Accelerating Software Acquisition Using Generative AI for Regulatory Compliance

 $\mathsf{MAY} \ 7, \ 2025$

John E. Robert, Deputy Director, Software Solutions Division, jer@sei.cmu.edu Carlos Olea, SEI Visiting Scientist, <u>colea@sei.cmu.edu</u> Yash Hindka, SEI Member of Technical Staff, <u>yhindka@sei.cmu.edu</u> Nanette Brown, Senior SEI Member of Technical Staff, <u>nb@sei.cmu.edu</u> Douglas C. Schmidt, Dean of Computing, Data Sciences & Physics at William & Mary, <u>douglas.c.schmidt@wm.edu</u>



Document Markings

The following markings MUST be included in work product when attached to this form and when it is published.

For purposes of double anonymous peer review, markings may be temporarily omitted to ensure anonymity of the author(s).

Carnegie Mellon University 2025

This material is based upon work funded and supported by the Department of Defense under Contract No. FA8702-15-D-0002 with Carnegie Mellon University for the operation of the Software Engineering Institute, a federally funded research and development center. NO WARRANTY. THIS CARNEGIE MELLON UNIVERSITY AND SOFTWARE ENGINEERING INSTITUTE MATERIAL IS FURNISHED ON AN "AS-IS" BASIS. CARNEGIE MELLON UNIVERSITY MAKES NO WARRANTIES OF ANY KIND, EITHER EXPRESSED OR IMPLIED, AS TO ANY MATTER INCLUDING, BUT NOT LIMITED TO, WARRANTY OF FITNESS FOR PURPOSE OR MERCHANTABILITY, EXCLUSIVITY, OR RESULTS OBTAINED FROM USE OF THE MATERIAL. CARNEGIE MELLON UNIVERSITY DOES NOT MAKE ANY WARRANTY OF ANY KIND WITH RESPECT TO FREEDOM FROM PATENT, TRADEMARK, OR COPYRIGHT INFRINGEMENT.

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution. Please see Copyright notice for non-US Government use and distribution.

This material may be reproduced in its entirety, without modification, and freely distributed in written or electronic form without requesting formal permission. Permission is required for any other use. Requests for permission should be directed to the Software Engineering Institute at permission@sei.cmu.edu.

DM25-0657



Carnegie Mellon University Software Engineering Institute

Motivation for Al-Augmented Acquisition

Software Acquisition Using Generative AI for Regulatory Compliance



Carnegie Mellon University Software Engineering Institute

The CMU SEI-led study Architecting the Future of Software Engineering: A National Agenda for Software Engineering Research & Development laid out a multi-year roadmap for engineering next-generation software-reliant systems.

Study available online at <u>https://www.sei.cmu.edu/go/nation</u> <u>al-agenda</u>

Software Engineering Research Roadmap (10-15 Year Horizon)



Software Engineering Research Roadmap (10-15 Year Horizon)

Carnegie Mellon



How Advances in Generative AI Affected Our Study Findings

• Based on extensive experience, we've refined our taxonomy to focus on the degree of AI-augmentation for the software development lifecycle (SDLC) & system operations

Degree of AI-augmentation

for System Operations

High Application of Large Language Models (LLMs) in Software Engineering: Overblown Hype or Disruptive Al-auamented Al-auamented Change? systems built using systems built using AIconventional SDLC augmented SDLC techniques IPEK OZKAYA, ANITA CARLETON, JOHN E, ROBERT, AND techniques **OCTOBER 2. 2023** Has the day finally arrived when large language models (LLMs) turn us all into better software engineers? Or are LLMs creating more hype than functionality for software development, and, at the same time, plunging everyone into a world where it is hard to distinguish the perfectly formed, yet sometimes fake and incorrect, code generated by artificial intelligence (AI) programs from verified and well-tested systems? LLMs and Their Potential Impact on the Future of Software Engineering Conventional systems Conventional systems This blog post, which builds on ideas introduced in the IEEE paper Application of Large Language Models to Software Engineering built using conventional built using Al-Tasks: Opportunities, Risks, and Implications by Ipek Ozkaya, focuses on opportunities and cautions for LLMs in software development, the implications of incorporating LLMs into software-reliant systems, and the areas where more research and SDLC techniques augmented SDLC innovations are needed to advance their use in software engineering. The reaction of the software engineering community to the accelerated advances that LLMs have demonstrated since the final guarter of 2022 has ranged from snake oil to no help for techniques programmers to the end of programming and computer science education as we know it to revolutionizing the software development process. As is often the case, the truth lies somewhere in the middle, including new opportunities and risks for developers using LLMs. Research agendas have anticipated that the future of software engineering would include an Al-augmented software Low development lifecycle (SDLC), where both software engineers and Al-enabled tools share roles, such as copilot, student, expert, and supervisor. For example, our November 2021 book Architecting the Future of Software Engineering: A National Agenda for High Software Engineering Research and Development describes a research path toward humans and Al-enabled tools working as Low trusted collaborators. However, at that time (a year before ChatGPT was released to the public), we didn't expect these Degree of Al-augmentation in the opportunities for collaboration to emerge so rapidly. The figure below, therefore, expands upon the vision presented in our 2021 book to codify the degree to which AI augmentation can be applied in both system operations and the software Software Development Lifecycle (SDLC) development lifecycle (Figure 1), ranging from conventional methods to fully Al-augmented methods.

