Carnegie Mellon University

PSAM 18 · Probabilistic Safety Assessment and Management

# A Traceable AI-Agent Operation Framework Using Graph-Based Procedures in Nuclear Abnormal Scenario Simulation

**Seongeun Park** · Christopher McComb · Jinghua Xiao · Joonsun Hwang · Pingbo Tang

Pittsburgh, Pennsylvania · July 19-24, 2026

BACKGROUND

# Why Is This Hard for Nuclear Operators?

In safety-critical infrastructure such as nuclear power plants, when an abnormal event occurs, operators respond by reading dense, text-based procedures.

1 **Read Current State** → 2 **Compare with Procedure** → 3 **Judge Branch Logic** → 4 **Act & Verify**

**This 4-step cycle has to be repeated by a human, over and over, under time pressure.**

Even reading the same procedure, different operators can interpret it differently

BACKGROUND

# Where Does Existing Research Stand?

## Computer-Based Procedures

*e.g., Rancor, RIPS*

- Display procedures and plant states
- Track procedural progress

→ **Procedure execution support**

## AI-Based Operator Support

*e.g., COSS, NAMAC*

- Diagnose and predict system conditions
- Recommend mitigation actions

→ **Recommendation quality**

Limited action-level traceability connecting each AI-selected action to its procedure context, validation result, and resulting state change.

PREVIOUS WORK

# BuildLink: Graph-Based Decision Support

**BuildLink** structures procedures and incident reports as **knowledge graphs** for AI-based retrieval.

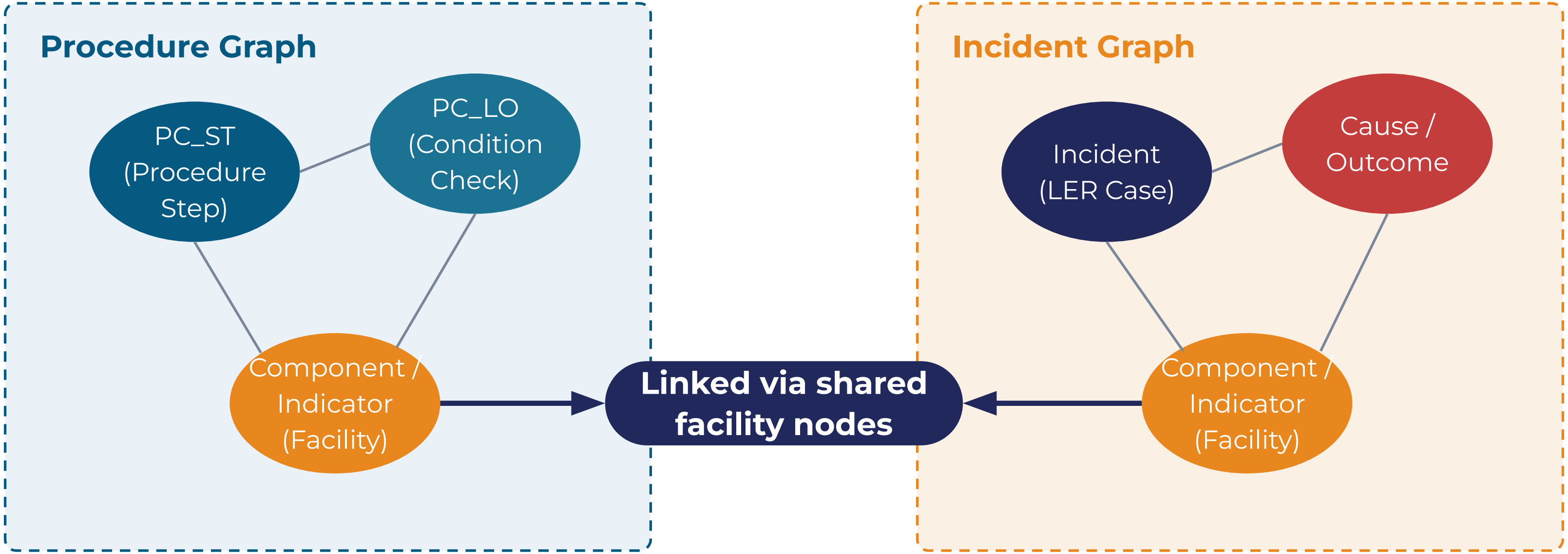
| | | |
|---|---|---|
| 1 | **Procedure Knowledge Graph** | Converts text procedures into explicit 'condition–state–action' relationships |
| 2 | **Incident Knowledge Graph** | Structures 600+ NRC Licensee Event Reports (LERs) into components, causes, outcomes, and human actions |
| 3 | **GraphRAG Advisor** | Retrieves related procedure and incident evidence from the current state and user query |

