AI-enabled Rapid Intelligent Systems Engineering

High Confidence Software and Systems 2024 (HCSS 2024)

Annapolis, MD, Monday, May 6th, 2024

LOCKHEED MARTIN

Mauricio Castillo-Effen, Ph.D.

© 2024 LOCKHEED MARTIN CORPORATION

Overview

- Background
- Systems Engineering Sociotechnical Model
- Complexity
- GenAI Tech Stack
- Use Cases and Experiments
- Additional Thoughts

Goal: Present a snapshot of current thinking and efforts at LM ATL

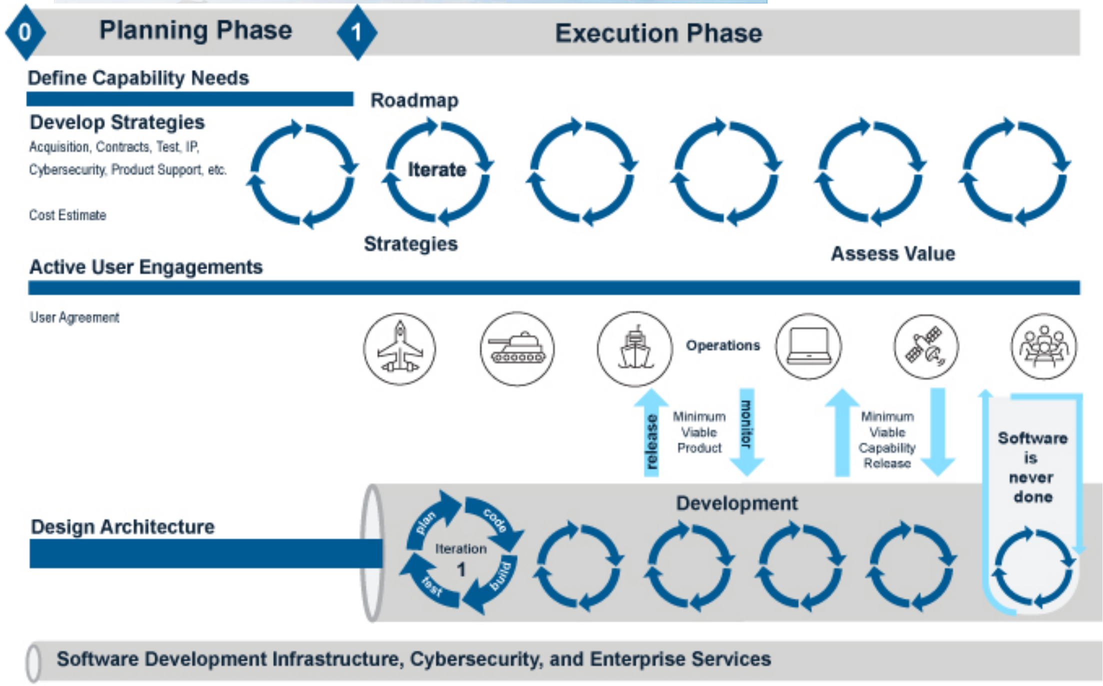
DISCLAIMER
The views and opinions presented in this presentation are solely the author's, and they do not represent the official policy or position of the Lockheed Martin Corporation or the Lockheed Martin Advanced Technology Laboratories.

© 2024 LOCKHEED MARTIN CORPORATION

Why “Rapid”?

- System acquisition is becoming “agile”
- Capabilities will be assessed and assembled **just-in-time** and **at the edge** to address specific mission needs
- Capabilities are added/composed and weaknesses addressed continually (“**software is never done**”) based on dynamic needs
- Stakeholders need to make risk-informed decisions and use the information obtained from SE activities
- Transition from traditional waterfall V-model systems engineering to Dev*Ops transforms the view of assurance from compliance- to value-driven

LM Project Hydra



From DoD instruction 5000.87: OPERATION OF THE SOFTWARE ACQUISITION PATHWAY

New acquisition regimes are characterized by **complexity** and **agility**. Can Generative AI help?

© 2024 LOCKHEED MARTIN CORPORATION
LOCKHEED MARTIN

Definitions

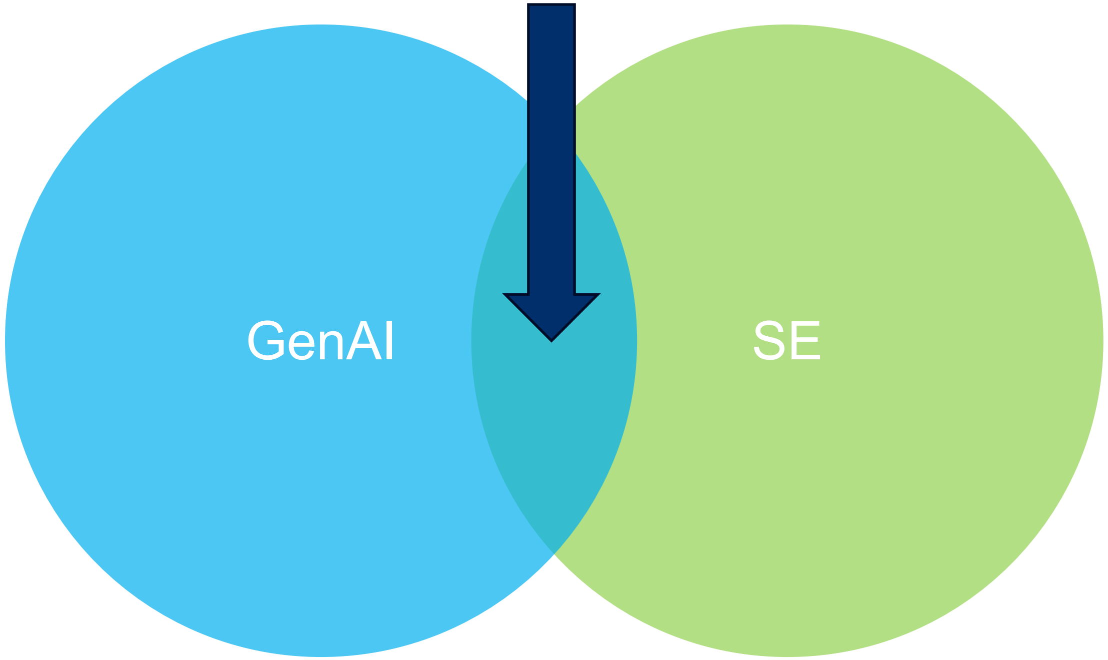
Rapid ("Agile"): A system development methodology based on iterative development where **requirements and solutions evolve** through collaboration between self-organizing cross-functional teams.

Generative AI (GenAI): a class of Artificial Intelligence technologies capable of **generating new content** ranging from text, images, and sound, to videos and code.

Systems Engineering: an **interdisciplinary** approach and means to enable the **realization of successful systems**. It focusses on defining customer needs and required functionality throughout the system's cycle, capturing requirements, then performing design synthesis and system validation while **considering the complete problem**: operations, cost and schedule, performance, training and support, test, disposal.

