Automating SE with AI and Natural Language Processing (NLP) Software

Brian P. Parrish, Xuan Chau

MITRE

Jul 26, 2025 - Jul 31, 2025

Approved for Public Release, Distribution Unlimited. Information derived from Public Release Case Number **25-2021**.
This technical data deliverable was developed using contract funds under Basic Contract No. W56KGU-18-D-0004. © **2025 The MITRE Corporation.**

Bios

Brian P. Parrish

Title

Principal Modeling and Simulation and Multi-Discipline Systems Engineer, The MITRE Corporation - 20 Years

Academics

- MS in Systems Engineering, Cyber Security
- BS Computer Science, Math
- BS Computer Information Systems, Business

Roles

- Synthetic Training Environment, Project Leader (PL)
- Development Command Soldier Center, PL
- Joint Modernization Command Multi-Domain Operations Simulation Center M&S Task Leader

Xuan Chau

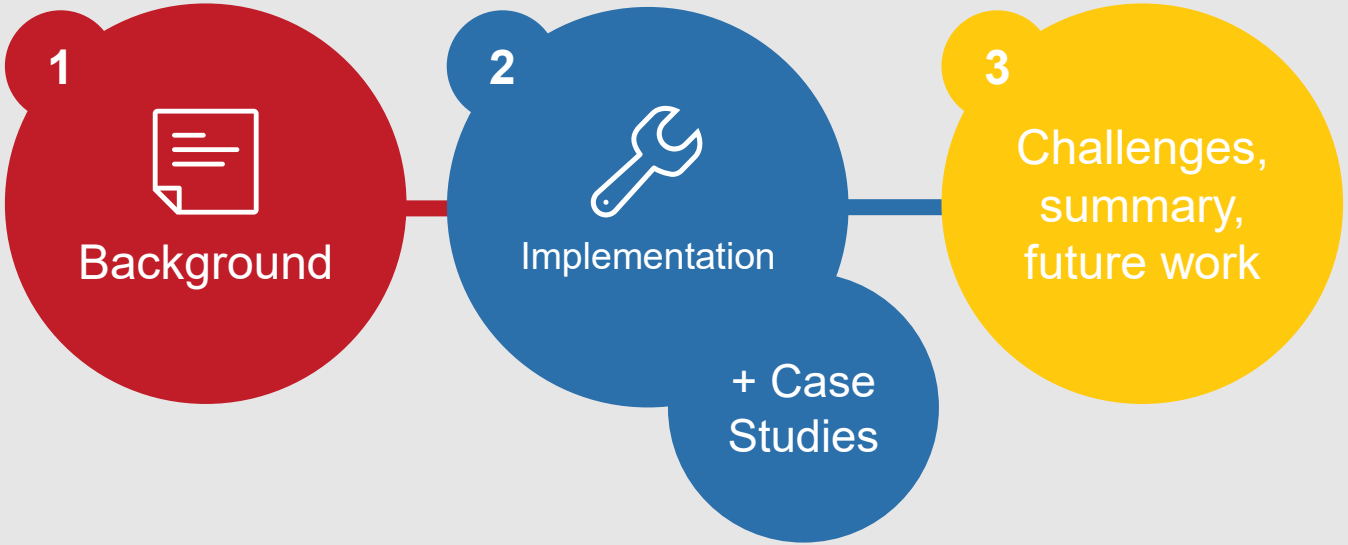
Title: Senior Software Engineer, The MITRE Corporation – 5 Years

