



ITALIAN NATIONAL AGENCY FOR
NEW TECHNOLOGIES, ENERGY AND
SUSTAINABLE ECONOMIC DEVELOPMENT



SAPIENZA
UNIVERSITÀ DI ROMA

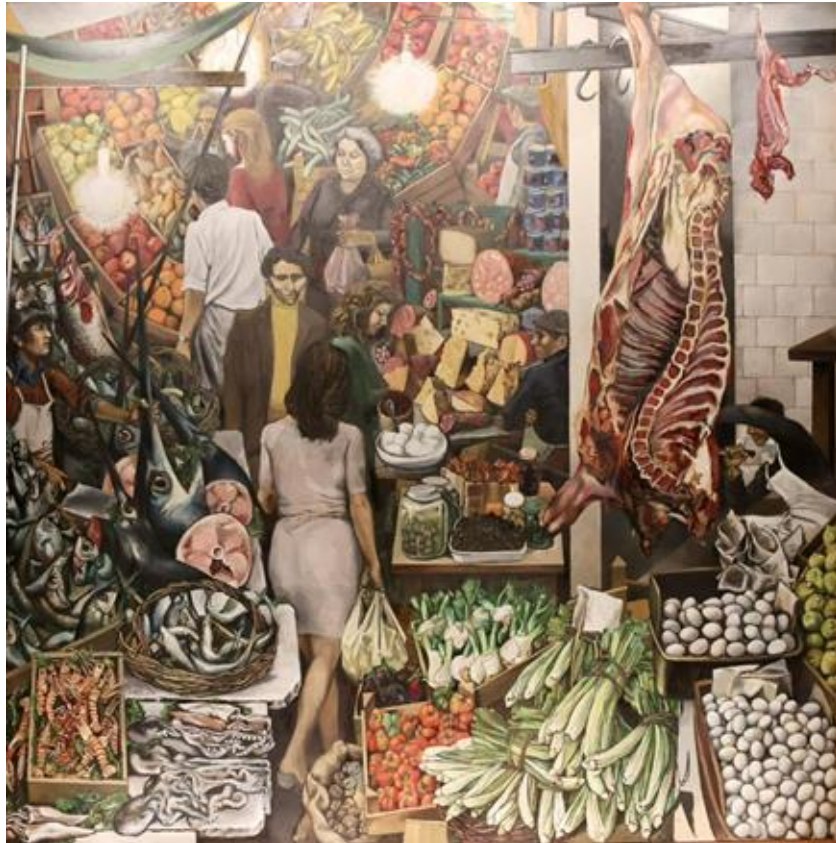
Simulating the Human in FRAM: A Practical Approach Using LLM-Based Agents

FRAMily 2026, Munich 15th-17th June, 2026

Gianluca Bottalico, Matteo Cesaroni, Antonio De Nicola,
Maria Guariglia Migliore, Riccardo Patriarca



Why Human Variability Matters?



La Vucciria (1974) - Renato Guttuso

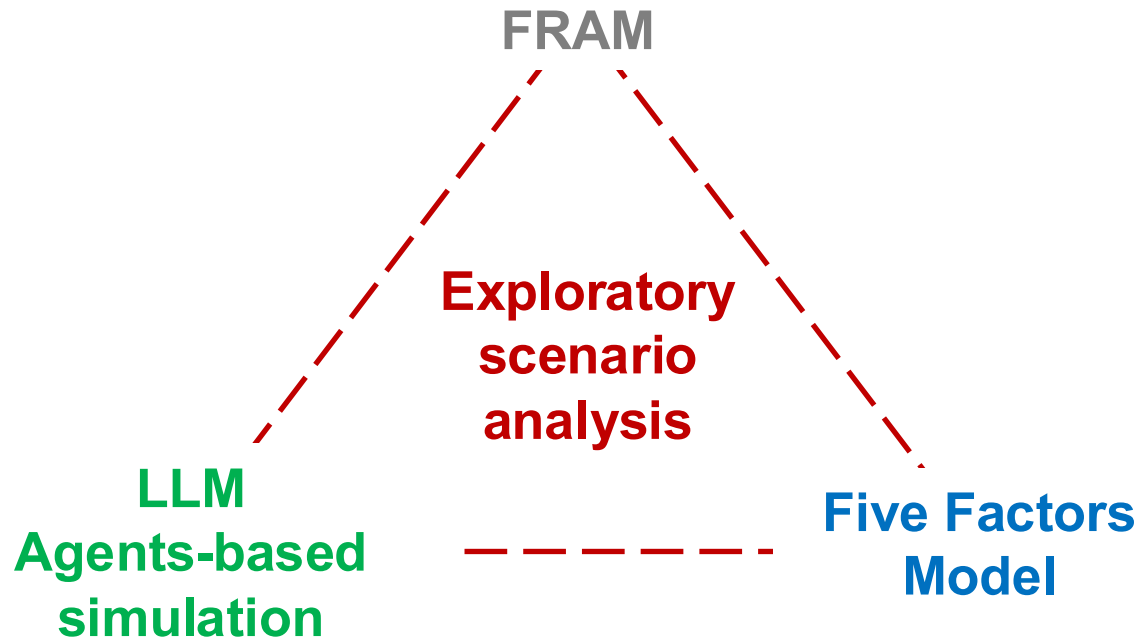
- ✗ Humans are often modelled as
 - Procedures
 - Probabilities
 - Average operators
- ✓ Reality is
 - Diverse individuals
 - Different goals
 - Different perceptions
 - Different decisions

