SYSTEMS ENGINEERING RESEARCH CENTER | AIRC ACQUISITION INNOVATION RESEARCH CENTER | SERC RESEARCH REVIEW 2026

CLEARED
For Open Publication
Nov 02, 2026
Department of War
DEFENSE OFFICE OF PREPUBLICATION AND SECURITY REVIEW

WRT-2515:Development of Tools and Methods for Secure Cyber Resilient Engineering (SCRE)

28 September 2026

Megan Clifford, Tom McDermott (Stevens)

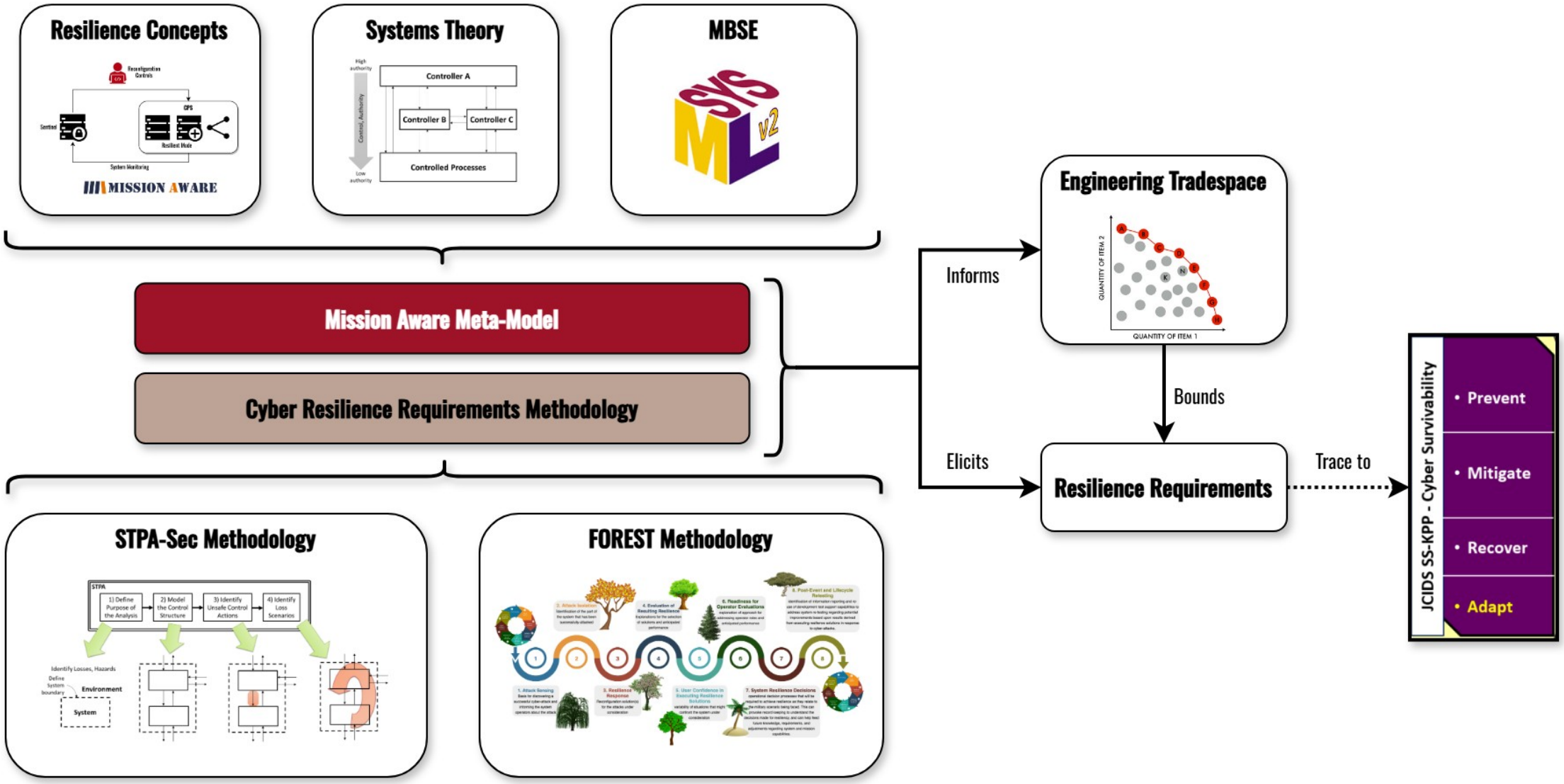
Kelli Essler, Stephen Adams, Mary Nerayo, Alex Kyer, Brian Walheim (VT)

