during which a management team can assess how trustworthiness-related events can potentially impact a business, and real options for addressing risks,
- implementation, during which the management team will establish trustworthiness targets and establish appropriate governance and
- iterate and maintain, is key since trustworthiness is not a static concept, and trustworthiness much be managed in the context of an overall changing landscape.

The overall approach is illustrated in Figure 4-1, below, and is discussed in more detail in the following subsections.

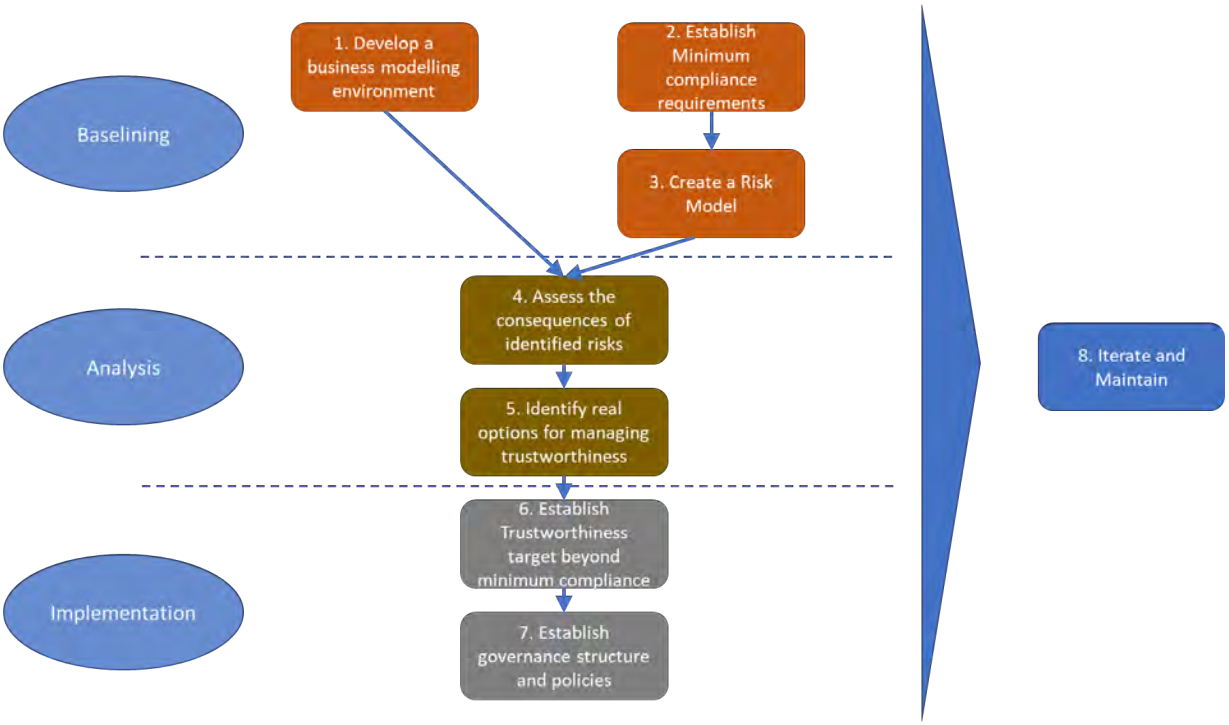


Figure 4-1: Trustworthiness Process flow

4.1 BASELINING

In the baselining phase of the best practice approach to managing trustworthiness, three stages should be undertaken in a closely coordinated way so that findings of any individual stage can inform all aspects of work in the baselining phase. The three stages included within the baselining phase are:

- develop a business modeling environment,
- establish minimum compliance requirements and
- establish trustworthiness drivers, including a risk model.

Each subsection below elaborates these stages.