© 2024 LOCKHEED MARTIN CORPORATION
LOCKHEED MARTIN

Systems Engineering (SE) ∩ Generative AI (GenAI)



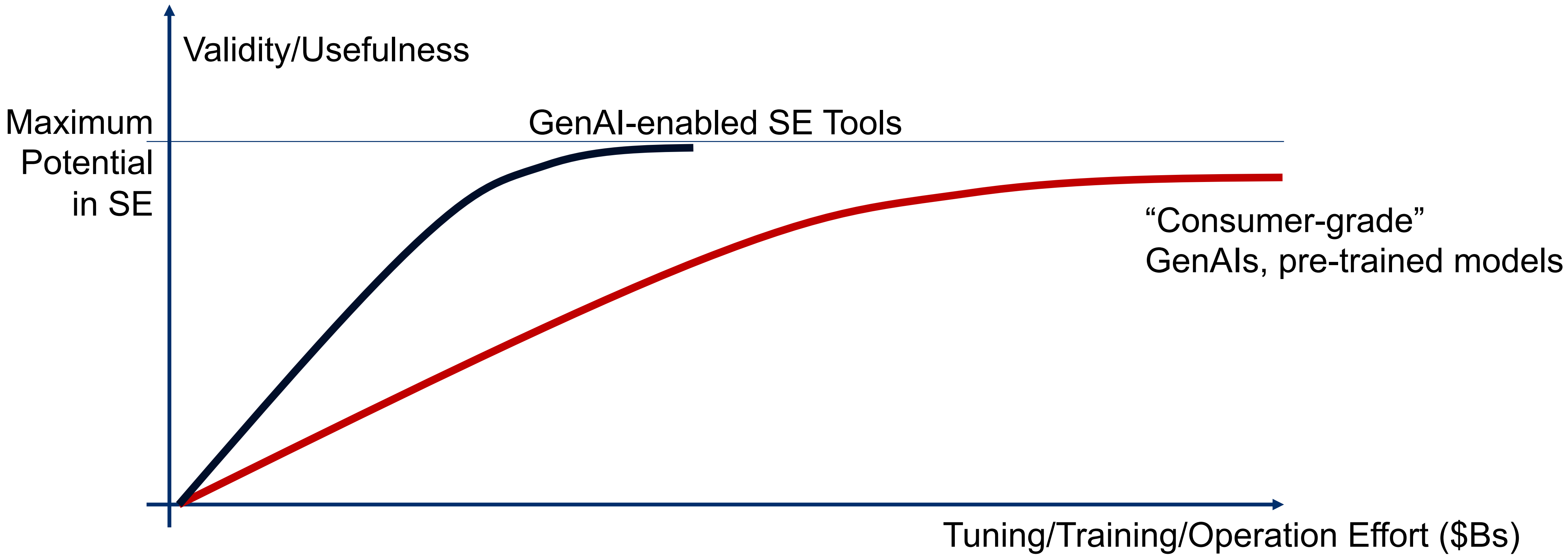
I. SE for GenAI. Exemplar concerns:
   - GenAI Lifecycle
   - Requirements for GenAI
   - GenAI architectures
   - TEVV of GenAI

II. GenAI for SE. Exemplar concerns:
   - Is GenAI useful for SE, despite weaknesses?
   - When/how is GenAI useful for SE?
   - How do we boost GenAI's utility and mitigate weaknesses?

Opportuntiy: the DIB employs one of the largest communities of Systems Engineers

© 2024 LOCKHEED MARTIN CORPORATION
LOCKHEED MARTIN

Hypothesis



Can we build useful capabilities for SE on pre-trained models?
Note: Usefulness does not imply validity or truthfulness

© 2024 LOCKHEED MARTIN CORPORATION

A Systems Engineering Approach

Research Questions:

- RQ1: If we want to introduce GenAI, how do we know what SE activities are most promising/suitable for augmentation?
  - What does the sociotechnical Systems Engineering process look like today? We need to create a "map" of SE activities, tasks, and roles and where AI has the best chances to help.
- RQ2: How would we measure improvements?
  - Intuitively: feed system with and without AI the same task, define acceptance criteria for generated outputs
  - Prerequisite: Qualify/quantify SE tasks by their "size" and complexity. Define what is a "big" task and what is a "hard" task.
- RQ3: What are possible GenAI configurations?
  - When we say "Generative AI," are we saying just GenAI solutions or possibly optimized prompt engineering, fine-tuning, combinations of GenAI with other forms of AI/ML, or other tools (simulation, verification, optimization, etc.)?
- RQ4: What experiments could demonstrate improvements empirically?
  - Formulate structured experiments and use cases

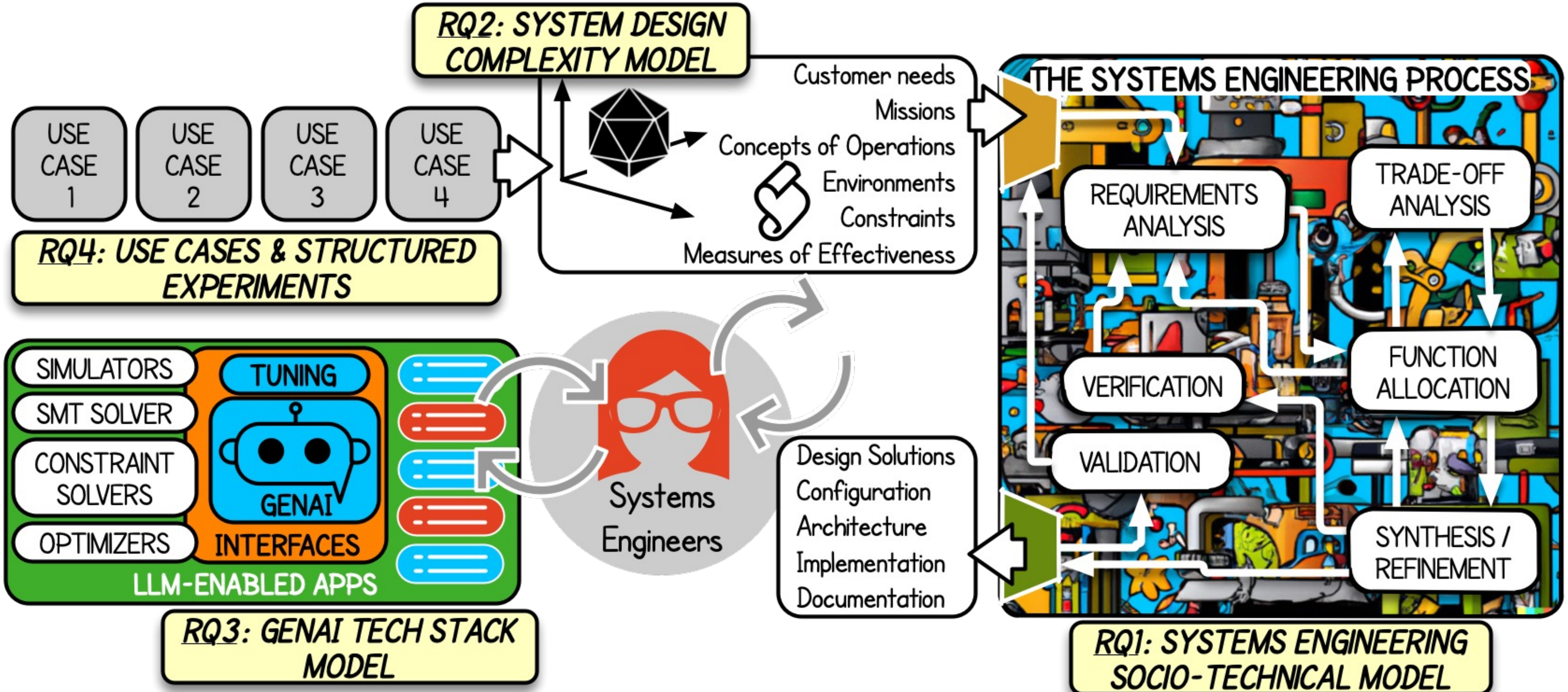
Structured quantitative approach to exploring these RQs is ongoing

