Towards A Good Automation of the Next Generation Man-Machine Interface System for Nuclear Power Plants – Systems Approach and New Safety Paradigm

Lee Yong-Hee
Advanced I&C Research Division, *Korea Atomic Energy Research Institute*
Daedeok-daero 989-111, Daejon, Korea, 34050
**Corresponding author: yhlee@kaeri.re.kr, yhlee0412@gmail.com*

***Keywords :** automation, Man-Machine Interface System, systems approach, positive safety, function analysis*

1. Background and Overview

This paper proposed the direction for automating an ideal instrumentation and control(I&C) system in next-generation nuclear power plants (NPPs). First, the development process and current issues of I&C systems and automation technology in NPPs were reviewed. By introducing the MMIS concept and reviewing several experiences with automation, the study discussed the desirable direction of automation. In particular, reflecting new paradigms such as *Safety II and Positive Safety*, recommendations and directions are discussed for development that can maximize the utilization of AI.

2. Brief on the History of Nuclear I&C

In NPPs, I&C system is the field that has developed most rapidly. From the outset, remote measurement and control functions for radiation safety were essential in NPPs, so measurement and control devices for various nuclear plant functions had to be installed in physically isolated spaces. The early I&C systems were merely an aggregation that physically gathered the instruments and controllers of NPP equipment in a separated location in.

The greatest change in nuclear power plant I&C systems was the evolution from an aggregation-type system, which physically gathered instruments and controllers, to a man-machine interface system (MMIS). MMIS is a paradigm introduced as a follow-up measure to the TMI accident, demanding a fundamental solution to the problem of human error. In most existing operating NPPs, improvements and relocations at the level of equipment upgrades and repositioning in the minimum feasible control rooms and control panels were limited.

However, at the new nuclear power plant, the I&C system could be designed and developed as an independent system. This is because, by designing not only from the perspective of the NPP equipment's functions but also from the operator's point of view, human error can be minimized. In particular, with the advancement of various computerization technologies, it has developed to provide an interface suitable for operators. This human factors perspective has enhanced user safety and convenience, leading not only to the rapid advancement of UI/UX technology in today's smartphones and other devices, but also to the starting point of the rapid development of I&C systems in NPPs.

Today, I&C systems have taken on the role of the brain and neural network of a nuclear power plant through the introduction of the MMIS concept. However, in order to become a fully functioning brain, its functional integration is still insufficient. It is said that, functionally, the many are still segmented, so they cannot be integrated and operated as a single neural network by a single brain. First, except for normal operation for power generation, it has not been integrated into the I&C system. In driving modes, there are many separately designed partial functions other than power generation, so integrated I&C are often not possible. Therefore, cases related to testing and inspection, as well as from the perspective of human error and safety, are still ongoing and are being relatively highlighted.

Second, many field control panels have not been integrated with the central control room. It is failing to provide a functionally fully integrated control space, such as a driver's seat in a car or a cockpit in an aircraft.

Nevertheless, through the application of advanced communication and computerization technologies, it will continue to develop into an integrated I&C system capable of realizing the MMIS concept.

3. Cases of Automations and AI Applications

Automation in NPPs has a long history. However, not all automation was successful or effective. To discuss the direction of next-generation MMIS, we aim to review case-based studies from two directions of automation.

First, experience in automating independent unit operations. Unit operations independent of the system are relatively small or easy to automate separately. For example, nuclear fuel management and refueling are carried out independently and separately from the output operation that produces electricity, so automation was attempted early on. However, the fatal accident at Saint Lauren NPP in France was an example that stained the nuclear industry from its early days with human error. This was because the driver exhibited a very serious maladaptive reaction (an active violation action) toward an automatic function that was not functionally sufficiently integrated. Even when automating functionally independent unit functions, it is

necessary to consider to what extent and in what manner they are integrated with the operator in order to enable effective automation.

Second, experience in automating unit auxiliary functions. Automating parts of complex functions in a system is beneficial for improving reliability and efficiency. In particular, automating functions that place a burden on drivers is beneficial as a way to support them. For example, there are many cases of automating the level control function of the steam generator (SG: Steam Generator), which is essential for output fluctuations during the startup and shutdown of a nuclear power plant. The function of automatically adjusting the steam generator's water level during the start-up and shutdown operations by the operator was a highly desirable form of automation, but in practice it often failed to produce significant results. There were many cases where switching had to be done manually, and during both manual and automatic switching, the system exhibited unexpectedly rapid behavior, which in many cases led to human errors (typical prediction and tracking errors).

