Design of resilient socio-technical systems by human-system co-creation

Tetsuo Sawaragi¹

Received: 18 July 2019 / Accepted: 26 March 2020 / Published online: 8 April 2020
© The Author(s) 2020

Abstract

The concept of resilience is commonly used to represent “an ability to adapt to a changing environment and to survive flexibly despite facing difficulties.” This definition of resilience emphasized its aspects of unpredictable responsiveness to external disturbances as well as self-organized phenomena, which seemed to be close to the definition of complex adaptive systems. This article summarizes the technical challenges to consider the implementation of a concept of resilience into systems from a variety of perspectives. Then, challenges to tackle with the resilience to variabilities in production plans, in work quality, in empirical knowledge, in human-automation systems, and in organizations are presented.

Keywords Resilience · Socio-technical systems · Steel production · Ecological interface design · Autonomous error recovery · Systemic accident analysis · Socially centered automation · Team situation awareness · Serious game

1 Introduction

Today, the scene of manufacturing in Japan needs increased flexibility to diverse “variabilities,” which are caused by the unexpectedness due to market transformation, overseas expansion of businesses, and changes in the energy and economic environments, and so on. In addition, variabilities caused by degraded product quality due to the retirement of skilled workers and the dysfunctions of old manufacturing equipment have caused incidents that may lead to severe accidents. To deal with various such variabilities, co-creation by human beings and systems is indispensable. Thus, it is needed to develop novel systemization technologies that (1) analyze vulnerability that systems have in the event of unexpected diseases, and (2) conducts stress tests aiming at prevention of severe accidents not only on hardware but also

on the entire activity conducted by human beings, machines, and organizations.

The concept of resilience has prevailed in diverse fields. Initially, it comes from physics of material science, where resilience is the ability of a material to absorb energy when it is deformed elastically and release that energy upon unloading. Then, psychological resilience comes out as the ability to cope with a crisis or to return to pre-crisis status quickly. Wherein, resilience exists when the person uses mental processes and behaviors in promoting personal assets and protecting self from the potential adverse effects of stressors. Holling, a Canadian ecologist and a former director of IIASA (International Institute for Applied Systems Analysis), proposed to use “resilience” in the field of social science due to that ecological systems had an ability to restore the original state upon environmental changes, which is similar to the homeostasis that living organisms have (Holling [1]). This definition of resilience emphasized its aspects of unpredictable responsiveness to external disturbances as well as self-organized phenomena, which seemed to be close to the definition of complex adaptive systems.

In every field, the term of resilience has been commonly used to represent “an ability to adapt to a changing environment and to survive flexibly despite facing difficulties, and to recover from a destabilizing perturbation in the work as it attempts to reach its primary goals.” Furthermore, the concept of resilience has received increased attention as abilities

This work was presented in part as a plenary speech at the 22nd International Symposium on Artificial Life and Robotics (Beppu, Oita, January 19–21, 2017). The authors submitted the modified draft in response to the invitation.

✉ Tetsuo Sawaragi
sawaragi@me.kyoto-u.ac.jp

¹ Department of Mechanical Engineering and Science, Graduate School of Engineering, Kyoto University, C Cluster C3 Kyoto Daigaku Katsura, Nishikyo, Kyoto 615-8540, Japan

for risk-aversion and crisis control in every level of societies, from individuals to entities such as businesses and governmental organizations.

In Japan, the realization of the ‘super-smart society’ of “Society 5.0” is being promulgated as the 5th Science and Technology Basic Plan by the Japanese Government (FY2016-FY2020) (CAO [2]). The following three issues characterize this initiative:

1. Realizing smart manufacturing by utilizing IoT and network technologies.
2. Connecting and fusing cyber space and physical space (Cyber-Physical System).
3. Systemization of service and enterprise integrating a variety of system elements.

Wherein, the key technologies are cyber security, IoT, Big Data, and Artificial Intelligence. Furthermore, as techniques having a strength that lies new value creation, robotics, sensors, biotechnologies, materials and nanotechnologies etc. are mentioned. Besides, the super-smart society is aiming to resolve various social challenges by incorporating the innovations of technologies and the sharing economy into every industry and social life. Such innovations will necessitate elaborating the SoS (System of Systems) (Selberg and Austin [3]) approach towards the value co-creation in the whole society integrating several kinds of systems, such as nature, society, biology, and humanity. Such mutually connected systems would face a phenomenon in which some small malfunctions cause large catastrophic failure chains. Thus, they are apt to become extremely responsive to random faults. Therefore, it would be significant to model and understand the dynamical complexity of the SoS and to develop a technique for guaranteeing their resilience.