4.1.1 DEVELOP A BUSINESS MODELING ENVIRONMENT

The first step in establishing trustworthiness management policies and an overall risk management approach is to develop a business case for the IIoT system in question that reflects the context and consequences of the IIoT system not operating within the expected trustworthiness characteristics. The business case should encompass all aspects of the business of which the candidate IIoT solution is a part. This is because increases (or reductions) in trustworthiness can have knock-on effects to a wider business, for instance via channels such as brand image or affecting shareholder value via the risk assessment of a business. Moreover, if the failure of a system might affect customer churn, or revenues per customer, across any of the

products or services offered by a company, then this needs to be taken into account. Further discussion of the construction of IIoT business cases can be found in the Business Strategy and Innovation Framework[1], Section 5.3 ("Evaluation").

The business case should also be sufficiently detailed to be able to reflect the knock-on financial and business consequences of any identified potential revision or adjustment to any trustworthiness dimension, and the financial and business implications of the current state of trustworthiness. In practice, this level of granularity and sophistication of analysis will only be achieved over time, and as analyses of a solution's trustworthiness environment evolves:

The construction of a business case is an iterative process that accounts for the evolving knowledge base around trustworthiness. The business model should also be capable of running multiple scenarios so that managers are able to understand the consequences of multiple combinations of potential threat events.

The level of sophistication of the model (and effort invested in developing it) should be calibrated to account for the overall scale of a business and the risks under analysis.

4.1.2 Establish Minimum Compliance Requirements

All IIoT solutions are subject to compliance requirements and it is likely that all trustworthiness domains will be affected by such requirements. The compliance requirements relevant to a given IIoT solution are driven by trade and governmental regulations and are therefore domain-specific.

Here are some mainstream examples of compliance requirements:

- General Data Protection Regulations (GDPR),
- Workplace health and safety regulations,
- Maintenance of records for auditing purposes and
- The Health Insurance Portability and Accountability Act (HIPAA).

The overall approach to establishing the long list of compliance requirements relevant to any specific IIoT solution is to engage with and seek input from a range of domain experts, including managers with relevant experience, domain-competent lawyers and any relevant regulators.

Clearly any IIoT system should be designed and deployed in a way that satisfies any relevant regulations or compliance requirements. Accordingly, for the purposes of subsequent analyses of risks pertaining to a candidate IIoT system, it should be assumed that the system meets any identified compliance requirements.

4.1.3 Establish Trustworthiness Drivers, Including a Risk Model

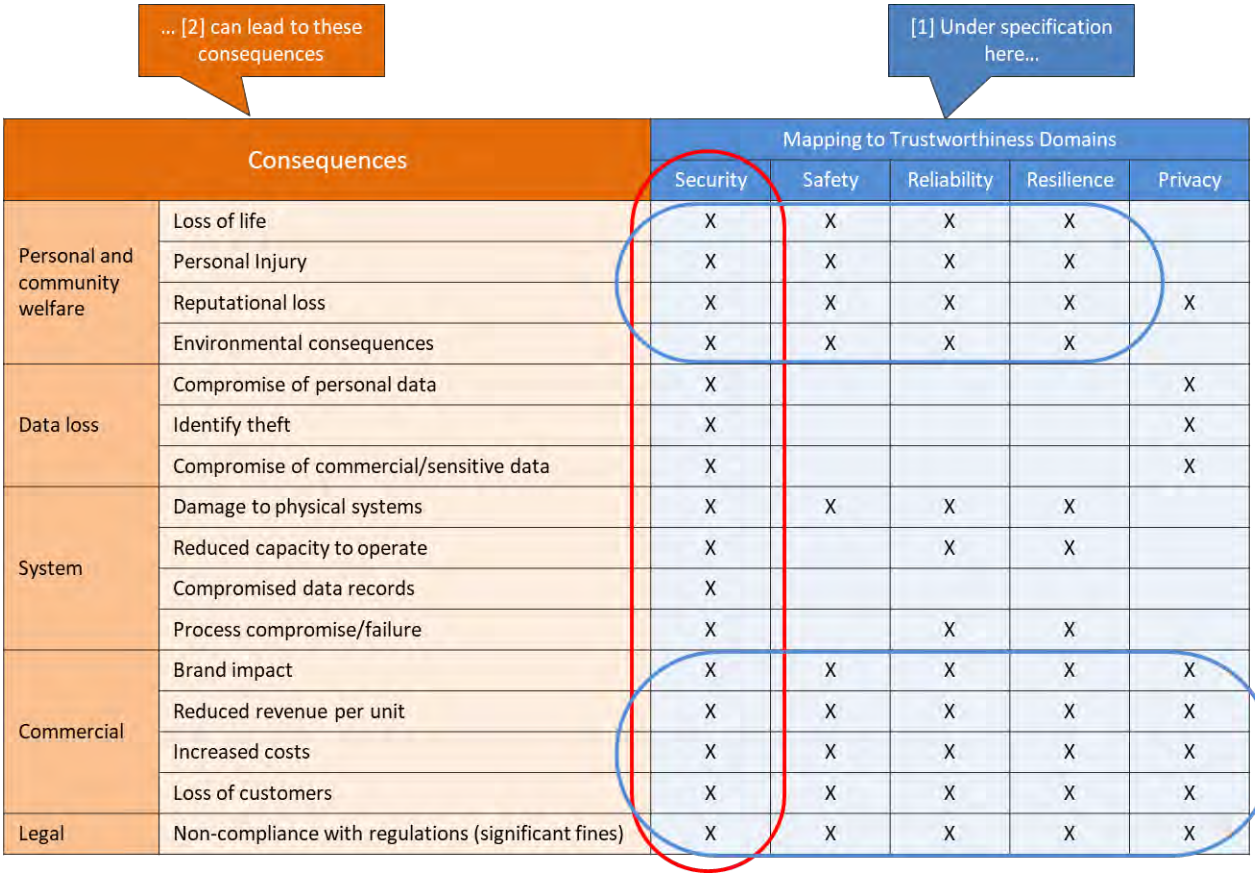
The third step to take when seeking to manage the trustworthiness of specific IIoT system is to generate a list of the potential risks associated with the system in question. Such risks will be specific events that can potentially affect any of the main trustworthiness domains.

