Human Factors Review of AI-Based Autonomous Operation in Nuclear Power Plants

Kun-Young Han [a*], Sa-Kil Kim [a]
[a] *Korea Atomic Energy Research Institute, 111, Daedeok-daero 989beon-gil, Yuseong-gu, Daejeon, 34057, Korea*
[*] *Corresponding author: kyhan-akane@Kaeri.re.kr*

Keywords : Artificial Intelligence, Autonomous Operation, Human Factors Engineering, Small Modular Reactors

1. Introduction

The development of next-generation nuclear reactors, including Small Modular Reactors (SMRs), is accelerating amid the global imperatives for energy security and the achievement of carbon neutrality [1-3]. In particular, SMRs aim for integrated multi-module operation [4] and high-level automation [5]. To ensure both safety and economic efficiency [6, 7], the application of Artificial Intelligence (AI) is being extensively reviewed across various domains, including anomaly detection [8-11], fault diagnosis [12-14], control [15-17], and decision support [18-20]. Furthermore, attempts to integrate key functions—such as monitoring, diagnosis, control, and decision-making—at the plant level through digital twin-based autonomous operation [21, 22] signify a fundamental paradigm shift. This transition moves beyond conventional operator-support systems toward AI-based autonomous operation concepts. These technological advancements are driving fundamental changes in the operational environment of nuclear power plants. Notably, operators' roles are shifting from direct manipulation of plant equipment to situation assessment and judgment, as well as high-level oversight, indicating a move toward reduced direct human intervention.

To ensure safety in the nuclear domain, it is important to evaluate how AI-induced changes in human and organizational factors (HOF) affect function allocation, personnel tasks, human–system interfaces, procedures, training, and human performance before AI deployment and integration. By referencing the Human Factors Engineering (HFE) Program Review Model documented in NUREG-0711 [23], this study identifies supplemental review considerations for the HFE review framework in support of the transition to AI-based autonomous operation concepts.

2. Notional AI and Levels of Autonomy

The U.S. Nuclear Regulatory Commission (NRC) defines five notional AI and autonomy levels to characterize potential uses of AI and autonomy in the nuclear industry [24]. This framework defines the scope of AI application, starting with simple decision support, progressing to a collaborative stage in which algorithms provide recommendations for human execution, moving further to a stage in which the machine (i.e., the AI-based system) makes decisions and conducts operations under human oversight, and finally reaching a fully autonomous stage in which no human intervention is presumed (Table I). The NRC notes that as the level of autonomy increases, reliance on human intervention or oversight decreases; therefore, higher autonomy levels may warrant greater regulatory scrutiny of the AI system. This suggests that, at higher autonomy levels, the evidentiary requirements for demonstrating safety may increase.

Meanwhile, NUREG-0700 [25], the existing guideline for nuclear Human-System Interface (HSI) design, distinguishes operational levels from manual to fully automated operation. However, it differs from the AI autonomy model in that it still assumes human monitoring and emergency intervention even at high levels of automation. Therefore, it is necessary to comparatively analyze the NRC's AI autonomy classification against existing automation levels to specifically redefine shifting human roles—such as monitoring, approval, intervention, and backup.

In particular, while direct operational errors may decrease as AI technology matures, the likelihood of new cognitive errors, such as reduced situation awareness or automation bias, is expected to rise. Accordingly, establishing an HFE review framework grounded in a human error prevention perspective may help mitigate these emerging risks.

3. Considerations Related to Human and Organizational Factors

Several characteristics of AI warrant supplemental consideration in the HFE review process. AI-enabled functions may be data-driven and may exhibit performance variation over time. The NRC's potential AI technical considerations for regulatory decision-making [24] identify key technical considerations relevant to AI-based systems, such as explainability, trustworthiness, data drift, and fielded performance degradation. Technical uncertainties arising from these characteristics, together with the complexities of human–AI teaming, are not yet fully addressed in the existing HFE review framework.