NATIONAL SECURITY INSTITUTE VIRGINIA TECH

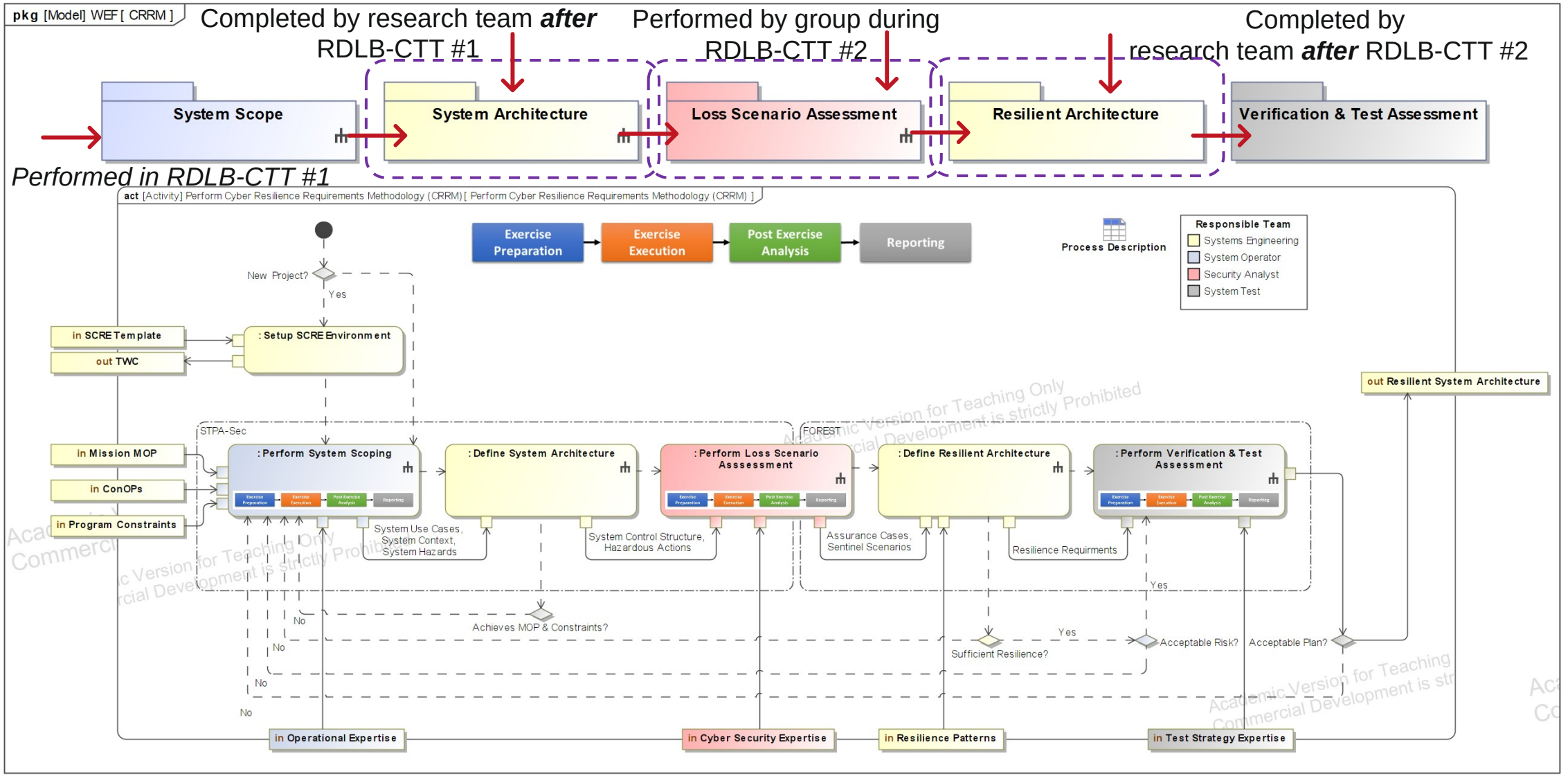
STEVENS INSTITUTE of TECHNOLOGY THE INNOVATION UNIVERSITY

Cleared for public release

SERC Foundational SCRE Methodologies



RESILIENCE DRIVEN LOSS-BASED – CYBER TABLE TOPS (RDLB-CTT)