According to these trends, in this article, resilience is discussed from the perspectives of safety management for the socio-technical systems, in which technical, human, and organizational factors do interact with each other emerging complicated behaviors leading to the severe accidents. This article summarizes the technical challenges to consider the implementation of a concept of resilience into systems from a variety of perspectives; (1) development of systemization technology for risk prediction, (2) development of systemization technology for the response to emergent situations, and (3) development of systems that foster awareness for safety and human resources. Then, challenges to tackle with the resilience to variabilities in production plans, in work quality, in empirical knowledge, in human-automation systems, and in organizations are presented.

2 Socio-technical systems as large-scale complex systems

2.1 Systemic model for socio-technical systems

Socio-technical systems (STSs) initially emerged from research at the Tavistock Institute, a British not-for-profit organization, in the 1950s on the effects of the introduction of powered machines on the work, management-labor relations, and the lives, families, and societies of coal miners (Hoffman and Militello [4]). Since then, socio-technical systems theory has enjoyed around 60 years of development and application internationally by both researchers and practitioners. The over-arching philosophy, embracing the joint design and optimization of organizational systems (incorporating both social and technical elements), has maintained its practical relevance and has seen increasing recognition and acceptance by audiences outside the social sciences (Eason [5]).

The socio-technical systems are now studied in more diversified fields. The following summarizes the socio-technical systems in the field of “cognitive engineering” that was proposed by Hollnagel and Woods, putting a strong focus on the trends in analyzing accidents caused by organizational factors as well as in safety design, within interacting systems (Hollnagel and Woods [6]).

At work sites, such as production sites and medical sites, workers have been experiencing drastic changes in their work, resulted from the introduction of new equipment and systems, as well as ever-changing clients’ needs. Such changes in work do occur from the interaction between the site and the surrounding environment and entity. Namely, entities have multiple communities of practice, in which some artifact objects such as procedures or mechanical systems have been developed omitting the communities’ practice. Then, such entities may be rejected by the communities or may disturb the communities’ activities, which sometimes can lead to severe accidents.

Thus, currently, safety assessment for the socio-technical system is needed as many incidents are induced by the interaction among factors of technology, individual, society, management, and organization, as well as the complex and dynamic aspects of these factors. Concerning with this, recently a new accident pattern, systemic accident, has arisen, in which accidents are not directly caused by a significant deviation from safety procedures, but by amplification and resonant interaction of accumulated performance variability in a particular work procedure. Accordingly, reliability assessment methodologies should evolve from traditional ones to advanced ones. That is, it is necessary to put a focus on the performance shaping factors that depend on the relevant context, the dynamic

aspect resulted from the interaction between the human being with the situation, the interdependency among the influencing factors, and the overall impacts, by including a modeled human internal mechanism. There proposed the three factors that characterize socio-technical systems, T (technology), O (organizations), and P(people), and almost any technology today is both designed and used with implicit or explicit relations between those three factors. These relations form a triangle, as shown in Fig. 1 (Brandt and Černetič [7]). This triangle is a useful reminder that almost any (advanced) technology today is embedded in a sort of socio-technical system.

Concerning this, Hollnagel has offered the eleven factors of the Common Performance Conditions (CPC) (Hollnagel et al. [8]) that is listed in Table 1. These factors are what commonly affect the formation of human performance in a socio-technical system, and when some of these factors are inappropriate, some deviances from the normal state may occur. They are the factors for variabilities causing performance variability. He has also proposed an analysis method of Functional Resonance Analysis Method (FRAM) to deal with the systemic accident, where the potential hazards in designed procedures can be detected against the possible disturbances caused by those CPCs (Hollnagel [9]).

2.2 Activity theory for socio-technical systems

Another approach to analyzing socio-technical systems exists in the field of social science, called activity theory. This theory originated from the psychology studied by Vygotsky et al. and Engeström developed further [10]. Wherein, the core principle exists in; “activity as a unit of analysis,” “object orientation,” and “mediational properties.” Namely, the activity is not defined by a bilateral relationship between a human being and an object, but by a trilateral relationship in which a tool mediates a human and an object

Fig. 1 T (technology), O (organizations), and P(people) quoted from Brandt and Černetič [7]

Table 1 CPCs as sources of variabilities

Common performance conditions	TOP
Availability of resources	(P, T)
Training and experiences	(P)
Quality of communication	(P, T)
HMI and operational support	(T)
Access to procedures and methods	(P)
Conditions of work	(T, O)
Number of goals and conflict resolution	(P, O)
Needed time/available time relation	(P)
Circadian rhythm	(P)
Crew collaboration quality	(P)
Quality and support of organization	(O)

T technology, *O* organization, *P* people

(Fig. 2a). Here, a tool includes not only a physical tool but also signs, language, and concepts.