No two people experience the same situation in exactly the same way!

Limitations of the current approaches

- Existing approaches struggle to explain
 - Why variability emerges
 - How different individuals react to the same context
 - How interactions amplify or dampen variability
- Key question
 - Can we model human variability as an emergent phenomenon?

Our Proposal



Our Goal

- Exploration not prediction
 - Exploration of plausible behaviors
 - Understanding variability generation mechanisms
 - Assessment of system resilience

Case Study: TAPIRO Research Nuclear Reactor

TAPIRO (Fast Flux Test Reactor at Zero Power)

- It is a specialized nuclear research reactor operated by **ENEA** near Rome, Italy.
- Designed to generate an **intense fast-neutron flux** for scientific research and experimental activities.
- It features a compact core fueled with **highly enriched uranium–molybdenum alloy**.
- The reactor is enclosed within a thick borated-concrete biological shield ensuring safe operation.

Why is it relevant?

- Safety-critical socio-technical environment
- Strong interaction between procedures and human decisions
- High operational reliability requirements



Main Technical Features of TAPIRO

Reactor type:

Fast neutron
research reactor

**Maximum
power:**

5 kW

Moderator:

None (fast reactor)

Fuel:

Highly enriched
Uranium–
Molybdenum alloy

Neutron flux:

up to 4×10^{12} n/cm²/s

**Experimental
capability:**

Extensive
experimental
irradiation channels



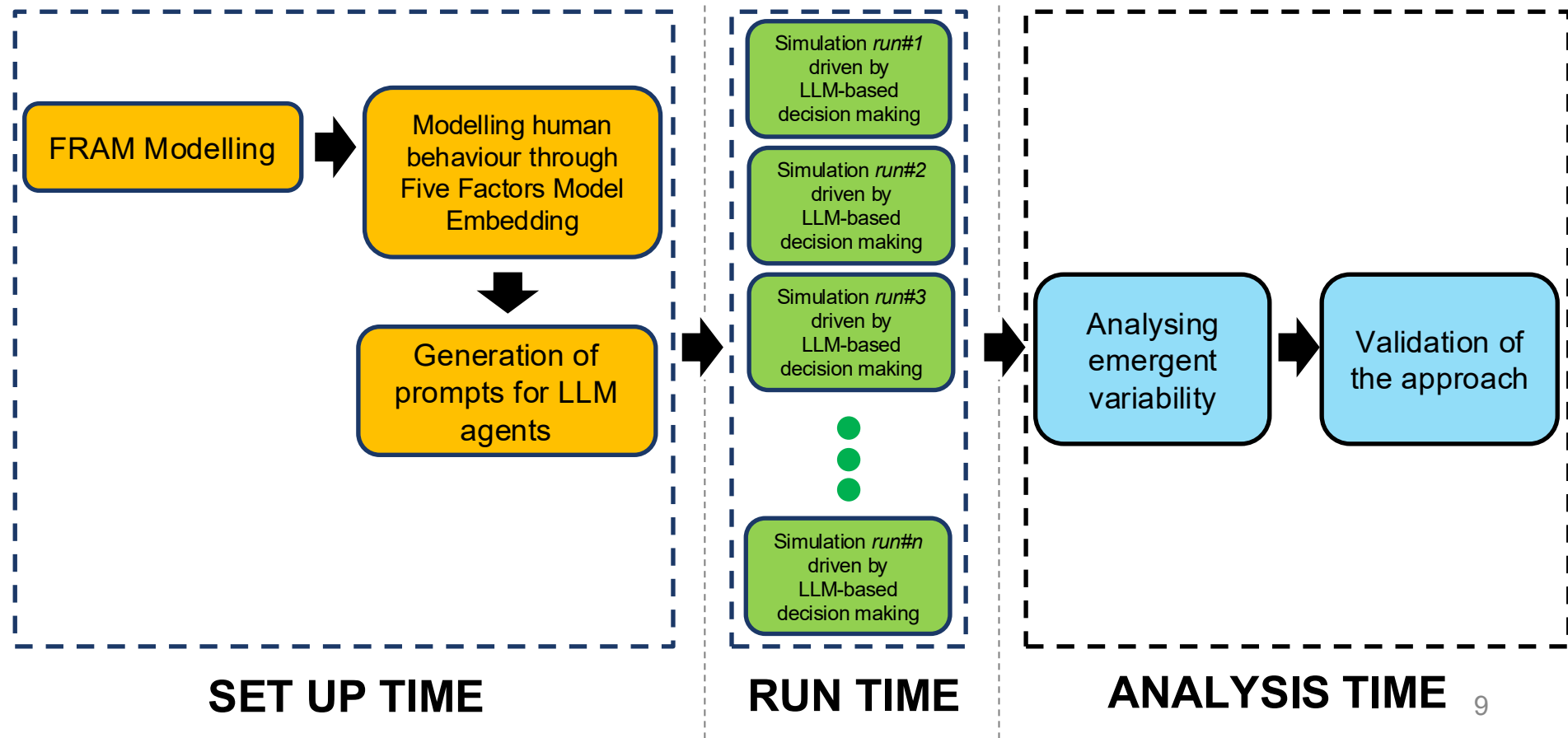
- European-level research infrastructure supporting nuclear safety, advanced reactor technologies and operator training.