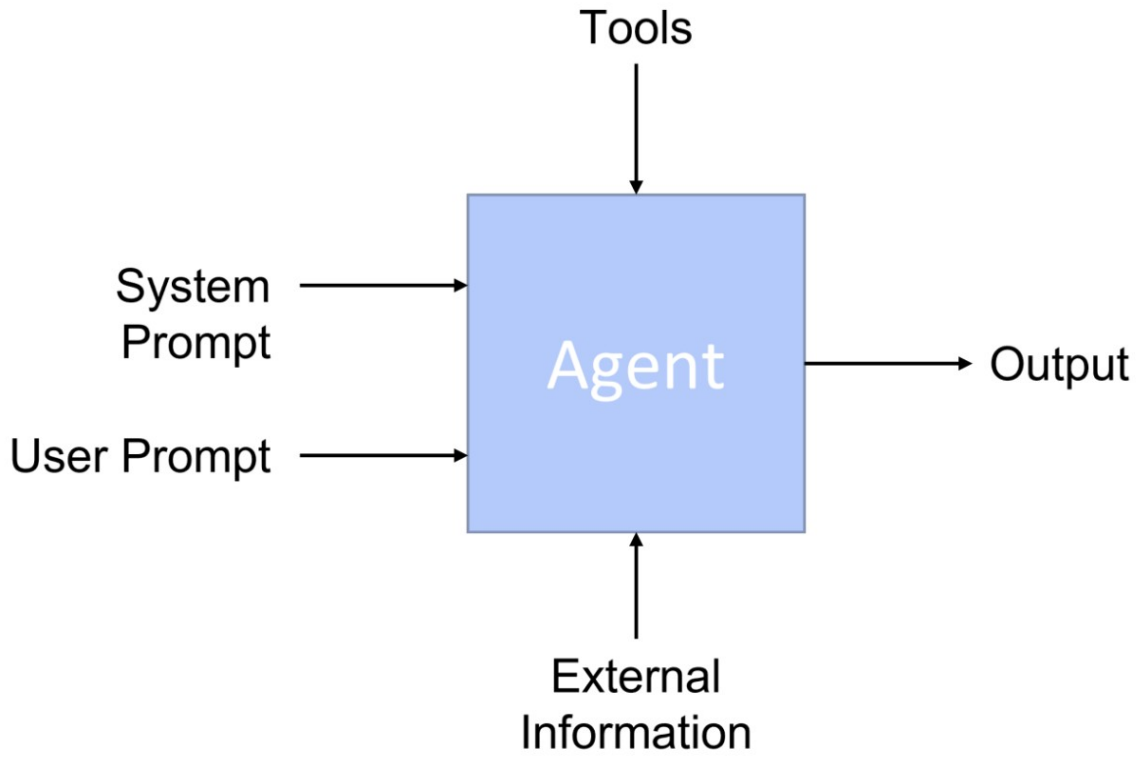
SCRE Tabletop Methods - Lessons Learned

- Model Walkthroughs must emphasize context and impact
  - Continuous reference to process steps and their purpose helped participants
  - Future exercises should strengthen connections between:
    - Modeling layers (use cases, control structures, hazard chains, etc.)
    - Mission assurance and system vulnerabilities
  - Improves engagement and clarity, especially for non-technical stakeholders
- Early and flexible engagement with stakeholders
  - Scheduling is a challenge, getting right set of domain experts (particularly SCADA level)
  - Need early alignment with communities of interest
  - Frequent domain expert involvement enriches scenarios and validates assumptions
- Visualization challenges with Cameo modeling tool
  - Live navigation in Cameo is cumbersome
    - Difficulty toggling views, navigating nested diagrams
  - But models informed narrative structure, initiated collaborative discussion, exercises, and learning

- Multiple complex processes
- System context is critical
- Graphical model visualization not interpretable by non-MBSE experts
- True deep system SME's need to be engaged differently
- Need simpler views for "context experts" not "deep design experts"
- Need to reduce cycle times: meetings to models to meetings (ideally hours)
- Opportunity to generative AI to make a significant difference

