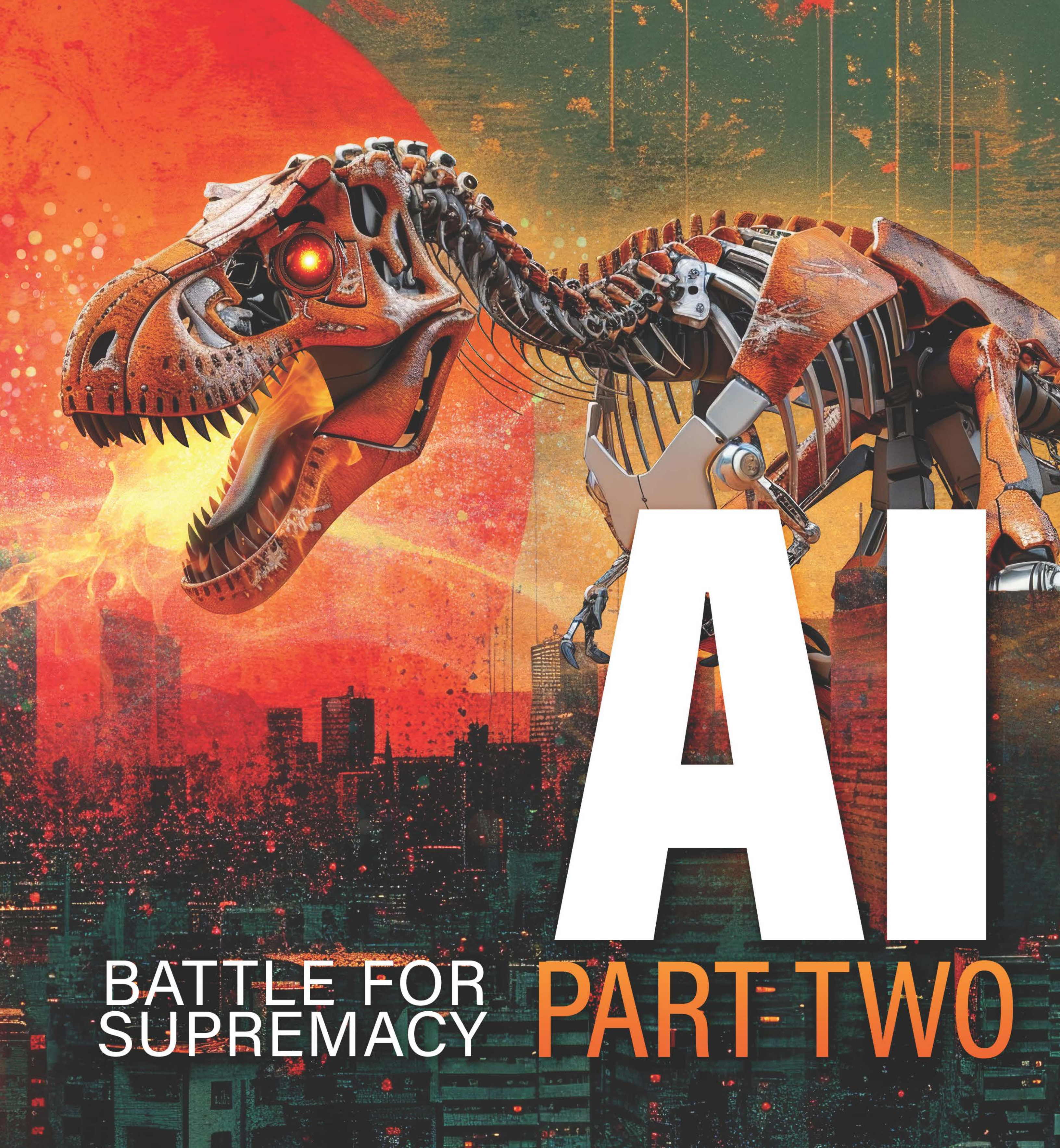
CrossTalk

The Journal of Defense Software Engineering August 2024



AI

BATTLE FOR SUPREMACY PART TWO

CrossTalk: The Journal of Defense Software Engineering

CrossTalk: The Journal of Defense Software Engineering is sponsored by the Air Force Sustainment Center Software Directorate (AFSC/SW). It is also supported by other partners within the Department of Defense (DoD), other United States Air Force (USAF) systems, and the software engineering community. Established by Gen. Richardson, the Air Force Materiel Command Commander (AFMC/CC), the AFSC Software Directorate supports the AFMC Strategic Plan by facilitating and delivering an integrated software ecosystem across the Department of the Air Force (DAF).