Furthermore, activities are understood to have a mutually linked structure, in which the three factors (communities, rules, and division of labor) are added to trace the alternation and development of activities as “a succession of contradictions” generated within this structure and among multiple activities. Figure 2b shows this framework. Each node represents factors making up the activity, where Object is assumed to produce Outcome. The essence of this scheme is that the essential task is always to grasp the systemic whole, not just separate

Fig. 2 Elements of socio-technical systems by activity theory (Engeström [10])

Fig. 3 Possible contradictions generated in each of elements and among those. Numbers in the figure correspond to the levels mentioned in the text (Engeström [10])

connections. Generated contradictions might appear at the following four levels that are shown in Fig. 3:

Level 1: Primary inner contradictions (dual nature) within each constituent component of the central activity.

Level 2: Secondary contradictions between the constituents of the central activity.

Level 3: Tertiary contradictions between the motive of the dominant form and the one of a culturally more advanced form.

Level 4: Quaternary contradictions between the central activity and its neighbor activities.

Wherein conflicts and misunderstandings may emerge among these activity systems. The difference between Level 3 and Level 4 might be unclear. Level 3 connection represents a gap between the two activities; the lower one is an activity that has already taken root and the upper one is a more culturally advanced activity that someone tries to introduce to the former. At this time, Level 3 conflict does emerge. This conflict may appear, for instance, when an advanced automaton is introduced to a work site, where the main activity has been done fully manually. Although the motive of introducing automation is the same as the one sought for by the manual task, this also force changes of the conventional connections with the other activities that are Level 4 conflicts.

For the discussion of the resilience, it is critically important to understand that a human activity evolves within an organization and work community through the alternation and development of activities triggered by a series of contradictions generated in this structure and among multiple activities.

2.3 Systemic aspects at manufacturing sites

When looking at the industry, many manufacturing sites are driven to bear significant vulnerabilities that can lead to system malfunctions due to that minute performance variability. Those malfunctions do emerge being triggered by external disturbances and being amplified by the complex interactions and coincidences of such external disturbances. Such a situation is a result of the excessively streamlined work processes and the subsequent loss of margin to the safety borderlines without being noticed. Current manufacturing sites are obliged to challenge their limits because of not only needs for quality and cost but also needs for low-carbon and high-efficiency production of fossil fuels.

For the management of such a situation, it is needed to monitor the deviancies and the variations of the system performance closely during the regular operation, before they are detected as errors, as well as to have the capability to predict the spread of the impact of such deviancies and variations appropriately in case they develop into abnormal variations. Furthermore, production plans should be flexibly changed during operation as well as that the actor governing the control of the entire equipment should be altered between the human being and the automated system adaptively to avoid system failures or breakdowns.

It is also needed to establish an adaptive system that is able to ensure enough durability for maintenance of the regular manufacturing capacity as far as possible despite facing (1) restrictions or external disturbances in out-factory logistics or in-factory manufacturing equipment due to natural disasters; or (2) abnormal inputs due to diverse skill level of individual workers. Such systems also should be able to rapidly restore the regular manufacturing capacity even when being temporally disturbed. These systems are precisely resilient systems.

Key concepts to discuss resilience are “Safety-I and Safety-II” (Hollnagel [11]). Safety-I refers to a stable state in which “there are no unacceptable risks” or “undesirable states never happen.” Safety-I has been a general concept of safety. On the other hand, Safety-II refers to a dynamic and proactive concept, which takes into account the response to external disturbances or failures, and in which “systems can continue to operate without falling to catastrophic state.” The absolute safety does not mean just a stable state in which there happen no failures, but also “ability to maintain the success under an ever-changing condition.” In this context, Safety-II is equivalent to a “dynamic non-event.”

2.4 Sources of the performance variability

Sources of the performance variability as fluctuation factors in manufacturing plants are summarized as follows.

(1) Variabilities in production plans:

Failure or deterioration of the manufacturing equipment, delay of delivery, changing stock, and customer churn, etc. are unavoidable.

(2) Variabilities in a complex production supply chain:

A vast supply chain across the diverse manufacturing processes constituted by the planning, operation, and maintenance systems, wherein multiple processes of each system are organically linked.

(3) Variabilities in empirical knowledge and work quality:

The occurrence of product quality variability as skilled workers leaves their workplaces because of their aging or labor shortage.

(4) Variabilities in organizations:

A large number of organizations are now deepening their prediction and preparation from top to bottom to address every event with facility designs and operation procedure documents. That is, this is a style of a traditional CCCI (Command, Control, Communication & Intelligence). On the other hand, as for safety an alternative style of CCTI (Craftsmanship, Connectivity, Trust & Inspira-

tion) does exist and is still in dispute which style is valid for responding to occurrences of unexpected events.

For the discussion of the implementation of the resilience into systems, the following section mainly summarizes the following technical challenges.

- (1) Predicting potential risks in the socio-technical systems,
- (2) Appropriately taking measures when an abnormality occurs, and
- (3) Nurturing human resources.

