METHOD FOR MITIGATING SYSTEM FAILURES

TAKAFUMI NAKAMURA
Fujitsu FSAS Inc., System Support Promotion Unit, Japan

ABSTRACT

In this paper, a method is proposed for mitigating system failures that recognizes the shortcomings of current state-of-the-art methodologies (i.e., crisis management, risk management, normal accident theory, and high reliability organization). The current majority of methodologies for ICT systems use a reductionist approach (i.e., one that lacks a holistic view). Therefore, we need more holistic methodologies to mitigate system failures. There are many system failure examples in our world. The Tokyo Stock Exchange crashed on the 1st of November 2005 because of an operational error, severely impacting the global economy. Such system failures severely impact not only ICT systems but also social systems (transportation systems, nuclear plant systems, etc.). A JR West train derailed and overturned in Kyoto, Japan on the 25th April 2005 because of driver misconduct and caused the loss of 106 passengers lives. On the 11th of March 2011, a massive earthquake fiercely shook eastern Japan; it was followed by a devastating tsunami. This caused the Fukushima No. 1 nuclear plant hydrogen explosion, which will take a long time to clean up. The progress of ICT technologies (i.e., cloud, virtual, and network technologies) inevitably shifts ICT systems into complexity with tightly interacting domains. This trend requires a novel way to mitigate system failures to promote system safety more than ever. Emergent properties should be dealt with in order to promote system safety. Crisis management should focus on holistic properties over partial components. This paper introduces a system failure framework to promote a holistic view to manage and therefore mitigate system failures. An application example of ICT system failures exhibits the effectiveness of this methodology.
Keywords: risk management, crisis management, normal accident theory (NAT), high reliability organization (HRO), information and communication technology (ICT).

1 INTRODUCTION

There are many examples of similar system failures repeating and of negative side effects created by quick fixes. Introducing safety redundant mechanisms does little to reduce human errors. As pointed out by Perrow ([1], p. 260), the more redundancy is used to promote safety, the greater the chance of spurious actuation; "redundancy is not always the correct design option to use." While instrumentation is being improved to enable operators to run their operations more efficiently, and certainly with greater ease, the risk would seem to remain about the same.

Weick and Sutcliffe ([2], p. 81) explained why traditional total quality management (TQM) has failed. "We interpret efforts by organizations to embrace the quality movement as the beginning of a broader interest in reliability and mindfulness. But some research shows that quality programs have led to only modest gain this might be the result of incomplete adoption. But we would go even further, and argue that the reason for incomplete adoption is the necessary infrastructure for reliable practice is not in place even where TQM success stories are the rule. The conclusion is consistent with W.E. Deming's insistence that quality comes from broad-based organizational vigilance for problems other than those found through standard statistical control methods."

There are six stages from the initial stage to cultural readjustment through catastrophic disasters ([3], pp. 88). They are Stage I: Initial beliefs and norms, Stage II: Incubation period, Stage III: Precipitating event, Stage IV: Onset, Stage V: Rescue and salvage, and Stage VI:

WIT Transactions on The Built Environment, Vol 174, © 2018 WIT Press
www.witpress.com, ISSN 1743-3509 (on-line)
doi:10.2495/SAFE170201

Full cultural readjustment. The second stage, or incubation period, is hard to identify because of the various side effects of quick fixes [3], [6]. Therefore, the second stage plays the crucial role that leads to catastrophic disaster. Many side effects due to quick fixes of information and communication technology (ICT) systems have been identified [4], [6]. There are two factors in particular that make it difficult to prevent ICT system failures: the lack of a common language for understanding system failures and the lack of a methodology for preventing future system failures. These shortcomings result in local optimization and the introduction of quick fixes as countermeasures.

This paper aims to mitigate system failures by promoting a holistic approach and by introducing a system failure framework. This approach is novel in that current methodologies tend to focus only on hard science, which leads to myopic management and fails to recognize invaluable ways to recognize the shortcomings of state-of-the-art methodologies (i.e., current methodologies fail to change the status quo).