AFSC/SW serves as the organic source of software engineering services for DAF weapon systems and equipment, and for all echelons of software development and sustainment, keeping pace with advancing technology, mission capability needs, and dynamics of the cyber environment.

The mission of CrossTalk is to encourage the engineering development and proper management of software to improve the reliability, sustainability, and responsiveness of our warfighting capability.

CrossTalk Online: Current and past issues are posted at the following locations: The Software Directorate website, All Partners Access Network (APAN), and Defense Technical Information Center (DTIC). The Software Directorate website houses the four most recent issues of Crosstalk while past issues can be found on APAN or DTIC.

https://afscsoftware.dso.mil/crosstalk/

https://community.apan.org/wg/crosstalk/

https://www.dodtechipedia.mil/dodwiki/x/HwDqFQ (Requires .mil domain for full support)

Subscriptions: Please send an email to the publisher to receive a notification when each new issue is published online. Readers can also sign up for notifications on APAN.

Article Submissions: We welcome articles of interest to the defense software community. Articles must be approved by the Technical Review Board (TRB) prior to publication. Please follow the Author Guidelines, available at any of the websites above. CrossTalk does not pay for submissions. Published articles remain the property of the authors and may be submitted to other publications. Security agency releases, clearances, and public affairs office approvals are the responsibility of the authors and their organizations. Potential articles can be emailed to: AFSC.SWSWDE.Crosstalk@us.af.mil

Reprints: Permission to reprint or post articles must be requested from the author or the copyright holder and coordinated with CrossTalk.

Trademarks and Endorsements: CrossTalk is an authorized publication for members of the DoD. Contents are not necessarily the official views of, or endorsed by, the U.S. government, the DoD, the sponsors, or co-sponsors. All product names referenced in this issue are trademarks of their respective companies.

Publishing Schedule and Back Issues: CrossTalk is currently being published quarterly. Please phone or email us to see if back issues are available, free of charge.

ISSN 2160-1577 (prior print versions); ISSN 2160-1593 (online)

CrossTalk Team

AFSC Software Directorate

Sponsor

Mr. Edward W. Ayer, SES
Director, AFSC Software Directorate

Managing Director

Gabbrieal C. Madkins
Chief, Directorate Services Division

Assistant Director

Megan G. Allen
Chief, Creative Content Branch

Managing Publishers

Lennis L. Burton
CrossTalk Publications Manager
Siria L. Snounou
CrossTalk Publications Manager
Destinie Comeau
CrossTalk Publications Manager

Technical Reviewer

David Webb
Software Subject Matter Expert

Contact us

Phone

Lennis L. Burton, (801) 775-3262

Siria L. Snounou, (801) 777-4734

Destinie Comeau, (801) 775-3246

E-Mail

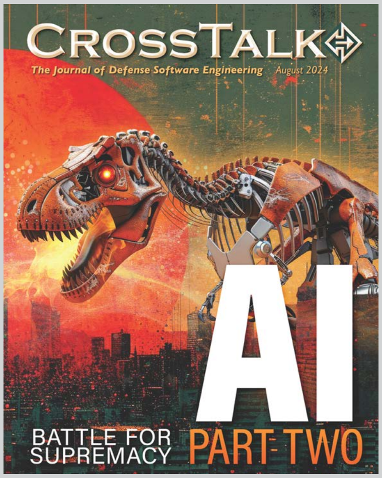
AFSC.SWSWDE.Crosstalk@us.af.mil

Connect with us on the AFSC Software Directorate LinkedIn page!

SOFTWARE
DIRECTORATE

CrossTalk: The Journal of Defense Software Engineering

Directory



Cover Design by Kent Bingham

AI Part 2

From the Sponsor

Battle of the Bots

Ritesh Patel
Deputy Chief Engineer
Systems Engineering Senior Scientific and Technical Manager,
Naval Information Warfare Systems Center

When I was in college, the show *BattleBots* was one of the most popular on television. I, along with millions of other fans, tuned into the *Discovery* channel and spent many hours weekly watching home-built robots from across the globe duel until both were destroyed or one emerged victorious.

