Achieving Trustworthiness

Through Risk Management, System Theory and Resilience

2022-07-27

Author:

Frederick Hirsch
Upham Security
hirsch@uphamsecurity.com

CONTENTS

FIGURES

TABLES

1 INTRODUCTION

Trustworthiness is defined[1] by the Industry IoT Consortium (IIC) as the "degree of confidence one has that the system performs as expected with characteristics including safety, security, privacy, reliability and resilience in the face of environmental disturbances, human errors, system faults and attacks".

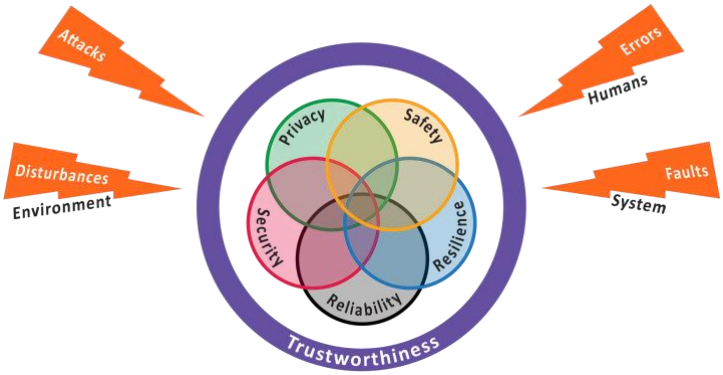

Figure 1-1: Resilience and trustworthiness. (Source: Industry IoT Consortium.[2])

One of the Trustworthiness characteristics, *Resilience*, is defined by the IIC as the "ability of a system or component to maintain an acceptable level of service in the face of disruption"[2]. The following more detailed definition is consistent with but goes further than the IIC definition:

> "A system is resilient if it can adjust its functioning prior to, during, or following events (changes, disturbances, and opportunities) and thereby sustain required operations under both expected and unexpected conditions."[3,4]

Change, errors, attacks, faults and disturbances create hazards for an organization that may need to be managed. Risk management is important for understanding consequences and having confidence in a system. Resilience is important for anticipating, addressing and learning from events in order to maintain confidence in the system and address concerns that cannot easily be

---

[1] Claude Baudoin et al., "Industry IoT Vocabulary, Version 3.0," March 22, 2022, *https://www.iiconsortium.org/vocab-01/Industry-IoT-Vocabulary.pdf.*

[2] Marcellus Buchheit et al., "The Industrial Internet of Things Trustworthiness Framework Foundations," July 15, 2021, *https://www.iiconsortium.org/pdf/Trustworthiness_Framework_Foundations.pdf.*

[3] Erik Hollnagel, "Resilience Engineering and the Future of Safety Management," in *Handbook of Safety Principles*, ed. Niklas Möller, Sven Ove Hansson, and Jan-Erik Holmberg, Wiley Essentials in Operations Research and Management Science (Hoboken, NJ: Wiley, 2018).

[4] An earlier version of this definition is also used in ARPANSA, "Regulatory Guide - Holistic Safety - Sample Questions (ARPANSA-GDE-1754WEB)", June 15, 2021, *https://www.arpansa.gov.au/regulation-and-licensing/licensing/information-for-licence-holders/regulatory-guides/regulatory-guide-holistic-safety-sample*.

addressed with traditional risk management. Resilience management includes a number of principles and approaches that can be used in conjunction with risk management and systems theory techniques in order to increase the trustworthiness of a system.

The goal is to ensure that unacceptable losses do not occur and that hazards that can lead to losses are eliminated, mitigated or controlled. A hazard is defined here as "A system state or set of conditions that, together with a particular set of worst-case environment conditions, will lead to an accident (loss)"[5]. Traditional risk management techniques can help achieve the goal, as can system theory analysis[6] and resilience engineering.

2 TRADITIONAL RISK MANAGEMENT

To understand the need for systems theory and resilience it is first necessary to understand traditional risk management and its limitations.

The IIC defines risk as the "effect of uncertainty on objectives", deriving the definition from ISO/IEC 27000[7]. Another definition of risk commonly used is that risk is the "combination of the probability of occurrence of harm and the severity of that harm"[8]. Risk analysis can be described as answering three questions[9,10,11]:

1. What can go wrong?
2. How likely is it?
3. What would be the consequences?

Quantitative risk analysis is an approach that associates consequences such as the death of a person, for example, with the likelihood or probability. This approach works best when it is possible to determine meaningful probabilities and to value consequences.

One way to determine probabilities is by frequency analysis. If an event occurs repeatedly over time, it is possible to use the historical data to obtain frequencies of occurrence and determine a probability distribution. This is possible for events for which there is data such as automobile

[5] Nancy G Leveson, *Engineering a Safer World.* (Cambridge: The MIT Press, 2016), *http://www.oapen.org/download?type=document&docid=1004042.*

[6] Leveson.

[7] "ISO/IEC 27000:2018(E) Information Technology — Security Techniques — Information Security Management Systems — Overview and Vocabulary, 5th Edition," 2018, *https://www.iso.org/standard/73906.html*.