In addition, the adoption of artificial intelligence (AI) technology is expected to open up an area where, through the introduction of AI, power generation systems can evolve into a new dimension. However, there are still significant barriers to introducing AI into NPP I&C systems.

Most AIs cannot guarantee sufficient safety either in advance or afterward. In particular, as seen in the increasingly prominent Generative AI, securing functional reliability is difficult, and even control remains fundamentally uncertain.

This fundamentally differs from the early AI (such as Alpha-Go), which was expected to be easily applicable in specific fields with clear technical foundations and acceptable limits. If AI is developed based on general knowledge, it appears to be inevitably prone to creating false statements (hallucination) and flattery of the questioner.

Although it proposes ethical requirements such as the fundamental need for AI to be managed within human control, monitoring or managing the development of AI based on such requirements seems unlikely to escape the hindsight effect from the outset. The verification of AI safety required for public, highly reliable systems such as nuclear power remains a distant prospect. Therefore, considering the unlimited liability related to safety required of NPPs, even if the functions of a nuclear power plant are to be partially or entirely independent functions, it has become difficult to entrust them to the AI technology currently being developed.

While automation is limited to fixed rules and logic, autonomy is sometimes distinguished in detail as the ability to adapt to circumstances and apply rules and logic in modified forms. However, in the broad sense, autonomy appears to fall within the scope of automation. Most I&C functions in existing NPPs can be automated and automated by AI. As with most systems, the adoption of AI in NPPs often ultimately aims for Level 5 unmanned automation without human intervention. However, it is questionable whether unmanned automation is technically feasible, and also whether it is desirable in terms of safety and effectiveness.

The aviation sector already has automated takeoff and landing capabilities that have been developed and applied for a long time, but they have not been regarded as successful. Because it was relatively not sufficiently superior to the pilot's average capability, it was left to be chosen as a kind of option. One limitation of this kind of automation can be said to be the problem of functional immaturity.

In the aviation sector, partial-function automation has recently led to even more serious consequences, as seen in the case of MCAS included in B-737 Max-8. In the strategic improvement of the B-737 promoted by the American Boeing company, an automatic altitude control system (so-called MCAS) was developed and applied to assist pilots during the aircraft's takeoff and landing processes, but there was a case that led to a crash disaster that resulted in the deaths of hundreds of people in succession. Since it was not a function sufficiently integrated into the pilot's acceptance and selection, it is presumed that it rather seriously interfered with takeoff and landing.

In contrast, in the automotive sector, various cases of AI application are developing relatively actively. For example, while the lane departure warning system is highly effective, the crash-avoidance automatic stop system is still controversial regarding its effectiveness. In addition, autonomous driving functions that replace driving-related decision-making through visual perception and judgment appear to have secured statistically convincing improvements in safety and efficiency, despite several fatal accidents (considering the level of practical application of driverless taxis).

4. New Safety Paradigms and Good Automation

A nuclear power plant is a power generation system, but it is based on a level of safety that cannot be compared with any other system. Because NPPs can only be realistically accepted if they can avoid the unfamiliar and still uncertain issue of radiation. This is evident in the situation where, even when faced with the global challenge of the climate crisis, nuclear power is not readily accepted as one of the good alternatives. Nuclear power plants have also been leading in the technological realization of safety and have achieved a new level of accomplishment. As seen in programs such as PRA, we have realized design and management measures that minimize NPP risks through meticulous risk assessment and systematic reliability improvements.

However, as seen in terms of public acceptance, the perceived risk of nuclear remains high, and the recognized safety level does not appear to be sufficient to accept in everyday life. In the meantime, ensuring safety through methods that minimize risk is not enough.

They may also put more effort into promotion, viewing this as a lack of public understanding or

cognitive distortion. However, from the perspective of behavioral science or behavioral economics, it is interpreted differently. Not only can engineering safety not fully represent immediate practical safety, but it may also be more effective, in practice, to realistically expand engineering safety rather than to make uncertain efforts that would change public perception. By introducing the recently advanced *Safety-II and Positive Safety* paradigm into the nuclear field, it is necessary to find ways to ensure more realistic safety in the design of next-generation NPPs.