Carnegie Mellon

See application-of-large language-models-Ilms-in-software-engineering-overblown-hype-or-disruptive-change

Carnegie Mellon University Software Engineering Institute

Degree of AI-augmentation for System Operations

High

AI-augmented systems acquired using conventional acquisition methods *AI-augmented systems acquired using AI-augmented acquisition methods*

Conventional systems acquired using conventional acquisition methods

Conventional systems acquired using AI-augmented acquisition methods

High

THE FUTURE OF SOFTWARE ENGINEERING AND ACQUISITION WITH GENERATIVE AI Software Engineering Institute (SEI)

J. Robert, J. Ivers, D. C. Schmidt, I. Ozkaya, & S. Zhang, "The Future of Software Engineering & Acquisition with Generative AI," Crosstalk, June 2024.

Low

Low

Degree of AI-augmentation in the Acquisition Activities

Carnegie Mellon University Software Engineering Institute

Degree of AI-augmentation for System Operations

AI-augmented systems acquired using conventional acquisition methods AI-augmented systems acquired using AI-augmented acquisition methods

Conventional systems acquired using conventional acquisition methods *Conventional systems acquired using AI-augmented acquisition methods*

High

Low

Low

High

Degree of AI-augmentation in the Acquisition Activities

J. Robert, J. Ivers, D. C. Schmidt, I. Ozkaya, & S. Zhang, "The Future of Software Engineering & Acquisition with Generative AI," Crosstalk, June 2024.

A military-grade GPS satellite system using traditional data transmission & encryption for operations & is developed using conventional acquisition processes without any AI-augmented tools or methods.



Carnegie Mellon University Software Engineering Institute

Degree of AI-augmentation for System Operations AI-augmented systems acquired using conventional acquisition methods AI-augmented systems acquired using AI-augmented acquisition methods

Conventional systems acquired using conventional acquisition methods *Conventional systems acquired using AI-augmented acquisition methods*

Low

High

Low

Degree of AI-augmentation in the Acquisition Activities

High

GPS-guided munition where the content is not AI-augmented, but the acquisition activities employ AIassistance in identifying & analyzing relevant regulations, standards, & potential security risks.



J. Robert, J. Ivers, D. C. Schmidt, I. Ozkaya, & S. Zhang, "The Future of Software Engineering & Acquisition with Generative AI," Crosstalk, June 2024.



J. Robert, J. Ivers, D. C. Schmidt, I. Ozkaya, & S. Zhang, "The Future of Software Engineering & Acquisition with Generative AI," Crosstalk, June 2024.

Degree of AI-augmentation

Carnegie Mellon University Software Engineering Institute

Degree of AI-augmentation for System Operations

AI-augmented systems acquired using conventional acquisition methods AI-augmented systems acquired using AI-augmented acquisition methods

Conventional systems acquired using conventional acquisition methods

Conventional systems acquired using AI-augmented acquisition methods

High

Low

High

Low

Degree of AI-augmentation in the Acquisition Activities

J. Robert, J. Ivers, D. C. Schmidt, I. Ozkaya, & S. Zhang, "The Future of Software Engineering & Acquisition with Generative AI," Crosstalk, June 2024.

automated regulatory compliance.

An autonomous vehicle or platform

that employs AI to navigate while also

using AI-augmented acquisition

processes, methods, & tools, such

as text summarization & semi-



Carnegie Mellon University Software Engineering Institute

As software-reliant systems become increasingly essential to the success of DoD missions there has been greater emphasis on codifying software acquisition strategies policies

https://aaf.dau.edu/aaf/aaf-pathways/



Challenges for acquisition professionals

 Heavily regulated environment with multiple levels of policy & legal responsibilities



https://aaf.dau.edu/aaf/aaf-pathways/

Carnegie Mellon



Challenges for acquisition professionals

 Heavily regulated environment with multiple levels of policy & legal responsibilities



https://aaf.dau.edu/aaf/aaf-pathways/

Carnegie Mellon



https://aaf.dau.edu/aaf/aaf-pathways/

16

Carnegie Mellon

University

Challenges for acquisition professionals

- Heavily regulated environment with multiple levels of policy & legal responsibilities
- Volume of policy, guidance, & standards documents
- Concurrency of activities within one program





https://aaf.dau.edu/aaf/aaf-pathways/

CMU SEI © 2025 Carnegie Mellon University

Challenges for acquisition professionals

- Heavily regulated environment with multiple levels of policy & legal responsibilities
 - Volume of policy, guidance, & standards documents
- Concurrency of activities within one program & across multiple programs in a system-of-system



[DISTRIBUTION STATEMENT A] Approv

Acquisition Example: Software Safety



- A. <u>Regulatory documents</u> Policies, standards, and guidance documents.
- B. <u>Safety Requirements</u> –Document that lists safety requirements derived from the regulatory documents that are relevant to a specific system.
- C. <u>Derived Safety Requirements</u> Document that captures derived safety requirements that decomposes the safety requirements from top system level to every software module or component in the system.
- D. <u>Safety Verification</u> Collection of safety verification documents and data, that summarize the verification evidence (test results, analysis results, etc.) when compared to the derived safety requirements.