Why TAPIRO?

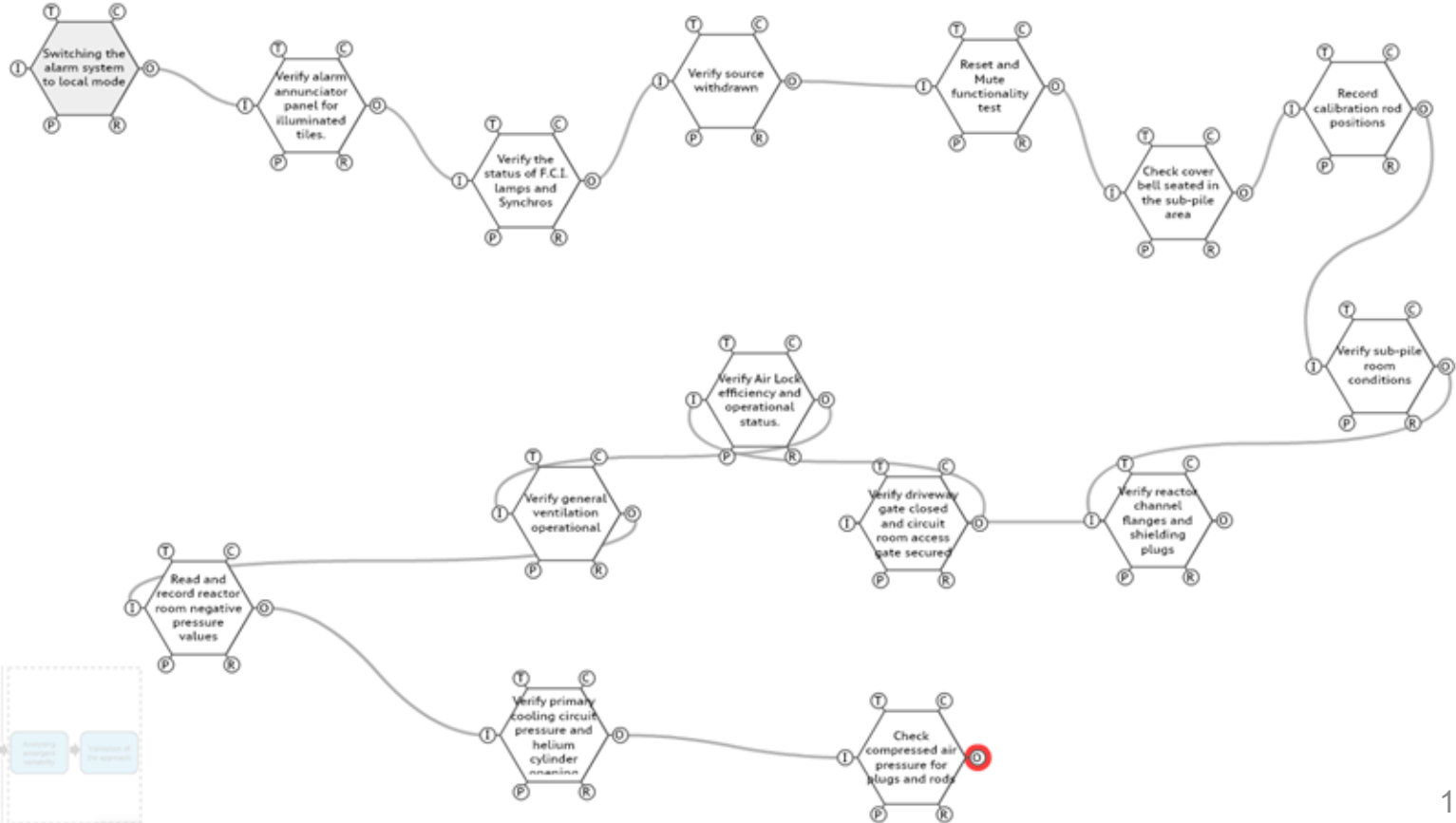
- A tightly controlled, high-reliability setting where human variability can significantly influence operational performance and resilience.
- Research nuclear reactor operations involve:
 - Strict procedures
 - Human decision-making
 - Coordination among actors
 - Safety-critical activities
- Focus Scenario
 - **Reactor startup procedure**
- Objective
 - Investigate how operator variability affects system performance and how variability propagates through FRAM functions.