While the arena of metal-clashing robots may be less popular today (except in Las Vegas), the battle for dominance rages on in a new form. Today, many nation-states are locked in a fierce race to develop cutting-edge generative Artificial Intelligence (AI) algorithms and Large Language Models (LLMs), vying for battlespace supremacy. This struggle could dramatically reshape the world order, transforming not only the technology and communication landscape, but also economics, security, and social norms.

Our ability to outmaneuver competition will be founded on private sector innovations. AI and Machine Learning (ML) applications have long existed, but LLMs brought the technology to the forefront by enabling easy access to foundational models and algorithms through an Application Programming Interface (API). Since OpenAI launched ChatGPT in late 2022, investments have poured in. This influx of capital has created a comprehensive ecosystem of tools and technology, fostering vibrant open-source communities. Developers now use platforms such as *Hugging Face* to create chatbots capable of human-like conversation and assistants that summarize vast amounts of text, while tools like *Stable Diffusion* and *DALL-E* (an acronym combining Dalí and WALL-E) generate images and videos from simple prompts.

LLMs and generative AI offer numerous opportunities and unique challenges for the Department of Defense (DoD). LLMs are invaluable for achieving decision superiority, particularly as the volume of data accessible to warfighters continues to grow. By fusing vast amounts of data, LLMs enable faster and more accurate decision-making, allowing users to leverage the power of machines to make appropriate battlefield decisions swiftly and effectively.

To address the corresponding challenges, the DoD established Task Force Lima in 2023 to explore responsible approaches to harnessing the power of generative AI, assigning the Chief Digital and Artificial Intelligence Office (CDAO) to lead the effort [1].

CDAO Officer Dr. Radha Plumb describes the problem space in three parts:

- Making data readily available, accessible, and usable
- Creating enabling infrastructure consisting of labeled data, development and test environments, sufficient computing, and test and evaluation tooling
- Implementing acquisition pathways to procure innovative technology and solutions from the private sector [2]

The services are assisting CDAO in their efforts, playing a critical role in scaling commercial and open-source LLMs, integrating them into Programs of Record, and optimizing the user experience for operational use by our nation's warfighters. The DoD's engineering and scientific workforce actively participates in AI communities, including Naval Applications of Machine Learning (NAML) workshops. They collaborate to develop and apply LLMs and generative AI to various products including Co-pilot coding assistants, enhancing developer productivity; *NIPRGPT*, a conversational AI securing communications; and *ChatIT*, a LLM chatbot pathfinder undergoing Fleet deployment to provide valuable lessons learned for future AI/ML applications.

Private entities are contributing as well. *Ask Sage* provides GovCloud-hosted LLMs and agents with API access to government organizations. *AskSage's* capabilities will automate once-complicated processes such as Risk Management Framework implementation using LLMs.

There is still much work to be done. Deploying AI and LLM capabilities requires an understanding of available data, sharing requirements, and data-sharing and availability limitations between organizations. This includes preparing and labeling data while ensuring appropriate data rights when using commercial and open-source models. Additionally, creating an open, federated data architecture is essential to leveraging labeled data at scale.

Crucially, discovering and promoting use cases to encourage end-to-end experimentation will help demonstrate the value of these technologies to the stakeholders and facilitate the investment in and creation of the necessary infrastructure.

In this issue of *CrossTalk*, we hear from professionals across the defense community about overcoming these challenges and showcasing AI and LLMs exemplars to tackle key use cases driving AI superiority and strategic advantage.

Dr. Lori Flynn and Dr. Will Klieber's article "Using LLMs to Automate Static Analysis Adjudication and Rationales" discusses a model for using LLMs to handle static analysis output, initial tooling developed and experimental results, related work by others, and additional work needed.

Rebel Brown's article "Understanding Today's Artificial Intelligence for Government Use" clarifies the state of AI and ML today, including how they are being and might later be used in the government sphere.