[8] "ISO/IEC Guide 51 Safety Aspects - Guidelines for Their Inclusion in Standards, Third Edition," April 1, 2014, https://www.iso.org/standard/53940.html, see definition 3.9.

[9] Jan-Erik Holmberg, "Quantitative Risk Analysis," in *Handbook of Safety Principles*.

[10] Risto Tiusanan, "Qualitative Risk Analysis," in *Handbook of Safety Principles*.

[11] Much of this section is derived from the previous two references.

accidents, for example. Another way to view probabilities is as the degree of belief of the party determining the value, a subjective interpretation.

In either approach the probability is used along with a measure of the consequence of the outcome to determine the risk. At an abstract-level consequences such as the death of a person or multiple people, harm to society, harm to the environment and so on are considered. In practice a "surrogate criterion" might be used, such as the release of radiation from a nuclear power plant to the environment, or the system failure of a railway control system[12].

The resultant risk is the sum of all the consequences (C_i) and their likelihoods (P_i):

$$\text{Risk} = \sum\_i \ [\![P_i * Ci]\!]$$

Calculating quantitative risk values is only part of the process since the most important aspect is to understand the risk and make appropriate decisions. It is impossible to remove all risk, so the question is which risks are important to address.

One approach to deciding which risks to address is based on a curve which plots the probabilities against the degree of consequence:

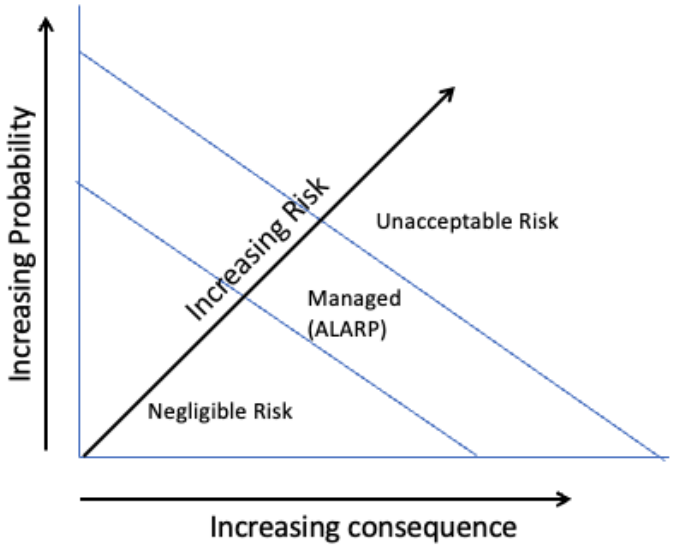

Figure 2-1: Risk regions and ALARP (Source: derived from Holmberg[12]).

There are three regions:

1. Negligible risk: Area where the consequences and likelihood are low enough to allow a judgement that the risk is not a concern.
2. Unacceptable risk: Area where the risk is judged to be too high due to significant consequences and/or high frequency.

[12] Holmberg, "Quantitative Risk Analysis."

3. Managed: Area where the risk is neither negligible nor unacceptable. This is where effort is taken to make the risk "as low as reasonably practical" (ALARP), aligning the risk reduction costs with the benefits.

Negligible risks can simply be accepted. Unacceptable risks are to be avoided at all costs, though in some cases it is decided they can never happen, so when they do it is a surprise. This is one area where resilience and systems theory come into play.

Managed risks are those which are expected to occur and for which techniques can be used to mitigate, transfer or eliminate them.

Organizations often use a risk matrix to categorize risks and to prioritize which managed risks to address. This can be based on quantitative or qualitative evaluations of frequency and consequence. The matrix generally looks like the following (with varying numbers of rows and columns, and sometimes quantitative measures on the axes):

Probability	**Harm severity**			
	Negligible	**Marginal**	**Critical**	**Catastrophic**
Certain	High	High	Very high	Very high
Likely	Medium	High	High	Very high
Possible	Low	Medium	High	Very high
Unlikely	Low	Medium	Medium	High
Rare	Low	Low	Medium	Medium
Eliminated	Eliminated			

Figure 2-2: Risk matrix. (Source: Wikipedia.[13])

[13] Image: Wikipedia contributors (2021, November 5). In *Wikipedia, The Free Encyclopedia*. Retrieved 14:13, April 25, 2022.

There are a number of detailed methods for performing risk analysis that can contribute to an understanding and to a risk matrix. One analysis method is to create an event tree which starts from some "initiating event" and leads to a variety of consequences based on successes or failures at various decision points (e.g. did an engine fail and, if so, did a second engine fail etc.). An event tree analysis can show the path to success or failure outcomes based on the various events that occur over time and the outcome of each event. This requires an understanding of the possible initiating events and the various event paths that can occur.