Academics: BS in Computer Science

Roles: RAAM Lead Developer

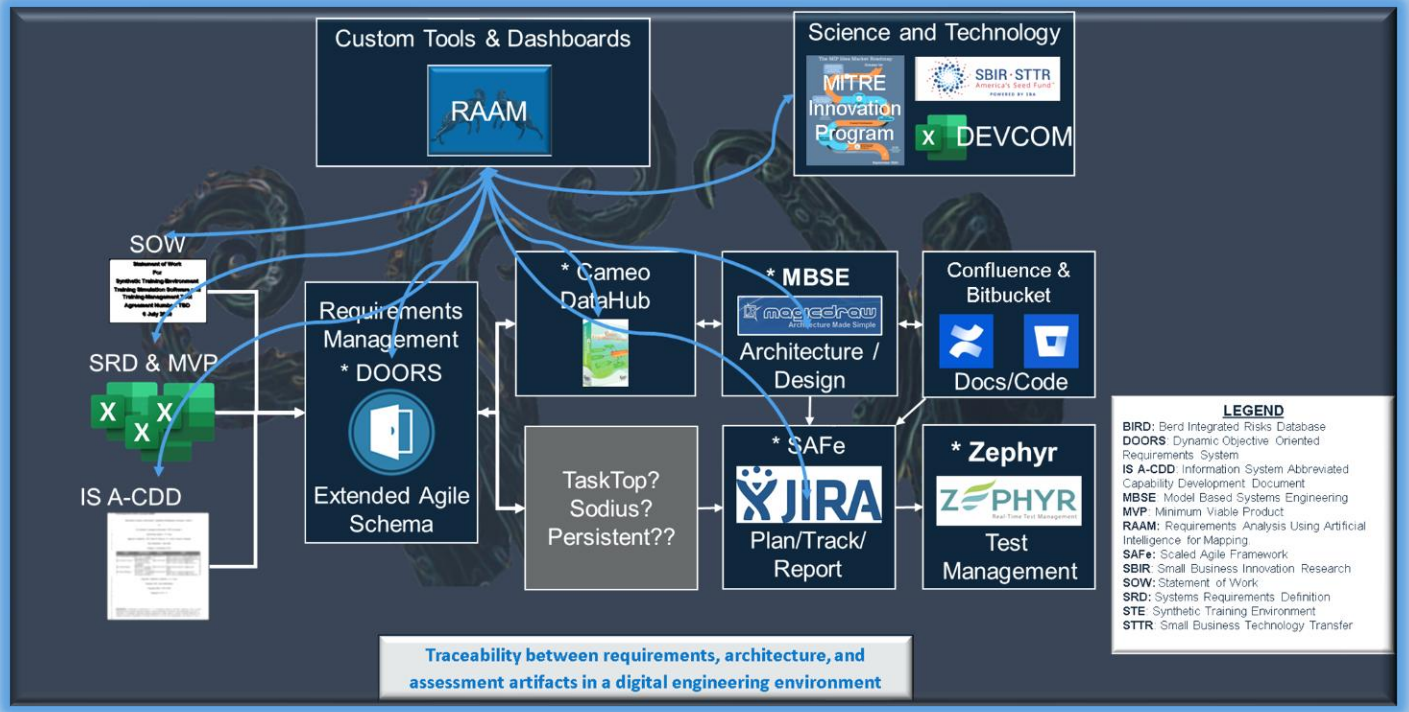
Approved for Public Release, Distribution Unlimited. Information derived from Public Release Case Number **25-2021**. This technical data deliverable was developed using contract funds under Basic Contract No. W56KGU-18-D-0004. © **2025 The MITRE Corporation.**

Today's Agenda



Approved for Public Release, Distribution Unlimited. Information derived from Public Release Case Number **25-2021**. This technical data deliverable was developed using contract funds under Basic Contract No. W56KGU-18-D-0004. © **2025 The MITRE Corporation.**

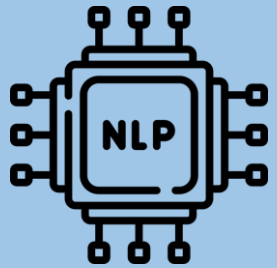
Digital Engineering Overview



Operational Requirements → Functional Requirements → Designs / Architecture → Existing S&T → Agile Development → T&E

Approved for Public Release, Distribution Unlimited. Information derived from Public Release Case Number **25-2021**. This technical data deliverable was developed using contract funds under Basic Contract No. W56KGU-18-D-0004. © **2025 The MITRE Corporation.**

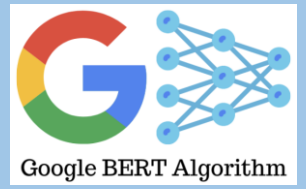
NLP, LLMs, & Generative AI



Natural Language Processing encompasses a machine's ability to interpret, manipulate, and comprehend human language.

Development of **large language models (LLMs)** enable for extremely powerful NLP capabilities.

- Question-answering
- Summarization
- Entity recognition
- …etc.



MISTRAL AI_

Each LLM is specialized for a different task. E.g.:

- *GPT - text completion, translation, conversation*
- *BERT - question answering, sentiment analysis*

Approved for Public Release, Distribution Unlimited. Information derived from Public Release Case Number **25-2021**. This technical data deliverable was developed using contract funds under Basic Contract No. W56KGU-18-D-0004. © **2025 The MITRE Corporation.**

Applying NLP to Systems Engineering…

Goal: Create software to accelerate / automate tedious SE tasks by leveraging modern NLP technology.

General challenges:

- LLMs must comprehend specialized text with domain-specific terminology & acronyms
- Not a miracle technology; human experts **must** validate machine output
- LLMs must be approved for sensitive information
- How do we ensure usability and adoption into existing DE workflows?

Approved for Public Release, Distribution Unlimited. Information derived from Public Release Case Number **25-2021**. This technical data deliverable was developed using contract funds under Basic Contract No. W56KGU-18-D-0004. © **2025 The MITRE Corporation.**

Addressing domain-specific text

Frequency of domain-specific terminology & acronyms in engineering artifacts significantly degrades performance of general purpose LLMs (Ling, et al., 2023).

Fine tuning

- Modifies base model directly with special knowledge
- High compute requirements, complexity, and cost

Retrieval Augmented Generation (RAG)

- Utilize external data to augment GenAI's context (Barron, et al., 2024)
- Lower compute / complexity / cost

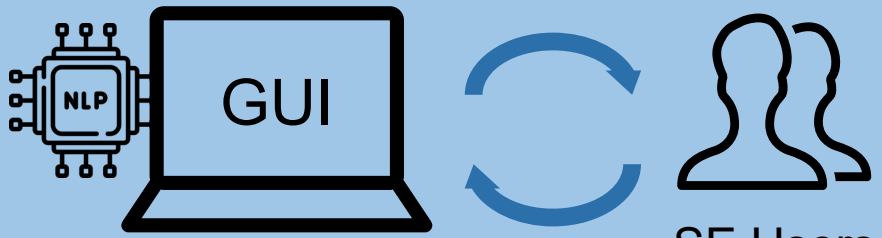
User-curated dictionaries

- RAAM implements "Domain models" – user curated dictionaries, to substitute terminology
- Lowest compute / complexity / cost