Tony Lau's article "Leveraging AskSage: An Intelligent Question-Answering System for Enhanced Efficiency in Government-Related Tasks" explains the background, current uses, future possibilities, security, and ethical considerations of AskSage, an AI software for the government and DoD.

As AI continues to evolve, the real question is who will achieve this first? Like a showdown of *Battle-Bots*, only the best and fastest can win.

Picture References

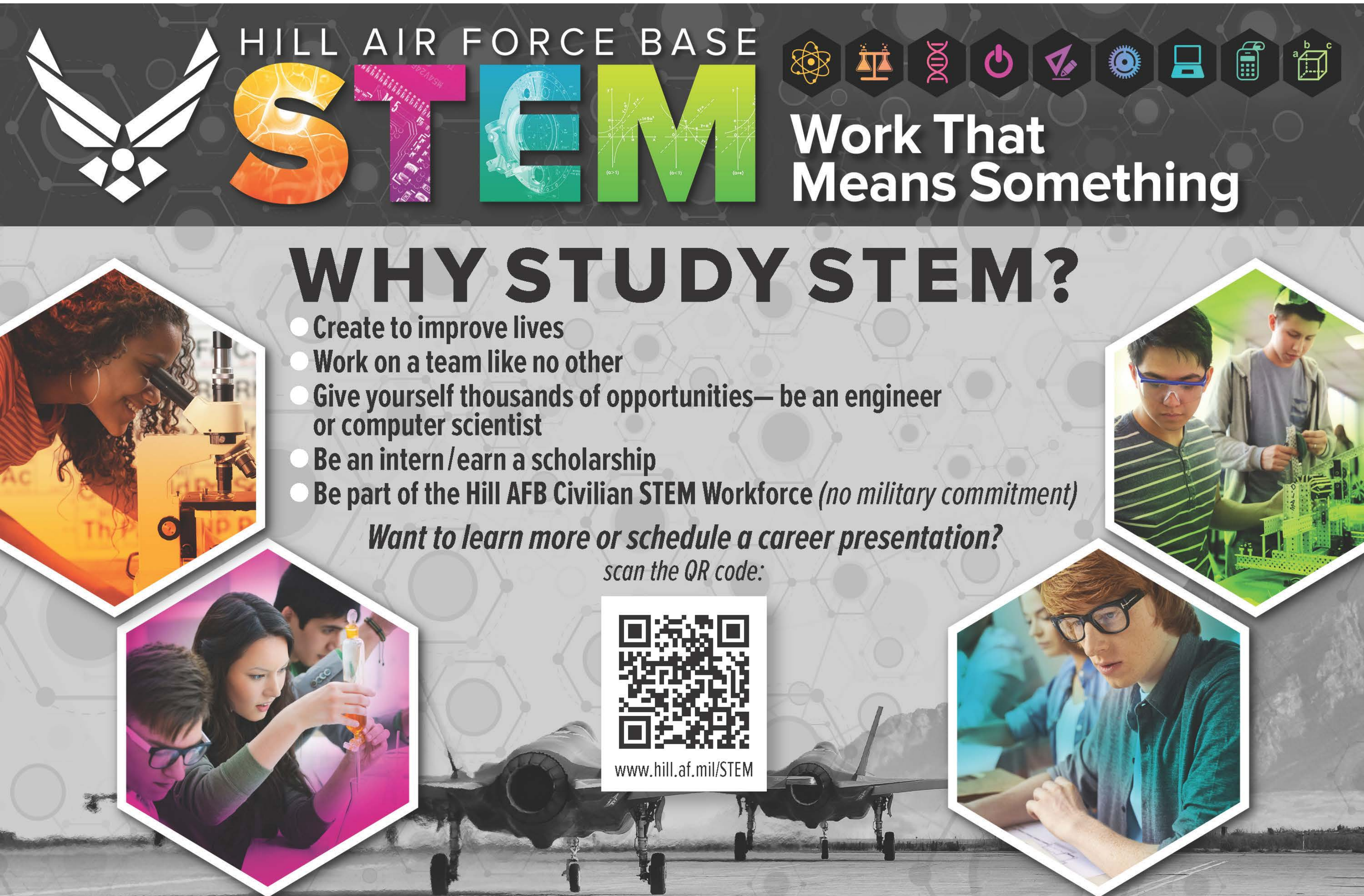
[A] “A kinetic war battlefield in a natural, mountainous region with unmanned drones actively identifying a target. In the sky, a few prominent quad drones.” ChatGPT, Version 0.4, OpenAI, 29 July 2024. chatgpt.com