※ *In BuildLink, AI supported human operators as an **advisor**.*

*Park, S., Hwang, J., Jeon, K., Boring, R. L., & Tang, P. (2026)*
*BuildLink: Graph-Structured Procedural Representation with Incident Integration. International Conference on Computing in Civil Engineering*

PREVIOUS WORK

# Linking Procedures and Incidents



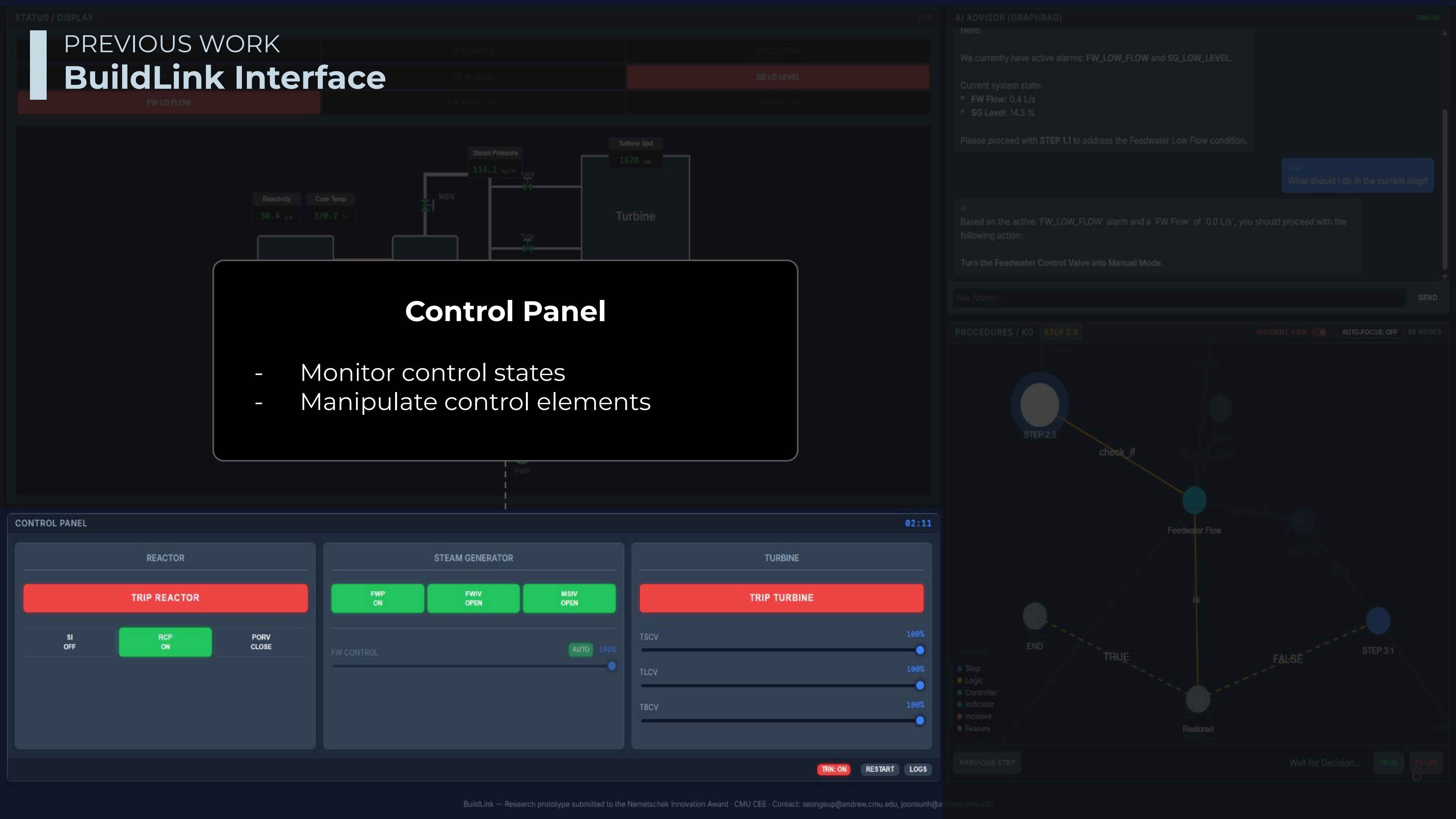

→ *An operator can move directly from a procedure step to a related past incident and back again.*