Approved for Public Release, Distribution Unlimited. Information derived from Public Release Case Number **25-2021**. This technical data deliverable was developed using contract funds under Basic Contract No. W56KGU-18-D-0004. © **2025 The MITRE Corporation.**

Case Study: Requirements Analyzer using AI for Mapping (RAAM)

RAAM



NLP powered web application

1. **Requirements Generation**
2. **Requirements Quality Analysis**
3. **Traceability Analysis**

Approved for Public Release, Distribution Unlimited. Information derived from Public Release Case Number **25-2021**. This technical data deliverable was developed using contract funds under Basic Contract No. W56KGU-18-D-0004. © **2025 The MITRE Corporation.**

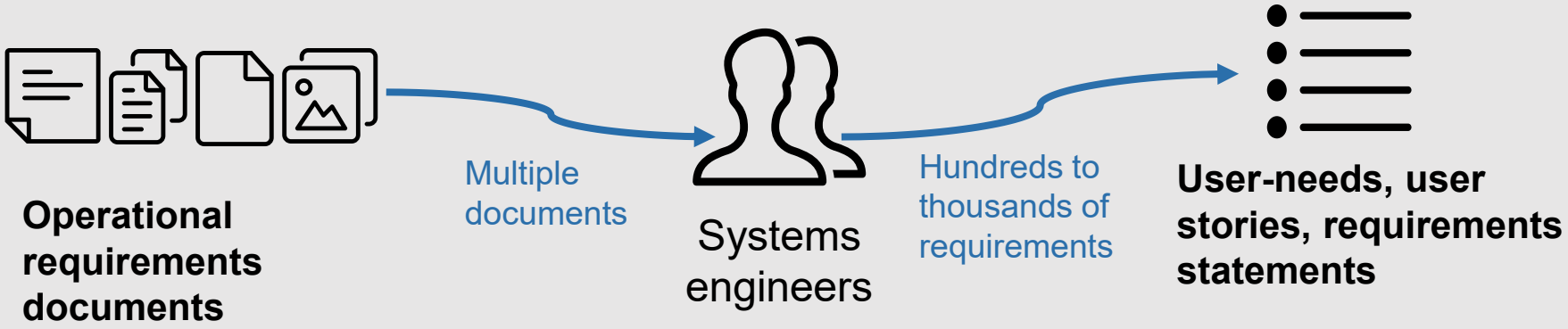
RAAM Capabilities

- Requirements Generation
- Requirements Quality Analysis
- Requirements Tracing

Approved for Public Release, Distribution Unlimited. Information derived from Public Release Case Number **25-2021**. This technical data deliverable was developed using contract funds under Basic Contract No. W56KGU-18-D-0004. **© 2025 The MITRE Corporation.**

Manual requirements generation

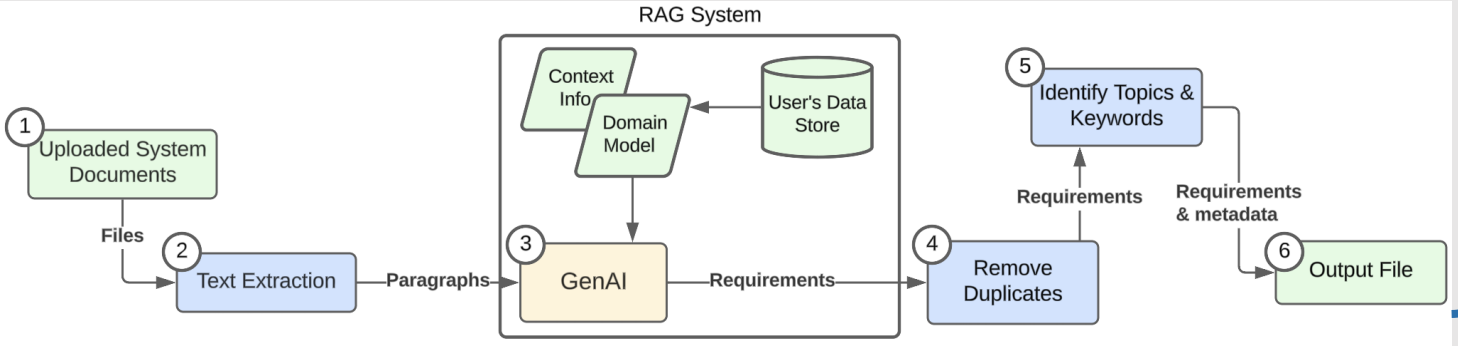
Traditional workflow: Derive traceable and well-written user-needs, user stories, and requirements statements from high-level operational requirements documents (e.g., capability development documents, doctrine, CONOPS, operational mode summaries / mission profiles [OMS/MPs], training manuals, etc.).