References

[1] United States Department of Defense. “DoD Announces Establishment of Generative AI Task Force.” *United States Department of Defense,* 10 Aug 2023. https://www.defense.gov/News/Releases/Release/Article/3489803/dod-announces-establishment-of-generative-ai-task-force/

[2] Allen, Gregory C. “AI Transformation at the DoD: A Conversation with Chief Digital and AI Officer, Dr. Radha Plumb.” *Center for Strategic and International Studies,* 15 July 2024. https://www.csis.org/analysis/ai-transformation-dod-conversation-chief-digital-and-ai-officer-dr-radha-plumb

- Ritesh Patel; Deputy Chief Engineer, Systems Engineering Senior Scientific and Technical Manager; Naval Information Warfare Systems Center



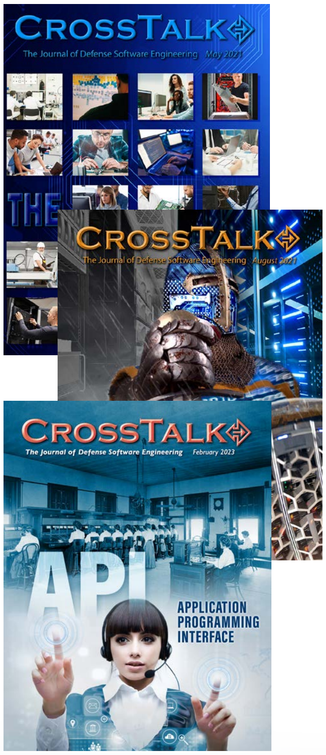
Call For Articles

If your experience or research has produced information that could be useful to others, Crosstalk can get the word out. We are specifically looking for articles on software-related topics to supplement upcoming theme issues. Below is the submittal schedule for the areas of emphasis we are looking for.

Big Data

Nov 2024 Issue

Submission Deadline:
September 15, 2023


CrossTalk
The Journal of Defense Software Engineering May 2021
THE
CrossTalk
The Journal of Defense Software Engineering August 2021
CrossTalk
The Journal of Defense Software Engineering February 2023
API
APPLICATION PROGRAMMING INTERFACE

Keeping up with the Cloud

February 2025

Submission Deadline:
October 31, 2024

The 3 C's of Agile

May 2025

Submission Deadline:
March 15, 2025

Please follow the Author Guidelines for Crosstalk, available at the APAN or DTIC site.

We accept article submissions on software-related topics at any time, along with Letters to the Editor, Open Forum, and BackTalk. To learn more about the types of articles we're looking for, please visit the above sites or contact us by email or phone

Contact Us

By phone

Lennis L. Burton, (801) 775-3262
Siria L. Snounou, (801) 777-4734
Destinie Comeau, (801) 775-3246

By email

517SMXS.CrossTalk.Articles@us.af.mil

Using LLMs to Automate Static-Analysis Adjudication and Rationales

Dr. Lori Flynn
Senior Software Security Researcher,
Software Engineering Institute

Dr. Will Klieber
Software Security Researcher,
Software Engineering Institute

Abstract

Software vulnerabilities are a serious concern for the Department of Defense (DoD). Software analysts use static analysis as a standard method to evaluate the source code, but the volume of findings is often too large to review in their entirety, causing the DoD to accept unknown risk. Large Language Models (LLMs) are a new technology that show promising initial results for automation of alert adjudication and rationales. This has the potential to enable more secure code, better measure risk, support mission effectiveness, and reduce DoD costs. This article discusses our model for using LLMs to handle static analysis output, initial tooling we developed and our experimental results, related work by others, and additional work needed. Beyond static-analysis alert adjudication, similar techniques can be used to create LLM-based tools for other code analysis tasks.