Image generated with assistance of AI

© 2024 LOCKHEED MARTIN CORPORATION

LOCKHEED MARTIN

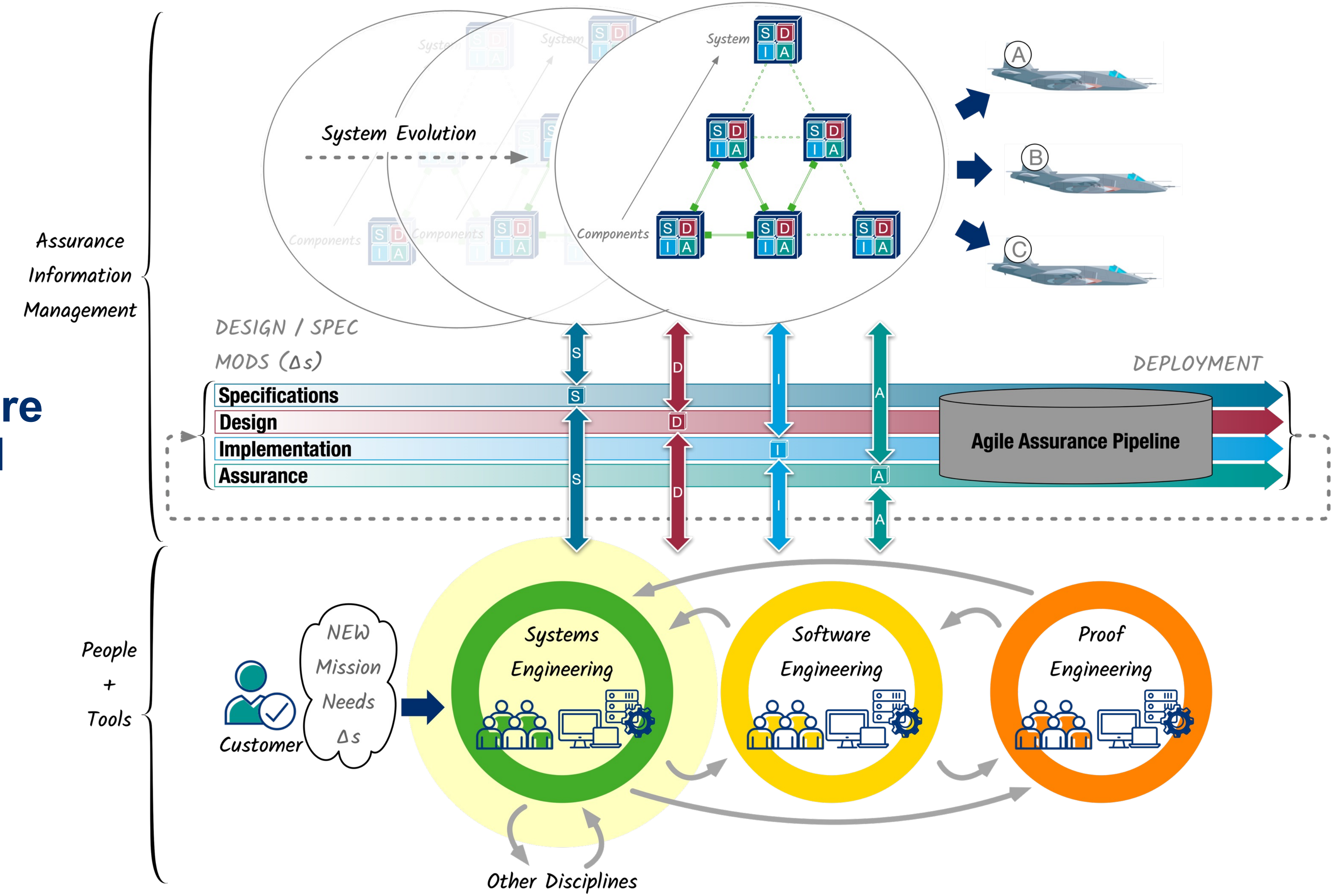
Overview



© 2024 LOCKHEED MARTIN CORPORATION
LOCKHEED MARTIN

SE Sociotechnical Model

Exemplar Future Sociotechnical System



© 2024 LOCKHEED MARTIN CORPORATION
LOCKHEED MARTIN

INCOSE's Competency Framework

Competency := observable, measurable set of skills, knowledge, abilities, behaviors, and other characteristics an individual needs to successfully perform work roles or occupational functions.

Integrating	Core	Professional	Management	Technical
Project Management	Systems Thinking	Communications	Planning	Requirements Definition
Finance	Lifecycles	Ethics and Professionalism	Monitoring and Control	System Architecting
Logistics	Capability Engineering	Technical Leadership	Decision Management	Design for...
Quality	General Engineering	Negotiation	Concurrent Engineering	Integration
	Critical Thinking	Team Dynamics	Business and Enterprise Integration	Interfaces
	Systems Modeling and Analysis	Facilitation	Acquisition and Supply	Verification
		Emotional Intelligence	Information Management	Validation
		Coaching and Mentoring	Configuration management	Transition
			Risk and Opportunity Management	Operation and Support

Which *competencies* could benefit the most from GenAI and how?

© 2024 LOCKHEED MARTIN CORPORATION

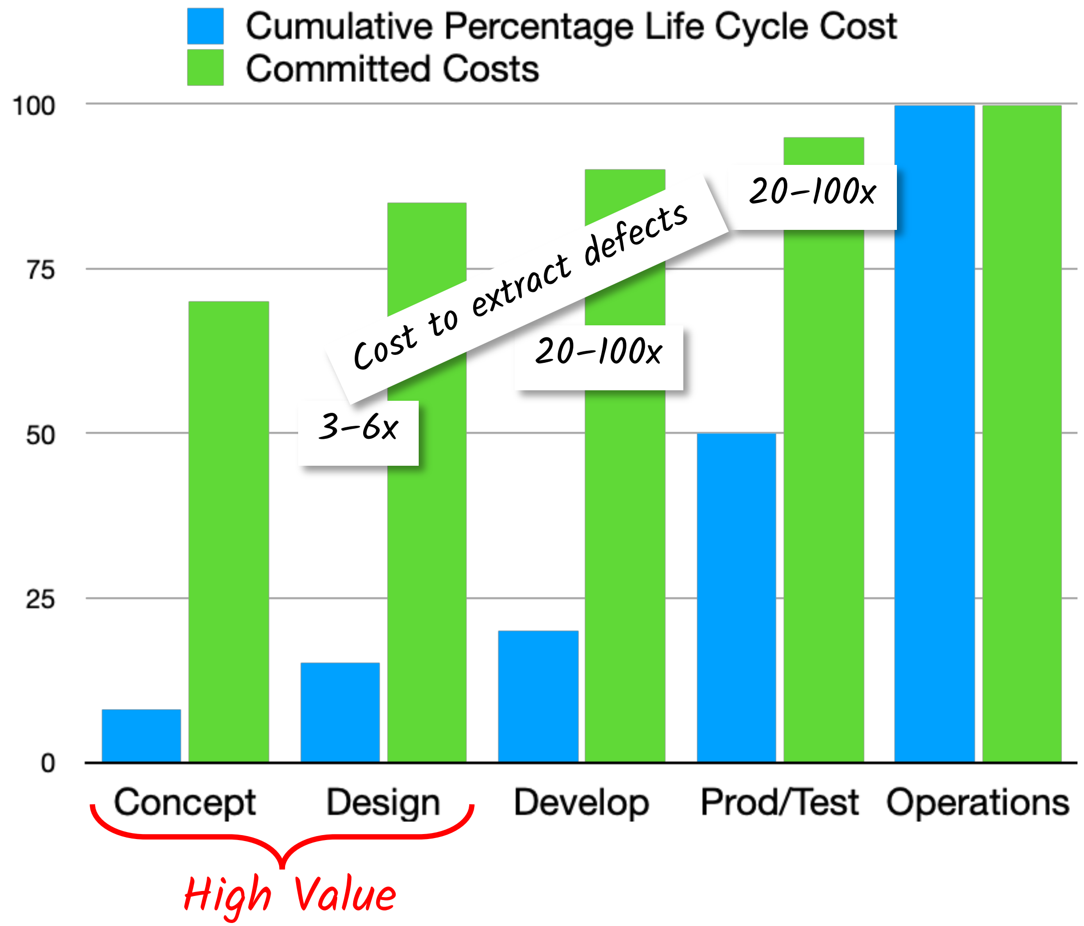
SE Processes & Value

Process := Set of interrelated or interacting activities that transforms inputs into outputs

Technical Processes
Business or Mission Analysis
Stakeholder Needs & Reqts Definition
System Reqts Definition
Architecture Definition
Design Definition
System Analysis
Implementation
Integration
Verification
Transition
Validation
Operation
Maintenance
Disposal

Focus

Technical Processes are most aligned with Technical and Core SE Competencies



Data from INCOSE SE Handbook 4th Ed, 2015

As we work on improving truthfulness and validity of GenAI outputs, early lifecycle stages seem promising