Overview of the Simulation Approach



FRAM Process Instance of the Reactor Startup



Modeling Humans with LLMs

Our Approach

Operators are represented by LLM agents that:

- Perceive context
- Interpret goals
- Make decisions
- Explain reasoning



Behavior emerges dynamically from context



Personality Modelling: Five Factors Model

It reflects creativity, curiosity, and a willingness to try new experiences.



Openness

It measures self-discipline, organization, and a strong sense of duty or goal-directed behavior



Conscientiousness



Extraversion

It is characterized by high energy, positive emotions, sociability, and assertiveness.

It reflects tendency to be compassionate, cooperative, trusting, and helpful toward others.



Agreeableness



Neuroticism

It assesses emotional stability and the tendency to experience negative emotions like anxiety, moodiness, or stress.

Paul T. Costa JR, Robert R. McCrae, *Four Ways Five Factors are Basic*, 1991



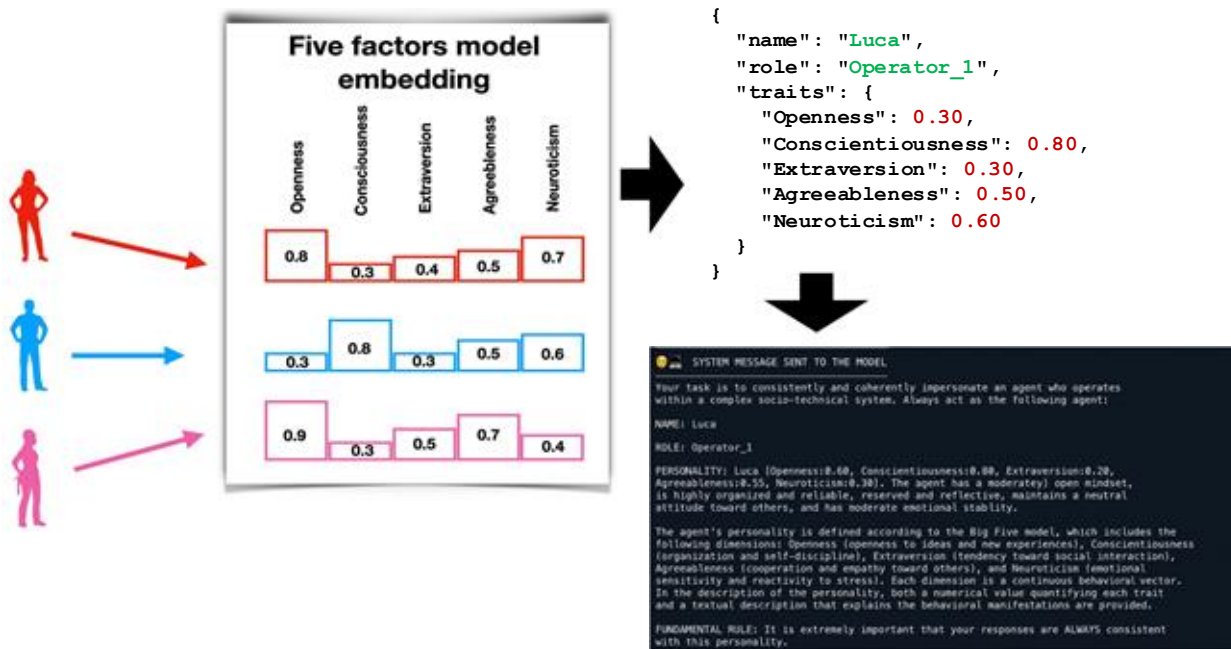
Five Factors Model Embedding and Prompt Generation