Approved for Public Release, Distribution Unlimited. Information derived from Public Release Case Number **25-2021**. This technical data deliverable was developed using contract funds under Basic Contract No. W56KGU-18-D-0004. © **2025 The MITRE Corporation.**

AI-assisted workflow for generating requirements

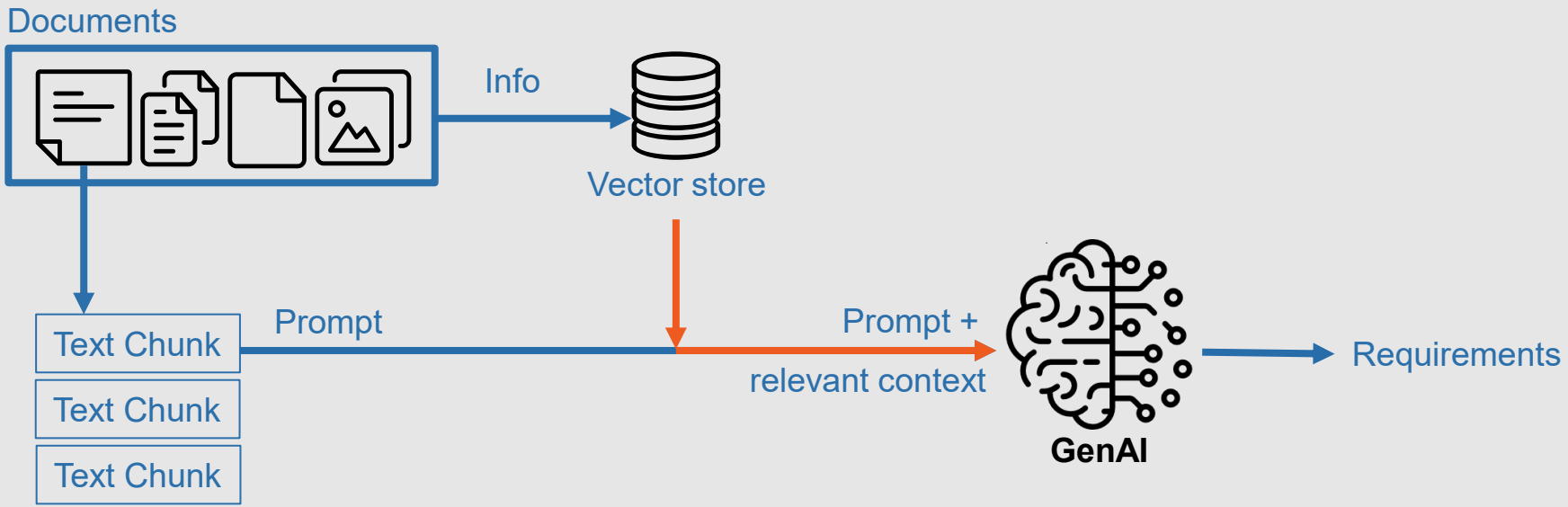
1. Engineer provides elicited system documents to software
2. Documents loaded as context documents into RAG database
3. Software decomposes documents into text chunks (by section or paragraph)
4. Generative AI (+RAG) is prompted to generate requirements from text chunks
5. Engineers receive and further iterate on generated requirements



Approved for Public Release, Distribution Unlimited. Information derived from Public Release Case Number **25-2021**. This technical data deliverable was developed using contract funds under Basic Contract No. W56KGU-18-D-0004. © **2025 The MITRE Corporation.**

Retrieval Augmented Generation

Generated requirements improved by GenAI getting context information through RAG database, which contains input documents.



Approved for Public Release, Distribution Unlimited. Information derived from Public Release Case Number **25-2021**. This technical data deliverable was developed using contract funds under Basic Contract No. W56KGU-18-D-0004. © **2025 The MITRE Corporation.**

RAAM sample output spreadsheet

UUID	Hierarchy ID	Paragraph #	Heading	Original Paragraph	LLM-generated requirements	Subject	Verb(s)	Object(s)	Sentiment
		15		d. This concept proffers how Army forces maneuver to gain positions of advantage in competition and conflict. As part of a joint and multinational force, Army forces will maneuver at echelon to deter adversaries and regional competitors during competition, dominate and win in armed conflict, and return to competition on favorable terms. During competition, Army forces will support regional engagement, shape the security environment, prevent conflict, and provide options for responding to and resolving crisis. When necessary, Army forces transition rapidly to armed conflict, continuously converging lethal and nonlethal effects, and maneuver to destroy or defeat adversary forces, control land areas and resources, and protect populations.					Neutral
40	1.1.38	15			Army forces will maneuver to gain positions of advantage in competition and conflict	Army forces	maneuver	gain positions	
41	1.1.39	15			Army forces will maneuver at echelon as part of a joint and multinational force	Army forces	maneuver at as	echelon, part	
42	1.1.40	15			Army forces will deter adversaries and regional competitors during competition	Army forces	deter during	adversaries, regional competitors	
43	1.1.41	15			Army forces will dominate and win in armed conflict	Army forces	dominate, win in	armed conflict	

Metadata columns

Headings

Original paragraphs

Generated requirements

Approved for Public Release, Distribution Unlimited. Information derived from Public Release Case Number **25-2021**. This technical data deliverable was developed using contract funds under Basic Contract No. W56KGU-18-D-0004. © **2025 The MITRE Corporation.**