2 SOSF META-METHODOLOGY AS COMMON LANGUAGE

2.1 System of system failure

The proposed system of system failure (SOSF) meta-methodology for covering all system failure models [4]–[6] is derived from the system of system methodologies (SOSM) [7], [8] and system failure classes. The system of system methodologies classifies the world of objects into two dimensions: systems and participants. The system dimension has two domains: simple and complex. The participant dimension has three domains: unitary, plural, and coercive. Therefore, SOSM classifies the world of objects into six (2×3) domains, and there is an appropriate methodology for each domain. The system of system failure complementarily covers these domains on the basis of this worldview to enable the viewing of object system failures. SOSF uses four domains (excluding the coercive domain because the main focus of this paper is technological systems rather than broader social domains) from SOSM. On top of these four domains, we add a third dimension to identify the person or factor responsible for the system failure. To identify the root causes of failures, we classify system failures on the basis of system boundaries and the responsible system level introduced with the viable system model (VSM) [9], [10]. Failures are classified in accordance with the following criteria [4], [6], [11].

Class 1 (Failure of deviance): The root cause is within the system boundary, and conventional troubleshooting techniques are applicable and effective.

Class 2 (Failure of interface): The root cause is outside the system boundary but is predictable in the design phase.

Class 3 (Failure of foresight): The root cause is outside the system boundary and is unpredictable in the design phase.

System safety can be achieved through the actions of various stakeholders. One such common language was developed by Van Gigch [12], for the taxonomy of system failures. There are six categories of system failures, i) technology, ii) behaviour, iii) structure, iv) regulation, v) rationality, and vi) evolution. In particular, SOSF was designed by allocating each type of failure from this taxonomy [12], into an SOSM meta-methodology space. Fig. 1 shows this space.

There are two widely used failure analysis techniques: failure mode effect analysis (FMEA: [13]) and fault-tree analysis (FTA: [14]). FMEA deals with single-point failures by taking a bottom-up approach, and is presented as a rule in the form of tables. In contrast, FTA

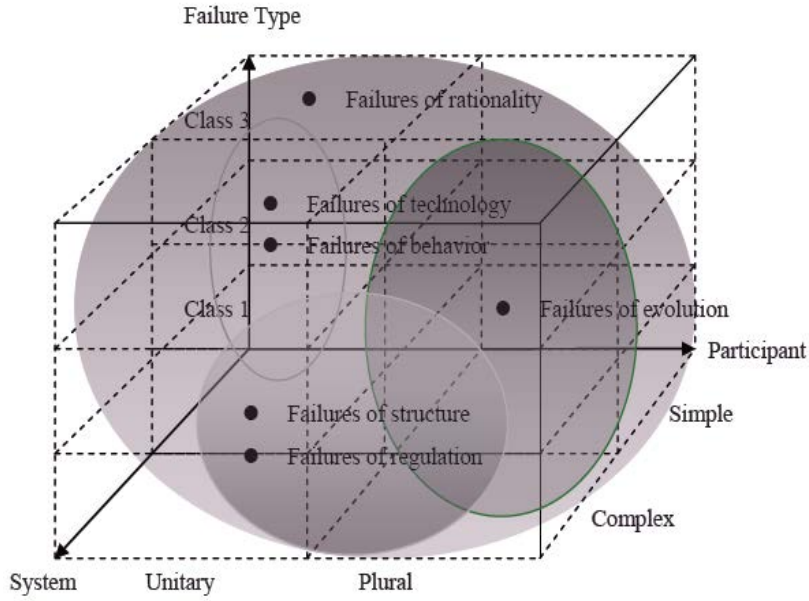

Figure 1: SOSF meta-methodology space.

analyses combinations of failures in a top-down way, and is visually presented as a logic diagram.

Both methodologies are mainly used in the design phase. However, these methodologies are heavily dependent on personal experience and knowledge, and FTA in particular has a tendency to miss some failure modes in failure mode combinations, especially emergent failures.

The major risk analysis techniques (including FMEA and FTA) are explained in ([15], pp. 24–27) ([16], Chapter 4) [17]. Most failure analyses and studies are based on either FMEA or FTA. FMEA and FTA are rarely both performed, though, and when both are done they will be separate activities executed one after the other without significant intertwining.

Current methodologies tend to fail to take a holistic view of the root causes of system failures. And a majority of them stay in the unitary-simple-class 1 domain. It is important to identify and cover the plural-complex-class 3 domain. In the next section, another method is introduced for understanding system failures holistically.