[1] See [IIC-BSIF2016]

The risk model extends to defining the potential consequences of not meeting expected levels of performance quality and predictability during operations. Such consequence may typically be mitigated by investment in reliability and resilience, but may be affected by other trustworthiness characteristics.

Finally, a risk model is defined for the entire IIoT system and its governing entity (an enterprise, an organization). There may be models at finer granularity—e.g. a department, a production unit. (See the "middle-out management strategy" in previous section.) Risks, and consequently trustworthiness requirements, may be specific to a sub-system or a business unit.

Figure 4-2 below illustrates a framework that managers can use to help identify consequences of events that may affect any given trustworthiness domain.[1] The 'X' marks signify a relationship between a specific trustworthiness domain and a range of potential consequential events. The values provided in that matrix are only illustrative and will depend on the IIoT system of interest.



Consequences		Mapping to Trustworthiness Domains				
		Security	Safety	Reliability	Resilience	Privacy
Personal and community welfare	Loss of life	X	X	X	X	
	Personal Injury	X	X	X	X	
	Reputational loss	X	X	X	X	X
	Environmental consequences	X	X	X	X	
Data loss	Compromise of personal data	X				X
	Identify theft	X				X
	Compromise of commercial/sensitive data	X				X
System	Damage to physical systems	X	X	X	X	
	Reduced capacity to operate	X		X	X	
	Compromised data records	X				
	Process compromise/failure	X		X	X	
Commercial	Brand impact	X	X	X	X	X
	Reduced revenue per unit	X	X	X	X	X
	Increased costs	X	X	X	X	X
	Loss of customers	X	X	X	X	X
Legal	Non-compliance with regulations (significant fines)	X	X	X	X	X

Figure 4-2: Mapping Consequences to Trustworthiness Characteristics

Within this risk-identification phase particular consideration should be given to the availability and integrity of data sourced from any third parties, including any SLAs associated with those data sources and also wider considerations around the business continuity of data providers.

[1] Although clearly there are many approaches that managers could adopt to generating a list of potential risks associated with a candidate IIoT system, and this framework is just an illustration of a single method.

Such counterparty risks can be key to the overall risk analysis of an IIoT system, and are often also readily changeable elements that can potentially improve an overall trustworthiness profile.