Requirements generation case study using RAAM

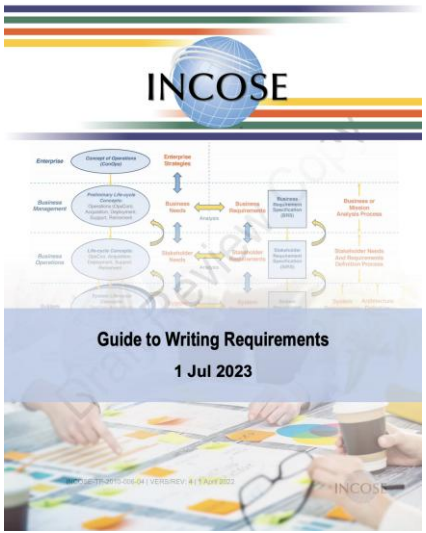
In 2023, MITRE & the Synthetic Training Environment Cross Functional Team (STE CFT) leveraged the RAAM's requirements generation capability to extract initial user stories / requirements from Army Futures Command Multi-Domain Operations (MDO) concept documents **(4 documents, ~400 pages).**

Manual extraction	**With RAAM NLP Automation:**
Extract and process contents of all 4 documents	Processing 4 documents (using tool, post-processing generated output)
20 hours (5 hours per document, estimated)	**~30 minutes**

Overall, it was estimated that the tool **saved 19.5 hours** of labor, accelerating the effort by **4,000%**.

Approved for Public Release, Distribution Unlimited. Information derived from Public Release Case Number **25-2021**. This technical data deliverable was developed using contract funds under Basic Contract No. W56KGU-18-D-0004. © **2025 The MITRE Corporation.**

Requirements quality analysis



(INCOSE, 2022)

Background: INCOSE defines 41 rules for writing high-quality requirements

Purpose: High-quality requirements ensures that the system will meet stakeholder needs and prevents costly reworks later in development.

Challenges:

Time: Manually checking for quality issues is time-consuming

Expertise: Requires familiarity with both the system and complex quality standards

Costly: Quality oversights are common and costly to fix (especially left unaddressed)

Approved for Public Release, Distribution Unlimited. Information derived from Public Release Case Number **25-2021**. This technical data deliverable was developed using contract funds under Basic Contract No. W56KGU-18-D-0004. © **2025 The MITRE Corporation.**

AI-assisted quality analysis

Goal: Automating the evaluation of shall requirements quality **increases speed** and **reduces missed quality issues**.

Based on our investigation, **32/41 (78%)** of INCOSE rules can be (at least partially) checked for automatically with various NLP techniques…

INCOSE Rules	NLP "Check" Method
R1 – Use a structured, complete sentence	Sentence structural analysis with statistical models
R20 – Avoid phrases that indicate the purpose of the requirement	Combined keyword matching & structural analysis with LLM pipelines
R12, 13, 14 – Use correct grammar, spelling, & punctuation	Grammar & spell-checking libraries
R40 – Group related requirements together	(Un)Assisted topic analysis (e.g. LDA)
R30 – Express each need and requirement once and only once	Document similarity analysis (e.g. cosine similarity & LLMs)

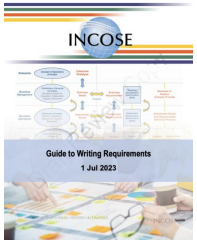
Approved for Public Release, Distribution Unlimited. Information derived from Public Release Case Number **25-2021**. This technical data deliverable was developed using contract funds under Basic Contract No. W56KGU-18-D-0004. © **2025 The MITRE Corporation.**

GenAI-assisted quality analysis

Generic LLMs like Llama3, Mistral, etc. can be prompted to evaluate quality and **rewrite** requirements. **RAG system** with INCOSE documents as context *significantly improves* results.

Other system documents

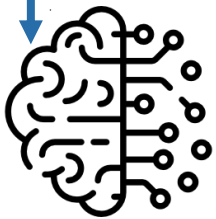
RAG



Output requires review (LLMs may hallucinate information)

Original

"The Army shall support secuirty operations that provide early and accurate warning of enemy attacks."