Figure 4 shows the relationship between the above three research challenges. The typical background idea is that establishment of resilience requires the co-creation between human beings and systems (Sawaragi [12]). It is needed to design systems that derive and cultivate the existing empirical knowledge, which requires to provide adequate constraints and supports with workers and to derive the workers' responses accordingly. Not by eliminating human beings or ultimately limiting their work, but by pursuing system

Fig. 4 An overview of the human-system co-creative design of resilience (Sawaragi [12])

design that allows some human judgment left in the loop, resilience can be established through the fusion of knowledge of human beings and automation.

3 Approaches to design resistant socio-technical systems

3.1 Issues to be addressed according to activity theory

In this section, the author provides the issues of resilience in manufacturing, which are instantiated from the general scheme of Engeström [10] of Fig. 2b.

- “Subject” is a human worker
- “Instruments” is an automated device
- “Object” is work quality attained (“Outcome” is qualified and accredited work quality)
- “Community” may include a team consisting of multiple workers and cognitive agents
- “Rules” are manuals and procedures
- “Division of Labor” is role assignments among multiple workers and between a human and automation.

The overall architecture represents an activity as a socio-technical system. Resilience should be discussed at all levels of contradictions, as mentioned in 2.2. According to this overall scheme, research topics ongoing at the author’s group are presented in the following subsections.

3.2 The resilience of subject: ecological interface design

The operation of trains, which transport mass passengers and cargos, requires high safety and accuracy. Because the demand level for the accuracy of train operation is high in Japan, even if the disruption to the schedule is slight, the driver of the train bears an absolute psychological pressure, which may result in the driver’s error in judgment.

Train drivers have to take into account not only the safety and accuracy of train operation but also other diverse factors. However, adherence to the operation schedule and train stopping position widely varies between individuals. One of the countermeasures against such variations in operational performance can be the automatization of train operation. However, when a train operation is automated, the train drivers become responsible for the response to situations that are not considered in the automated operating systems. On the other hand, the automatization of the train operation deprives the drivers of the valuable opportunities to acquire the appropriate situational awareness, by excluding the drivers from the control loop of the train operation.

As what replaces such automatic operation systems, the author and colleagues successfully developed a system that could decrease variabilities in train operation performance by promoting the drivers’ capabilities for judgment in complex situations with the adequate visualization of information (Kudo et al. [13]). This system assumes for use in an operation support interface, which helps the drivers cognitively. This interface is based on an idea of ecological interface design (EID) (Burns et al. [14]). To guarantee an operator’s intuitive activities of monitoring, estimating and responding, specifically for complex sociotechnical, real-time, and dynamic systems, the goal of EID is to make constraints and complex relationships in the work environment perceptually evident (e.g., visible, audible) to the user (i.e., the operator). EID allows more of users’ cognitive resources to be devoted to higher cognitive processes such as problem-solving and decision making.

EID proceeds through the work domain analysis and the task analysis, then the constraints and relationships of the work environment in a complex system are identified, which are reflected perceptually (through an interface) to shape user behavior. The author’s group designed two kinds of interfaces, one of which was based on the constraint-based approach, and the other is on the instruction-based approach. As shown in Fig. 5, the former interface is to visualize the input-out relationships of the train driver’s operations as well as the constraints of the boundaries of the speeds for the safe operation of the train. On the other hand, the latter interface guides the train driver by instructing what operations to take and in which timing to take sequentially. That is, a run curve of the driving profile optimizing the train operation is calculated and is displayed to the driver, and the driver is forced to operate the train so that his/her operation history could coincide with that curve. As compared with the former, with this instruction-based approach it would be easier for the train driver to directly know what to do next, while there left less freedom for the train operator to operate differently from what is instructed.

Comparative analysis of the performances using both types of interfaces is done from the viewpoint of the acquisition skill of the driver and the ability to cope with variabilities. That is, the influence of the suspension of the support by the interface and the delay of the departure time on the driving performance was investigated. As a result, in the instruction-based approach, it was confirmed that the judgment of the train driver strongly depends on the presented information in the display. On the other hand, there is no such strong dependency in the constraint-based approach, and it is suggested that the train driver has an opportunity to make an operational judgment independently, and may make up his/her judgment criteria. In addition, it was revealed that there is a danger that the instruction-based approach promotes driving at a higher speed than necessary and

Fig. 5 **a** Constraint-based interface, **b** instruction-based interface, **c** a scene of the experiments (Kudo et al. [10])

reduces the safety of operation if a departure delay occurs. On the other hand, it was confirmed that the constraint-based approach promotes flexible responses that use slower speeds to adjust for on-time arrival by referring to the displayed lower speed limit.

3.3 The resilience of instruments: autonomous error recovery

Introduction of robots in manufacturing plants has been attempted actively to build an automated production line for high mix/low volume production. Wherein, it is crucial for industrial robots to prevent “temporary stop” caused by various factors to improve productivity as well as to reduce the production costs. Under the present circumstances, automated error recovery of robotic manipulators is restricted only to “reactive” re-execution of failed commands. However, the validity of this recovery strategy is quite restricted, since the work state is sometimes changed significantly caused by an error especially in an occasion of cooperative work by multiple robots when simply iterating the command cannot recover from the error to a normal procedure.