3 IC CHART AND QUANTIFICATION OF RISK FACTORS

3.1 Normal accident theory and IC chart

It is not unusual that several failures happen sequentially or simultaneously. Each is not a catastrophic failure in itself; however, the complex (i.e., unexpected) interaction of those failures may have catastrophic results. Tight coupling of a component involves a cascade of single-point failures that quickly reach a catastrophic end before safety devices come into effect. This is called system failure or a normal accident as opposed to a single-point failure.

Perrow [1], analyzed system failures using the interaction and coupling of system components. This is called normal accident theory.

The IC chart is a table for classifying object systems by interaction and coupling. Fig. 2 shows the IC chart developed by Perrow [1]. Topological expression was done subjectively by Perrow [1]. By combining the two variables in this way, a number of conclusions can be made. It is clear that the two variables are largely independent. Examine the top of the chart from left to right. Dams and nuclear plants are roughly on the same line, indicating a similar degree of tight coupling. But they differ greatly on the interaction variable. Whereas there are few unexpected interactions possible in dams, there are many in nuclear plants. Or, looking across the bottom, universities and post offices are quite loosely coupled. If something goes wrong in either of these, there is plenty of time for recovery, and things do not have to be in a precise order. But in contrast to universities, post offices do not have many unexpected interactions - it is a fairly well laid out (linear) production sequence without a lot of branching paths or feedback loops. The IC chart defines two key concepts, the types of interaction (complex and linear) and the types of coupling (loose and tight). They are laid out so that we can locate organizations or activities that interest us and show how these two concepts, interaction and coupling, can vary independently of each other.

The next section introduces a system failure example that should be dealt with systemically and holistically. In other words, "safety is a system problem."

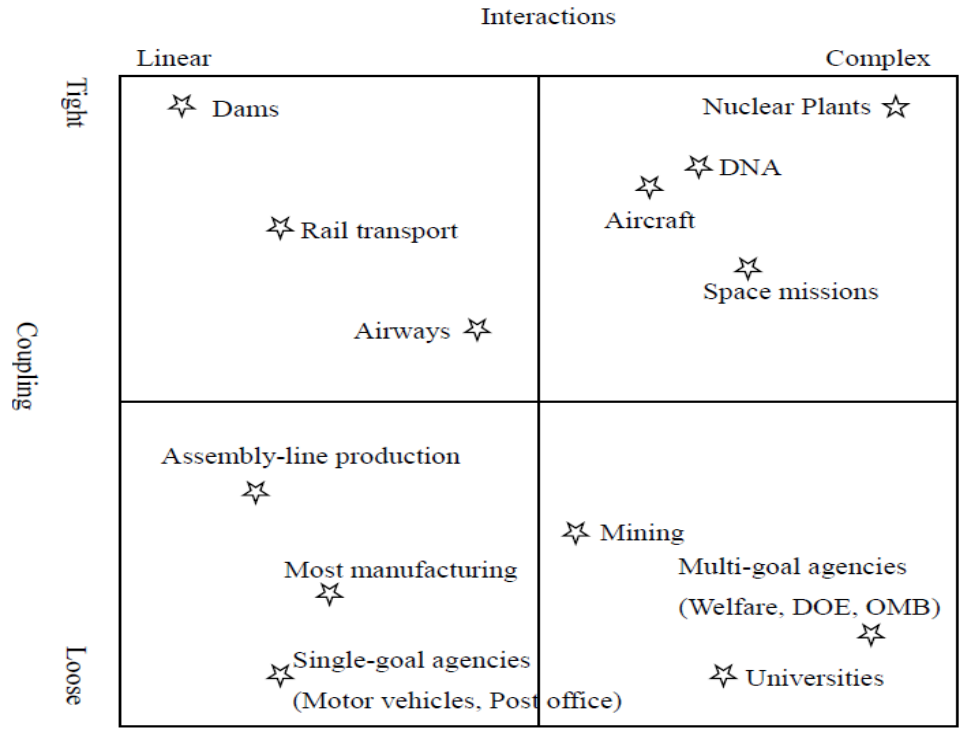

Figure 2: Interaction/coupling chart.