In addition, the nuclear regulatory bodies of the U.S., the U.K., and Canada—namely the NRC, ONR, and CNSC—note in their joint guidance [26] that many AI capabilities may be used to augment human decision-making rather than replace it. Because of this blending of machine and human decision-making, organizations and end users may be presented with unique HOF

Table I: Comparison of Conventional Automation Levels and Notional AI Autonomy Levels

Levels of Automation [25]	**Notional AI and Autonomy Levels** [24]
Level 1: Manual Operation No automation Operators manually perform all tasks.	Level 0: AI Not Used No AI or autonomy integration in systems or processes.
Level 2: Shared Operation Automatic performance of some tasks Operators perform some tasks manually.	Level 1: Insight (Human decision-making assisted by a machine) AI integration in systems is used for optimization, operational guidance, or business process automation that would not affect plant safety/security and control.
Level 3: Operation by Consent Automatic performance when directed by operators to do so, under close monitoring and supervision Operators monitor closely, approve actions, and may intervene to provide supervisory commands that automation follows.	Level 2: Collaboration (Human decision-making augmented by a machine) AI integration in systems where algorithms make recommendations that could affect plant safety/security and control are vetted and carried out by a human decisionmaker.
Level 4: Operation by Exception Essentially autonomous operation unless specific situations or circumstances are encountered Operators must approve of critical decisions and may intervene.	Level 3: Operation (Machine decision-making supervised by a human) AI and autonomy integration in systems where algorithms make decisions and conduct operations with human oversight that could affect plant safety/security and control.
Level 5: Autonomous Operation Fully autonomous operation. System cannot normally be disabled but may be started manually Operators monitor performance and perform backup if necessary, feasible, and permitted	Level 4: Fully Autonomous (Machine decision-making with no human intervention) Fully autonomous AI in systems where the algorithm is responsible for operation, control, and intelligent adaptation without reliance on human intervention or oversight that could affect plant safety/security and control.

challenges. While humans may deploy AI systems with a clear objective, it may be difficult to determine whether the intended objective has been met, and the human may or may not be responsible for the behavior of the machine. Such issues can have profound impacts on reactor operator licensing, concepts of operation, and rulemaking. Therefore, consideration should be given to how AI failure affects human–machine teaming, particularly as reliance on AI-driven decision-making increases. Accordingly, the joint guidance specifies key HOF considerations, including the capability the AI component or element is intended to provide; which functions, roles, or responsibilities will be allocated to the AI as opposed to humans; novel functions, roles, or responsibilities arising from the inclusion of the AI capability; how humans interact with the AI; and humans' ability to intervene, or not, with the operation of the AI (level of autonomy).

Informed by these considerations, this study identifies supplemental review considerations for the NUREG-0711 HFE Program Review Model [23] to support the safe and effective transition to AI-based systems, including AI-based autonomous operation concepts.

3.1 Planning and Analysis Phase

Early-stage AI integration may warrant establishing a multidisciplinary team—including AI specialists—as part of HFE program management to address technical uncertainties. Provisions for performance monitoring during operation and for tracking AI-related HFE issues that affect human performance may also need to be established. An Operating Experience Review may help surface key considerations and potential AI-related risks early on, such as defining the operational envelope, addressing alarm-related issues, and anticipating maintenance concerns. In particular, Functional Requirements Analysis and Function Allocation may need to be updated to reflect changes in human–AI role assignments across operational conditions and system states in ways that support operators' situation awareness. Task Analysis may also need to be updated accordingly to address AI-related monitoring and intervention tasks and their associated task support needs. Determining appropriate staffing levels commensurate with the level of autonomy and reviewing qualification requirements that reflect human–AI collaboration capabilities may also be warranted in this phase. In addition, tasks associated with AI monitoring and intervention may warrant re-evaluation to determine whether they constitute Important Human Actions (IHAs) and, where applicable, may need to be explicitly identified and evaluated within the HFE review process.

3.2 Design Phase

HSI features should be designed to support operators' understanding of AI-generated outcomes and to provide clear handover interfaces and guidance, and associated procedures should be developed accordingly. In addition, training requirements associated with effective human–AI interaction, including higher-level supervisory and management competencies needed to mitigate the "out-of-the-loop" phenomenon during transitional stages and provisions for periodic retraining, may need to be identified as inputs to training program development. Such measures are important for fostering calibrated trust and mitigating issues—such as over-trust or distrust—arising from the "black-box" nature of AI.