4.2 RISK ANALYSIS

Risk avoidance and mitigation is a main driver for trustworthiness as previously stated in the first part. The second phase of the best practice approach to managing trustworthiness is the 'risk analysis' phase. It comprises two further stages, as follows:

- assess the consequences of identified risks and
- assessing maturity and identify real options for managing risk

Each stage describes best practices and tools for managing risk in the trustworthiness assessment and management cycle.

4.2.1 ASSESS THE CONSEQUENCES OF IDENTIFIED RISKS

Having undertaken the baselining phase of the best practice approach to managing trustworthiness, the next step is to assess the probability of each of the potential trustworthiness events identified earlier, using the following grades:

- none, or negligible probability,
- very low probability,
- significant probability,
- significant likelihood and
- near certainty.

These grades should be assessed against an appropriate timeline. For personal injuries, this might be an assessment of the probability of an event taking place within a one-year timeframe, while for errors of data integrity it may be more appropriate to measure the probability of an event taking place within a one-day timeframe.

For each identified potential trustworthiness event, there should be a high-level assessment of the expected consequences of any event, since this will to some extent affect the calibration of probability grades for each potential trustworthiness event. For personal injury, a 1% chance of an 'event' within a year could be classed as a 'significant probability', while a 1% chance of an erroneous data reading within a year in the context of many IIoT systems is of 'negligible probability'.

The high-level assessment of the expected consequences of any trustworthiness event should account for a range of extenuating and mitigating factors that may be relevant to the specific trustworthiness event in question. Such factors include:

Scale of breach: Is the breach very limited in scale, compared to the overall solution or is it complete and fundamental?

Reversibility: Can the breach be reversed with a definite cost, or will it become an ongoing and open-ended exposure.

Downstream effects: Is there potential for a trustworthiness event to impact other IIoT solutions (or real-world events) that are potentially influenced by any outputs of the IIoT system in question?

Potential criticality: Do specific trustworthiness events have different levels of impact for different user groups?[1]

Secondly, it is necessary to estimate the financial impact of any potential trustworthiness breach (and the costs of any trustworthiness measures) using the business model developed in the first stage. The expected business risk associated with any specific trustworthiness event is simply a product of the probability of that event and the impact of that event, and so these are the two critical inputs to any approach to optimizing trustworthiness within an IIoT system.

Risk is not always measured financially. For example, governments, not-for-profits, and other public institutions may attempt to quantify the impact on their mission and use a measure for delivering the mission. An example is the amount of food provided for those in need and how a trustworthiness failure could reduce such deliveries from a norm.

For dire consequences to specific trustworthiness events there is often a higher requisite level of confidence in the ability of the counter-measures and mechanisms used to mitigate or avoid the events. Having a body of evidence that supports the choices is a normal approach followed in industries that have high-impact consequences to trustworthiness events and are shaped by their stakeholders.

On a point of methodology, if the potential real-world consequences of a specific trustworthiness event are not clear, then either the 'worst case' scenario consequences must be assumed, or the definition of the trustworthiness event in question must be refined and redefined as multiple potential events, each with a clear expected real-world consequence.

4.2.2 Assess Maturity in Trustworthiness and Identify Options for Managing Risk

Managers of a company have choices in regards to the investments in tangible assets. These options have consequences to the organizational business continuity, health, strategy and results. The investments to be made can relate to the "business as usual", such as increasing production and expanding the business. These investments can mitigate trustworthiness risks through direct investment or, for example, by purchasing insurance.

Decisions on investment should be based on a quantitative analysis of financial scenarios, risks and consequences. Such analyses work best for high-frequency low-impact events since the probabilities and expected impact can be more easily quantified. This approach is less effective for analyzing the impact of low frequency (or unexpected) and high impact events, such as downtime due to a meteorite strike. In this case, the risk assessment may be best left to expert third parties via an insurance policy.

[1] For instance, information about use of drugs by a rock star is a non-story, whereas for a politician it could be career ending.