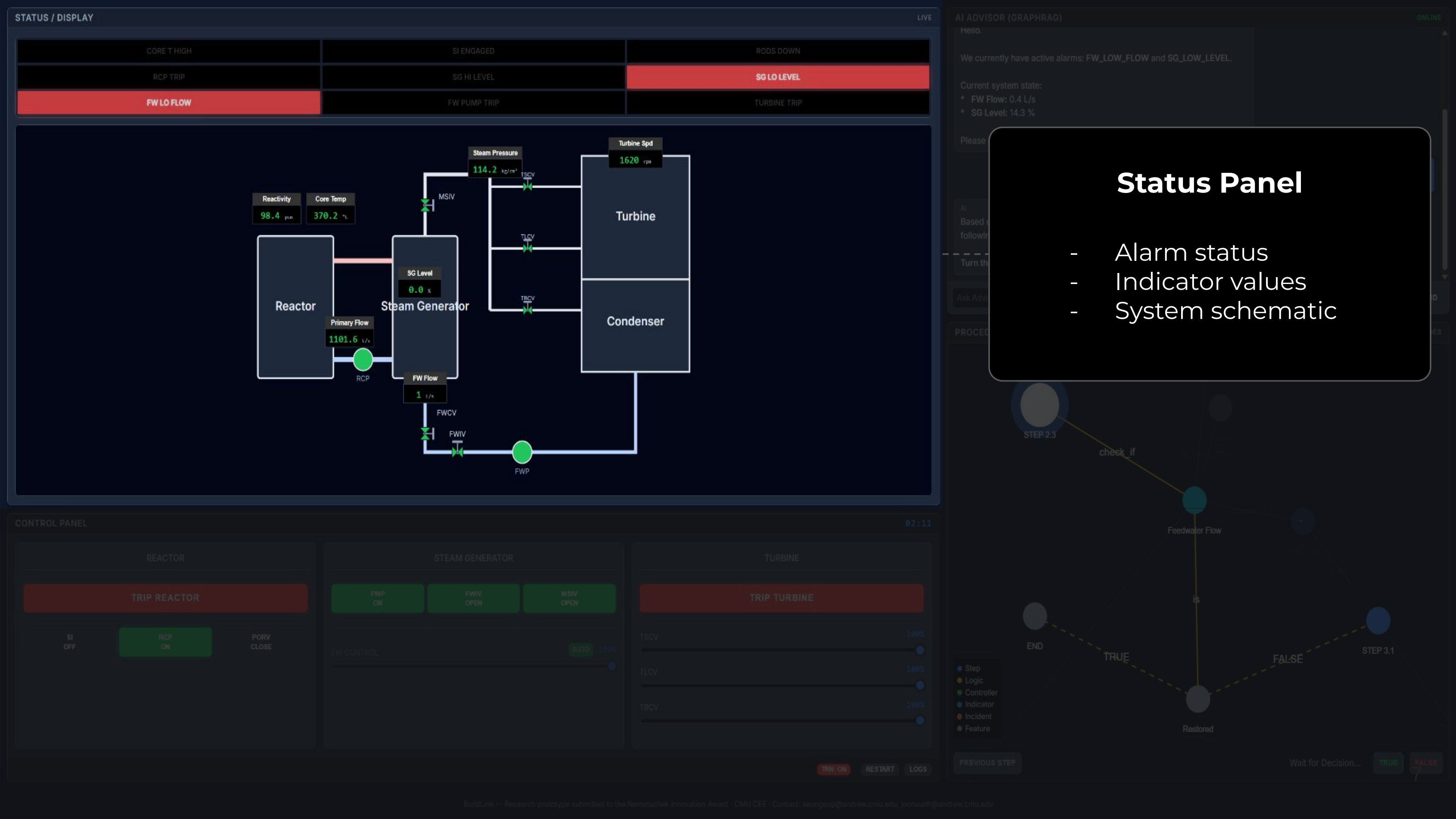
PREVIOUS WORK

# BuildLink Interface

## Control Panel

- Monitor control states
- Manipulate control elements



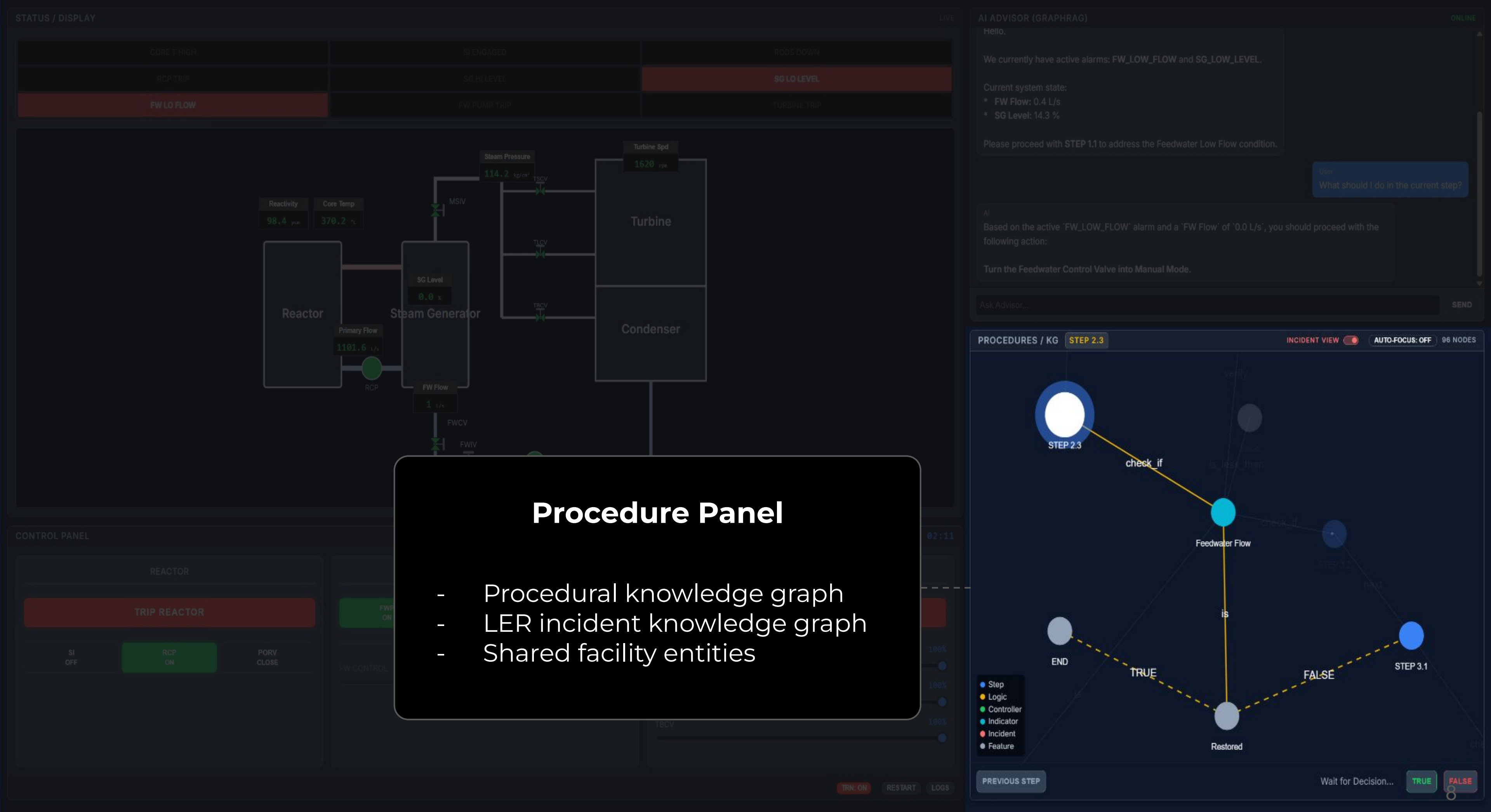