Motivation for Improving Static Analysis on DoD Software

The Authorization-to-Operate (ATO) process assesses risks that software may introduce [1]. During Test & Evaluation (T&E) and Independent Verification & Validation (IV&V), software analysts evaluate source code for security weaknesses to measure risk and enable code improvement in preparation of ATO and fielding. Static analysis (SA) is widely used and is one of the best techniques available: it is much more practical than full formal verification, and it can catch vulnerabilities that can evade dynamic analysis. But static analysis still requires significant manual effort and is inherently difficult, time-consuming, and expensive (See page 16 for more information about these methods). Manual effort is required for each SA alert to adjudicate whether it is a true or false positive, since SA tools sometimes produce false positives. There are many types of code flaws identified in taxonomies such as the Common Weakness Enumeration (CWE), and SA tools produce alerts for many types. Human analysts must be able to analyze each kind to be able to adjudicate the alert, which requires great expertise.

The most common strategy to adjudicate alerts given finite time and resources is to prioritize potential vulnerabilities by a combination of likelihood (with static-analysis tools usually pre-filtering out unlikely ones in their default configurations) and severity (e.g., Security Technical Implementation Guide (STIG) Category 1 [2]) and then manually review only the top alerts. However, even code weaknesses in lower-severity Application Security and Development (ASD) STIG categories can also cause costly mission failure. Many types of code flaws can lead to vulnerabilities that common attack patterns use; e.g., the Common Attack Pattern Enumeration and Classification (CAPEC) [3] describes an attack pattern [4] that takes advantage of a lower-category weakness [5]. As another example, the Ariane flight V88 rocket explosion (which resulted in a loss of more than $370 million) was caused by code flaws that static analysis tools can detect (integer overflow and improper exception handling) [6] [7] but that often aren't put in STIG Category 1.

Latest LLMs as Breakthroughs for Automating Static Analysis Alert Adjudication

Large Language Models, such as GPT-4 (OpenAI's latest Generative Pre-trained Transformer (GPT))) [8], present a significant breakthrough, for two major reasons:

1. They produce a detailed explanation to support their final answer, in contrast to older machine learning (ML) techniques [9] which involve statistical algorithms that can learn from data and generalize to unseen data. These older ML techniques lacked interpretability and often pivoted on irrelevant details that merely correlated with vulnerabilities in their training data. The generated explanation can be double-checked by both humans and the LLM itself.

2. They can generate and use function summaries, function preconditions, and other intermediate results to enable LLM-based tools to adjudicate alerts whose adjudication requires analyzing multiple functions spread across the codebase.

Chan et al. use LLMs to detect over 250 vulnerability types in code being edited. They deployed their model as a Visual Studio (VS) Code extension with ~100K daily users, with a 90% reduction in the rate of vulnerabilities in developer code [10]. Fan et al. developed an intelligent agent that responds to queries by processing code and interactions with LLMs, SA tools, code retrieval tools, and web search tools to check intentions of code segments and detect bugs [11]. LLMAO (LLM fAult lOcalization) [12] is an LLM-based approach for localizing program defects at line level, outputting bug probabilities for each line of code. It localized more bugs in the same set of benchmark codebases than the previous best deep-learning fault localizer, and it doesn't require any additional training or test cases to handle unseen projects. It uses a bidirectional language model, allowing it to consider both preceding and following lines.

Two recent papers have attempted to quantify the benefits of applying LLMs to the problem of pruning false positives from static-analysis alerts [13][14]. Both have explored the benefits of designing a prompting algorithm, highlighting the importance of chain-of-thought, task decomposition, and progressive prompting strategies. One found that an LLM-enabled system demonstrated high precision and recall in a real-world scenario and even identified 13 previously unknown use-before-initialization (UBI) bugs in the Linux kernel [14]. While these studies provide useful templates for system design, they do not fully address the DoD's challenges because they both focused on the relatively narrow application of UBI bugs in the Linux kernel as a single case study.

Sherman found that LLMs often perform poorly when asked to find all security issues in a snippet of code[16]. We have found that LLMs do much better when asked to adjudicate a specific type of issue on a specific line. Li et al also found that GPT-4 works well for this task [11].