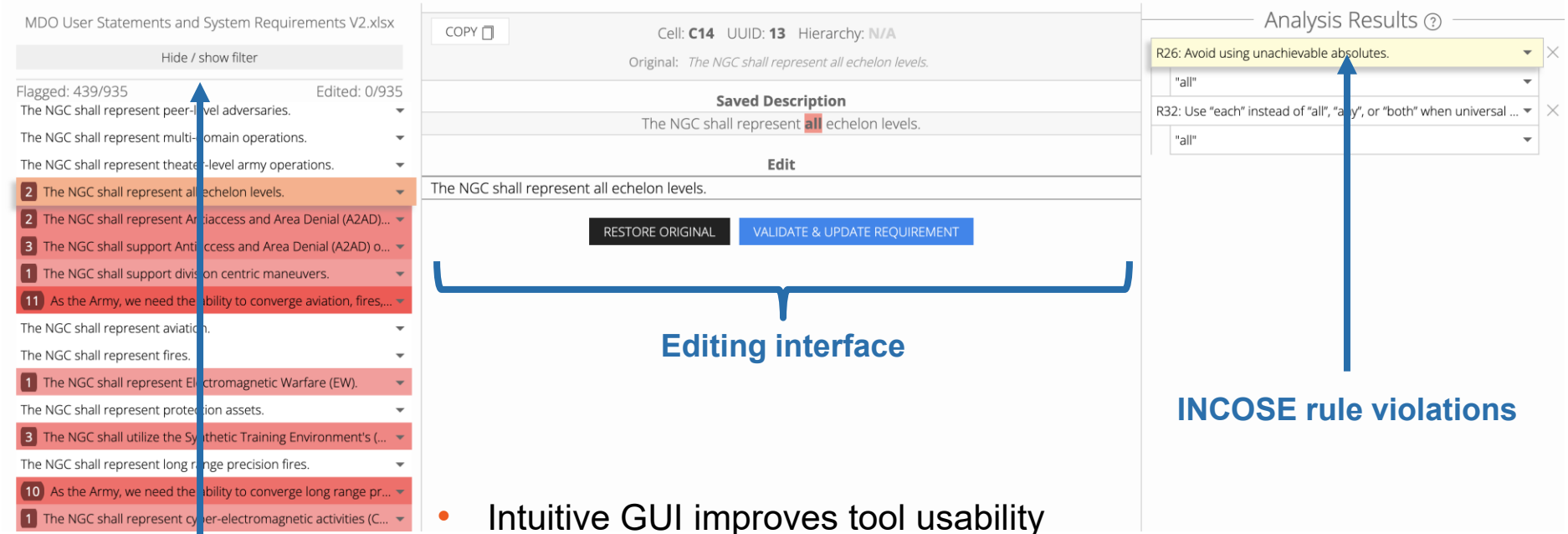
GenAI

Rewritten with GPT

"The Army shall support security operations that provide warnings of enemy attacks with a detection time of less than 5 minutes and an accuracy rate of greater than or equal to 95%, as defined in the Security Operations Data Dictionary."

Approved for Public Release, Distribution Unlimited. Information derived from Public Release Case Number **25-2021**. This technical data deliverable was developed using contract funds under Basic Contract No. W56KGU-18-D-0004. © **2025 The MITRE Corporation.**

RAAM RQA Editing GUI



- Intuitive GUI improves tool usability
- Embraces human-in-the-loop development

Approved for Public Release, Distribution Unlimited. Information derived from Public Release Case Number **25-2021**. This technical data deliverable was developed using contract funds under Basic Contract No. W56KGU-18-D-0004. © **2025 The MITRE Corporation.**

RAAM sample analysis result output

Taken from RAAM's Requirements Quality Analyzer (RQA)

Column for each INCOSE rule

R1	R2	R5	R7	R8	R9	R10	R16	R17	R18	R19	R20	R21
Ok	Ok	Ok	Ok	Ok	Ok	Ok	Ok	Ok	Ok	Ok	Ok	Ok
Ok	Ok	Ok	Ok	Ok	Ok	Ok	Ok	Ok	Ok	Ok	Ok	Ok
Ok	Wrong	Ok	Wrong	Ok	Ok	Ok	Ok	Ok	Ok	Ok	Ok	Ok
Ok	Ok	Ok	Wrong	Wrong	Ok	Ok	Wrong	Ok	Ok	Ok	Ok	Ok
Ok	Ok	Ok	Ok	Ok	Ok	Ok	Ok	Ok	Ok	Ok	Ok	Ok

- One row per shall requirement
- Good visualization of overall requirements quality
- Brings attention to frequently-broken rules

Approved for Public Release, Distribution Unlimited. Information derived from Public Release Case Number **25-2021**. This technical data deliverable was developed using contract funds under Basic Contract No. W56KGU-18-D-0004. © **2025 The MITRE Corporation.**

Case Studies

In 2012, MITRE support to the TRADOC Capability Manager for the Integrated Training Environment (TCM ITE) created the first RAAM RQA prototype to automate the evaluation of **several thousand requirements** (clarify vague requirements).

Manual requirements evaluation

Optimistic assumption: Engineers take 5 seconds to evaluate for each rule (4 minutes total per requirement).

Hundreds of hours needed to evaluate several thousand requirements.

With RAAM RQA

- Automatic pipeline returned a detailed evaluation of thousands of requirements in ~one minute of processing time.
- Enabled working groups to initially focus on only the requirements that failed evaluation

Approved for Public Release, Distribution Unlimited. Information derived from Public Release Case Number **25-2021**. This technical data deliverable was developed using contract funds under Basic Contract No. W56KGU-18-D-0004. © **2025 The MITRE Corporation.**

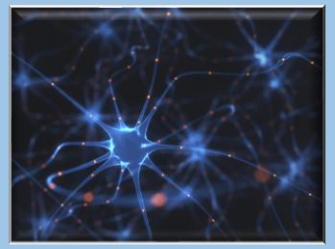
Requirements traceability analysis

The tracing and comparison of large numbers of engineering artifacts occurs for a multitude of use cases throughout the systems engineering lifecycle.