## Status Panel

- Alarm status
- Indicator values
- System schematic

# Procedure Panel

- Procedural knowledge graph
- LER incident knowledge graph
- Shared facility entities



PREVIOUS WORK

# From Text Procedure to Knowledge Graph

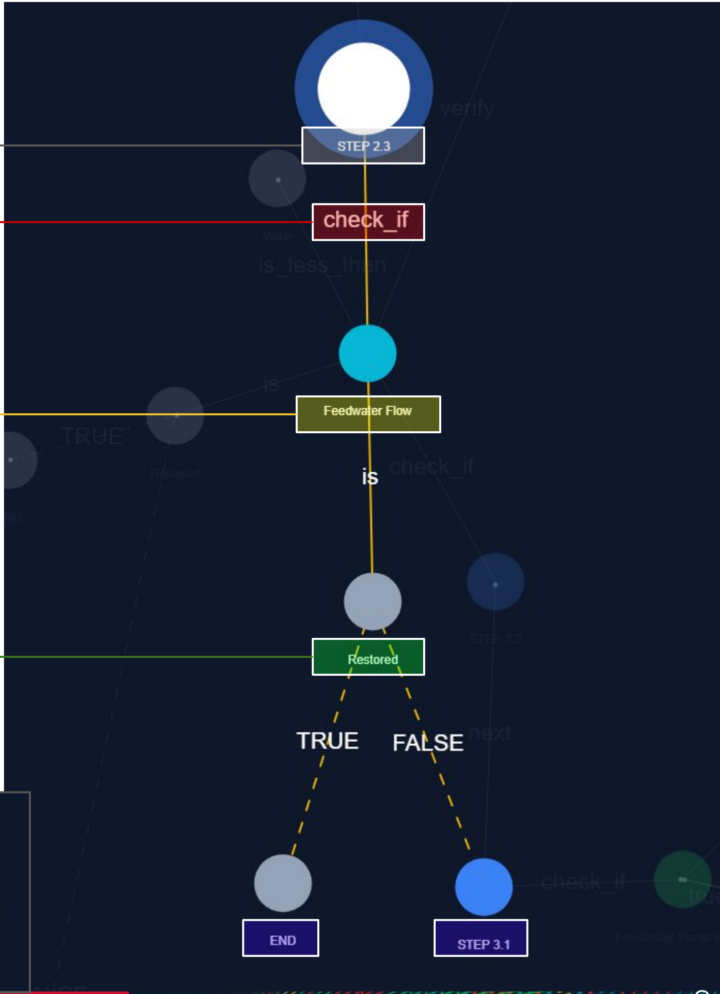
Step 2.3 Check if the Feedwater Flow has been restored