To tackle this problem, the author’s group considers a more intelligent robotic system which has the capability of “deliberate” error recovery. They are building a hierarchical planning system for autonomous error recovery of multiple

robotic manipulators (Matsuoka et al. [15]). The proposed deliberate layer can modify an execution plan taking account of the effects of detected errors. They are applying classified Conceptual Graph to represent semantic information about work state. They are proposing a repair strategy to find out the required operations to recovery. Simulations of the proposed system embedded into virtual assembly robots have shown that the system can modify and update an execution plan to achieve a goal state during work against unexpected errors. As an example, Fig. 6 illustrates how the algorithm works. Figure 6a, b represents the initial state and the goal state, respectively. In Fig. 6, subscripts f, i, and t represent a feeder position, an insertion position, and a work-transfer position of the individual workpieces of (a), (b) and (c), respectively.

For the workpiece (b), there is a place b_i for transferring between robots. For the workpiece (c), since there is little room for insertion space and the posture is required to be accurate, the two robots collaboratively support it at the time of insertion. It is necessary to achieve the three sub-goals of inserting parts (a), (b), and (c). The delivery is realized by synchronizing [release] and [grasp] of the two robots and inserting the workpiece (c) by synchronizing the two commands of [release]. Two robots cannot enter at a time except for b_i and c_i . Furthermore, there is a restriction that only arm #1 reaches b_i and arm #2 only reaches b_f . From the spatial

Fig. 6 An autonomous error recovery system (Matsuoka et al. [15])

constraints involved in this work, robot #1 does not have a command corresponding to “move(b_f),” and robot #2 places the workpiece (b), and robot #1 receives and inserts it.

For this planning problem, the system derives the initial planning as a schedule, as shown in Fig. 6c, and according to this schedule, the system started to execute this plan, but when robot #2 tried to execute “move(b_f),” it failed in, and an error occurred as shown in a red arrow in Fig. 6d. In response to this, the system inferred the cause of the error, and according to the repair strategy, it derived an alternative plan shown in Fig. 6d autonomously. The resulted optimal schedule generated by the system is shown in Fig. 7, where the original error-free optimal plan is also shown. In this way, the expansion of the automated recoverable range of a robotic system and reduction of manual intervention is expected. Figure 8 illustrates a testbed system, into which our error recovery algorithm is embedded.

3.4 The resilience of rules: stress check for work procedures

To achieve Safety-II, it is needed to clarify the root cause underlying each accident. The previous safety analyses have focused on unsafe acts such as human errors and breach of rules by front-line workers. These unsafe acts directly affect the system safety and are characterized by that the negative impacts of such acts become conspicuous relatively quickly (active failure). On the contrary, there are potential factors, of which impacts do not become apparent quickly and lay hidden without bringing any harm, but afterward can destroy the system defense by interacting with the local environment (latent condition). People who work within a complex system commit unsafe acts, such as errors or breach of rules, by some reasons that cannot be explained with the psychology studying an individual. Such reasons continue to be hidden within organizations and never become apparent unless

time	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48
robot #1	move(a)			grasp(a)					move(a)			release(a)				move(b)					<u>grasp(b)</u>		move(b)		release(b)			move(c)		grasp(c)																		
robot #2	move(b)				grasp(b)			move(b)													<u>release(b)</u>		move(c)		grasp(c)			move(c)																				

(a)

time	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	
robot #1	move(a)			grasp(a)				<u>move(a)</u>			<u>release(a)</u>				move(a)					release(a)		move(b)		<u>grasp(b)</u>		move(b)		release(b)																					
robot #2	move(b)				grasp(b)			<u>move(b)</u>													move(b)																												

error detection → reset

(b)

Fig. 7 a Original error-free optimal plan, b error-recovery optimal plan (Matsuoka et al. [15])

Fig. 8 A picture of the testbed system

facing abnormal conditions resulted from unsafe acts or others. Such latent conditions do not lay behind statically but change with time by interacting with each other below the surface. The processes of such changes may lead to changes in front-line works, as well as to accidents in the worst scenario. That is, unexpected variabilities cause systemic accidents during work processes that are accumulated and amplified by interacting with each other in a resonant manner. There are many factors in the environment surrounding human–machine systems, and each factor is fluctuating. By being added the variabilities in the environmental factors, the variabilities in the collaboration work between human beings and machines become resonant with these added variabilities. As a result, the variabilities that are initially not large enough to emerge as signals may be developed to come out as deviations in human operation or machine failures, and eventually become amplified to serve as direct accident causes (Sawaragi et al. [16]).