3.2 Safety is a system problem

The safety phenomenon occurs at the organizational and social levels above the physical system as illustrated by Rasmussen's analysis of the Zeebrugge ferry mishap [18] shown in Fig. 3. In this accident, those independently making decisions about vessel design, harbor design, cargo management, passenger management, traffic scheduling, and vessel operation (shown at the left of Fig. 3) were unaware of how their design decisions might interact with decisions made by others, which lead to the ferry accident. Each local decision may be "correct" (and "reliable," whatever that might mean in the context of decisions) within the limited context within which it was made but can lead to an accident when the independent decisions and organizational behaviours interact in dysfunctional ways (portrayed by the intersecting rightward arrows in Fig. 3). As the interactive complexity grows in the systems we build, accidents caused by dysfunctional interactions among components become more likely. Safety is a system property, not a component property, and must be controlled at the system level rather than at the component level [19]. In this situation, modelling activity in terms of task sequences and errors is not very effective for understanding behaviour, so we have to dig deeper to understand the basic behaviour shaping mechanisms [18].

In the next chapter, a system failure framework is introduced for understanding and revealing the current methodologies' blind spots.

4 SYSTEM FAILURE FRAMEWORK

Partial solutions are not enough to promote safety, as explained in the ferry accident example in the previous section. To solve the safety issue, this paper introduces a system failure framework to accommodate the holistic perspective. It consists of two basic dimensions: the horizontal, which pertains to the scope or size of a problem or situation that a person is inherently (instinctually) comfortable in dealing with, and the vertical, which pertains to the kind of decision-making processes that a person inherently (instinctually) brings to bear on a problem or situation. The framework is important because it shows that, for the how and why on any issue or problem of importance, there are at least four very different attitudes or stances with regards to the issue or problem. None of them is more important or right, so we need to check all perspectives intentionally in order to overcome psychological blind spots. Fig. 4 shows the framework.

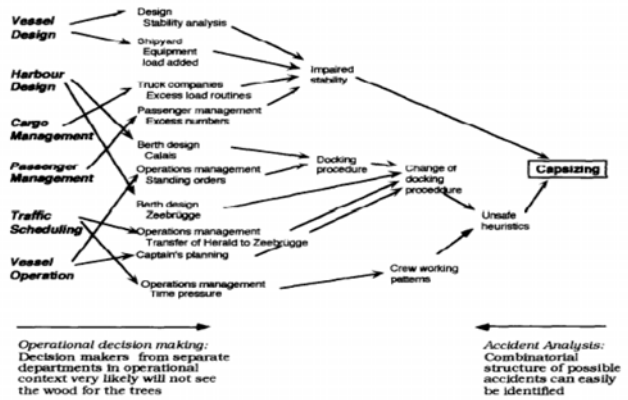

Figure 3: Complex pattern of the Zeebrugge accident.

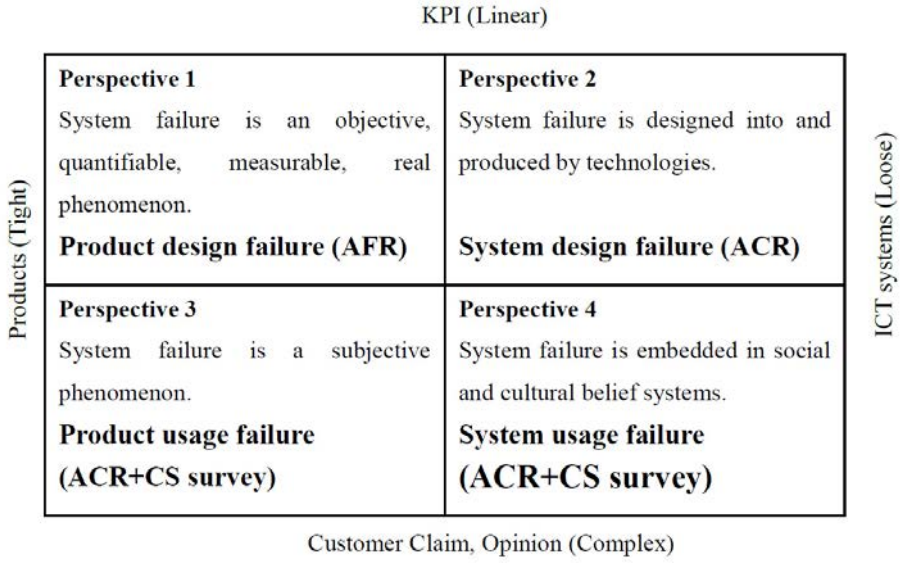

Figure 4: System failure framework.