Carnegie Mellon

Acquisition Example: Software Safety



- A. <u>Regulatory documents</u> Policies, standards, and guidance documents.
- B. <u>Safety Requirements</u> –Document that lists safety requirements derived from the regulatory documents that are relevant to a specific system.
- C. <u>Derived Safety Requirements</u> Document that captures derived safety requirements that decomposes the safety requirements from top system level to every software module or component in the system.
- D. <u>Safety Verification</u> Collection of safety verification documents and data, that summarize the verification evidence (test results, analysis results, etc.) when compared to the derived safety requirements.

Carnegie Mellon University

Acquisition Example: Software Safety



- A. <u>Regulatory documents</u> DIID can occur between multiple policy, standards, and guidance documents.
- B. <u>Safety Requirements</u> DIID can occur between regulatory documents and safety requirements document that lists safety requirements derived from the regulatory documents that are relevant to a specific system.
- C. <u>Derived Safety Requirements</u> DIID can occur between safety requirements documents and derived safety requirements document that extends the safety requirements from top level to every software module or component in the system.
- D. <u>Safety Verification</u> DIID can occur with safety verification documents that summarize the verification evidence (test results, analysis results, etc.) when compared to the derived safety requirements.

Carnegie Mellon

AI-Augmented Acquisition

Challenges for acquisition professionals

- Heavily regulated environment with multiple levels of policy & legal responsibilities
- Volume of policy, guidance, & standards documents
- Concurrency of activities within one program & across multiple programs in a system-of-systems





AI-Augmented Acquisition

- Al tools to augment people
- Intelligent automation to scale

An overarching challenge for software acquisition professionals is detecting & remediating incompleteness, inconsistencies, & discrepancies within/between documents and software engineering artifacts Carnegie Mellon

University

Carnegie Mellon University Software Engineering Institute

Find DIID in multiple Find DIID in multiple policy or standard policy or standard documents and multiple documents compared to one software artifact software artifacts Find DIID from one policy Find DIID in one policy or standard document document compared to multiple software artifacts compared to one software artifact

Artifact Scale

Policy Scale

Software Acquisition Using Generative AI for Regulatory Compliance

Role of DIID in Regulatory Risk and Non-Compliance



Types of DIID

Incompleteness	Inconsistency	Discrepancy
• <u>Incomplete</u> :	 <u>Terminology</u>: Using 	<u>Factual discrepancies</u> :
Important	different terms	Conflicting factual
context or terms	interchangeably	information.
are missing.	without clear definitions or consistency.	• <u>Policy or procedural</u> <u>discrepanci</u> es: Deviations from established protocols.
	 <u>Structural</u>: Lack of uniform structure in presenting information. 	• <u>Narrative discrepancies</u> : Different user stories fail to align.
		• <u>Theoretical discrepancies</u> : Actual results conflict with theoretical predictions.

AI-Augmented DIID Detection in DoD Safety Standards



The Role of Prompt Engineering in DIID Detection



LLM Testing Frameworks for DIID Use Cases



Measuring the Impact of AI-Augmented DIID Detection

Carnegie Mellon University Software Engineering Institute

Compliance Time savings & Accuracy gains Variance & dashboards & analyst workload vs. manual consistency reduction baselines aggregated metrics **Tuned LLM** YE5 95% -> Risk met LLM Manual LLM + TRAD. TRI vector DB review vector DB Tunued Untuned Originate **FASTER REVIEW** from in-scope HIGHER ISSUE X CYCLE Minimized LLM causes RECALL output variance **REDUCED REVIEWE** PRECISION FATIGUE Minimized **INSIGHTS** from category prompt MORE DOCS PER engineering assignment PERIOD variance Absence of critical details can lead to **Risk categories** false confidence cited

Future Research for AI-Augmented DIID Detection

DIID

Dataset Expansion

Domain Benchmarking



- Benchmark LLMs across DoD domains
- Test general-purpose vs. fine-tuned vs. on-prem LLMs
- Metrics: adaptability, accuracy, Integration

- Annotated document pairs (DIID issue label:)
- Expand domain-specific examples via LLM self-generation
- Semi-supervised augmentation with

- Inputs: UMLs, ICDs, source code
- Use vision-language models or diagram parsers

Multi-Modal

DIID Detection

 Compare diagram logic vs. narrative text for mismatc-

- Al flags ambiguous/ conflicting clauses in policies
- Suggest edits or clarifications
- Feedback loop to UX & policy authors

Human-in-the-Loop Validation



Carnegie Mellon University

Thank You!



Carnepicture Mello(optional) University Software Engineering Institute