The author and his colleagues have proposed an analysis methodology to resolve such systemic accidents using the Functional Resonance Analysis Method (FRAM) proposed by Hollnagel [9]. Their newly proposed method was developed by integrating the method of FRAM and Fuzzy CREAM, which was applied to actual air crash accident accompanied by the automation of high-tech airplanes near Cali Airport, Colombia in 1995, which was the first fatal accident of high-tech aircraft B757 in 13 years of whose exemplary service at that time (Hirose et al. [17]). The algorithm for safety analysis is shown in Fig. 9.

American Airline flight 965 was about to land at the airport, which is close to Cali, Colombia in 1995. The flight was already 2 h behind due to the departure delay at Miami, and it was dark outside. During the approach, the flight crew was proposed the runway change for landing by ATC (Air Traffic Controller) and accepted it. However, just after accepting the proposal of runway change,

Fig. 9 An algorithm for safety analysis proposed in Hirose et al. [17]

the crews of flight 965 became busy in identifying a new approach course. Moreover, they entered the wrong course for landing to the Flight Management Computer (FMC), and flight 965 went off the correct course, which the crew did not notice for a while. Finally, flight 965 flew into the mountains of Andes and finally crashed into the terrain. The details of the accident are in public in (Simons [18]).

Automation is introduced to reduce the workload of humans and improve the accuracy of task performance. However, it also brought about the change of operator's tasks, and this might cause accidents which have never experienced before. These accidents are mainly due to the discrepancy between the behavior of equipment and human cognition or deviation from the standard operation procedures (SOPs). For the prevention of these accidents, the feasibility of procedures must be rechecked and analyzed.

One of the most critical points of the Cali accident was that the crew of flight 965 entered the wrong course after they accepted the proposal of runway change from ATC. Though runway change is regarded as a regular event during the operation, it seems to have led to the fatal error of the crew in this case. In Hirose et al. [17], they focused on what had happened from the time when the crew accepted the proposal of runway change from ATC to the time when the crews entered the wrong course to the FMC. They analyzed these events using their proposed method and simulated how the deviation of SOPs started and grew in the cockpit. Figure 10 shows the FRAM models of the analysis.

Wherein, function is defined as what has to be done to achieve a specific goal such as each items in a procedure. It is defined with six aspects;

- Input (I): Input to the functions, trigger
- Output (O): Outcome of functions

Fig. 10 Simulation outputs of safety assessment for the functions **a** initial states **b** resulted states after a variability given (Hirose et al. [17])

- Precondition (P): Conditions that must be satisfied before functions are carried out
- Resource (R): What is consumed during the process (fuel, energy, labor force...)
- Control (C): What supervises or restricts the function
- Time (T): Time required to accomplish the process

By specifying the values of the above aspects for each function and by checking their identities, we can model the holistic model as a dependency network among functions, where functions are connected via their shared values of the aspects.

As shown in Fig. 10a, once the variability in some functions of a specific instance are generated, they propagate and interact with the variabilities in other functions, which influences the performance of each functions. In addition, they could resonate in a specific context, which changes the situation significantly. Figure 10b illustrates the final results of the simulation, where the FRAM can analyze and represent this, which is one of the most characteristic points of this method.

The final results of the simulation are shown by displaying how the stability statuses of each function are evolving according to the chronological progress of the procedure. The stability status is evaluated as a value of PAF (probability of action failure), which is classified into one of the four qualitative (discrete) statuses of strategic, tactical, opportunistic, and scrambled in the decreasing order of the degrees of stability. Details of the simulation is shown in [17].

They also made another simulation assuming when the SOPs were designed differently and verified that such a severe accident could be avoided based on the simulation. From these results, they concluded that the dependencies of functions have to be taken into account for the analysis of the feasibility of procedures, and this method is expected to contribute to not only the analysis of the accident but also the pre-analysis of the safety of designed procedures.

3.5 The resilience of division of labor: socially centered automation

Recently, the division of labor between the human agent and the automation might be causing some novel problems for attaining their goal. Wherein, automation can be defined as the execution by a mechanism or a machine of a function that was previously carried out by a human (Parasuraman and Riley [19]). The development of automation, mainly since the 1950s, has been a significant influence on the relations between humans and machines, and, therefore, also on how we view human-machine systems. Discussions of automation have been central in human factors engineering since the late 1940s, usually under the topics of function allocation or automation strategies for aviaional automation.

Over the years, several different principles for automation design have been suggested (Hollnagel and Woods [6]). According to the most straightforward approach, commonly known as the left-over principle, the technological parts of a system should be designed to do as much as feasible (usually from an efficiency point of view), while the rest should be left for the operators to do. Since the determination of what is left over reflects what technology cannot do rather than what people can do, the inevitable result is that humans are faced with two sets of tasks. One set comprises tasks that are either too infrequent or too expensive to automate. The other set comprises tasks that are too complicated, too rare, or too irregular to automate.