To this end, the concept of automation required for next-generation NPPs and its application methods were discussed under the concept of "good automation." The following is a summary of the direction of AI-driven automation applying a new safety paradigm.

First, good automation is the development of new functions and functional structures, rather than the exclusion and replacement of humans.

By leveraging existing experience in automation and AI in nuclear and other fields, it is necessary to review the desirable direction of automation. In deriving the sub-detailed functions that achieve the nuclear power plant's highest-level goal of "***safe and efficient power generation***", the existing functional items and functional structure may be completely changed. The existing safety functions, configured according to the safety barriers established around the light-water reactor, must be completely reconsidered and reorganized in next-generation plants.

Second, good automation must be integrated with the purpose and objectives of the entire system. The automation of a particular area or partial function should be examined in terms of how it affects the goals of the overall system. Unlike the complementary approach at NuScale NPP, the function analysis and allocation of the Human Factors Program Plan (HFPP) must be carried out completely anew, incorporating new technologies such as AI. In particular, it is important to consider the problems and limitations of the adoption of automation and AI discussed earlier during this process.

Third, good automation should aim not only to the designed functions to minimize known risks but also to foster the emergent expansion of extended role for more safety. As a key lesson from the Fukushima accident, the IAEA presented new safety challenges in preparing for unknown risks and unpreparedness. This means that the safety required of NPPs is difficult to achieve through a minimization approach that merely prepares for known risks such as DBA, beyond-DBA, and so on. For enhancement of preparedness, IAEA introduced the notion of *Safety Culture*. AI could serve as a scrutinized monitoring of plant personnel and their performances for more concrete measures to safety culture. Since resilience proposed in Safety-II is an important guiding principle, it is possible to explore the functional architecture needed for the development of the *RAG* (*Resilience Assessment Grid*). In addition, as proposed by Positive Safety, it may also be worth considering actively expanding into new areas of safety, such as demand control and response management within the power grid. AI might positively nudge the plant personnel and extend the capabilities in known tasks and unknown situations. AI might be powerful to support safety culture by monitoring humans rather than machines by automated I&C.

5. Conclusions and Discussions

The advancement of I&C systems, triggered by the introduction of the MMIS concept in NPPs, can now blossom through the adoption of AI. However, in NPPs where safety is the highest priority, the technical nature of AI means that safety concerns cannot be resolved all at once, and AI-based automation is blocked by major barriers. In the case of next-generation nuclear power, it is not desirable to ignore the limitless potential of artificial intelligence and simply stand by. We should actively pursue automation using AI rather than waiting for it, until we can secure the functional perspective of safety required by nuclear power and verify it according to the nuclear manner required currently.

References

- Hollnagel, E., Synesis, : The Unification of Productivity, Quality, Safety and Reliability, Routledge, 2020
- Hollnagel, E., From Safety to Safely: Principles and Practice of Systemic Potentials Management, Routledge 2025
- IAEA, Fukushima Accident Report, 2016
- IAEA, Considerations on Performing Integrated Risk Informed Decision Making, TECDOC-1909, 2020
- Lee, Y.H., A Critical Revisit to Human Error Studies for Advances in Human Factors Engineering: Critiques to Traditional Safety Paradigms & Safety II, ESK-2025 Spring, 2025
- Lee, Y.H., A Discussion on the Development Direction of the Next Generation I&C System in Nuclear : To Cope with the Human Error Uncertainty Remaining after Fukushima Accident, KNS-2025 Spring, 2025
- Lee, Y.H., A Positive Safety Paradigm Proposed for Advances in Human Factors Engineering and Human Error Studies in Nuclear, KNS-2025 Autumn, 2025
- Lee, Y.H., A New Safety Paradigm for Successful Next Generation Nuclear Systems Integrating AI and Automation – For Continuing the Nuclear Leadership for Safety Realization, KNS-2026 Spring, 2026
- NuScale, Human Factors Engineering Project Management Plan, NP-RP-0813-4248-NP, 2013
- NuScale, Human Factors Engineering Functional Requirements Analysis and Function Allocation Implementation Plan, TR-124333-NP, 2022
- OECD, Behavioral Insights and Public Policy: Lessons from Around the World, OECD, Paris, 2017
- Rasmussen, J., The Role of error in organizing behavior, Qual. Saf. Health Care;12, 377-383, 2003
- U.S. Nuclear Regulatory Commission, Human Factors Engineering Program Review Model, NUREG-0711, Rev.3, 2012