If the Feedwater Flow has been restored

- End procedure

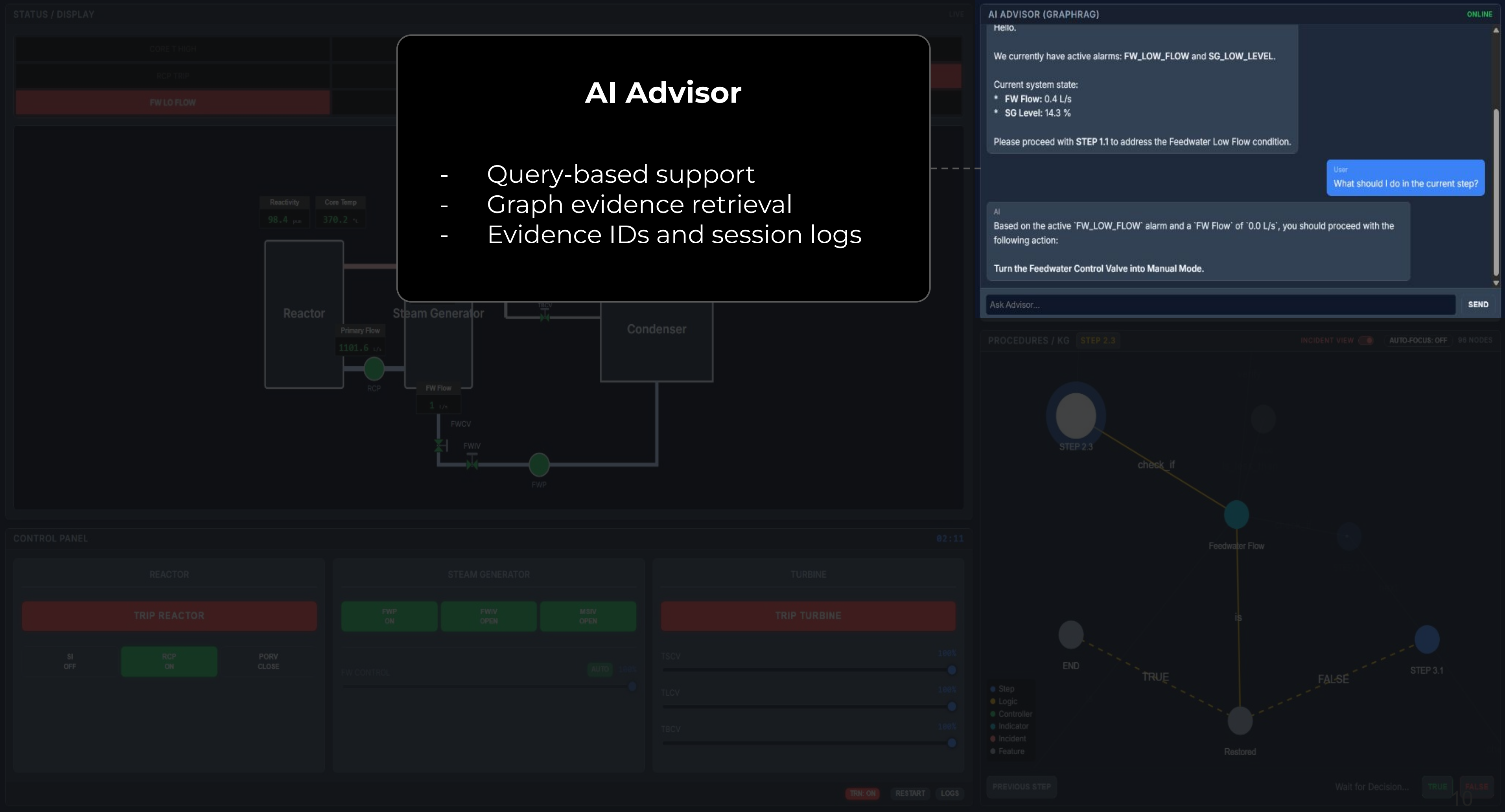
If the Feedwater Flow had not been restored