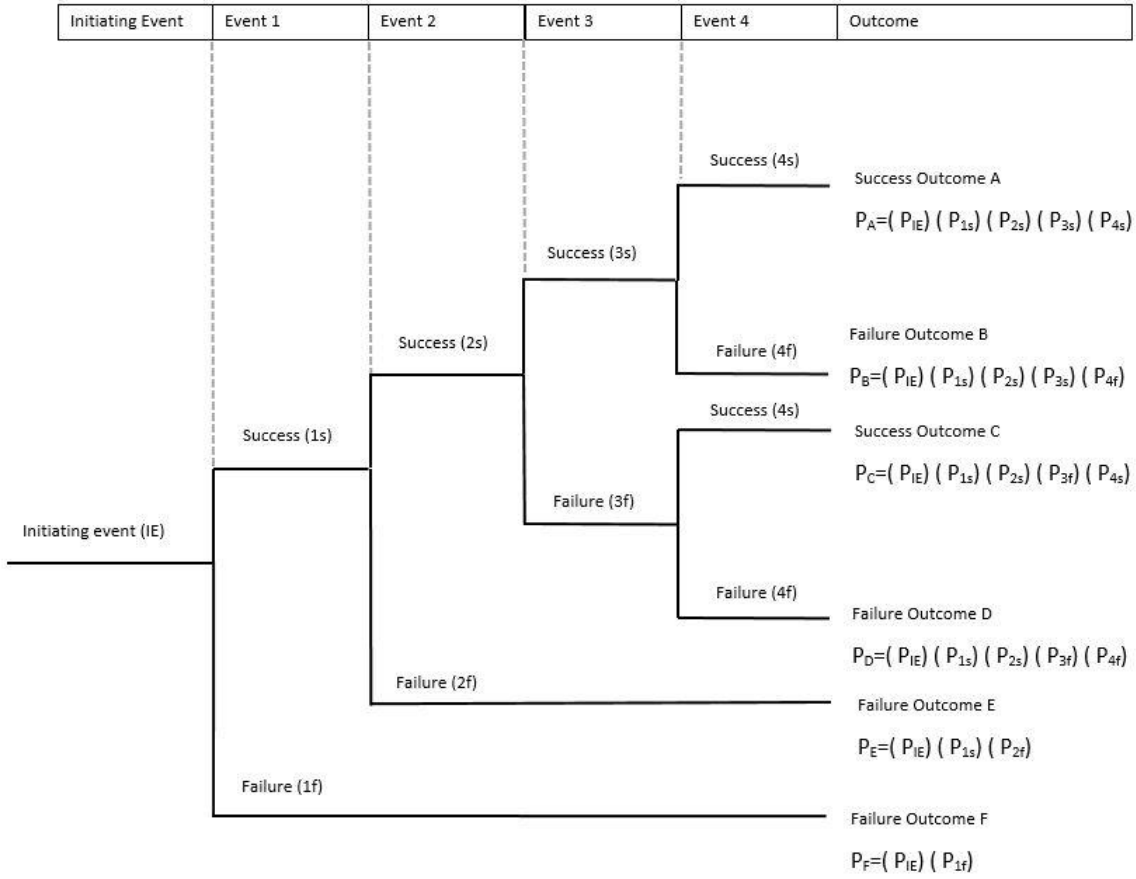

Figure 2-3: Event tree. (Source: Wikipedia.[14])

Another approach is a fault tree, as used in reliability analysis. This can be used to divide a system into components and sub-components to understand dependencies, how failures may propagate, and which failures are critical.

[14] Image: By 570SJR - Own work, CC BY-SA 3.0
https://commons.wikimedia.org/w/index.php?curid=29334798

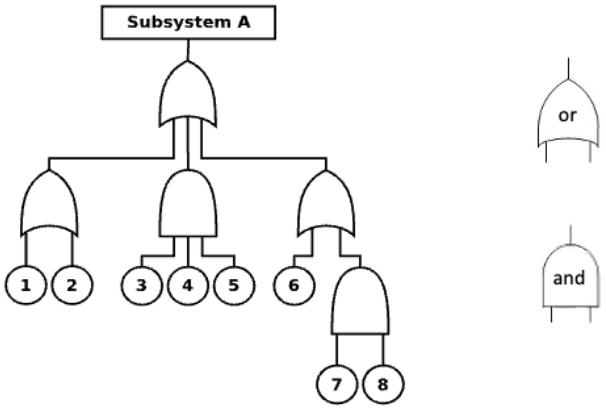

Figure 2-4: Fault tree. (Source: Wikipedia.[15,16])

Safety integrity levels (SIL)[17] provide requirements for functional safety and include quantitative requirements on failure frequency and probability, tolerance for failures, software quality and governance and process management. This is supposed to aid the analysis.

The analysis depends on understanding the initiating events, hazards. A hazard is a "condition that can cause injury or death, damage to or loss of equipment or property, or environmental harm"[18] thus potentially leading to negative outcomes.