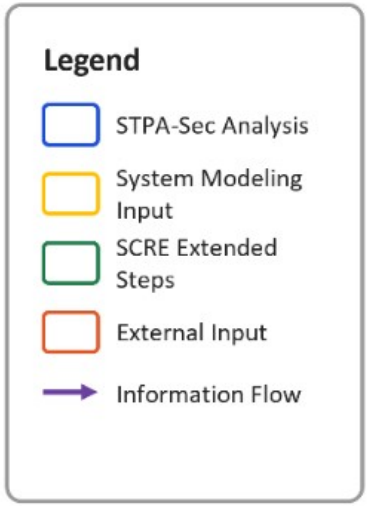
Basic Agent Structure

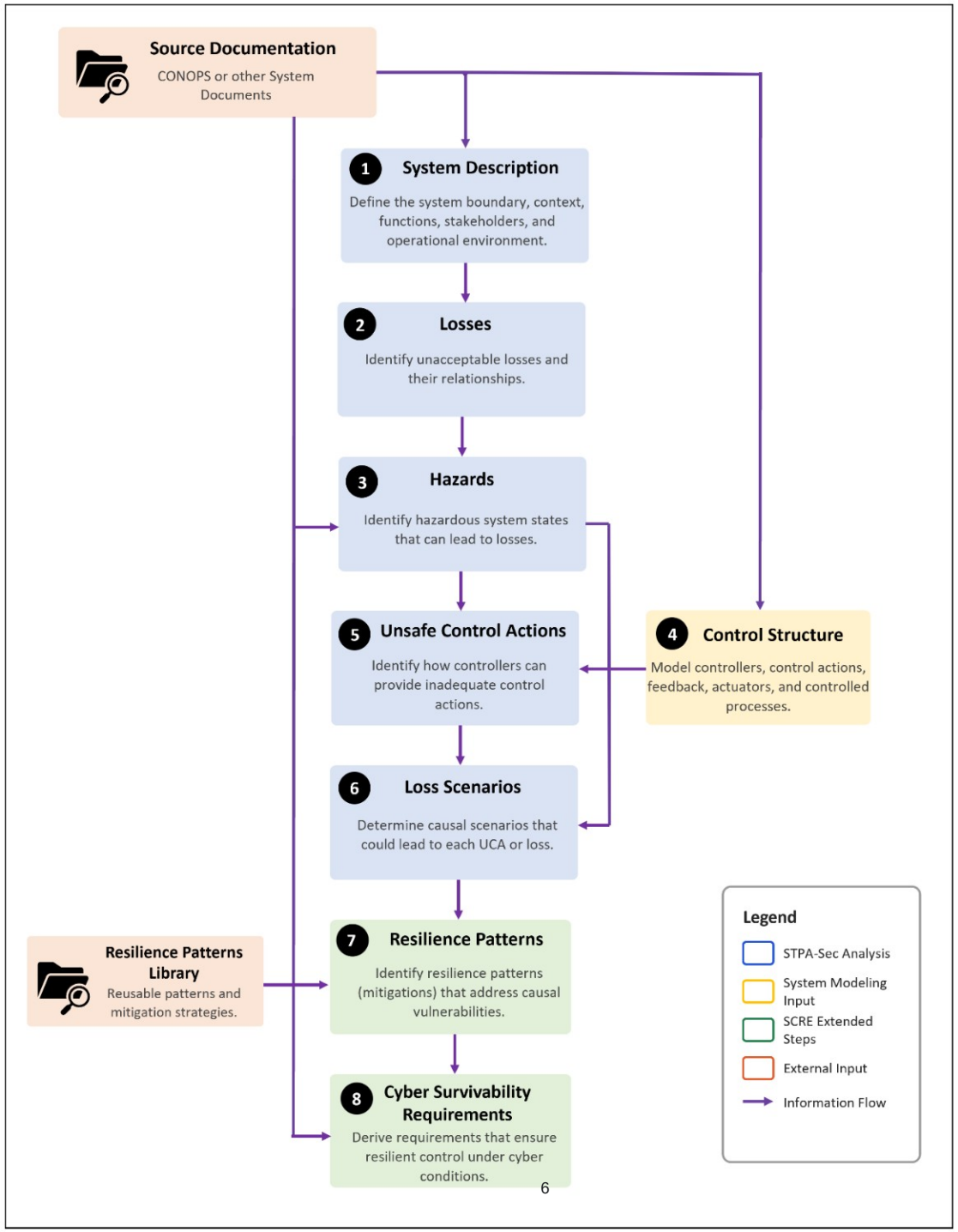
- An agentic approach offers many advantages:
  - **Task decomposition and constraint** reduce the likelihood of hallucinations by limiting each agent's scope to well-defined responsibilities.
  - **Modular design** enables individual agents to be developed, tested, and validated independently, improving reliability and maintainability.
  - **Scalability and extensibility** allow system capabilities to be expanded incrementally through the addition of new agents without requiring significant redesigning of existing components.