© 2024 LOCKHEED MARTIN CORPORATION
LOCKHEED MARTIN

Exemplar Competencies and Use Cases for Applying GenAI

Competency	INCOSE Definition	Exemplary Use Cases
Requirements Definition	To analyze the stakeholder needs and expectations to establish the requirements for a system	• Extract candidate requirements from a Concept of Operations • Represent stakeholders • Extract tentative formal representations from natural language requirements
System Architecting	The definition of the system structure, interfaces and associated derived requirements to produce a solution that can be implemented	• Formulate many alternative structures and allocate requirements to components • Formalize informal architecture definitions
Design for...	Ensuring that the requirements of all lifecycle stages are addressed at the correct point in the system design	• Identify *ility tradeoffs • Formulate design and optimization problems
Interfaces	The identification, definition and control of interactions across system or system element boundaries	• Formulate potential contracts (vertical, horizontal) • Identify modularity, scalability, and interoperability issues

We have identified several potentially high-value use cases

© 2024 LOCKHEED MARTIN CORPORATION

Complexity Model

Fundamental Questions

- Are problems inherently “hard” or complex (objective complexity) or is complexity a human-ascribed quality (subjective complexity)?
  - Moravec’s Paradox: “Things that are very easy for humans turn out to be very hard for machines and vice versa.”
  - Are there certain types of problems that are easier for machines? What are the main reasoning modalities?
    - Deductive, Inductive, Abductive
    - Causal, spatial, problem solving, decision making, planning, etc.
  - How do we know when something is hard for humans?
    - Effort, Mistakes?
- There are definitions for complexity when the system is realized but what about design problems when there’s no system?
  - A holistic view implies considering the system and its context, and since context is dynamic, complexity evolves!

Image generated with assistance of AI

© 2024 LOCKHEED MARTIN CORPORATION

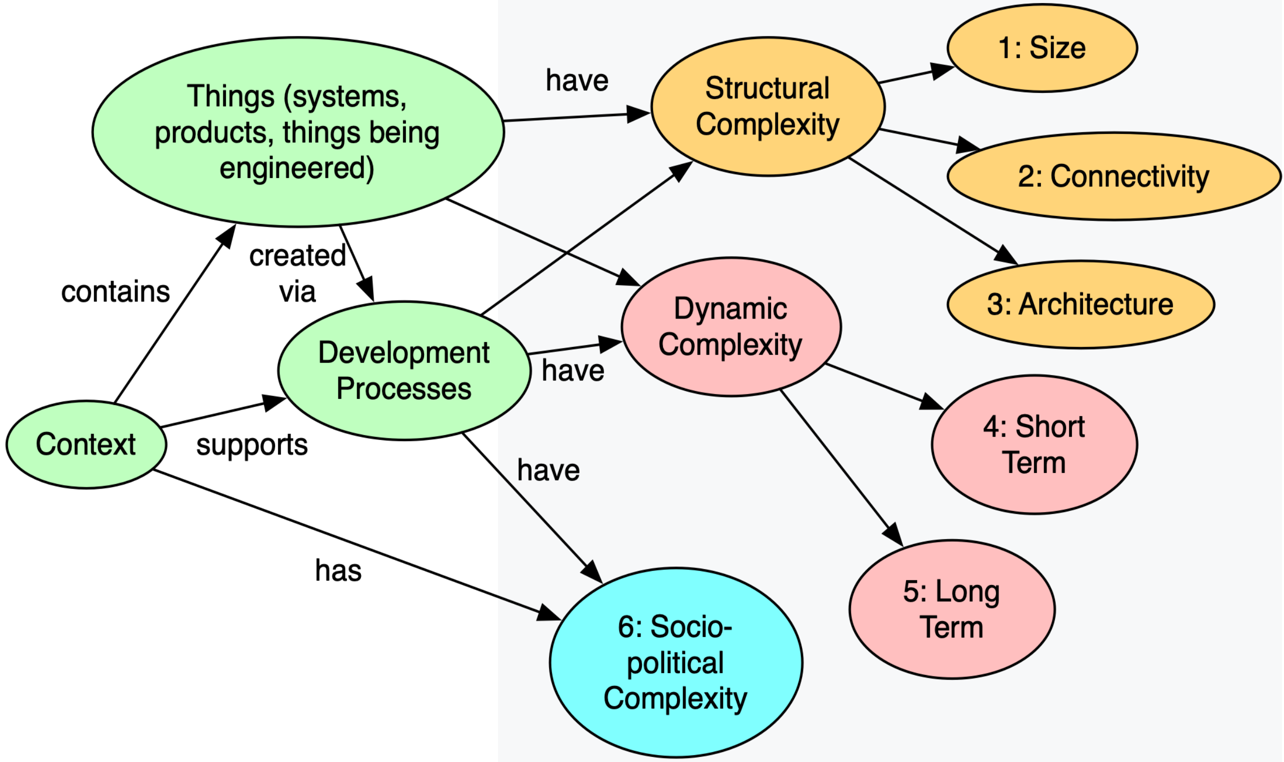
Standard Definitions and Types of Complexity

INCOSE:

- Complexity ≠ complicated
  - Complicated: Interactions are governed by fixed interrelationships. It may be understood by recursive decomposition. Compositional.
  - Complexity: Interactions give rise novel emergent patterns. System properties disappear when analyzing components in isolation
- A **complex system** is a system in which there are non-trivial relationships between cause and effect: each effect may be due to multiple causes; each cause may contribute to multiple effects; causes and effects may be related as feedback loops, both positive and negative; and cause-effect chains are cyclic and highly entangled rather than linear and separable.
- **Complexity** is the state or quality of being complex, which is characterized by the number and diversity of parts or elements, the intricate nature of their relationships, and the high level of interdependence among them

SEBOK (Systems Engineering Book of Knowledge)

- **Complexity** is a measure of how difficult it is to understand how a system will behave or to predict the consequences of changing it.



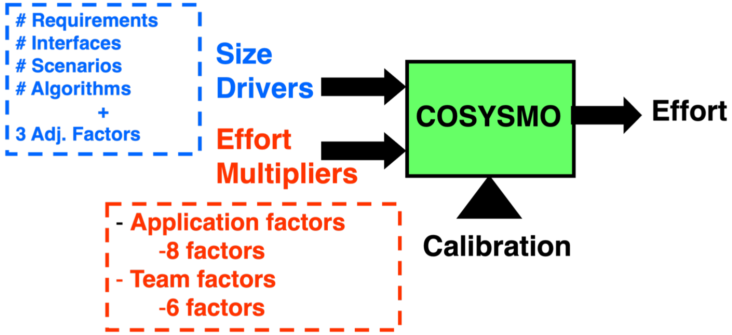
Complexity Typology by Sheard &Mostashari (2010)

Good for explaining complexity but not for measuring complexity

© 2024 LOCKHEED MARTIN CORPORATION

Surrogates of Complexity

- Cost/cost



Constructive Systems Engineering Cost Model (COSYSMO) by Valerdi & Boehm (2010)

- Human Errors
  - Woods, David, et al. *Behind Human Error*. CRC Press, 2017. Examples of human mistakes:
  - Confuse the model for the real thing
  - Confirmation bias
  - Underestimate complexity: Failure to anticipate interactions / emergent properties
  - Miscommunication: poor documentation of assumptions, poor requirements

Can GenAI help Systems Engineers make less mistakes?

(…despite being far from infallible)

"American medical cost of remediation of hip metallosis complications will be on the order of $50 billion dollars."

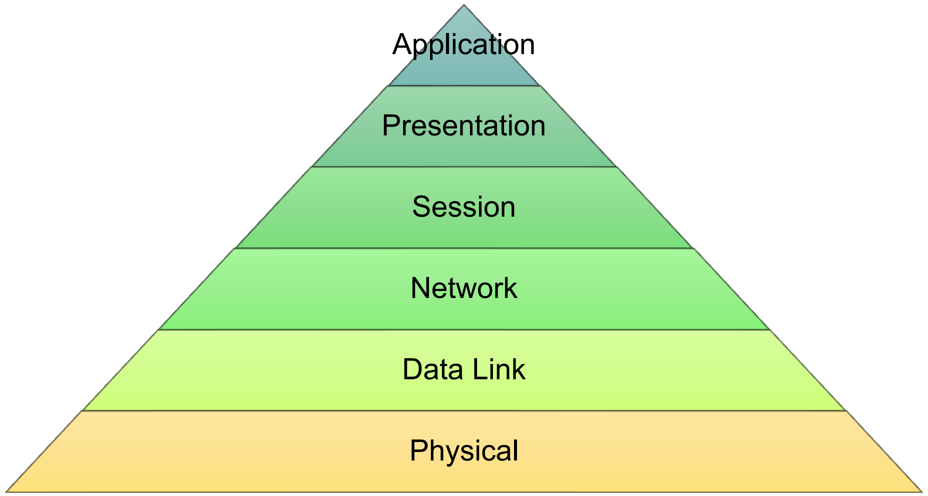
Tower, Stephen. "Hip metallosis and corrosion—a million harmed due to FDA inaction." Journal of Patient Safety 15.3 (2019): 257-259.