- Go to step 3.1




# AI Advisor

- Query-based support
- Graph evidence retrieval
- Evidence IDs and session logs



# From AI Advisor to AI Actor

### BuildLink (Prior Work)

## AI = Advisor

AI only **'presents'** procedure & incident evidence, final judgment & action remain with the **human**

→ *Explaining 'why this answer' is enough*

### This Study

## AI = Actor

AI directly **'selects and executes'** actions the simulator state actually **changes**

→ *Also record 'why this action, in this state'*

**Key question :** A scenario's final success or failure alone tells us nothing about why the AI took a given action, or how it connects to the procedure.

# Research Question

*How can AI-agent actions in nuclear abnormal operations*

***be constrained and recorded as an analyzable decision path?***

### Constrain

Limit actions to a predefined action space

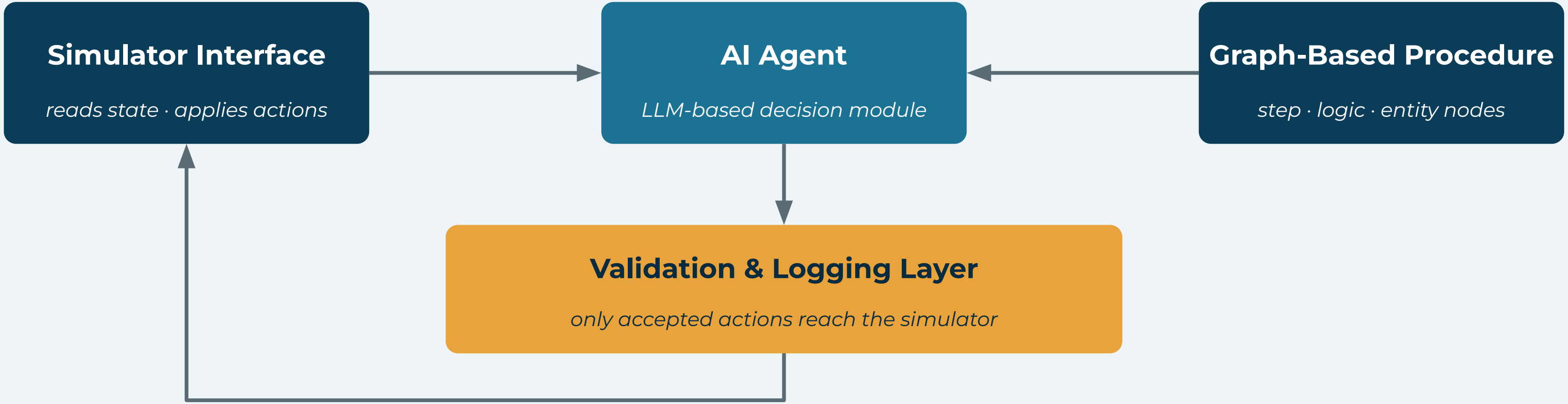
### Validate

Check procedure context and simulator rules

### Log

Record actions, validation, and state transitions

# How the Framework Works




*Every action is logged with its procedure context, validation result, and pre/post state.*

# From State Interpretation to Decision-Path Logging

1. **State** — Interpretation
2. **Procedure** — Logic
3. **Action** — Selection
4. **Validate** — Check feasibility
5. **Log** — Record context, result, transition

*Repeats until the scenario completes.*

**Nothing reaches the simulator unvalidated.**

Candidate actions are held until validated, enabling a split between accepted sequences and rejected attempts.