Simulation *run*
driven by
LLM-based
decision making

Analysing
emergent
variability

Hypothesis

Different personality profiles
lead to different variability
patterns



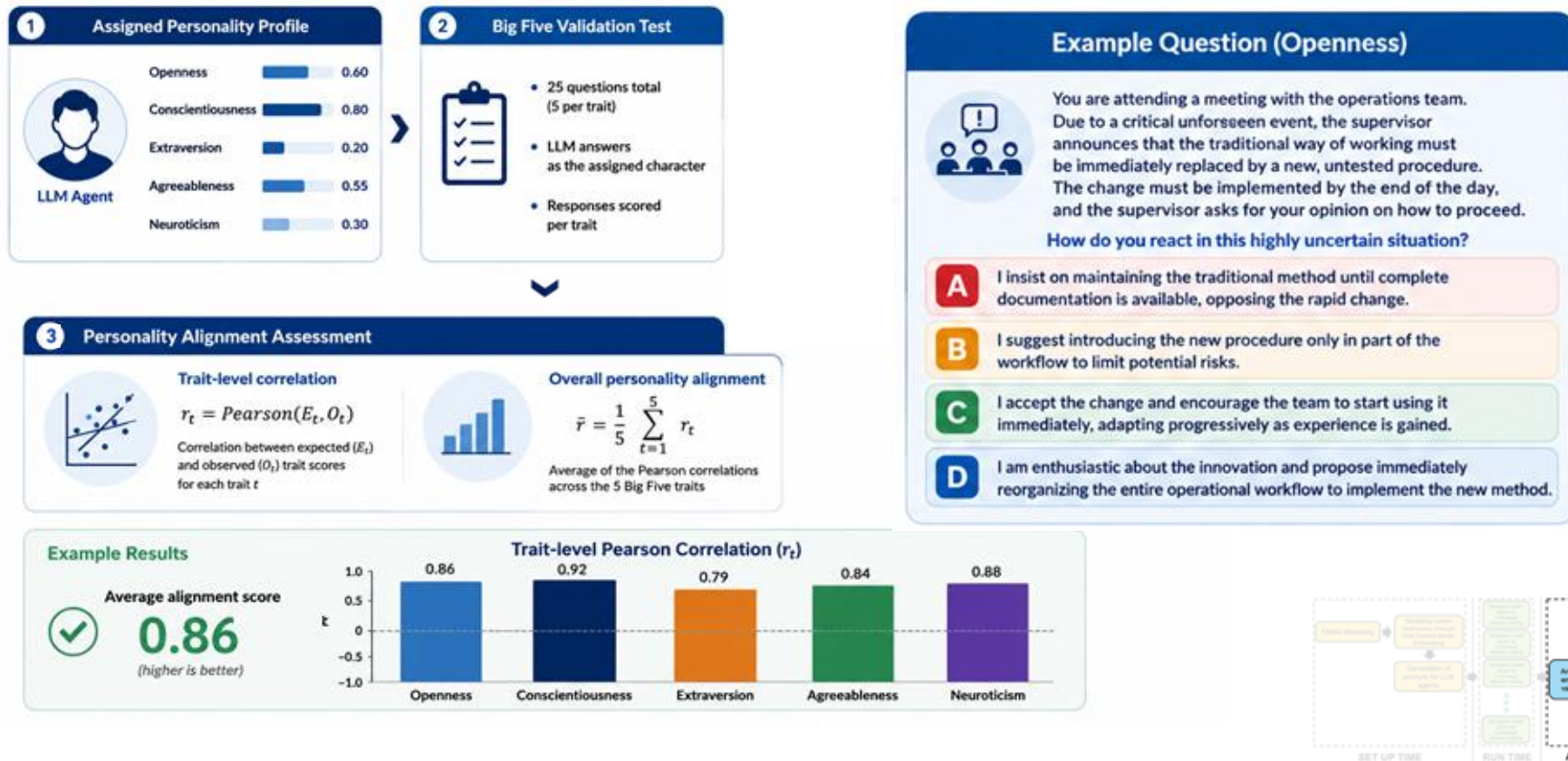
Personality traits are used to generate heterogeneous decision-making behaviors across simulation runs.