Trustworthiness investment decisions should not be based solely on quantitative analysis but also include judgments and investments based on understanding of potential high impact events. Care must be taken that both the data and the analysis used to make decisions are appropriate and that the confidence in the data quality and analysis is appropriate to the concerns.

Specifically, within the security domain, the IIC IoT Security Maturity Model[1,2] provides a framework to enable an organization to better match its investment in technology and organizational changes by determining the gap between the current state and the appropriate target state. Security maturity as defined in the IIC SMM is the degree of confidence that the current security state meets all organizational needs and security-related requirements. Security maturity level is a measure of the understanding of the current security level, its necessity, benefits and cost of its support. Initially targeted at security, the Security Maturity Model is designed to extend to the various trustworthiness characteristics.

The IoT Security Maturity Model illustrated in Figure 4-3 includes a hierarchy of domains, such as *governance*, *enablement* and *hardening*, as well as sub-domains and practices.

There are two dimensions to the evaluation of maturity—comprehensiveness and scope. *Comprehensiveness* captures the degree of depth, consistency and assurance of security measures that support security maturity. *Scope* reflects the degree of fit to the industry or system needs. The Security Maturity Model also defines a process to improve the maturity, through a plan-do-act cycle of continuous improvement around the evaluation of the needed target as shown in Figure 4-4, the current state and changes to address the results of the gap analysis.

[1] See [IIC-SMMPG2019]
[2] [IIC-SMMD2019]