Hazards can be dealt with by (1) making sure they do not exist by designing them out, (2) by limiting their impact or (3) by training operators to manage them. For example, outlets have openings designed to match the plugs that go into them, making it hard to mismatch what is plugged in against the circuit, thus designing out a particular hazard. Circuit breakers can limit the impact of an overloaded circuit by breaking the circuit upon detection of an overload condition. Finally, training can be used to advise people to not replace fuses with larger capacity fuses incompatible with the wiring.

One way to manage risk is to create barriers to hazards, adopting an approach of defense in depth, to prevent hazards from causing harm. This was modeled by Reason as the "Swiss Cheese Model" – a loss occurs only if the holes in the various barriers line up (some holes are latent conditions, others are active failures, others unsafe acts, for example):

[15] Image: By Offnfopt, modeled after image create by U.S. Military - Own work created from scratch using File:Fault tree.png as a reference, Public Domain, *https://commons.wikimedia.org/w/index.php?curid=52420454*

[16] Images: Stand alone symbols By Rich Baldwin - Own work, CC0, *https://commons.wikimedia.org/w/index.php?curid=13535039*

[17] *https://en.wikipedia.org/wiki/Safety_integrity_level*

[18] Harold E. Roland and Brian Moriarty, *System Safety Engineering and Management*, 2nd ed (New York: Wiley, 1990).

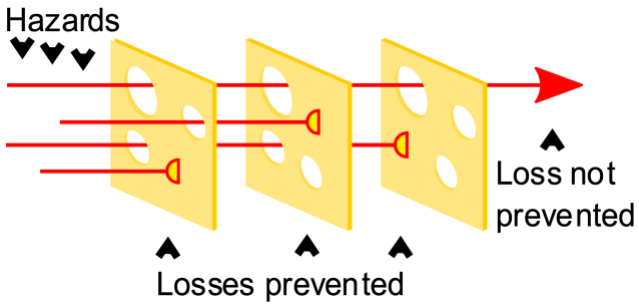

Figure 2-5: Swiss Cheese model. (Source: BenAveling CC BY-SA 4.0[19] via Wikimedia Commons.[20])

By understanding hazards and performing risk analysis it is possible to categorize and prioritize risks, determine which risks are practical and cost effective to manage, and to communicate and address hazards with a defense in depth strategy of elimination and mitigation.

3 Difficulties with Risk Management

Despite the benefits of risk analysis in various industries there still have been catastrophic events that were not anticipated or prevented through the use of such traditional risk management. The reason is that it is very difficult to anticipate all possibilities in complex and dynamic systems, especially when considering events that are low probability and high consequence. Complex systems also allow losses to occur for which there is no single root cause or initiating event.

There are many new pressures on organizations creating new hazards as businesses rely on computer technology, data, and analytics as never before. These pressures increase the difficulty of performing timely risk analysis and raise the risk of low probability and high impact events. The pressures include:

1. The increasing pace and interdependencies of business, increasing complexity and the speed with which harmful consequences can occur.
2. Digital transformation which also increases complexity and scale. This includes new technologies (data analytics, cloud, big data, AI, digital twin and IoT technologies).
3. The increasing reliance on data and algorithms to create 'smart' systems that operate at speed and scale.
4. The rapid pace and necessity for change and adaptation, including new business models, intense global competition, climate and sustainability concerns.
5. Emerging systemic economic pressures including those related to supply chain reliability.

There are also technical difficulties:

[19] *https://creativecommons.org/licenses/by-sa/4.0*
[20] Image: *https://upload.wikimedia.org/wikipedia/commons/0/07/Swiss_cheese_model.svg*

1. Complex, dynamic and open systems are difficult to understand and model.
2. Infrequent events can have significant impact, but frequency data is not available.
3. Estimating probabilities and severities can be subject to bias and inaccuracy, especially when choosing which outcomes and events to consider.
4. Meaningful and pragmatic metrics can be hard to create.
5. Tradeoffs are hard to make.

In dynamic and complex systems root cause analysis no longer suffices since many factors can contribute to outcomes[21]. "Normal Accidents" can be expected to occur[22]. Probabilities are often not meaningful or appropriate. Other approaches, such as using a systems model and using an understanding of losses, hazards, unsafe control actions and necessary constraints can offer an approach that can identify more problems, beyond component failure scenarios, find them earlier and at lower cost since they can be understood before design, development and implementation[23,24,25].

It is hard to go from an understanding of risks to indicators that allow measurement and improvement. For a workplace safety example, safety metrics such as 'days since the last accident' look backward and correspond to the concept of 'free from unacceptable risk' and are not necessarily useful for "safety in the future". They do not address chance, do not incorporate learning about the system and may give rise to complacency if there have not been any accidents over time[26].

The use of probabilities in risk analysis can themselves raise questions since they can be hard to understand and establish. Randomness is variability that is fundamental while uncertainty can reflect a lack of knowledge and can be reduced. There can be uncertainty with the modeling, with lack of evidence to estimate probability distribution parameters, and with assumptions. Incompleteness includes "known unknowns" and "unknown unknowns." All of these uncertainties indicate that traditional risk management is not a complete solution and even with best effort incidents and accidents may still occur. This can be addressed with a systems model approach as well as resilience engineering.