Validation of the approach

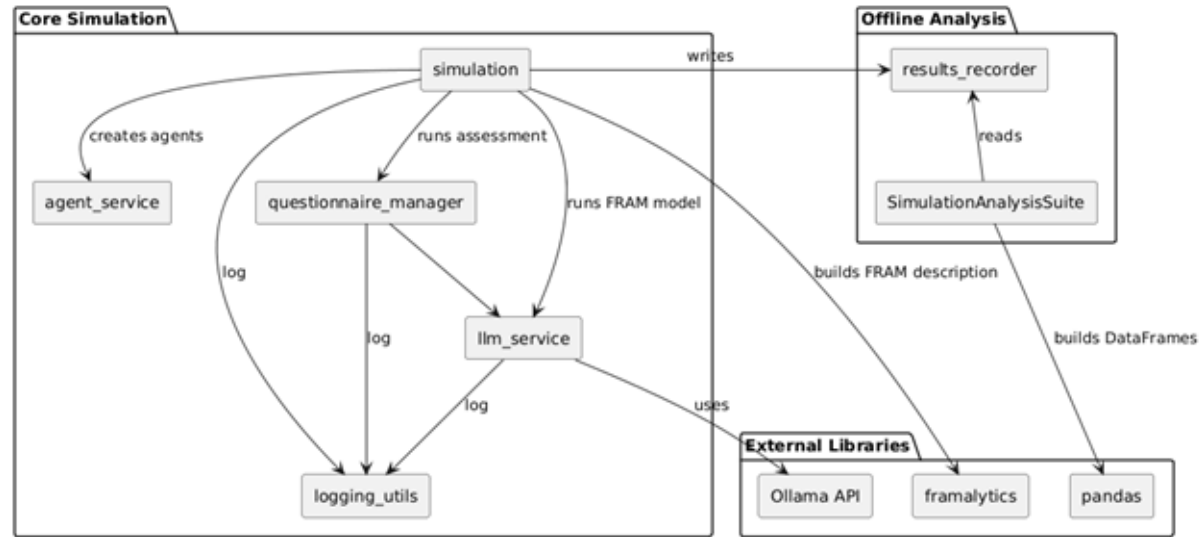
14

We validate whether LLM agents behave consistently with the assigned personality profile



Simulator Implementation: Technical Architecture

- The simulator has been implemented in Python
- LLM: OpenAI's open-weight gpt-oss:20b with 21 billion parameters and 3.6 billion active parameters
- GPU: NVIDIA GeForce 4080 SUPER (16 GB GDDR6X)
- FRAMEWORK: Ollama



Running the Experiments

Same scenario, many runs, emergent variability

1. Single operational scenario

TAPIRO Research Nuclear Reactor



FRAM model of the scenario

Fixed Elements Across Runs



Procedures



System Configuration



Goals & Constraints



2. Many simulation runs (varying factors)

Each run uses a different combination of:

Agent Personalities



Initial Conditions



Run 1



Run 2



Run 3



...

Run N



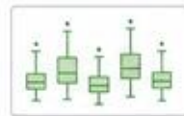
N simulation runs

LLM-based agents make decisions at runtime.
Different decisions lead to different trajectories

3. Outcome analysis



Function Completion Times



Action Choices



Variability Indicators



System Performance Measures

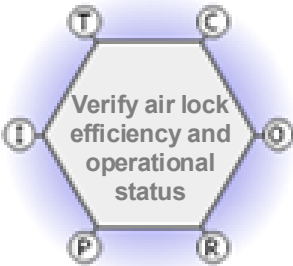


Goal

Observe patterns across many simulation runs and understand how variability emerges