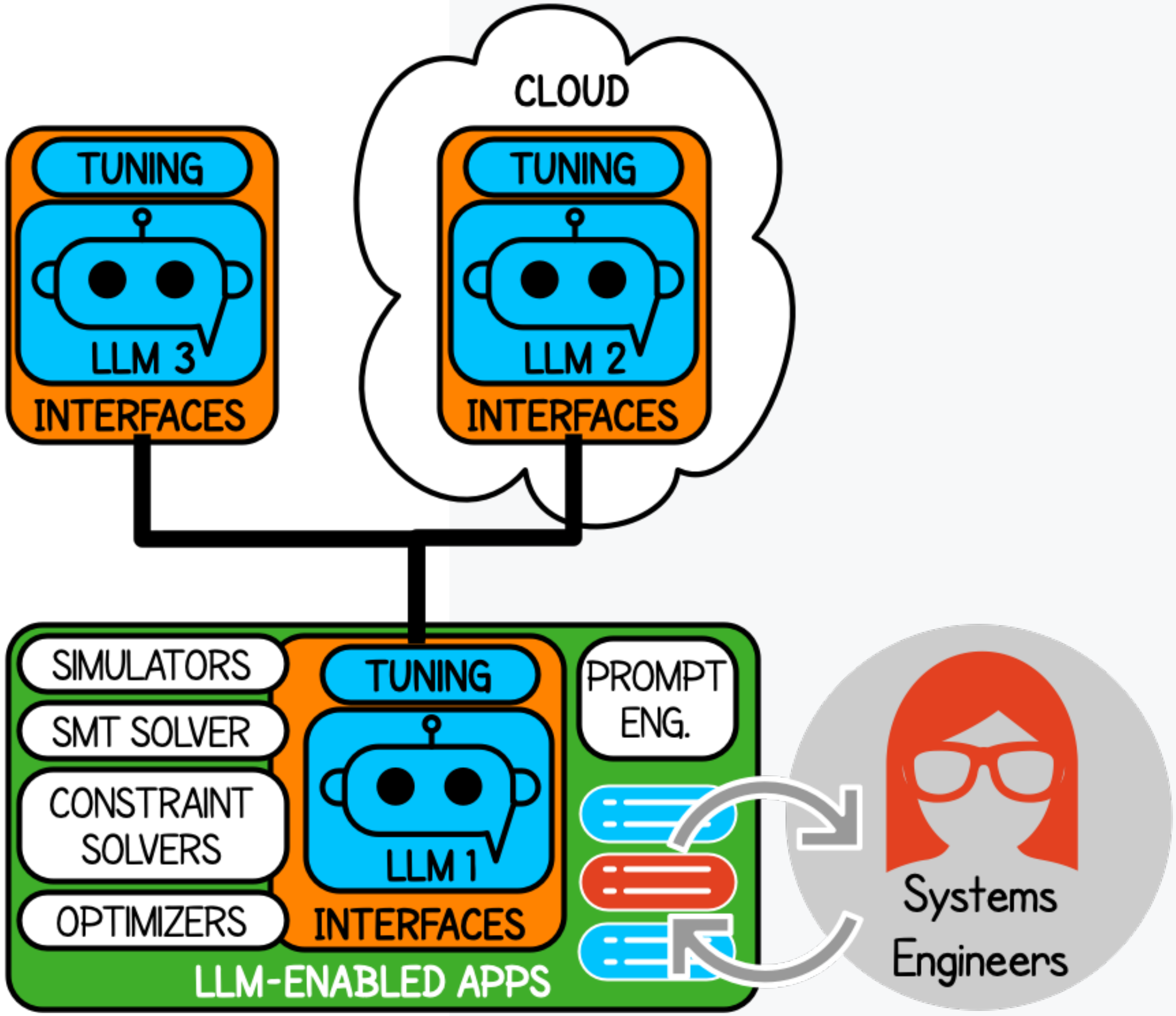
© 2024 LOCKHEED MARTIN CORPORATION
LOCKHEED MARTIN

GenAI Tech Stack Model

Questions

- What does a “Technology Stack Model” of GenAI solutions look like?
- How do we establish a vocabulary of configurations?
- Example: Open Systems Interconnect (OSI) Reference Model

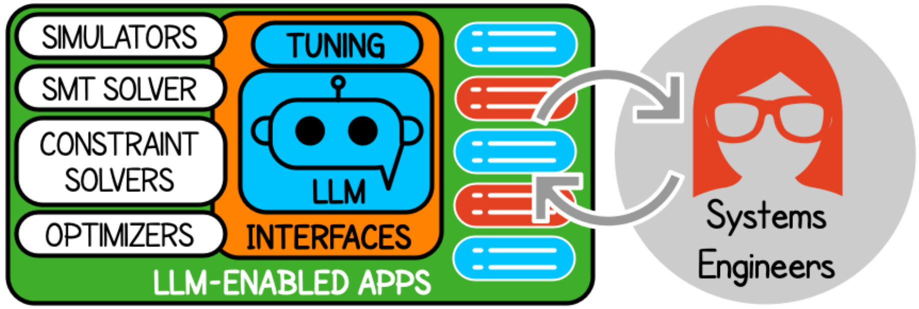




GenAI solutions are not monolithic

© 2024 LOCKHEED MARTIN CORPORATION

Integration with External Tools



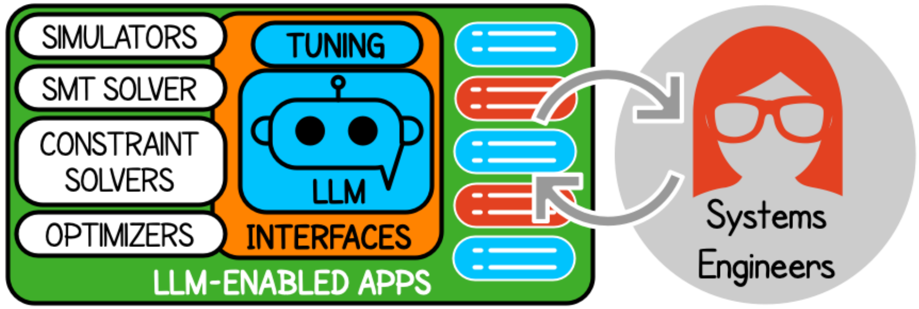
Configuration Elements	Examples

LLM Integration with External Tools	• External tools provide grounding and verification functionality that complement *LLMs' abilities*. E.g.:
	  • Formal verification and falsification tools capable of processing the LLM's intermediate results.
	  • LLM's outputs as scenario specifications executable in an appropriate simulation environment, the LLM can guide physics-grounded "what if" analyses.
	• Exemplar LLM interfaces:
	  • Langchain
	  • LlamaIndex
	  • Function Calling

LLMs talk to external tools

© 2024 LOCKHEED MARTIN CORPORATION
LOCKHEED MARTIN

LLM Utilization & Prompt Engineering



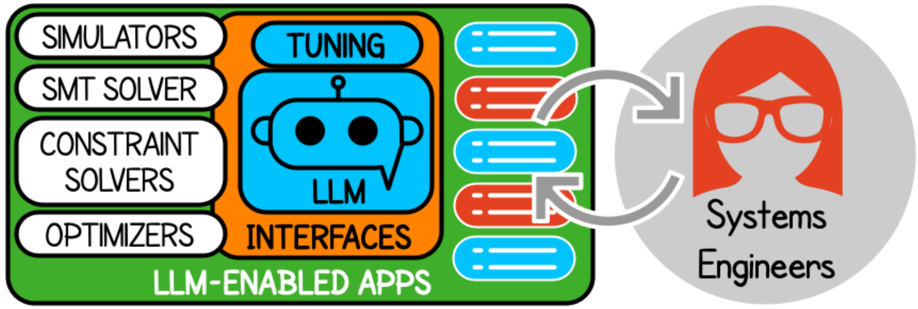
Configuration Elements	Examples

LLM Utilization & Prompt Engineering	• Exemplar Utilization Modalities
	  • In-context Learning
	  • Chain-of-Thought
	• LLM interfacing:
	  • Langchain
	  • LlamaIndex
	  • Data sources for embeddings, e.g.: vector databases
	  • External Knowledge Retrieval (RAG)