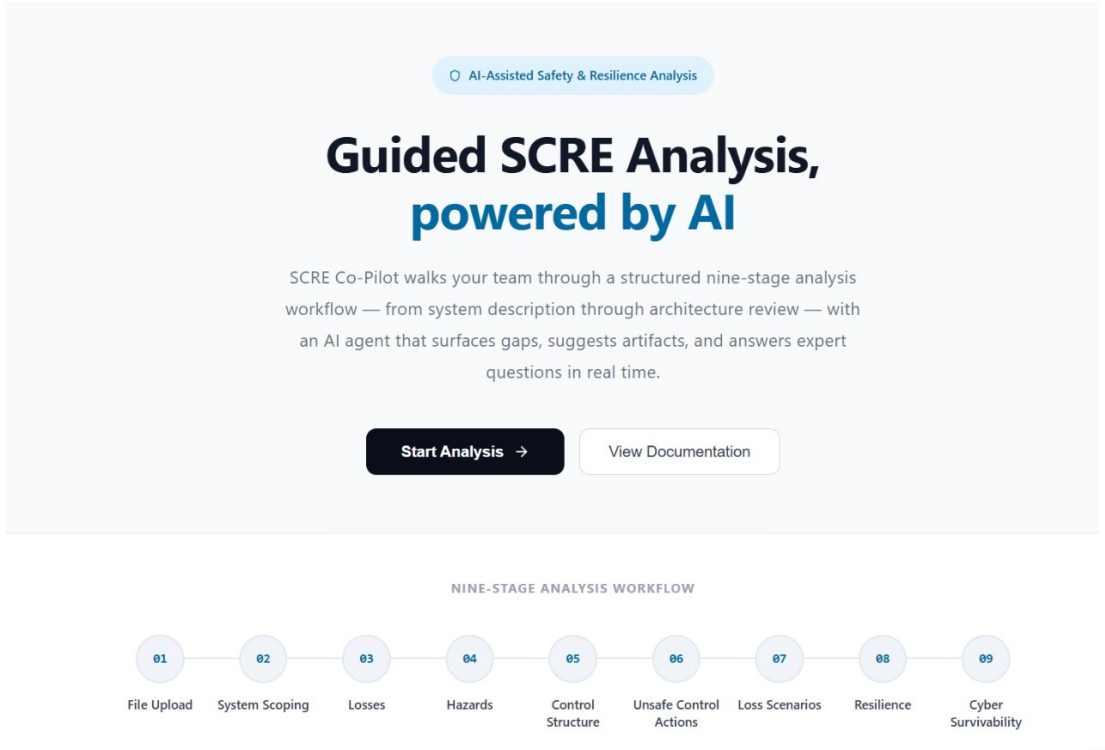
To support reliable handoffs between these specialized agents, each stage produces structured JSON outputs using standardized identifiers for major analytical artifacts, including SYS-###, L-###, H-###, and UCA-###. These identifiers establish explicit relationships between upstream and downstream results rather than relying on natural-language similarity to reconstruct those relationships.

Multi-agent architecture and information flow of the SCRE assistant





SCRE Co-Pilot



HOW TO USE THIS TOOL

Human expertise, AI assistance

The Co-Pilot is designed around a collaborative model — the analyst drives the analysis, and the tool provides targeted support at each stage. Neither replaces the other. The quality of the analysis depends on both.

THE ANALYST'S ROLE

The analyst brings domain knowledge, operational context, and engineering judgment that the tool cannot supply. This includes knowledge of the specific system under analysis, awareness of real-world threat actors and operational constraints, and the ability to evaluate whether an artifact is technically sound and appropriately scoped.

Throughout the workflow, the analyst retains full authority over every artifact. Suggestions from the Co-Pilot are reviewed and either approved, rejected, or modified before they are accepted. Nothing enters the analysis without explicit human confirmation.

THE CO-PILOT'S ROLE

The Co-Pilot provides two forms of support. The Stage Agent assists directly within the active stage — suggesting artifacts, checking that entries are correctly formed and traceable, and flagging gaps or inconsistencies the analyst may not have noticed. It is stage-locked, meaning its context and suggestions are always specific to the work at hand.

The STPA-Sec Expert is available at any point in the workflow for methodology questions, cross-stage reasoning, and review of completed artifacts. It reads all uploaded documents and analysis artifacts but does not modify them — its role is to inform, not to act.