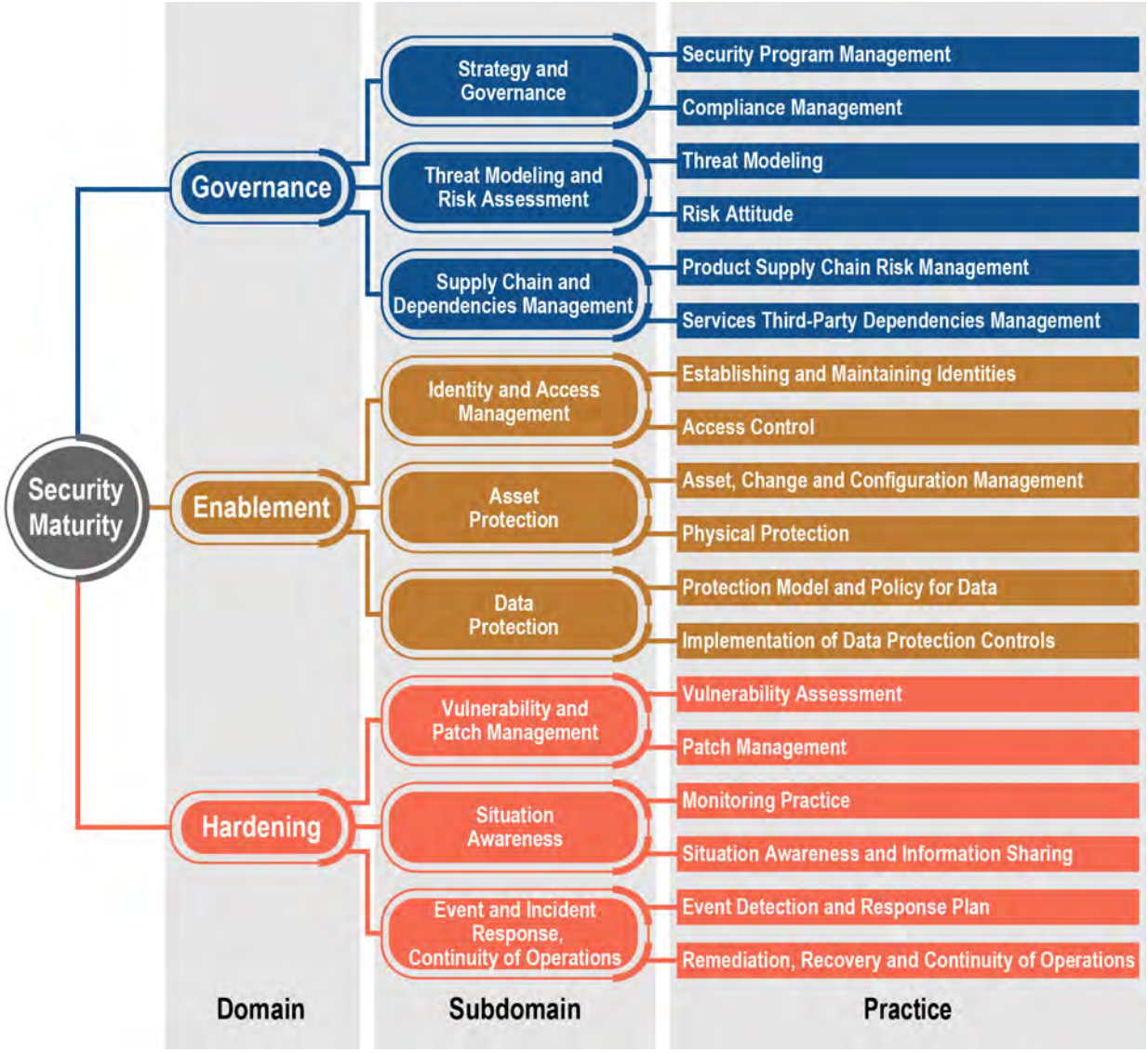


Figure 4-3: Security Maturity Model domains

The net result of the process, structured hierarchy and evaluation of comprehensiveness and scope allows increased confidence in the needed match of investment to the requirements. This can be used to address both financial, mission and low-frequency event concerns.

Similarly, analytic-based approaches should also be adopted to the assessment of real options for managing trustworthiness associated with the remaining four aspects of trustworthiness, keeping in mind that many specific events have the potential to affect more than one aspect.

In general, trustworthiness-related risks can be mitigated through organizational changes (e.g. training staff) as well as technology deployment (e.g. using identity management). Other traditional risk management approaches may also be used, such as purchasing insurance, effectively transferring risk. Risks may also be accepted as a part of normal business operations, but this should only be done if the consequences are understood.

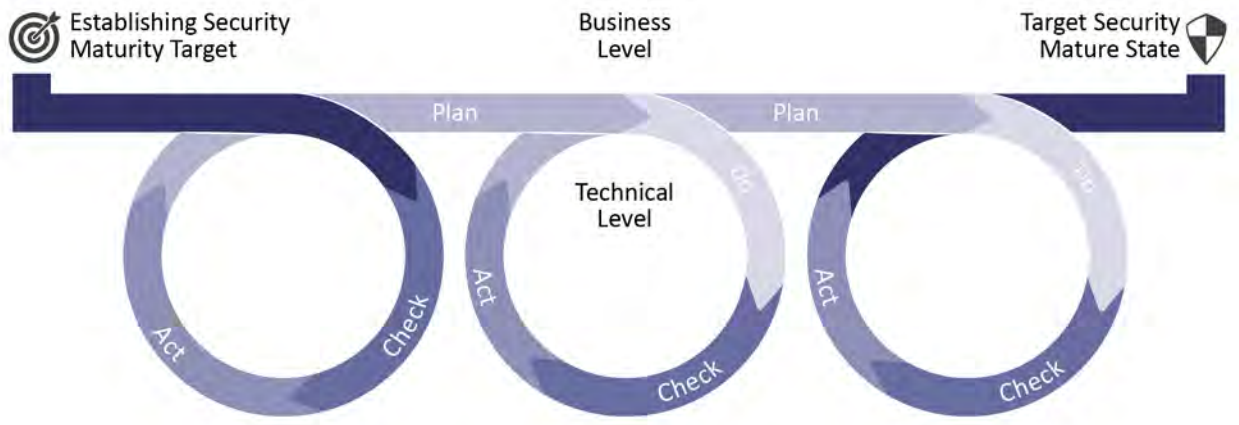


Figure 4-4: Security Maturity Model lifecycle

4.3 IMPLEMENTATION

The third phase is implementation, comprising two more stages,

- establish trustworthiness target beyond minimum compliance and
- establish policies and risk approach

as discussed below.