# Defining a Human-Interpretable Action Space

AI actions are limited by action type, target, and value range.

### Too Fine-Grained

Mouse clicks

Raw variable writes

### Action-Level Unit

Start / stop a component

Adjust a control valve

Select a branch

Advance a step

### Too Coarse

One action

per procedure step

*Human-readable and directly executable*

**Prototype: 25 predefined actions**

# Decision-Path Log Contents

| Logged Information | Examples |
|---|---|
| **Context** | Time, procedure step |
| **Action** | Type, target, value |
| **Validation** | Accepted/rejected, reason |
| **State Transition** | State before and after |
| **Agent Explanation** | Supporting information only |

*※ Reasoning is only auxiliary information. The primary basis for analysis is the evidence chain: procedure context → validation result → state transition.*

# Prototype Verification: Scale & Results

**99**
Graph nodes

**85**
Graph edges

**25**
Action definitions

**2 / 4**
Accepted / Rejected

## Representative log entries (excerpt)

| Procedure Step | Proposed Action | Result | Reason / State Transition |
|---|---|---|---|
| Step 1.1 | advance_procedure | ✔ **Accepted** | current_step changed to Step 1.2 |
| Step 1.2 | trip_reactor | ✘ **Rejected** | Not linked to current procedure context |
| Step 2.2 | feedwater valve = 1 | ✔ **Accepted** | feedwater_control_valve changed to 1 |
| Step 2.2 | feedwater valve = 1.5 | ✘ **Rejected** | Value must be between 0 and 1 |
| Step 2.3 | branch select = FALSE | ✘ **Rejected** | No valid graph transition found |

DISCUSSION

# Current Scope & Future Work

1. Extended BuildLink's procedure knowledge graph into a constraint on AI-agent execution
2. Reconstructable as an action-level Decision-Path Log, not just final success/failure
3. The core is an evidence chain: procedure context → validation result → state transition

## Current Scope

- Initial prototype focused on framework feasibility
- Validation depends on the available procedure graph and action definitions
- Agent reasoning is used as supporting information

## Future Work

- Refine action definitions and validation constraints
- Add a human review interface for accept/reject decisions
- Explore incident-knowledge integration and dynamic HRA frameworks

# Thank You

**Seongeun Park**

seongeup@andrew.cmu.edu

Carnegie Mellon University · Dept. of Civil and Environmental Engineering

## Acknowledgements

Pennsylvania Infrastructure Technology Alliance (PITA)

CUAHSI Hydroinformatics Innovation Fellowship (NSF EAR-1849458)

Chishiki-AI SCIPE Graduate Fellowship (NSF 2321040)

# References


Boring, R. L., Lew, R., & Ulrich, T. A. (2025). Rancor Integrated Procedure System (RIPS): A computer-based procedure platform for advanced reactor research. Idaho National Laboratory.

Boring, R. L., Thomas, K. D., Ulrich, T. A., & Lew, R. T. (2015). Computerized operator support systems to aid decision making in nuclear power plants. Procedia Manufacturing, 3, 5261–5268.

Lin, L., Athe, P., Rouxelin, P., Avramova, M., Gupta, A., Youngblood, R., Lane, J., & Dinh, N. (2021). Development and assessment of a nearly autonomous management and control system for advanced reactors. Annals of Nuclear Energy, 150, 107861.

Park, S., Hwang, J., Jeon, K., Boring, R. L., & Tang, P. (2026). BuildLink: Graph-structured procedural representation with incident integration. Proceedings of the International Conference on Computing in Civil Engineering 2026, ASCE.

Ulrich, T. A., Lew, R., Werner, S., & Boring, R. L. (2017). Rancor: A gamified microworld nuclear power plant simulation for engineering psychology research and process control applications. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 61, 398–402.