[21] Hollnagel, "Resilience Engineering and the Future of Safety Management."
[22] Charles Perrow, *Normal Accidents: Living with High-Risk Technologies*, Princeton Paperbacks (Princeton, N.J: Princeton University Press, 1999).
[23] Leveson, *Engineering a Safer World.*
[24] Nancy G. Leveson, "CAST Handbook," 2019, *http://psas.scripts.mit.edu/home/get_file4.php?name=CAST_handbook.pdf.*
[25] Nancy G Leveson, "Safety III: A Systems Approach to Safety and Resilience" (MIT Engineering Systems Lab, Aeronautics and Astronautics Dept., MIT, July 1, 2020), *http://sunnyday.mit.edu/safety-3.pdf*.
[26] John Wreathall, "Monitoring – A Critical Ability in Resilience Engineering," in *Resilience Engineering in Practice: A Guidebook*, ed. Erik Hollnagel et al., Ashgate Studies in Resilience Engineering (Farnham, Surrey, England ; Burlington, VT: Ashgate, 2011).

Assurance cases, of which safety cases are one example, can be used to better understand and communicate risk and how it is managed. These cases can be used to produce an *argument* that a system meets safety goals (e.g. in the safety case), by providing *evidence* that supports the argument that the goal is achieved while taking the system *context* into account. Such assurance cases are most useful when used throughout the system development cycle. The cases also need to be actively maintained to reflect change, be challenged and communicated in order to better support the goals. They also need to address possible confirmation bias. One approach is to enhance the assurance case to include not only the risk analysis, but also explicitly an analysis of the confidence in the case itself[27]. Assurance cases in conjunction with a systems approach offer a way to improve trustworthiness.

Although risk management can improve outcomes, resilience management can contribute by having a positive impact when the unexpected does occur.

4 Resilience Management

4.1 The Resilience Lifecycle

When a disruption occurs despite the efforts to mitigate the risk, resilience comes into play and enables the system to absorb the shock and to continue operating, perhaps in a reduced capacity, yet one that meets fundamental needs. Resilience includes the ability to take actions before a disruption occurs to anticipate it, to recover from a disruption and to learn and improve in anticipation of the next disruption. Resilience is not just technical – it includes organizational aspects as well. The resilience lifecycle is shown in Figure 4-1:

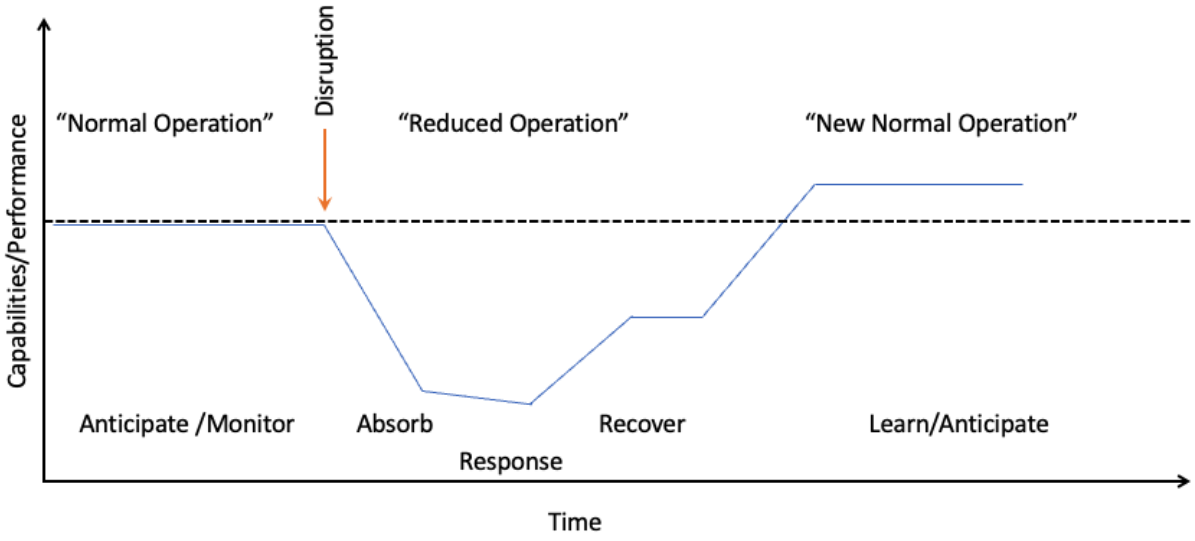

Figure 4-1: Resilience lifecycle.

[27] Tim Kelly, "Safety Cases," in *Handbook of Safety Principles*.