Mission Objectives to System Capabilities

Requirements to Architecture

Impact Analysis



Requirements to S&T

Requirements to Test & Evaluation

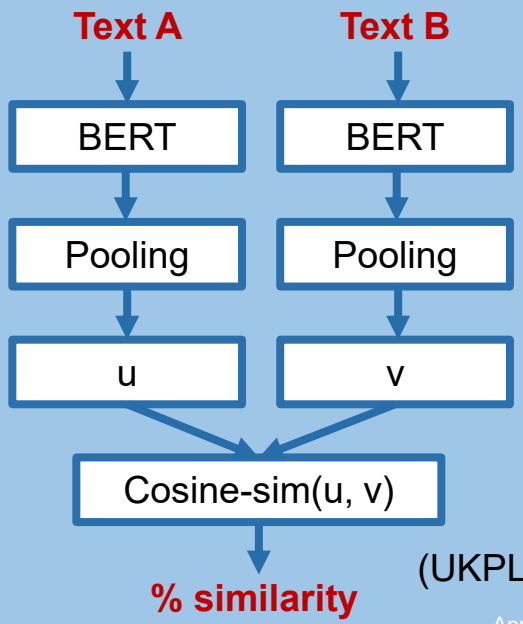
And much more…

Manual comparison is time-consuming and unintuitive with industry software, which usually aids with keyword searching or fuzzy matching at best.

Approved for Public Release, Distribution Unlimited. Information derived from Public Release Case Number **25-2021**. This technical data deliverable was developed using contract funds under Basic Contract No. W56KGU-18-D-0004. © **2025 The MITRE Corporation.**

AI-assisted traceability analysis

Goal: Accelerate identifying and establishing relationships between two related sets of engineering artifacts.

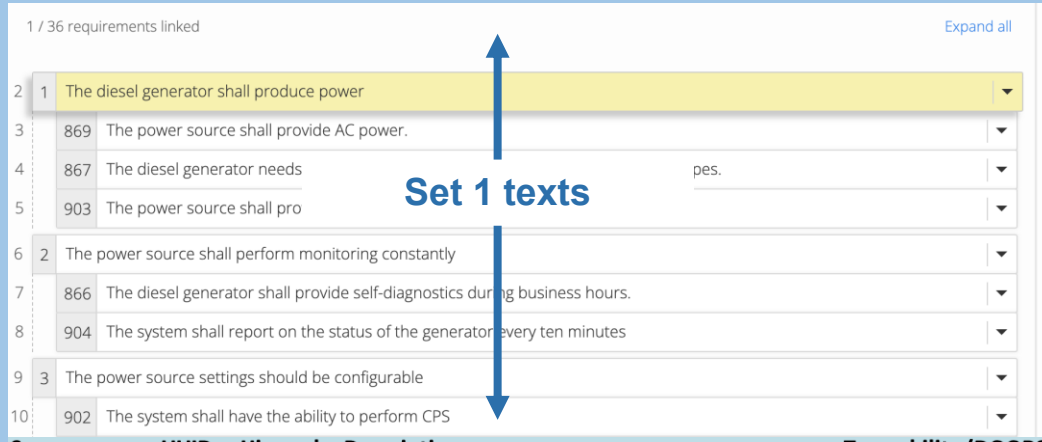


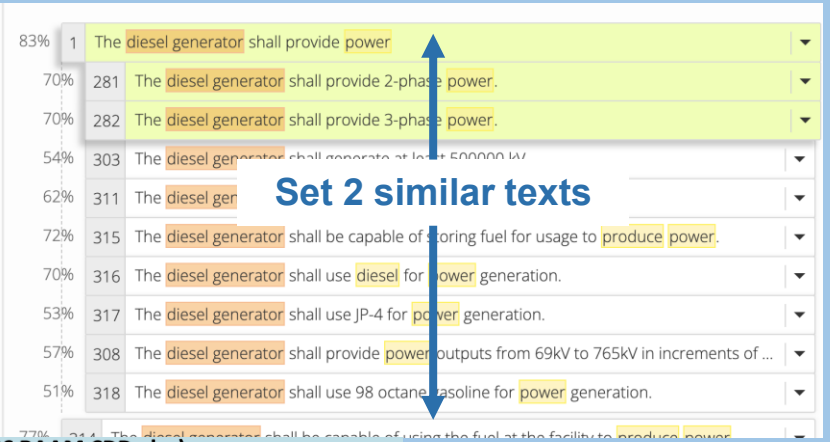
(UKPLab, 2025)

1. User uploads two sets of artifacts to compare
2. Software extracts and vectorizes each artifact as sentence embeddings (with **Sentence Transformers / SBERT**) + TF-IDF vectorization (Chandrasekaran & Mago, 2021)
3. Semantic similarity of each pair of artifacts is calculated with **cosine similarity function**
4. User reviews list of potential high-similarity "matches" and identifies relationships in GUI…

Approved for Public Release, Distribution Unlimited. Information derived from Public Release Case Number **25-2021**. This technical data deliverable was developed using contract funds under Basic Contract No. W56KGU-18-D-0004. © **2025 The MITRE Corporation.**