A second principle to the automation design is known as the compensatory principle. In this approach, the capabilities (and limitations) of people and machines are compared to many salient dimensions, and function allocation is made to ensure that the respective capabilities are used optimally. The determination of which functions to assign to humans and machines is, however, not as simple as the compensatory principle assumes. Function allocation cannot be achieved

simply by substituting human functions by technology, nor vice versa because of fundamental differences between how humans and machines function and because of dependencies among functions in ways that are more complex than a mechanical decomposition can account. In this approach, humans are supplemented—or replaced—by automation to ensure that the task can be accomplished under optimal conditions of work, but tasks are often designed so that the minimum demands require humans' maximum capacity, and if task demands exceed the minimum level, humans cannot respond to that.

A third principle, which has emerged during the 1990s, is called the complementarity principle. This approach aims to sustain and strengthen the human ability to perform efficiently and, therefore, considers the work system in the long term, including how the work routines change as a consequence of learning and familiarization. The main concern is not the temporary level of efficiency (or safety), but rather the ability of the system to sustain acceptable performance under a variety of conditions. The effects of automation may be transitory, such as variations in workload or task complexity; they may be short-term, such as effects on situation awareness or fatigue; and they may be in the medium to long terms such as trust, self-confidence and the level of skills. While automation design mostly has considered transitory or short-term effects, it is also necessary to consider the consequences over a more extended period. Thus, the complementarity principle emphasizes the conditions provided by the overall socio-technical system.

All of the above three principles are leaving brittleness of division of labors between the automation and the human agent, and how to divide labors between them should be done taking account of co-agency, or the coupling between man and machine (automation) in the context of socio-technical systems.

An automated driving system is becoming a hot topic and is a complex combination of various components that can be defined as systems, where perception, decision making, and operation of the automobile are performed by electronics and machinery instead of a human driver. The US Department of Transportation National Highway Traffic Safety Administration (NHTSA) provided a standard classification system in 2013 which defined five different levels of automation, ranging from level 0 (no automation) to level 4 (full automation) (NHTSA [20]). Although the success in the automated driving system has been known to be successful enabling the level 4 of full automation, commercially the available automation levels are up to level 3 of conditional automation, where the system is in complete control of vehicle functions, and a human driver must be ready to intervene when requested by the system to do so. This level is typically based on the left-over principle mentioned above, and the confuse experienced by the driver when the turn-over

request (TOR) is suddenly coming from the systems, and this would possibly be sources of variabilities preventing the driver-automation co-agency (e.g., loss of surveillance and automation-induced surprise). Concerning the second compensatory principle, determination of which functions to assign to humans and automation has not been thoroughly investigated, and a view of technology-centered design has concentrated to leaving driving tasks to automation as far as possible, while the tasks left to a human driver were not discussed from the viewpoint of human-centeredness. The third complementarity principle has been challenged as well, and the shared control scheme was proposed. Even in this framework, it is still ambiguous which is responsible for the decision on the override (i.e., the automation or the human driver). Therefore, a novel idea of socially-centered automation is required, where the automation and the human agent should be aware of their partner's situations and the nonverbal communication should be realized between the two agents, as shown in Fig. 11 (Sawaragi [21]). This will improve the driver-automation interactions assumed in level 3 of the conditional automation.

3.6 The resilience of community: team situation awareness

Current industries are forced to act according to the rapid changes of circumstances such as market changes and to the change of economic environments. Moreover, the reduction of expert workers may bring about a drop of the work quality, and the occurrences of malfunctions caused by aging of equipment may bring about a variety of unexpected variabilities frequently during the production. Taking measures suited to those occasions is realized through the

Fig. 11 Socially-centered automation

developing a novel approach to resilient team design. These are expected to contribute to clarifying the factors needed to achieve resilience by observing and analyzing the maturing processes of the decision-making processes.