4.3.1 ESTABLISH TARGET BEYOND MINIMUM COMPLIANCE

The prior two steps have dealt with identified trustworthiness events and their potential consequences and measures that can be taken to mitigate any risks. However, trustworthiness measures can themselves generate increased value for a business. This can happen through a number of mechanisms, including:

Brand impact: companies can seek to differentiate on the basis of trustworthiness, and become recognized as 'more trustworthy' than competitors.

Increased revenues: potentially products and services that are underpinned by higher levels of IIoT trustworthiness (or QoS) can be sold for higher unit revenues.

Market access: potentially new markets for products and services may become addressable if a company maintains higher levels of trustworthiness.

Accordingly, before setting trustworthiness targets, we should identify how any of the real options for enhancing trustworthiness as identified in the previous step might result in an opportunity to enhance revenue. This analysis should be undertaken with reference to the business case developed in step 4 of this overall process.

4.3.2 ESTABLISH POLICIES AND RISK APPROACH

The establishment of a governance structure for trustworthiness is discussed in section 3.4 (page 26), and the discussion here assumes that such governance structures are in place.

Under that assumption, the management team responsible will need to determine the precise combination of risk-mitigating actions to be adopted. As yet, there is no purely analytic approach they can use to determine the mitigating actions that should be adopted.

Industry norms will likely emerge that will provide guidance and input to the kinds of trustworthiness risk-mitigating approaches that should be deployed in a range of situations[1] but today the deployment of such measures remains a decision for the management team. The analytical stages described here provide the inputs required for them to take those decisions.

4.4 ITERATE AND MAINTAIN

The final phase of the best practice approach to managing trustworthiness is *iterate and maintain*.

Trustworthiness for a particular IIoT system will evolve in its definitions and its objectives. It is an evolving concept, and trustworthiness measures that are appropriate for the overall context of any IIoT system at any one time may not be suitable in the future, as developed in section 3.5 ("The Continuing IoT Trustworthiness Journey").

Accordingly, trustworthiness and trustworthiness measures should be reassessed periodically. A prudent management team will ensure that trustworthiness measures are reviewed as follows:

- *periodically*: depending on the criticality and the overall levels of risk exposure.
- *reactively*: when the trustworthiness environment changes, potentially through the introduction of new regulations (for instance, GDPR) or changes in the trustworthiness associated with any upstream processes or data sources (for instance the use of a new supplier to provide weather data inputs to a system). Additionally, new information may become available related to existing IIoT solutions.
- *on request:* potentially when downstream uses of data (or other outputs of the IIoT system in question) change, and in response to requests from relevant stakeholders.

A company's approach to maintaining trustworthiness and updating any associated analyses and impact assessments should be the subject of a documented and managed security policy.

[1] We also expect to see the emergence of professionals who focus specifically on auditing the trustworthiness of IIoT solutions, in much the same way that today's accountants audit the financial health of any business.

5 CONCLUSIONS AND OUTLOOK

Key to confidence is the belief in the trustworthiness of a system. This includes an understanding and justified confidence in the key system characteristics of safety, security, privacy, resilience and reliability. It is possible to rely on intuition when dealing with people, but this is not adequate for increasingly automated systems, where explicit trustworthiness criteria must be established, contracted, measured and managed. We outlined here some practices to assess and manage trustworthiness in an IIoT system specific to its context and the consequences that could occur from failing to reach appropriate levels of trustworthiness characteristics. These practices include how to balance the appropriate tradeoffs with business concerns and objectives. This in turn requires an understanding of the interactions between these characteristics and of their impact on business operations.

The business benefits of analyzing and managing the risk of a system and of increasing its trustworthiness include ensuring compliance with regulations, reducing potential future financial and reputational costs due to failures, and improving the brand and ability to work with partners.

To develop and deploy a trustworthy system will require attention to the stages previously defined, including periodic reassessment of these. The trustworthiness of an IoT system is a concern that must be addressed and tracked throughout the lifecycle of the system—from business model definition--and throughout the lifecycle of the data about the system.

