Nuclear Energy University Program
U.S. Department of Energy

Probabilistic Validation and Risk Importance Ranking Methodology for Automation Trustworthiness and Transparency in Nuclear Power Plants

PI: Zahra Mohaghegh (Dept. of Nuclear, Plasma, & Radiological Engineering [NPRE], UIUC).

Program: RC-6 (Reactor Concepts).

Collaborators: Ahmad Al Rashdan & Ronald Boring (INL); Rizwan Uddin & Seyed Reihani & Tatsuya Sakurahara & Ernie Kee (NPRE, UIUC); Vinh Dang and Luca Podofillini (Paul Scherrer Institute); Jonghyun Kim (Chosun University); Richard Kesler (University of Illinois Fire Service Institute); Jeffrey Stone (Exelon Generation); Kristin Kaspar (STPNOC).

ABSTRACT:

To improve efficiency and ensure safe, reliable operation, the U.S. nuclear industry is working to leverage automation as much as possible. Introduction of automation technologies, however, still presents challenging issues for most Nuclear Power Plants (NPPs). These issues include defining an appropriate end-state automation architecture, developing business cases for automation implementation, assuring automation trustworthiness, improving automation transparency, and addressing licensing process burden. These challenging issues contribute to higher costs and schedule uncertainties for automation deployment, creating hurdles for the use of automation in NPPs. Before committing to a significant investment in the deployment of an automation technology, decision-makers need methodologies that can generate sufficient evidence to verify that the automation would be explainable, trustworthy, and operationally acceptable. Existing methodologies for evaluating automation trustworthiness and improving automation transparency have three main limitations preventing them from being widely adopted for the nuclear industry. First, there is a lack of consensus in definitions for automation terms, such as trustworthiness, trust, and transparency across different domains (e.g., nuclear, transportation, information security, communication infrastructure, autonomous vehicles, explainable artificial intelligence), leading to the absence of widely accepted methodologies for automation trustworthiness and transparency. Second, existing methodologies for improving automation transparency are context-specific and lack a technical basis that is sufficient to support the generality of these approaches, making them unjustifiable for contexts other than those being tested. Third, existing methodologies for evaluating the trustworthiness of automation models rely heavily on qualitative/semi-quantitative approaches and/or empirical validation. To address these limitations and help nuclear utilities make better decisions on automation deployment, the project team will make four key contributions in this project: (1) Conduct a cross-disciplinary review of literature on automation trustworthiness and transparency to (a) establish definitions of these two terms for use in the nuclear domain and (b) generate a scientific relationship between an “automation trustworthiness evaluation” and “uncertainty analysis.” This scientific foundation helps creating a “generic” (rather than technology-specific) methodology for trustworthiness evaluation; (2) Develop a Probabilistic Validation (PV) methodology and advance an Integrated Probabilistic Risk Assessment (I-PRA) methodological framework for automation technologies to evaluate and improve automation trustworthiness; (3) Develop an advanced risk monitoring methodology at the interface of human and automation to improve automation transparency; and (4) Demonstrate the feasibility and practicality of the proposed methodologies by conducting two case studies that address practical interest in automation technology expressed by our NPP partners. The project will make use of the INL Human Systems Simulation Laboratory, the experimental facilities at the University of Illinois Fire Service Institute, the Virtual Education and Research Laboratory at UIUC, and plant historical data to support the case studies.