Same FRAM Function, Different Behaviors

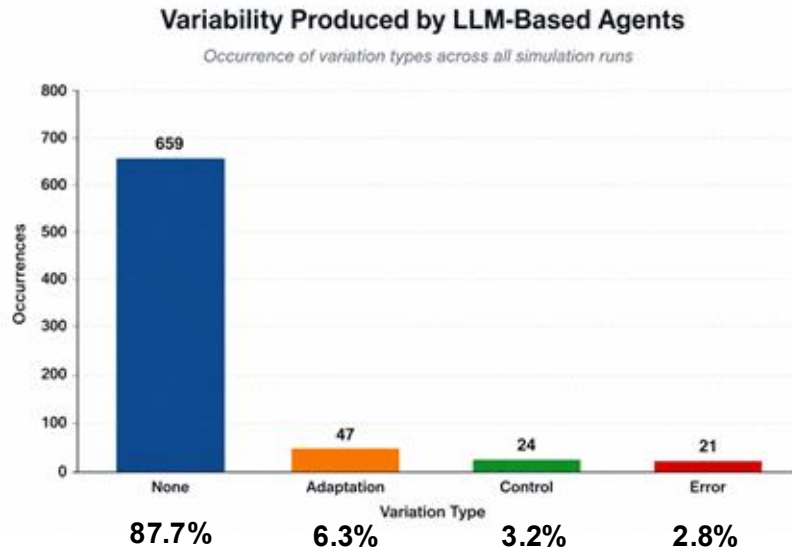
Verify air lock efficiency and operational status



	ACTION	TRAITS	VARIATION	IMPACT	DURATION
Operator 1 	I VERIFIED AIRLOCK EFFICIENCY AND OPERATIONAL STATUS	LOW CONSCIENTIOUSNESS → CROSS-CHECKS, PRESSURE	NONE	NONE	12 s 
Operator 2 	I VERIFIED AIRLOCK EFFICIENCY AND OPERATIONAL STATUS	HIGH CONSCIENTIOUSNESS → ENSURES INTEGRITY	NONE	NONE	18 s 
Operator 3 	I VERIFIED AIR LOCK-EFFICIENCY AND OPERATIONAL STATUS BY CHECKING THE GATE STATUS AND MONITORING THE PRIMARY COOLING CIRCUIT PRESSURE	HIGH OPENNESS, HIGH CONSCIENTIOUSNESS → TREND ANALYSIS	NONE	NONE	48 s 
Operator 4 	I VERIFIED THE AIR LOCK EFFICIENCY AND OPERATION AFTER THE COOLING CIRCUIT GATE CLOSES	HIGH NEUROTICISM → RISK CONCERN	ADAPTATION	INEFFICIENT AIR LOCK COULD ALLOW AIR INGRESS	24 s 

Variability Produced by LLM-Based Agents

Aggregate Results



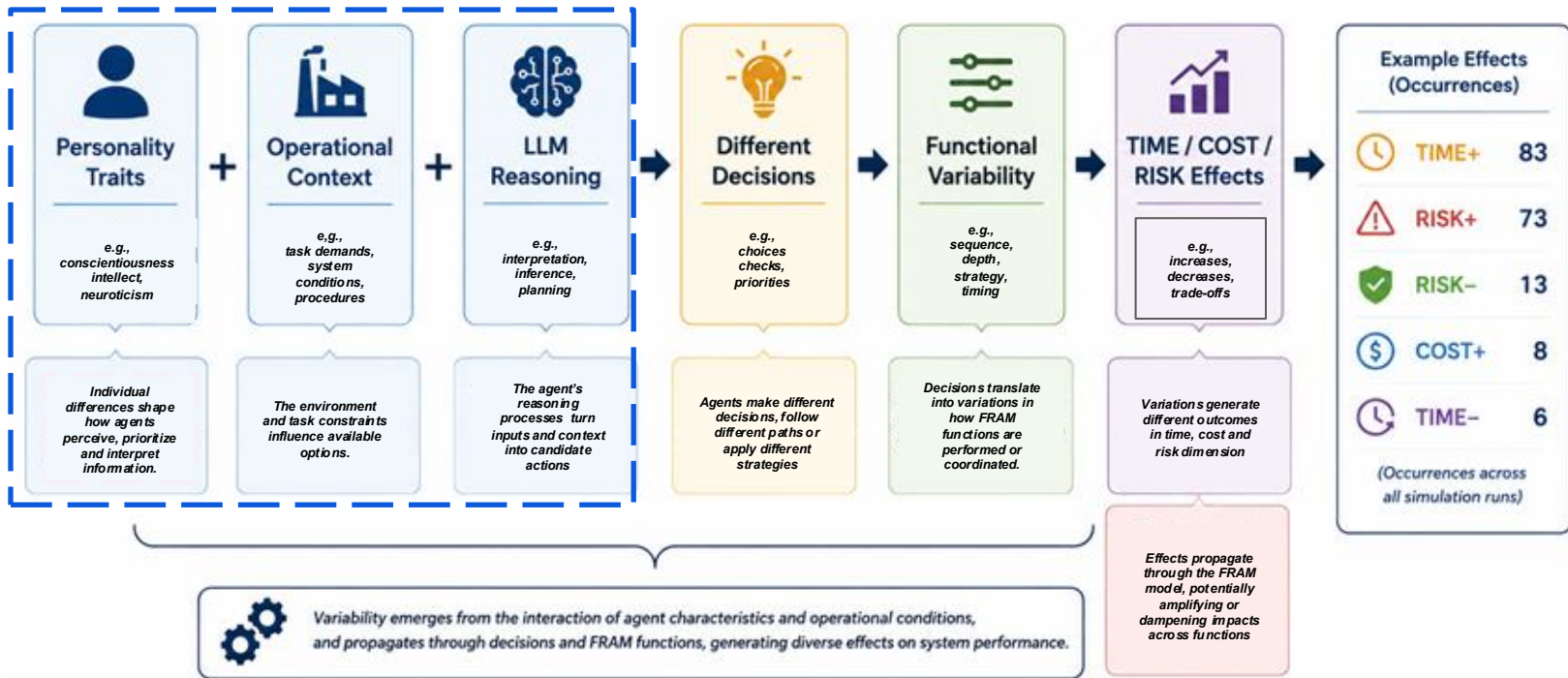
12.3% of actions resulted in functional changes due to adaptations, additional controls, or errors.