Testing

- Each STPA-Sec agent was tested in isolation with repeated runs and evaluated against expert-derived gold standards, so results reflect the singular agent under test.

What we tested

Two systems: Silverfish and the Oil & Gas Pipeline

STPA-Sec agents only; the Resilience Pattern and Cyber Survivability agents had no benchmarks

How we isolated each agent

Upstream inputs came from the gold standard, not from real agent output

This removes cascading errors from the measurement

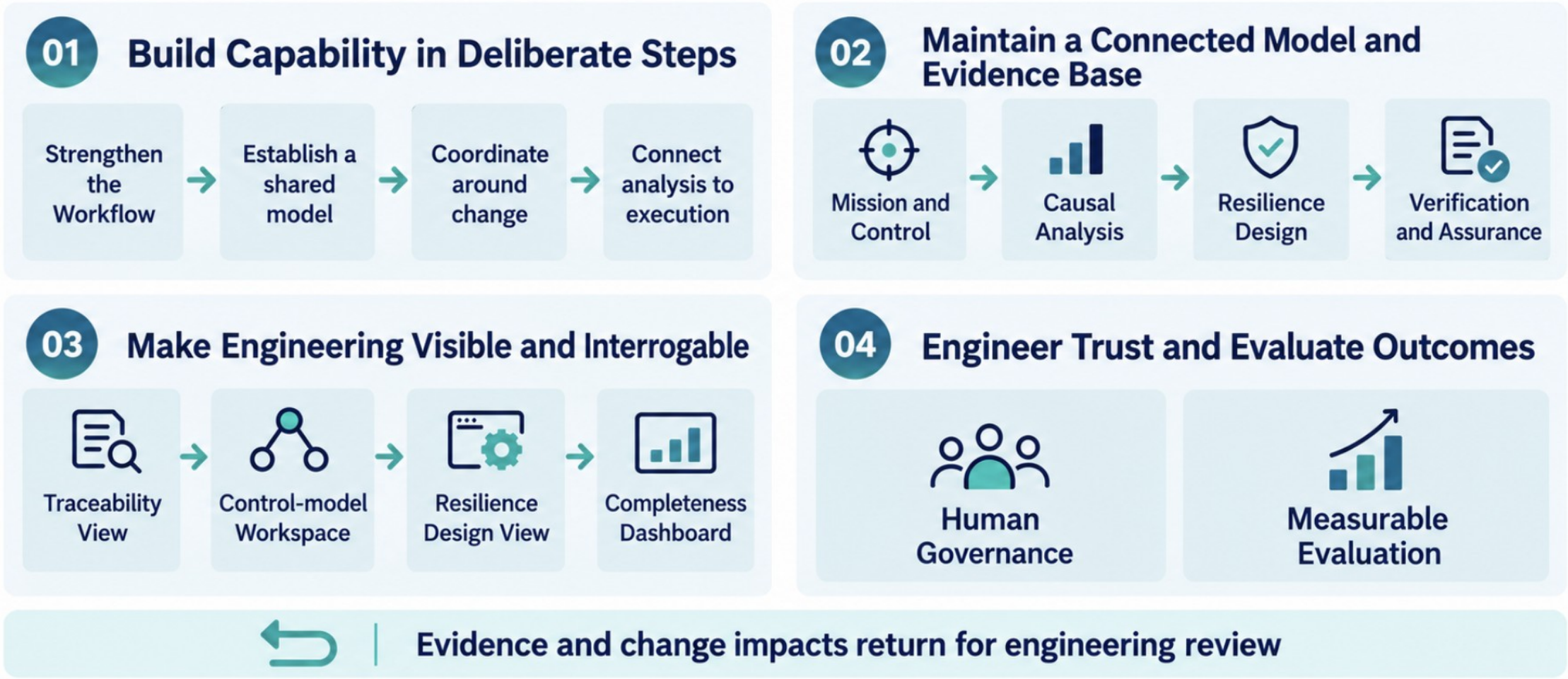
How we controlled the rest

Every condition was run 5 times and averaged, because LLM output is non-deterministic

The tests focus on input sensitivity (upstream agents and knowledge base)

Future Possibilities - Adaptive Partnership

- Engineering Cyber Resilience with Assurance at Mission Speed



DEFENSIBLE RESILIENCE DECISIONS SUPPORTED BY CONTINUOUSLY ASSESSABLE EVIDENCE