4 Conclusions

This article summarized the trends in studies on variabilities and resilience, as well as research activities conducted by the author and his colleagues. The occurrence of variabilities is inevitable, and it is impossible in practice to eliminate the variabilities only with technology or to automatize the recovery process from deviations due to variabilities. However, this difficulty means that the establishment of resilience requires the co-creation between human beings and systems. The most significant factor of the vulnerability to variabilities is the lack of the visualization of knowledge and information required for work completion. In the past, the lack of visualization was covered by the empirical knowledge of skilled workers to achieve an emergency response and quality assurance. However, such an approach now faces its limitations. It is needed to design systems that derive and cultivate the existing empirical knowledge, which requires to provide adequate constraints and supports with workers and to derive the workers' responses accordingly. Not by eliminating human beings or ultimately limiting their work, but pursuing a system design that allows some human judgment, resilience can be established through the fusion of knowledge of human beings and machines.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Holling CS (1973) Resilience and stability of ecological systems. *Annu Rev Ecol Syst* 4:1–23
- CAO (2015) Report on the 5th science and technology basic plan, https://www8.cao.go.jp/cstp/kihonkeikaku/5basicplan_en.pdf. Accessed 30 June 2019
- Selberg S, Austin M (2008) Toward an evolutionary system of systems architecture. 2010-07-15. Archived by WebCite® at <https://www.webcitation.org/5rEnx8BLfSoS>. Accessed on 15 May, 2010.
- Hoffman RR, Militello LG (2008) Sociological and ethnographic perspectives, perspectives on cognitive task analysis: historical origins and modern communities of practice. Taylor & Francis, London, pp 251–302
- Eason K (2008) Sociotechnical systems theory in the 21st century: another half-filled glass? In: Graves D (Ed.) *Sense in social science: a collection of essays in honour of Dr. Lisl Klein*, pp. 123–134
- Hollnagel E, Woods DD (2005) *Joint cognitive systems: Foundations of cognitive systems engineering*. CRC Press, Boca Raton
- Brandt D, Černecký J (2000) Automated systems based on human skill: joint design of technology and organisation. *Proceedings volume from the 7th IFAC symposium, Aachen, Germany, 15–17 June 2000*
- Hollnagel E (1998) *Cognitive reliability and error analysis method-CREAM*. Elsevier Science, Oxford
- Hollnagel E (2012) *FRAM: The functional resonance analysis method*. Ashgate, London
- Engeström Y (1987) *Learning by expanding: an activity-theoretical approach to developmental research*. Orienta-Konsultit Oy, Helsinki
- Hollnagel E (2015) Safety-I and Safety-II, the past and future of safety management. *Cogn Technol Work* 17:461–464
- Sawaragi T (2016) Human-system co-creative design of resilience. *IFAC Proc Vol* 49(19):468–473
- Kudo H, Horiguchi Y, Sawaragi T, Nakanishi H (2015) Train driver support interface based on visualization of running time constraints (in Japanese). *Trans Soc Instrum Control Eng* 51:171–180
- Burns CM, Hajdukiewicz JR (2004) *Ecological interface design*. CRC Press, Boca Raton
- Matsuoka S, Sawaragi T, Horiguchi Y, Nakanishi H (2016) Hierarchical planning for error recovery in automated industrial robotic systems. *Proc. of the 2016 IEEE international conference on systems, man, and cybernetics (SMC 2016)*
- Sawaragi T, Tsukamoto S, Horiguchi Y, Nakanishi H (2007) Semiotic process analysis of dynamic transformations in work procedures with-in organizational activities (in Japanese). *Oukan* 1:106–114
- Hirose T, Sawaragi T, Horiguchi Y (2016) Safety analysis of aviation flight-deck procedures using systemic accident model. In: Sawaragi T et al. (Eds) *Proceedings 13th IFAC/IFIP/IFORS/IEA Symposium, IFAC-HMS (Kyoto 2016): analysis, design, and evaluation of human-machine systems*. Elsevier Science, Pergamon
- Simmon D (1998) Boeing 757 CFIT accident at Cali Colombia becomes focus of lessons learned. *Flight Saf Digest* 17:1–31
- Parasuraman R, Riley V (1997) Humans and automation: use, misuse, disuse, abuse. *Hum Factors* 39:230–253
- US Department of transportation releases policy on automated vehicle development (2016) National highway traffic safety administration. Retrieved 9 December 2016.
- Sawaragi T (2018) Design of productive socio-technical systems by human-machine co-creation (in Japanese). *Meas Control* 57:79–84
- Nonaka T, Miki K, Odajima J, Mizuyama H (2016) Analysis of dynamic decision making underpinning supply chain resilience: a serious game approach. In: Sawaragi T et al. (Eds) *Proceedings 13th IFAC/IFIP/IFORS/IEA symposium, IFAC-HMS (Kyoto 2016): analysis, design, and evaluation of human-machine systems*. Elsevier Science, Pergamon
- Endsley MR (1995) Toward a theory of situation awareness in dynamic systems. *Hum Factors* 37:32–64
- Hackman JR (1987) The design of work teams. *Handbook of organizational behavior*, pp. 315–342
- Salas E, Prince C, Baker DP, Shrestha L (1995) Situation awareness in team performance: implications for measurement and training. *Hum Factors* 37:123–136

26. Sawaragi T, Fujii K, Horiguchi Y, Nakanishi H (2016) Analysis of Team Situation Awareness Using Serious Game and Constructive Model-Based Simulation. In: Sawaragi T et al. (Eds) Proceedings 13th IFAC/IFIP/IFORS/IEA symposium, IFAC-HMS (Kyoto 2016): analysis, design, and evaluation of human-machine systems, Elsevier Science, Pergamon
27. Cohen MD, March JG, Olsen JP (1972) A garbage can model of organizational choice. *Adm Sci Q* 17:1–25

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.