Maximize utilization of base LLMs without "touching" them

© 2024 LOCKHEED MARTIN CORPORATION
LOCKHEED MARTIN

LLM Adaptation and Tuning



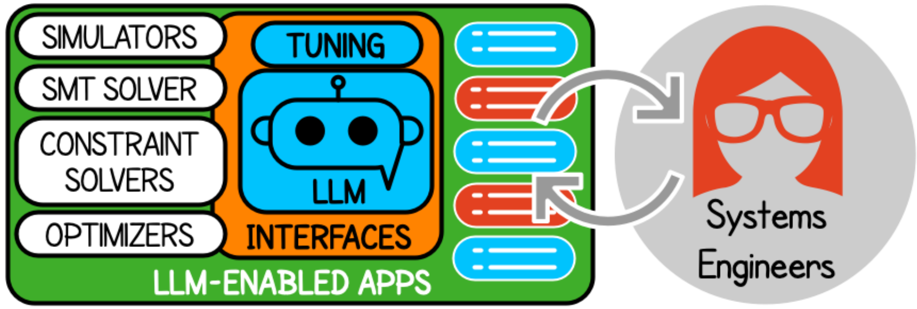
Configuration Elements	Examples

LLM Adaptation and Tuning	• Exemplar Adaptation and Tuning Modalities
	  • Instruction Tuning
	  • Addition of NN modules to LLM's base transformers
	  • Low-rank Adaptation (LORA)
	  • ~~Full fine-tuning~~

Augment base LLM

© 2024 LOCKHEED MARTIN CORPORATION
LOCKHEED MARTIN

LLM Base Model selection

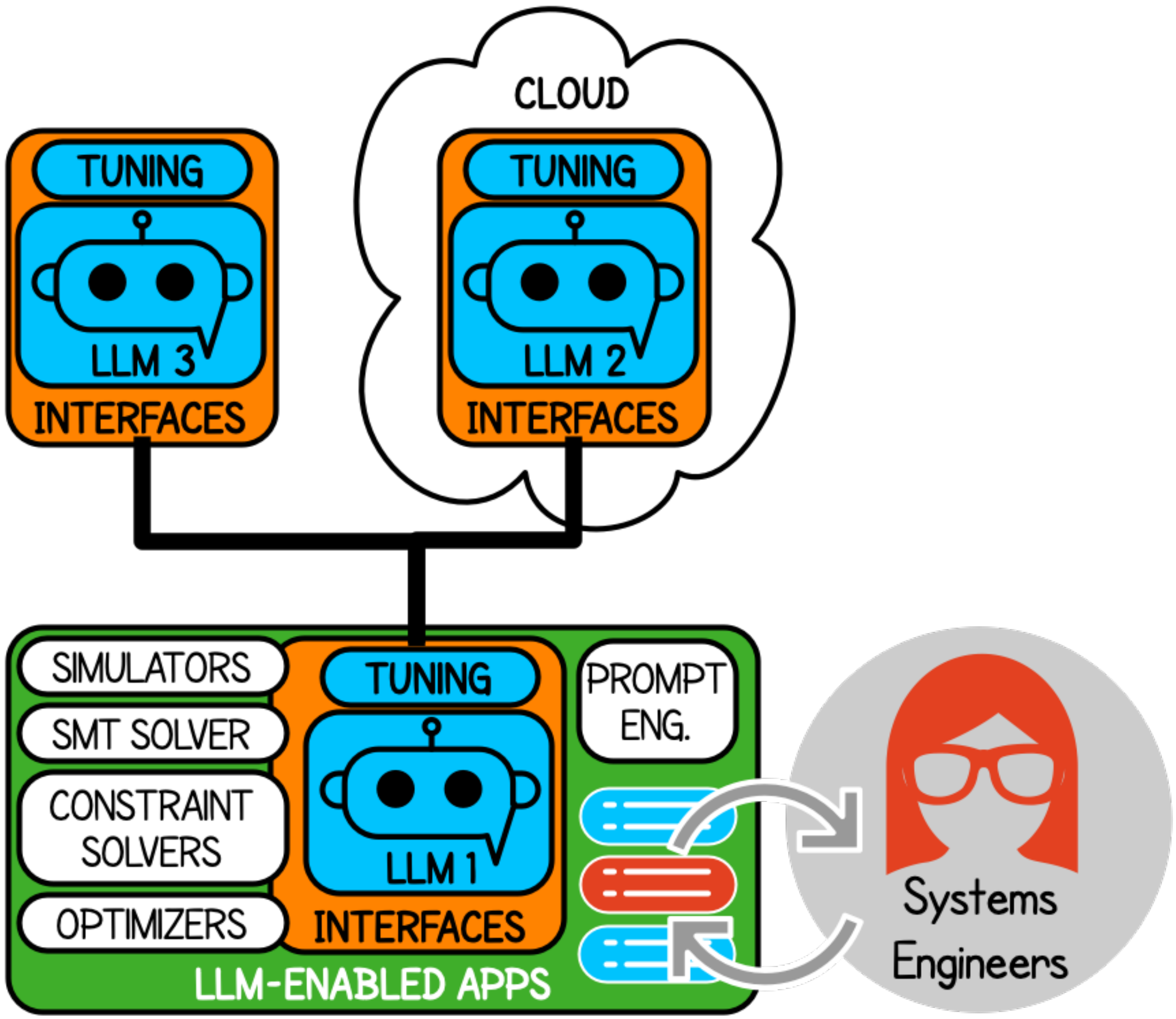


Configuration Elements	Examples
LLM Base Model selection	• Closed source: ChatGPT, Claude, Gemini, etc. • Open-source: LlaMA, Mixtral, Mistral, Falcon, etc. • ~~Training~~

Vary base LLM

© 2024 LOCKHEED MARTIN CORPORATION
LOCKHEED MARTIN

Advanced Configurations



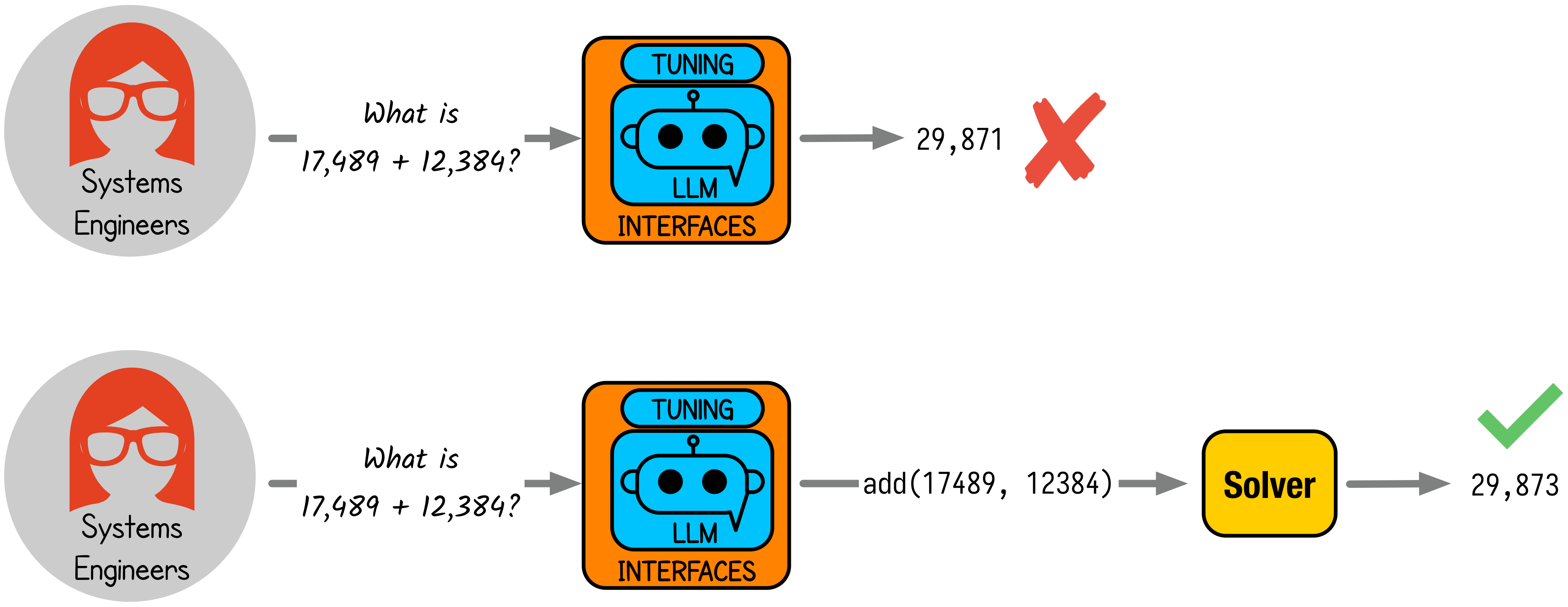
"The Shift from Models to Compound AI Systems," Zaharia et al. Feb. 2024