RAAM post-analysis GUI for establishing linkages





Source	UUID	Hierarchy	Description	Traceability (DOORS RAAM CDD.xlsx)
DOORS RAAM SRD.	1	1	The diesel generator shall provide power	DOORS RAAM CDD.xlsx - 869 - The power source shall provide AC power. DOORS RAAM CDD.xlsx - 867 - The diesel generator needs to provide multiple power types and multiple phase types. DOORS RAAM CDD.xlsx - 903 - The power source shall provide an alternating current at all times.
DOORS RAAM SRD.	282	1.2	The diesel generator shall provide 3-phase power.	DOORS RAAM CDD.xlsx - 886 - The diesel generator needs to provide consistent and reliable power at all times.
DOORS RAAM SRD.	315	1.5	The diesel generator shall be capable of storing fuel for usage to pr	DOORS RAAM CDD.xlsx - 1 - The diesel generator shall produce power DOORS RAAM CDD.xlsx - 875 - The diesel generator needs to store generated power in multiple types of devices (batteries, flywheels, etc.).
				DOORS RAAM CDD.xlsx - 1 - The diesel generator shall produce power DOORS RAAM CDD.xlsx - 888 - The diesel generator needs to be able to

“Linking table” summarizes linkages

Approved for Public Release, Distribution Unlimited. Information derived from Public Release Case Number **25-2021**. This technical data deliverable was developed using contract funds under Basic Contract No. W56KGU-18-D-0004. © **2025 The MITRE Corporation.**

Case Studies

Reviewing metrics collected from government project case studies between 2021 and 2023, the RAAM consistently accelerated traceability work by **400-500%.**

Case study: In 2023, the MITRE Corporation used the RAAM to consolidate, merge, and analyze **three requirements sets** to create user needs for the ArmyIgnitED's Government off the Shelf (GOTS) education management system.

Manual requirements evaluation

Engineer estimate: **85 hours**

Estimated based on the number of requirements and an estimated time per requirement.

With RAAM Tracing Tool

Measured time: 14 hours

This resulted in an **84% reduction in time** needed to complete the task (percent change).

Approved for Public Release, Distribution Unlimited. Information derived from Public Release Case Number **25-2021**. This technical data deliverable was developed using contract funds under Basic Contract No. W56KGU-18-D-0004. © **2025 The MITRE Corporation.**

Summary

Comparing the efficiency of SE work conducted with and without the aid of RAAM in case studies shows positive impacts of AI automation in several stages of the SE lifecycle:

1. Prototyping requirements and specifications
2. Identifying and resolving quality issues in requirements
3. Establishing traceability between engineering documents

Providing intuitive interfaces (GUIs):

- Improves usability and efficiency
- Enables human-in-the-loop workflow with AI systems

Approved for Public Release, Distribution Unlimited. Information derived from Public Release Case Number **25-2021**. This technical data deliverable was developed using contract funds under Basic Contract No. W56KGU-18-D-0004. © **2025 The MITRE Corporation.**

Future Work

- Explore further integration of genAI into existing capabilities
- Additional applications of NLP & genAI:
  - Converting waterfall shall requirements to Agile user stories
  - Generating architecture diagrams

Approved for Public Release, Distribution Unlimited. Information derived from Public Release Case Number **25-2021**. This technical data deliverable was developed using contract funds under Basic Contract No. W56KGU-18-D-0004. © **2025 The MITRE Corporation.**

References

Ling, C., Zhao, X., Lu, J. D., Zheng, C., Wang, J., Chowdhury, T., . . . Chen, Z. (2023). Domain Specialization as the Key to Make Large Language Models Disruptive: A Comprehensive Survey.

Barron, R. C., Grantcharov, V., Wanna, S., Eren, M. E., Bhattarai, M., Solovyev, N., . . . Alexandrov, B. S. (2024, October 3). Domain-Specific Retrieval-Augmented Generation Using Vector Stores, Knowledge Graphs, and Tensor Factorization. Los Alamos, New Mexico, United States of America.

INCOSE. (2022). INCOSE Guide to Writing Requirements V3.1 – Summary Sheet. Retrieved from INCOSE: https://www.incose.org/docs/default-source/working-groups/requirements-wg/rwg_products/incose_rwg_gtwr_summary_sheet_2022.pdf?sfvrsn=a95a6fc7_2

Chandrasekaran, D., & Mago, V. (2021). Evolution of Semantic Similarity—A Survey. ACM Computing Surveys, 54.

UKPLab. (2025). Semantic Textual Similarity. Retrieved from SBERT: https://sbert.net/examples/sentence_transformer/training/sts/README.html

Approved for Public Release, Distribution Unlimited. Information derived from Public Release Case Number **25-2021**. This technical data deliverable was developed using contract funds under Basic Contract No. W56KGU-18-D-0004. © **2025 The MITRE Corporation.**

Contact Info

Brian P Parrish: bparrish@mitre.org

Xuan Chau: xchau@mitre.org

MITRE

Approved for Public Release, Distribution Unlimited. Information derived from Public Release Case Number **25-2021**. This technical data deliverable was developed using contract funds under Basic Contract No. W56KGU-18-D-0004. © **2025 The MITRE Corporation.**

35th Annual INCOSE international symposium

hybrid event

Ottawa, Canada
July 26 - 31, 2025

Approved for Public Release, Distribution Unlimited. Information derived from Public Release Case Number **25-2021**. This technical data deliverable was developed using contract funds under Basic Contract No. W56KGU-18-D-0004. **© 2025 The MITRE Corporation.**