So far, this paper has introduced three perspectives to promote system safety. An isomorphic structure is depicted between the three perspectives to promote further holistic views. table 1 shows the isomorphic structure. The combination of these three perspectives promotes various perspectives to learn from previous system failures. The following hypotheses are derived from the above discussion.

- A loose domain has fewer system failures than a tight domain.
- A complex domain has a greater ratio of class 2 and 3 failures than a linear domain.

4.1 Application for ICT system failures

The Information-technology Promotion Agency, Japan (IPA) issued a guide book to promote the system reliability of mission critical ICT infrastructures in 2011 on the basis of an analysis for 43 system failure examples and provided a system check list to enhance system reliability [20]. The above hypotheses were confirmed by applying the system failure framework to the output of the study from IPA. The application results are discussed in the next chapter. Tables 2 and 3 are the 43 system failure examples and 54 check items and their classifications according to the system failure framework.

5 RESULTS

The 43 system failure examples can be classified in an IC chart. Figs 5, 6, and 7 show the results of this classification. All of the system failures that have severe social impacts belong to the tight domain (see Fig. 5).

1. Perspective 2 has the greatest number of failure causes (29) followed by Perspectives 1, 4, and 3.
2. The majority of the failure causes in perspective 1 are class 1 (88%).
3. The majority of the failure causes in perspective 2 are class 1 (34%) and .2 (55%).
4. All of the failure causes in perspective 3 are class 2.
5. Perspective 4 has the greatest number of class 3 failure causes (33%) followed by perspective 2 (10%). In other words, there are no class 3 failure causes in the tight domain.
6. The majority of the check items from perspectives 3 and 4 (i.e., complex domain) are qualitative check items (8 items out of a total of 12 items are qualitative in Table 3). (See Fig. 7.)

Table 1: Isomorphic structure between three perspectives.

	IC chart	System failure framework	System of system failure
Vertical Axis	Linear	KPI	Simple
	Complex	Customer Claim, Opinion	Complex
Horizontal Axis	Tight	Product	Unitary
	Loose	ICT systems	Plural

Table 2: 43 system failure examples and their classifications.

#	System Area Quadrant	Root Cause Quadrant	Root Cause Class
1	1	2	2
2	1	2	1
3	1	2	2
4	1	2	1
5	1	2	1
6	1	1	1
7	1	1	1
8	1	1	1
9	1	2	2
10	1	1	1
		3	2
11	1	2	2
12	1	1	1
		2	2
13	1	1	1
14	1	1	1
15	1	1	1
16	1	1	1
17	1	2	2
18	1	1	1
		2	2
19	1	1	1
20	1	4	3
21	1	1	1

Table 2: Continued.

#	System Area Quadrant	Root Cause Quadrant	Root Cause Class
22	1	1	2
23	2	1	2
		2	1
		2	2
		2	2
		2	2
		2	3
		4	1
24	1	1	1
25	1	2	1
26	1	2	2
		4	3
27	1	3	2
28	2	2	1
29	1	2	2
30	1	2	3
31	1	2	2
32	1	1	1
33	2	4	2
34	1	2	1
35	1	2	3
36	2	4	1
37	2	2	1
38	1	2	2
		4	3
39	1	4	2
40	1	2	1
41	1	1	1
42	1	2	1
43	2	2	2
		2	2
		4	1
		4	1

Table 3: 54 check items and their classifications.

Check #	Quadrant #	Quantitative KPI
1	2	○
2	2	
3	1	○
4	1	
5	2	
6	2	
7	2	
8	2	
9	2	
10	2	
11	2	
12	4	

Quadrant #	# of items	# of ratio
1	23	43%
2	19	35%
3	4	7%
4	8	15%

Table 3: Continued.