Also: multi-modal networks

© 2024 LOCKHEED MARTIN CORPORATION

Use Cases and Experiments

Basic Patterns

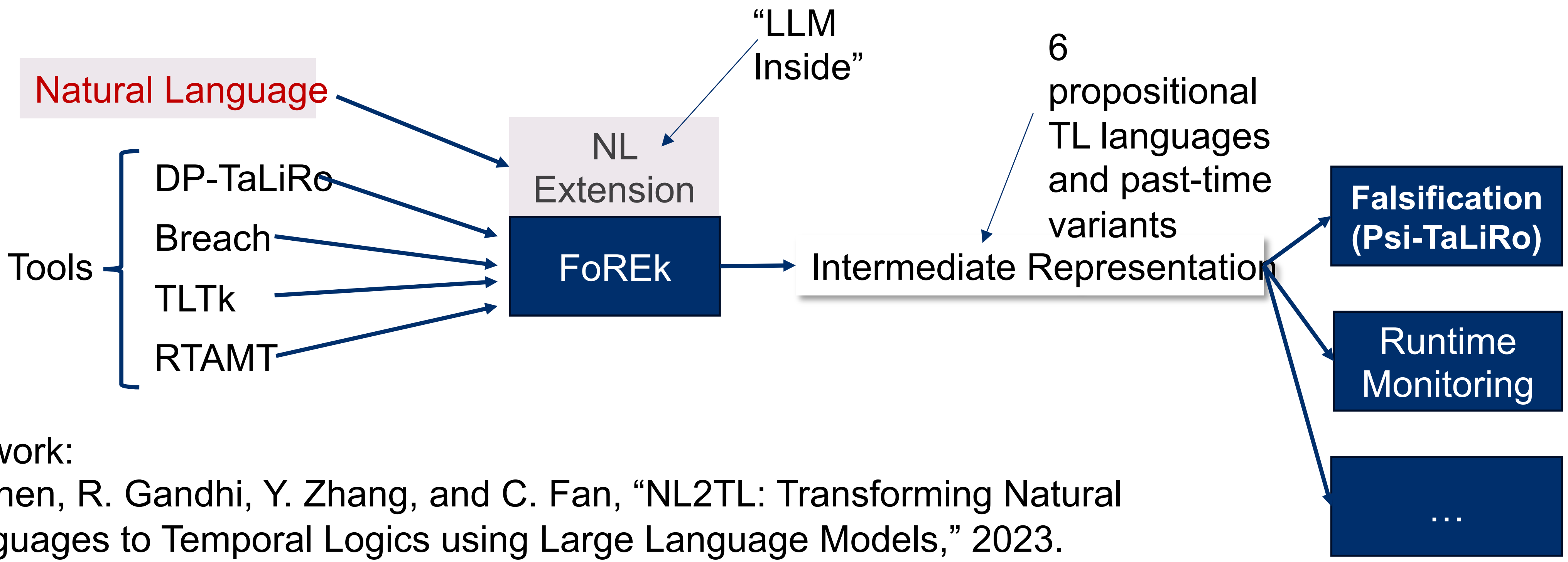


Another alternative: LLM generates “good guesses” that are sent to a checker/verifier

Avoid monolithic solutions

© 2024 LOCKHEED MARTIN CORPORATION

AUTO-FORMALIZATION OF REQUIREMENTS EXPRESSED IN TEMPORAL LOGIC



Close work:

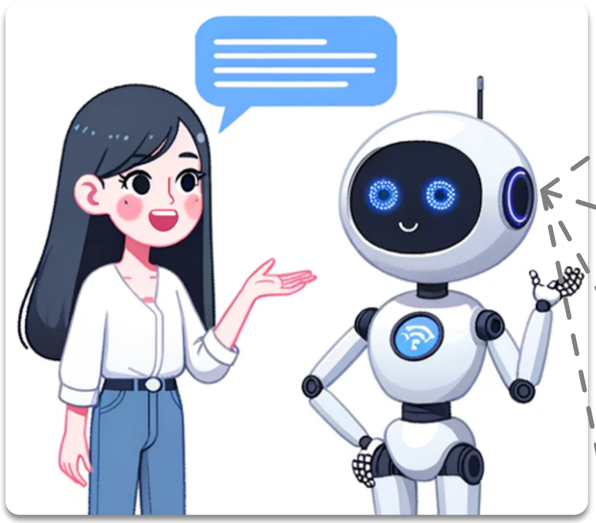
- Y. Chen, R. Gandhi, Y. Zhang, and C. Fan, "NL2TL: Transforming Natural Languages to Temporal Logics using Large Language Models," 2023.
- Approach: GPT-3 + dataset creation + fine tuning of T5 models (text-to-text transformer)

NEW:

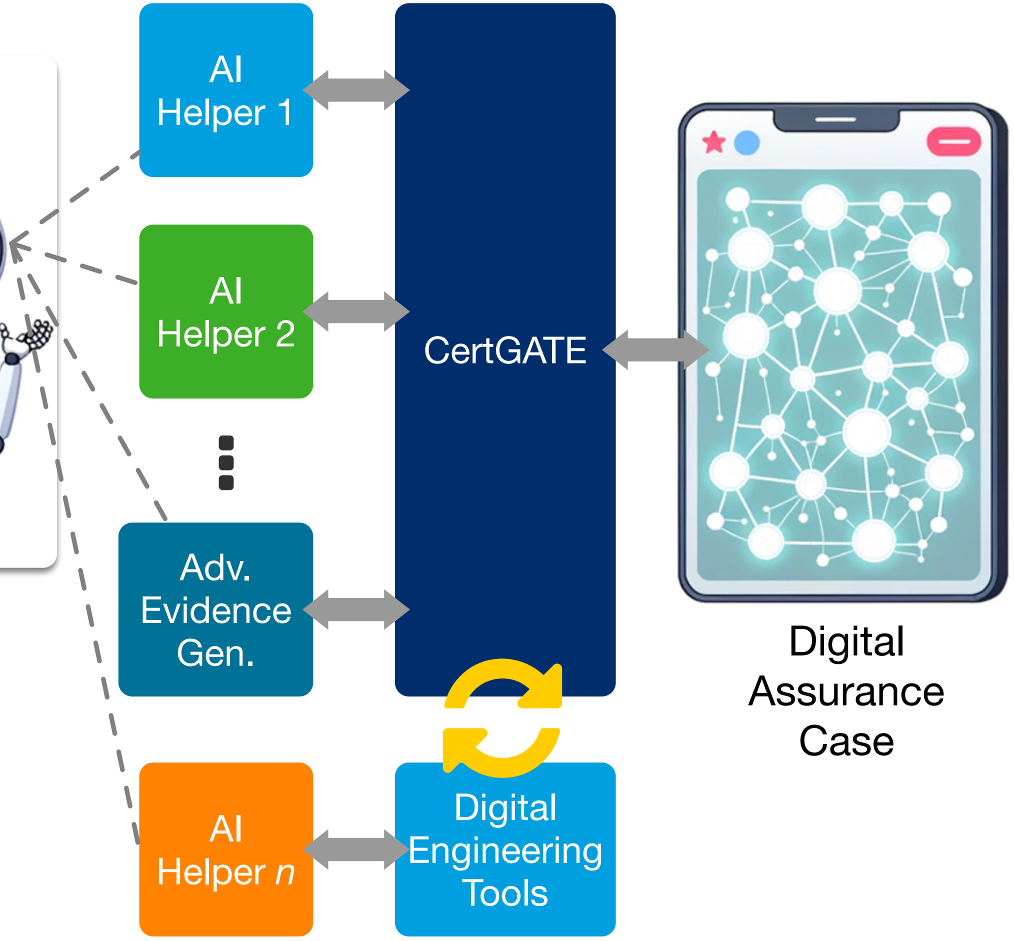
- ASU team has ARCH benchmark datasets
- Executable Intermediate Representation that could be used for automated error detection/correction