We have focused on the business concerns and process of managing a trustworthy system but intentionally ignored implementation details, such as practical aspects of trustworthiness assurance, and metrics definition. We intend to follow this work with a white paper outlining the importance and use of metrics for trustworthiness. The material developed here will also influence a separate and more ambitious Trustworthiness Framework, focusing in particular on the business viewpoint (as outlined in the IIC Reference Architecture[1]).

[1] See [IIC-IIRA2019]

Annex A REFERENCES

[GINT2016] SCADA security—what's broken and how to fix it, Andrew Ginter, Abterra Technologies, Inc., 1st edition, 2016-09-03, retrieved 2019-06-26
http://www.scada-security.ca/

[IIC-FOVI] Industrial Internet Consortium: Factory Operations Visibility & Intelligence Testbed, retrieved 2019-06-26
https://www.iiconsortium.org/fovi.htm

[IIC-BSIF2016] Industrial Internet Consortium: The Industrial Internet of Things, Volume B01: Business Strategy and Innovation Framework, version 1.0, 2016-11-15, IIC:PUB:B01:V1.0:PB:20161115, retrieved 2019-07-23
https://www.iiconsortium.org/BSIF.htm

[IIC-IIRA2019] Industrial Internet Consortium: The Industrial Internet, Volume G1: Reference Architecture Technical Report, version 1.9, 2019-06-19, retrieved 2019-07-23
https://www.iiconsortium.org/IIRA.htm

[IIC-IISF2016] The Industrial Internet of Things, Volume G4: Security Framework Version 1.0, 2016-09-26, IIC:PUB:G4:V1.0:PB:20160926, retrieved 2019-06-26
https://www.iiconsortium.org/IISF.htm

[IIC-IIV2018] Industrial Internet Consortium: The Industrial Internet of Things, Volume G8: Vocabulary Version 2.1, 2018-08-22, retrieved 2019-06-26, IIC:PUB:G8:V2.00:PB:20180822
https://www.iiconsortium.org/vocab/index.htm

[IIC-JOI20182] The IIC Journal of Innovation, 9th Edition: Trustworthiness. 2018-September, retrieved 2019-06-26
https://www.iiconsortium.org/news/journal-of-innovation-2018-sept.htm

[IIC-SFWP2017] Key Safety Challenges for the IIoT, An Industrial Internet Consortium Technical White Paper, IIC:WHT:IN6:V1.0:PB:20171201, 2017-12-01, retrieved 2019-06-26
https://www.iiconsortium.org/pdf/Key_Safety_Challenges_for_the_IIoT.pdf

[IIC-SMMD2019] IoT Security Maturity Model: Description and Intended Use, Version 1.1, 2019-02-15, retrieved 2019-06-26
https://www.iiconsortium.org/pdf/SMM_Description_and_Intended_Use_FINAL_Updated_V1.1.pdf

[IIC-SMMP2019] The Internet of Things (IoT) Security Maturity Model (SMM) Practitioner's Guide, Version 1.0, 2019-02-15, retrieved 2019-06-26
https://www.iiconsortium.org/smm.htm

AUTHORS AND LEGAL NOTICE

This document is a work product of the Industrial Internet Consortium Business Strategy and Solution Lifecycle Working Group, co-chaired by Jacques Durand (Fujitsu), Jim Morrish and Dirk Slama (Bosch Software Innovations).

Authors: The following persons contributed substantial written content to this document: Jacques Durand (Fujitsu), Frederick Hirsch (Fujitsu), Jim Morrish (Co-Chair Business Strategy and Solution Lifecycle Working Group), Bassam Zarkout (IGnPower), Marcellus Buchheit (Wibu-Systems).

Contributors: The following persons contributed valuable ideas and feedback that significantly improved the content and quality of this document: Andrew Ginter (Waterfall Security), Jesus Molina (Waterfall Security), Takao Mizutani (Fujitsu), Jürgen Neises (Fujitsu), Thomas Walloschke (Fujitsu).

Technical Editor: Stephen Mellor (IIC staff) oversaw the process of organizing the contributions of the above Authors and Contributors into an integrated document.