Check #	Quadrant#	Quantitative KPI
13	2	○
14	2	○
15	2	○
16	2	
17	2	
18	2	
19	1	
20	2	
21	2	
22	2	
23	1	
24	4	○
25	4	
26	1	○
27	1	○
28	1	○
29	1	○
30	1	○
31	3	○
32	4	
33	1	
34	1	
35	4	
36	1	
37	1	○
38	1	
39	3	○
40	3	○
41	1	
42	3	
43	1	
44	1	○
45	1	
46	1	○
47	1	○
48	1	
49	2	○
50	1	○
51	1	○
52	4	
53	4	
54	4	

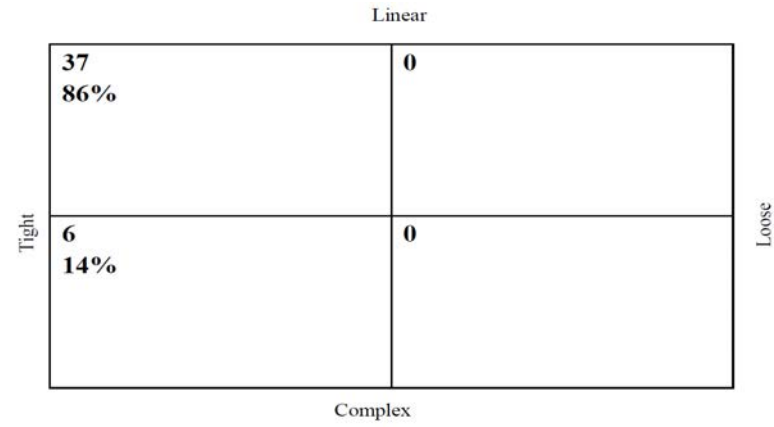

Figure 5: IC chart analysis of system failure.

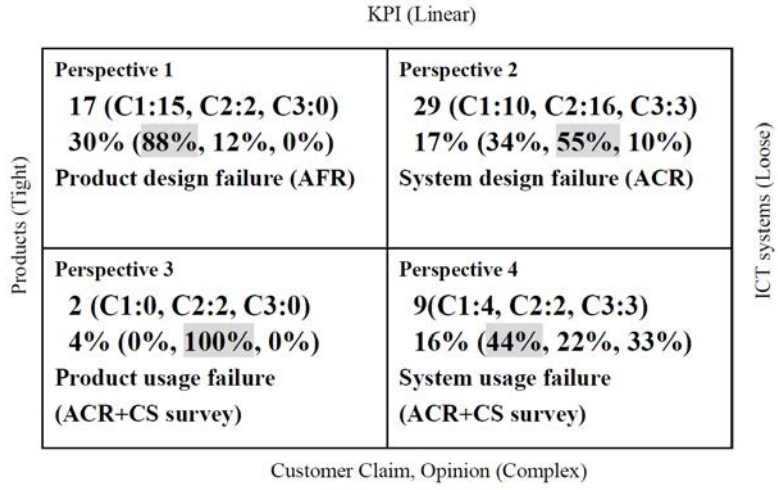

Figure 6: System failure framework (root causes).

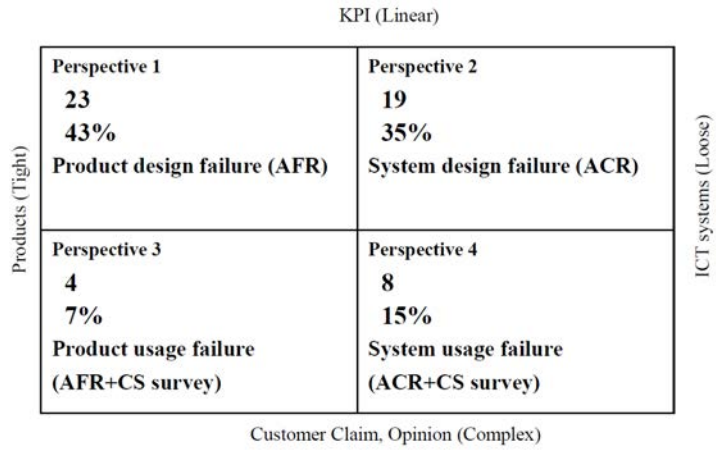

Figure 7: System failure framework (check items).

Fig. 6 shows the results of the analysis of system failure root causes. Cn represent a class n failure of SOSF (n: 1~3).

Fig. 7 shows the results of the analysis of the system failure check list. The majority of check items are located in perspectives 1 (23 items in Table 3) and 2 (19 items in table 3) (i.e. linear domain). And the check items have a quantitative KPI (17 items out of a total of 21 items are perspectives 1 and 2 in Table 3).