3.3 Verification and Validation Phase

It is important to confirm that the HSIs, procedures, and training developed for AI-based systems provide the alarms, information, controls, and task support identified through task analysis, conform to applicable HFE design guidance, and support safe operation of the integrated system, including its AI-enabled functions, under representative operating conditions. In the context of AI-based systems, these verification and validation activities may also need to examine whether AI-related features, such as explainability, performance variation over time, and human–AI interaction, support safe and effective operator performance. In addition, human engineering discrepancies identified during these evaluations may need to be assessed and resolved before implementation. Furthermore, it may be beneficial to strengthen the linkage between the Verification and Validation phase and the Implementation and Operation phase so that findings from verification and validation can be reflected in design implementation and subsequent human performance monitoring.

3.4 Implementation and Operation Phase

It may be beneficial to verify that the as-built design remains consistent with the verified and validated HFE design intent and to establish a lifecycle management framework for monitoring human and AI performance during operation. Given that AI performance may continue to evolve, such a framework may be needed to monitor whether the conclusions drawn from integrated system validation remain valid over time, whether changes to HSIs, procedures, and training adversely affect human performance, and whether important human actions can continue to be accomplished within applicable time and performance criteria. These insights should be used to identify and address degradation in both human and AI performance in a timely manner, to inform corrective actions in design implementation, and to support subsequent human performance monitoring to determine whether human performance remains acceptable after corrective actions. This feedback loop is also essential for maintaining the supervisory and management competencies identified in the design phase to support plant safety in beyond-design-basis emergencies under fully autonomous operation (Level 4).

4. Conclusions

As AI adoption in nuclear facilities and the shift toward AI-based autonomous operation concepts are under active discussion, this study compared the U.S. NRC's framework for notional AI and autonomy levels with the automation-related operational levels described in NUREG-0700 and summarized the HOF considerations presented by the nuclear regulatory bodies of the United States, the United Kingdom, and Canada. Based on these findings, supplemental review considerations for the NUREG-0711 HFE Program Review Model were identified to support the safe and effective transition to AI-based systems with characteristics distinct from those of conventional automated systems, including AI-based autonomous operation concepts.

Acknowledgement

This work was supported by the Nuclear Safety Research Program through the Regulatory Research Management Agency for SMRS (RMAS) and the Nuclear Safety and Security Commission (NSSC) of the Republic of Korea (No. 1500-1501-409)

REFERENCES

[1] International Energy Agency, Nuclear Power and Secure Energy Transitions: From Today's Challenges to Tomorrow's Clean Energy Systems, IEA, Paris, 2022.
[2] OECD Nuclear Energy Agency, Small Modular Reactors (SMRs) for Net Zero, OECD, Paris, 2023.
[3] International Atomic Energy Agency, Small Modular Reactors: Advances in SMR Developments 2024, IAEA, Vienna, 2024.
[4] J. O'Hara and J. Higgins, NRC Reviewer Aid for Evaluating the Human Factors Engineering Aspects of Small Modular Reactors, Brookhaven National Laboratory, BNL-96809-2012, 2012.
[5] D. E. Holcomb, U.S. Department of Energy Advanced Small Modular Reactor R&D Program: Instrumentation, Controls, and Human-Machine Interface (ICHMI) Pathway, Oak Ridge National Laboratory, ORNL/TM-2012/253, 2012.
[6] OECD Nuclear Energy Agency, Small Modular Reactors: Challenges and Opportunities, OECD, Paris, 2021.
[7] OECD Nuclear Energy Agency, The NEA Small Modular Reactor Dashboard, NEA No.7650, OECD, Paris, 2023.
[8] D. Chang, M. Liu, and Y. Lee, Accident Diagnosis of a PWR Fuel Pin During Unprotected Loss of Flow Accident With Support Vector Machine, Nuclear Engineering and Design, Vol.352, pp.1084-1094, 2019.
[9] C. Chai, C. Ramirez, M. Maceira, and O. Marcillo, Monitoring Operational States of a Nuclear Reactor Using Seismoacoustic Signatures and Machine Learning, Seismological Research Letters, Vol.93, No.3, pp.1660-1672, 2022.
[10] N. S. Arunraj, R. Hable, M. Fernandes, and K. Leidl, Comparison of Supervised, Semi-Supervised and Unsupervised Learning Methods in Network Intrusion Detection System (NIDS) Application, Anwendungen und Konzepte der Wirtschaftsinformatik, Vol.6, No.6, pp.10-19, 2017.
[11] M. E. Villa-Pérez, M. A. Álvarez-Carmona, O. Loyola-González, M. A. Medina-Pérez, J. C. Velazco-Rossell, and K. K. R. Choo, Semi-Supervised Anomaly Detection Algorithms: A Comparative Summary and Future Research Directions, Knowledge-Based Systems, Vol.218, Art.106878, 2021.
[12] S. Tolo, X. Tian, N. Bausch, V. Becerra, T. V. Santhosh,