There are four key resilience activities[28]:

1. *Anticipate* – An organization must anticipate the future, understand potential losses and hazards and scenarios that can lead to them, establish requirements and an implementation, both technical and non-technical, to avoid those losses. This might be done with an STPA analysis[29] ("System-Theoretic Process Analysis"), for example. An organization can also put measures in place to manage response and recovery.
2. *Monitoring* – An organization must know what to look for in the environment and itself, to monitor change to detect possible additional hazards.
3. *Response/Recovery* - An organization must know what to do when an unanticipated loss scenario occurs and be able to recover.
4. *Learn* – An organization must learn, remember, and modify its behavior based on events. This can be done with an analysis of a loss, by using a CAST analysis[30] ("Causal Analysis based on System Theory"), for example.

The governance process for these activities forms a cycle[31] (Figure 4-2):

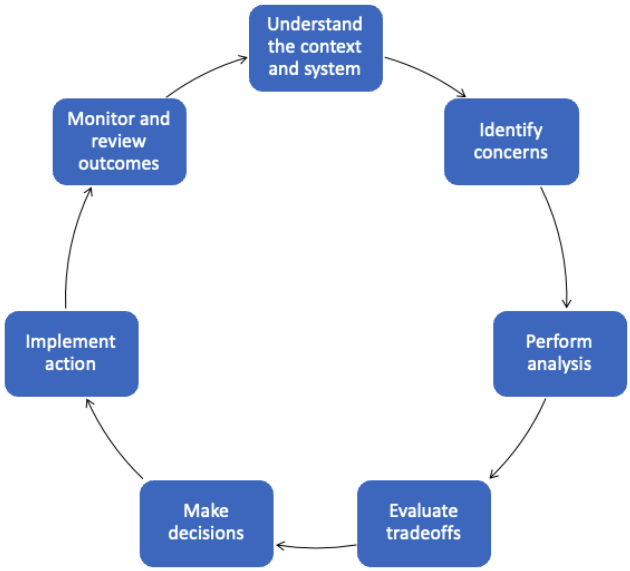

Figure 4-2: Governance lifecycle.

[28] Hollnagel, "Resilience Engineering and the Future of Safety Management."
[29] Nancy G. Leveson and John P. Thomas, "STPA Handbook," March 2018, *http://psas.scripts.mit.edu/home/get_file.php?name=STPA_handbook.pdf*.
[30] Leveson, "CAST Handbook."
[31] Ivo Häring et al., "Towards a Generic Resilience Management, Quantification and Development Process: General Definitions, Requirements, Methods, Techniques and Measures, and Case Studies," in *Resilience and Risk: Methods and Application in Environment, Cyber and Social Domains*, ed. Igor Linkov and José Manuel Palma-Oliveira, 1st ed. 2017, NATO Science for Peace and Security Series C: Environmental Security (Dordrecht: Springer Netherlands : Imprint: Springer, 2017), *https://doi.org/10.1007/978-94-024-1123-2*.

A process alone is not enough to achieve resilience. Success in achieving a resilient organization follows the same principles as success in establishing an organization that values safety or security. It requires true leadership and commitment at the board and executive level. True leadership means decisions to support resilience are reflected in budgets and the organizational culture. Such a culture allows employees to raise issues regarding resilience and safety without repercussions, supports communication and a willingness to allow anyone to make the hard tradeoffs (e.g. stop the production line, sacrificing production for safety).

To achieve resilience an organization needs certain competencies[32]:

- *Information Management* – ability to handle the large amount of data associated with an escalating situation and the ability to determine what is important.
- *Communication and Coordination* – the ability for people know each other's roles and tasks and communicate clearly.
- *Decision Making* – the ability to make timely and appropriate decisions. Consensus won't work since it is too slow. Top down directives won't work either since the workload would be too high to manage from the top. Thus decision making needs to be distributed and delegated (just as in the military in the 'fog of war', local decisions are made aligned with the overall objective).
- *Effect Control* – the ability to monitor and update the process.

Every phase in the resilience lifecycle is important to the overall resilience outcome and depends on actions an organization takes. The first lifecycle planning phase requires architecting, designing and building systems that are resilient.

4.2 ANTICIPATING

The primary aspect of anticipation is to understand potential losses, whether they be of life or injury, financial, reputation, damage to the environment, or mission. Once these potential losses are understood it makes sense to review the hazards that could lead to these losses and the scenarios where such losses could occur. For example, loss of life or injury near a chemical plant could occur from the hazard of an accidental release of chemicals, in the context of a scenario where the wind blew them in the direction of people (this might not always be the case). Once these scenarios are understood with an understanding of the system and unsafe actions it is possible to produce requirements and an implementation.

[32] John Bergström et al., "Training Organisational Resilience in Escalating Situations," in *Resilience Engineering in Practice: A Guidebook*.