Different agent profiles generated distinct execution patterns, producing adaptations, controls, and occasional errors under the same operational conditions.

Examples of agents exhibiting high variability rates:

Agent	Varied Functions / Executed Functions	Variation rate
Operator_47	14/15	0.933
Operator_9	9/16	0.563
Operator_42	7/17	0.412
Operator_25	7/17	0.412
Operator_23	5/5	1.000

How variability emerges

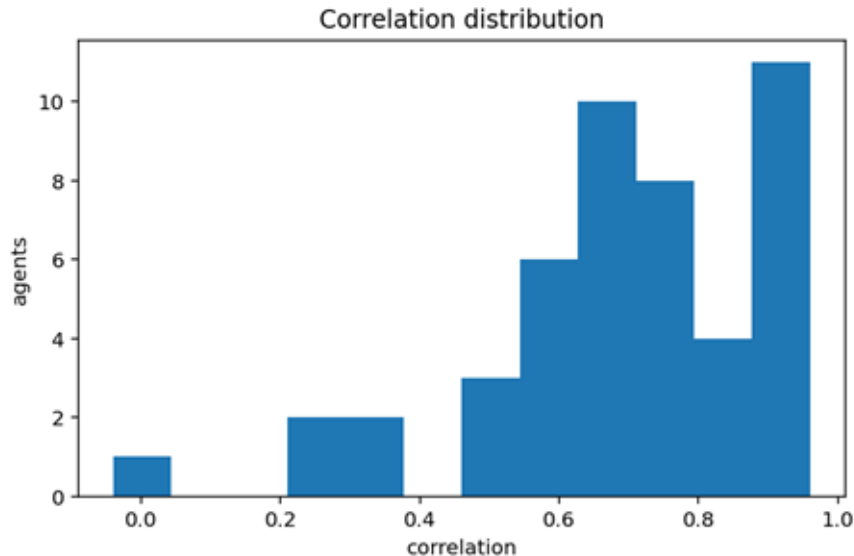


Behavioral Validity Assessment in the Case Study

Key results

- Median correlation = 0.71
- 51% of simulation runs achieved correlation ≥ 0.70

Metrics	Values
Analyzed runs	47
Mean correlation	0.68
Median	0.71
Correlation ≥ 0.70	51.1%



Results suggest that LLM-based agents can reproduce plausible behavioral patterns while preserving variability across executions.

Limitations and Future Work

Limitations

- Prompt sensitivity
- Hallucinations
- Limited transparency of LLM reasoning
- Dependence on predefined output formats

Future Work

- Improve transparency of LLM decision-making
- Extend to collaborative multi-agent scenarios
- Integrate neurophysiological data

Conclusions

Main Findings

1. LLM agents generate behavioral diversity.
2. Variability emerges non-uniformly.
3. Small decisions can propagate.

Thank you!

Acknowledgements

This work has been conducted in part within the project HARMONISE (Human neurophysiological Assessment for Risk Modelling in complex sOcio-techNical SystEms) funded by Sapienza University of Rome, Bando per la ricerca scientifica di Ateneo 2023.