G. Vinod, and E. Patelli, Robust On-Line Diagnosis Tool for the Early Accident Detection in Nuclear Power Plants, Reliability Engineering & System Safety, Vol.186, pp.110-119, 2019.
[13] S. Li, J. Chen, H. Lin, and W. Wang, A Composite Fault Diagnosis Model for NPPs Based on Bayesian-EfficientNet Module, arXiv preprint arXiv:2411.17707, 2024.
[14] Y. Mu, X. Hong, and Y. Liu, Fault Diagnosis Method for Nuclear Power Plant Based on Ensemble Learning, Atomic Energy Science and Technology, Vol.46, No.10, pp.1254-1258, 2012.
[15] K. Xiao, Q. Wu, J. Chen, X. Pu, Y. Zhang, and P. Yang, A Neural Network Predictive Control Method for Power Control of Small Pressurized Water Reactors, Annals of Nuclear Energy, Vol.169, Art.108946, 2022.
[16] H. F. Al Masri, Adaptive Neural Network Algorithm for Power Control in Nuclear Power Plants, Journal of Physics: Conference Series, Vol.781, No.1, Art.012052, 2017.
[17] A. Nuerlan, P. Wang, Rizwan-Uddin, and F. Zhao, A Neural Network Based Inverse System Control Strategy to Decouple Turbine Power in Multi-Reactor and Multi-Turbine Nuclear Power Plant, Progress in Nuclear Energy, Vol.129, Art.103500, 2020.
[18] S. Takaya, A. Seki, M. Yoshikawa, N. Sasaki, and X. Yan, Proposal of a Novel AI-Based Plant Operator Support System for the Safety of Nuclear Power Plants, Mechanical Engineering Journal, Vol.11, No.2, Art.23-00408, 2024.
[19] L. Tunkle, K. Abdurrahman, L. Liu, and M. J. Radaideh, Nuclear Microreactor Control With Deep Reinforcement Learning, arXiv preprint arXiv:2504.00156, 2025.
[20] G. Zhu, D. Guo, J. Li, Y. Xie, and D. Zhang, Intelligent Steam Turbine Start-Up Control Based on Deep Reinforcement Learning, Energy, Vol.320, Art.135335, 2025.
[21] L. Lin, P. Athe, P. Rouxelin, M. Avramova, A. Gupta, R. Youngblood, J. Lane, and N. Dinh, Development and Assessment of a Nearly Autonomous Management and Control System for Advanced Reactors, Annals of Nuclear Energy, Vol.150, Art.107861, 2021.
[22] R. Ponciroli, A. Dave, T. Nguyen, H. Wang, B. Kochunas, A. Cilliers, and R. Vilim, The Challenge of Autonomous Operation and Automated Reasoning as an AI Enabler, Proceedings of the GIF Industry Forum 2022, Toronto, Canada, 2022.
[23] U.S. Nuclear Regulatory Commission, Human Factors Engineering Program Review Model, NUREG-0711, Rev. 3, Washington, DC, 2012.
[24] U.S. Nuclear Regulatory Commission, Artificial Intelligence Strategic Plan: Fiscal Years 2023–2027, NUREG-2261, Washington, DC, 2023.
[25] U.S. Nuclear Regulatory Commission, Human-System Interface Design Review Guidelines, NUREG-0700, Rev. 4, Washington, DC, 2026.
[26] Canadian Nuclear Safety Commission, U.S. Nuclear Regulatory Commission, and Office for Nuclear Regulation, Considerations for Developing Artificial Intelligence Systems in Nuclear Applications, Joint Regulatory Report, 2024.