In addition it is useful to architect and design systems to reduce and manage complexity. The following principles support designing a system that "performs its intended functions in the manner intended" and is as resilient as possible[33,34]:

1	Absorption	The system can absorb the impact of a disruption.
2	Physical Redundancy	There are two or more independent redundant components.
3	Functional Redundancy	There are two or more ways to perform a task.
4	Layered Defense (Defense in depth)	There are two or more independent approaches address a single vulnerability.
5	Human in the Loop	There should be a human as part of the system where human thinking is required. The premise is that humans are better at dealing with unprecedented situations.
6	Reduce Complexity	A system should be no more complex than necessary by reducing the number of complex components (including humans) and interfaces.
7	Reorganization Principle	A system, either organizational or technical, should be capable of changing its entire architecture, or structure, in the face of a threat.
8	Repairability	A system should have the capability of being brought up to partial or full functionality over a specified period of time and in a specified environment.
9	Localized Capacity (Modularity)	The functionality of a system should be distributed through various nodes of that system so that if a single node is damaged or destroyed, the remaining nodes will continue to function.
10	Loose Coupling	The system should have the capability to limit the ability of failures to propagate from one component to the next in a system of many components.
11	Drift correction	If the system is drifting towards failure this is detected and measures can be taken to avoid the threat, or it can be diminished through corrective action.
12	Neutral state	Humans delay taking action when there is an opportunity to survey the situation and make a more reasoned judgement.
13	Inter-node interaction	Every node, or element, of a system should be capable of communicating, cooperating, and collaborating with every other node.
14	Reduce Hidden Interactions	Efforts should be made to assure that potentially harmful interactions between nodes of the system should be reduced.

Table 4-1: Resilience implementation principles

[33] Scott Jackson and Timothy L. J. Ferris, "Resilience Principles for Engineered Systems," *Systems Engineering* 16, no. 2 (June 2013): 152–64, *https://doi.org/10.1002/sys.21228*.

[34] Scott Jackson, "Principles for Resilient Design - A Guide for Understanding and Implementation," in *IRGC (2016) Resource Guide on Resilience*, v29 ed. (Lausanne: EPFL International Risk Governance Center, 2016), *https://beta.irgc.org/wp-content/uploads/2018/09/Jackson-Principle-for-Resilient-Design.pdf.*

4.3 MONITORING

Once a system has been created it is necessary to monitor for internal and external change. Doing so makes it possible to recognize hazards that can lead to problems early enough to take action to address them. Monitoring may need to pick up on 'faint signals', early signals of problems in a process or project[35]. An organization that can anticipate may also be more successful in monitoring by paying more attention to signals.

Incident reporting is a systematic activity to monitor the state of resilience. The airline industry, for example, has mechanisms for pilots to report incidents (which are not accidents, but which could result in accidents). This enables learning to take place and allows lessons to be shared and action to be taken to prevent accidents.

A typical system model has a controller managing processes using a model of the processes to understand the outcomes based on feedback. This also is a form of monitoring the processes against the model. Digital twins should make such monitoring possible for a system, thus enabling the detection of potential issues before they become emergencies.

The following patterns[36] are useful for monitoring:

1. *Recognize that adaptive capacity is falling or inadequate to the contingencies and squeezes or bottlenecks ahead.*
   This is related to the ability to monitor the environment and itself.
2. *Recognize the threat of exhausting buffers or reserves.*
   This includes the fact that economic pressures drive organizations to become 'lean' and 'just in time' to the point that they may no longer be resilient. We are now seeing supply chain issues related to this.
3. *Recognize when to shift priorities across goal tradeoffs.*
   Being able to sacrifice short term production goals in order to prioritize longer-term safety goals reflects the organization goals and culture.
4. *Make perspective shifts and contrast diverse perspectives that go beyond their nominal system position.*
   The ability to change the organization itself and its approach is important to a resilience response, learning and anticipating.
5. *Navigate interdependencies across roles, activities, levels.*
   The abilities of members of the organization to not work at cross-purposes or work purely locally is important to response and adaptation (e.g. avoiding local optimizations that are harmful organizationally in the larger sense)

[35] R Westrum, "Faint Hearts and Faint Signals – How Organizations Manage Signs of Trouble.," 1999. quoted in Wreathall, "Monitoring – A Critical Ability in Resilience Engineering."

[36] David D. Woods, "Resilience and the Ability to Anticipate," in *Resilience Engineering in Practice: A Guidebook*.

6. *Recognize the need to learn new ways to adapt.*

4.4 ABSORBING AND RECOVERING

When a disruption occurs, the ability to absorb it is related to the robustness and reliability of the system as well as how the system was architected, designed and built. Following resilience principles such as loose coupling can improve resilience, for example by reducing the possibility of cascading failures.

Much of absorbing and recovering from a disruption can depend on people, whether they can correctly interpret the situation and understand when to escalate and bring in additional resources. Asking for and obtaining help can make a big difference, whether it is obtaining the help of additional fire stations or requesting new viewpoints regarding anesthesia in an operating room[37].

Being able to accept new evidence and being willing to revise ones plans and behaviors is important to allow change and effective action to be taken even if not according to the previous experience[38]. This is where "humans in the loop" can be helpful, since they can 'think outside the box' and break the rules when needed, using creative thinking.

4.5 LEARNING

Learning is key to improvement and improving resilience and can happen by reviewing loss scenarios (accidents) with a systems analysis methodology such as provided in CAST[39]. This includes modeling the control structure, analyzing each component in the loss, identifying control structure flaws, and creating an improvement program which can include using this analysis to determine additional constraints and requirements to improve the system.