© 2024 LOCKHEED MARTIN CORPORATION

LLM-BASED ASSURANCE CASE INTERROGATION



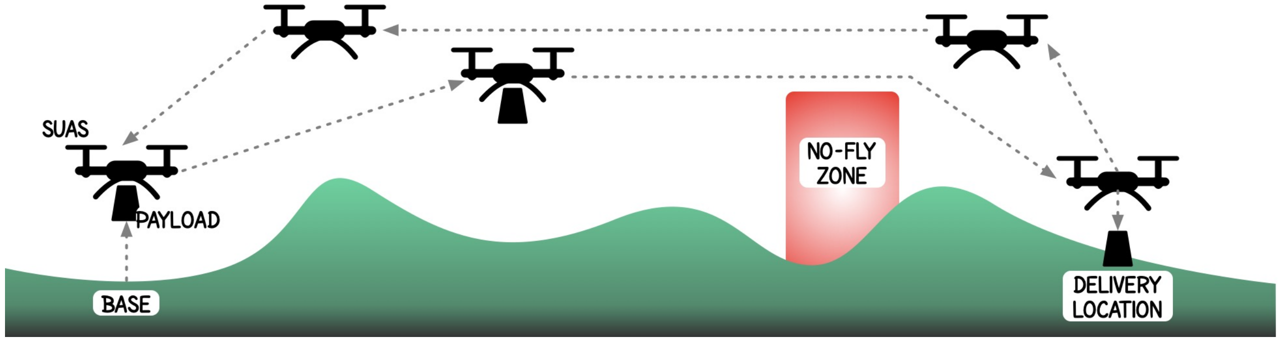
AI-enabled Interface



CertGATE envisioned to provide a solid high-assurance foundation that can be enhanced in usability by advances in AI.

© 2024 LOCKHEED MARTIN CORPORATION

Ongoing Experiments



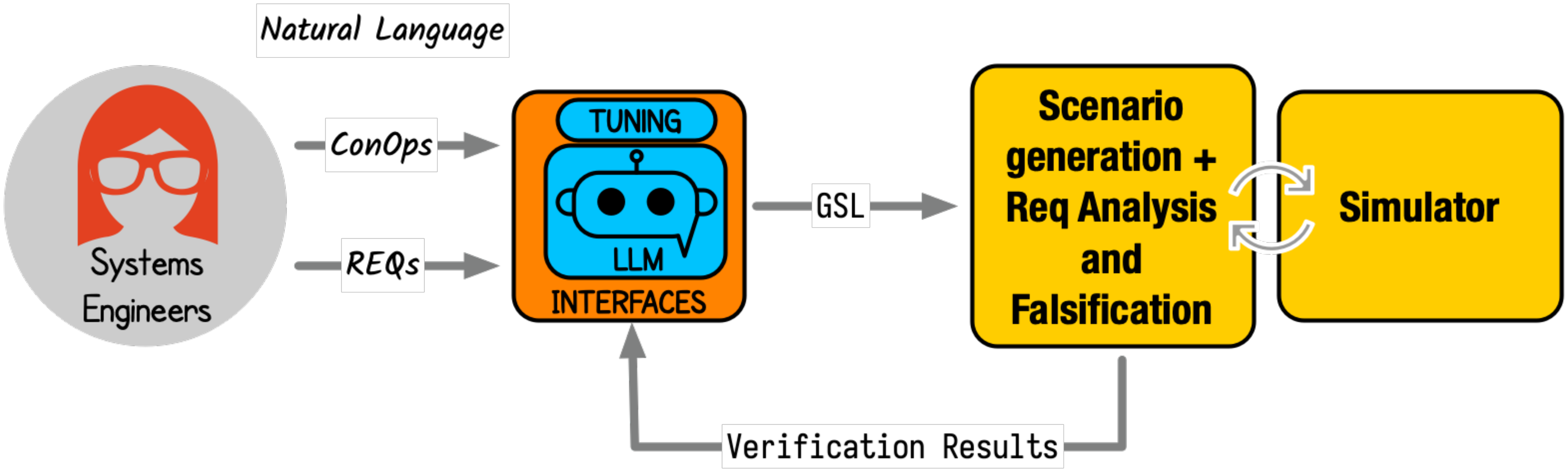
ConOps: A small Unmanned Aerial System (sUAS) is tasked with a mission that involves flying over enemy territory. The sUAS starts from a Base to a specified Delivery Location carrying a Payload and avoiding designated No-fly Zones. Upon reaching the delivery location, the sUAS hovers about two feet above the ground for 2 seconds and releases its Payload. Immediately after the payload is delivered, the sUAS returns to its Base. Both the Base and Delivery Locations are given in Latitude-Longitude coordinates. The distance between these two locations can be as much as 13 miles, and the Payload can weigh up to 35 pounds. The sUAS is expected to deliver the payload with a location accuracy of 5m.

Experiment 1: We want to derive high-level requirements (HLRs) and a basic functional decomposition from the ConOps description above. In this context, we will integrate the LLM with an aircraft performance simulation engine. The simulation results will validate the LLM's outputs.

Experiment 2: After establishing a basic set of requirements and functions, we want to use the LLM to perform hazard analysis To provide a path for Chain-of-Thought reasoning, we will employ a methodical approach such as the Systems-Theoretic Process Analysis (STPA)

© 2024 LOCKHEED MARTIN CORPORATION

From ConOps and Reqs in NL to Verification Results



Check Feasibility

© 2024 LOCKHEED MARTIN CORPORATION

Additional Thoughts

Ongoing Focus Areas

The impact of tools and representations understood by people and machines—Model-Based Systems Engineering (MBSE) (multimodality!)

Consider the transformation of the traditional Systems Engineering lifecycle model, to the agile lifecycle envisioned by current acquisition policy

User Experience (UX) / Human Integration concerns

- AI as assistant vs AI embedded into tools
- AI as an assistant to individuals or as an assistant to the team

Development of a "working model" of GenAI that can be used to develop AI-based tools

- Mathematical models: "untrusted oracle," "probabilistic translator," "sequential predictor," "probabilistic knowledge base," "discrete stochastic dynamical system," etc.
- Human metaphors: unreliable but creative assistant, Digital Librarian, etc.

GenAI for SE is an active area of exploration at LM

Image generated with assistance of AI

© 2024 LOCKHEED MARTIN CORPORATION

Conclusions

Positive

- GenAI seems promising in the early conceptualization and design stages. It could play multiple roles: facilitator, unreliable "low-cost" domain expert, creativity instigator, devil's advocate/contrarian, etc.
- GenAI can serve as a translator when combined with other tools and fine-tuning. It could be applied to high-criticality situations as long as accuracy can be guaranteed or if mechanisms for corroboration/validation are introduced. Hopefully, at cost!

Watch for

- The user interface to LLMs (e.g., ChatGPT) is "broken." It does not modulate assertiveness according to its validity like humans do.
- Concerns with integrating capabilities stemming from the commercial and consumer sectors (IP, cybersecurity)

Question:

- How do you keep up with developments?

Image generated with assistance of AI

© 2024 LOCKHEED MARTIN CORPORATION

LOCKHEED MARTIN

LOCKHEED MARTIN