The examples of quantitative KPIs are a human error ratio, bug numbers within 1 k steps of an application program, and so on. They can be measured by numerical values. On the other hand, the examples of qualitative KPIs are an involvement of wider stakeholders at the time of the root definition phase, an agreement on a goal of ICT systems with relevant stakeholders, and so on. They cannot be measured by numerical values.

6 CONCLUSIONS

The results of this application suggest that further research should be done to verify the necessity and sufficiency of current methodologies, especially in perspectives 3 and 4, by mapping them onto the system failure framework. This provides us with holistic views to promote system safety as well as to understand current methodologies' blind spots and to overcome myopic management.

Further research should be conducted with the following approaches.

1. Define the research area.
2. Identify the system failure incidents with unknown causes.
3. Identify the current state-of-the-art methodologies to mitigate system failures.
4. Map them onto the system failure framework.
5. Evaluate necessary and sufficient conditions of the current methodologies to cover 2).
6. Develop appropriate methodologies to overcome current blind spots.

REFERENCES

[1] Perrow, C., Normal Accidents: Living with High-Risk Technologies. Princeton Paperbacks: New York, 1999.
[2] Weick, K.E. & Sutcliffe, K.M., Managing the Unexpected: Assuring High Performance in an Age of Complexity (J-B US non-Franchise Leadership), 2001.
[3] Turner, B.A. & Pidgeon, N.F., Man-Made Disasters 2nd ed., Butterworth-Heinemann, UK. 1997.
[4] Nakamura, T. & Kijima, K., System of system failures: Meta methodology for IT engineering safety. *Systems Research and Behavioural Science*, **26**(1), pp. 29–47. 2009.
[5] Nakamura, T. & Kijima, K., Meta system methodology to prevent system failures. Proceedings of the 51st Annual Meeting of the ISSS in Tokyo, 2007.
[6] Nakamura, T. & Kijima, K., Total system intervention for system failures and its application to ICT systems, 2010. http://journals.isss.org/index.php/proceedings54th/article/viewFile/1436/519; https://www.igi-global.com/article/total-system-intervention-system-failure/58369; http://www.irma-international.org/chapter/total-system-intervention-system-failure/76230/.
[7] Jackson, M.C., Systems Thinking: Creative Holism for Managers. John Wiley & Sons: London and New York, 2003.
[8] Jackson, M.C., Creative Holism: A critical systems approach to complex problem situations. *Systems Research and Behavioural Science,* **23**(5), pp. 647–657. 2006.
[9] Beer, S., The Heart of Enterprise. John Wiley & Sons: London and New York. 1979.
[10] Beer, S., Brain of the Firm, 2nd ed., John Wiley & Sons, London and New York. 1981.
[11] Nakamura, T. & Kijima, K., A methodology to prolong system lifespan and its application to IT systems. *Proceedings of the 53rd Annual Meeting of the ISSS in Brisbane*, Jul. 2009.
[12] Van Gigch, J.P., Modelling, metamodeling, and taxonomy of system failures. *IEEE Trans. on Reliability*, **35**(2), pp. 131–136, 1986.
[13] IEC 60812. Procedure for failure mode and effect analysis (FMEA),2006.
[14] IEC 61025. Fault tree analysis (FTA), 2006.
[15] Bell, T.E., ed., Special Report: Managing Murphy's law: engineering a minimum-risk system", IEEE Spectrum, pp. 24–57, 1989.
[16] Wang, J.X. & Roush, M.L., What every engineer should know about risk engineering and management. Marcel Dekker, Inc, 2000.
[17] Beroggi, G.E.G. & Wallace, W.A., Operational Risk Management: A New Paradigm for Decision Making, *IEEE Transactions on Systems, Man and Cybernetics*, **24**(10), pp. 1450–1457. 1994.
[18] Rasmussen, J., Risk management in a dynamic society: a modelling problem, *Safety Science*, **27**(2),(3), pp. 183–213.1997.
[19] Leveson, N., Dulac, N., Marais, K. & Carroll, J., Moving Beyond Normal Accidents and High Reliability Organizations: A Systems Approach to Safety in Complex Systems. http://sunnyday.mit.edu/papers/HRO-final.doc, 2009.
[20] Information-technology Promotion Agency, Japan (IPA), System reliability promotion guide book. Retrieved 5th Jun. 2017. http://www.ipa.go.jp/files/000004556.pdf.