All of these stages of the resilience lifecycle relate to each other, and learning contributes to anticipation.

[37] John Cuvelier and Pierre Falzon, "Coping with Uncertainty. Resilient Decisions in Anesthesia," in *Resilience Engineering in Practice: A Guidebook*. John Cuvelier and Pierre Falzon, "Coping with Uncertainty. Resilient Decisions in Anesthesia," in *Resilience Engineering in Practice: A Guidebook*.

[38] D Mendonça and W A Wallace, "Adaptive Capacity: Electric Power Restoration in New York City Following the 11 September 2001 Attacks," 11, accessed May 17, 2022, *https://www.resilience-engineering-association.org/download/resources/symposium/symposium-2006(2)Mendonca_Wallace.pdf*2006(2)/Mendonca_Wallace.pdf.

[39] Leveson, "CAST Handbook."

5 CONCLUSION

Achieving trustworthy operation requires an understanding of a system, the context in which it operates, and the potential losses and the hazards that can contribute to those losses. Designing and building a trustworthy system requires an understanding of system design, necessary constraints and requirements and use of principles to reduce complexity and enable resilience. Traditional risk analysis, systems theory analysis, and resilience management are all necessary. Using these together allows an organization to deal with hazard scenarios.

An organization can assess its resilience by using the Resilience Analysis Grid[40] as well as using guidelines such as the Australian Radiation Protection and Nuclear Safety Agency[41] safety guidelines which explicitly mention resilience. Realizing that the concerns affecting safety and security are related to resilience and that governance, controls and operations matter in all instances, assessing the system using the IIC IoT Security Maturity Model[42,43] or with safety assessments can be useful. One of the goals of the IIC work in trustworthiness is to break down the siloes among the communities working with different trustworthiness characteristics, with an understanding of the commonality of the need to prevent losses by addressing the associated hazards.

Achieving resilience requires effective governance for the monitoring and anticipation, response, recovery, and learning phases of resilience. This requires leadership, management support and commitment, and a culture supporting trustworthiness. It also requires systems architecture, design and operations personnel to understand resilience principles, indicators, and actions. There is no silver bullet, but system design and resilience engineering can enhance risk management enabling safer and more trustworthy systems.

6 REFERENCES AND FURTHER READING

In addition to the specific resources quoted in the footnotes this paper also drew upon the following resources.

[40] Erik. Hollnagel, "RAG - The Resilience Analysis Grid," in *Resilience Engineering in Practice: A Guidebook*.

[41] ARPANSA, "Regulatory Guide - Holistic Safety - Sample Questions (ARPANSA-GDE-1754WEB)."

[42] Sandy Carielli et al., "IoT SMM Practitioner's Guide Version 1.2," May 5, 2020, *https://www.iiconsortium.org/pdf/IoT_SMM_Practitioner_Guide_2020-05-05.pdf.*

[43] The SMM offers many domains and practices directly applicable to resilience, such as the governance domain, monitoring and continuity practices, to give some examples. Learning and anticipation are accounted for in the Level 4 comprehensiveness levels. Although targeted at security there is thought about extending the SMM to trustworthiness in general, see Frederick Hirsch et al., "Extending the IIC IoT Security Maturity Model to Trustworthiness," IIC Journal of Innovation, 2018, 16, *https://www.iiconsortium.org/news/joi-articles/2018-Sept-JoI-Extending-the-IIC-Security-Maturity-Model-to-Trustworthiness.pdf*.

Bahr N., 2015, *System Safety Engineering and Risk Assessment*, Second Edition, CRC Press

Dekker S, 2019, *Foundations of Safety Science*, Sidney, CRC Press

Hollnagle E, Woods D, Leveson N, *Resilience Engineering: Concepts and Precepts*, CRC Press, 2006

Hollnagel E, Pariès J; Woods D; Wreathall J, *Resilience Engineering in Practice : A Guidebook* (Ashgate Studies in Resilience Engineering) , 2011

Hollnagel E, *Barriers and Accident Prevention*, Ashgate, Farnham, 2004

Jackson S, Ferris T, *Resilience Principles for Engineered Systems*, Wiley, 2012

Linkov I, Palma-Oliveira J, *Resilience and Risk*, Springer, 2016

Möller, N., Hansson, S., Holmberg, J., Rollenhagen, C., *Handbook of Safety Principles*, Wiley Essentials in Operations Research and Management Science, 2018

Ross et al, Developing Cyber-Resilient Systems: A Systems Security Engineering Approach, NIST Special Publication 800-160, Volume 2, Dec 2021 *https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-160v2r1.pdf*

NIST Cybersecurity Framework, *https://www.nist.gov/cyberframework*

7 ACKNOWLEDGEMENTS

The views expressed in the *IIC Journal of Innovation* are the contributing authors' views and do not necessarily represent the views of their respective employers nor those of the Industry IoT Consortium®.

- Return to *IIC Journal of Innovation landing page* for more articles and past editions.