Our Initial Results Using LLMs for SA Alert Adjudications

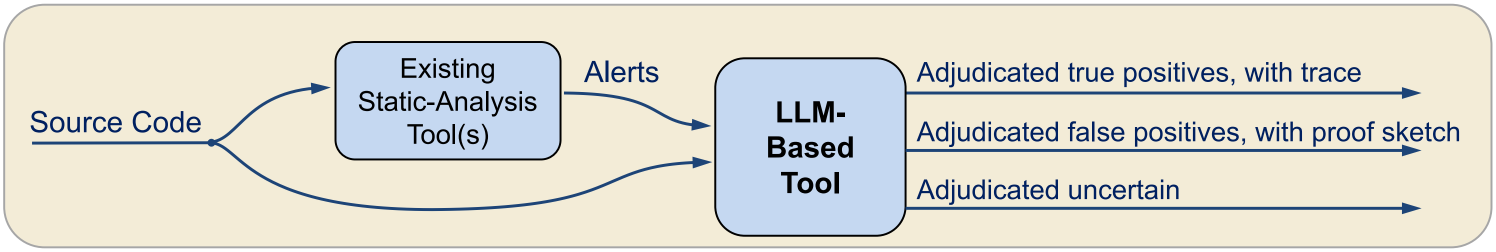

Source Code
Existing Static-Analysis Tool(s)
Alerts
LLM-Based Tool
Adjudicated true positives, with trace
Adjudicated false positives, with proof sketch
Adjudicated uncertain

Figure 1. *Using LLMs for SA Alert Adjudications.*

We developed a model of how an LLM-based tool could be used for SA alert adjudications, shown in **Figure 1.** The LLM-based tool ingests source code and SA alerts, then creates a query (a “prompt”) to the LLM for each alert. The LLM ideally outputs adjudicated true positives along with a trace, adjudicated false positives with a proof sketch, or it adjudicates as “uncertain.”

We developed partial automation to test this concept. A script inserts “// ALERT” to the code line that the SA alert identifies. The script creates an LLM prompt that includes the source code of the function that contains the alerted-about line, the type of code flaw to adjudicate (e.g., “integer overflow”), and additional data from the alert.

In this article, the GPT-4 links (meaning all the links that start with “https://chat.openai.com”) go to webpages that show tests that we conducted. They show the exact input that we provided to GPT-4. Each page also shows the full text of the responses from GPT-4, which often includes extensive step-by-step analysis of the code and the possible code flaw. We provide summaries and encourage those interested in additional detail to look at the full interactions shown at those links.

We note that GPT-4 is reliable at correctly following instructions to produce JSON (JavaScript Object Notation) output in a specified schema, making it relatively easy to write a script to parse the output from a GPT-4 API (application programming interface) call. In the rare case that it fails to produce output in the correct format, we simply try again until it produces output in the correct format.

Strategies for Mitigating Context-Window Limits

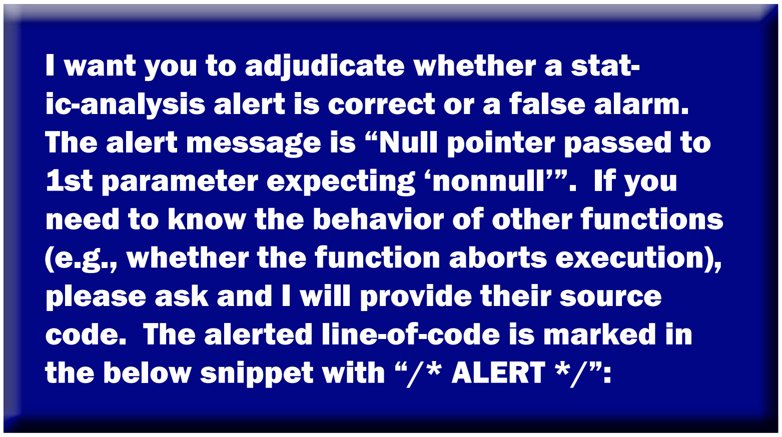
LLMs have a limited context window, which means that an LLM can usually ingest a single function but not an entire codebase. Sometimes, the LLM can make an adjudication based only on the function that contains the flagged line of code, but in other cases, additional context is needed. To overcome the context-window limit, we must summarize the relevant parts of the codebase enough so that the LLM can digest it.

Some strategies for this have been documented in the literature:

- Use traditional static analysis to produce required information, as in [17].
- Use the LLM itself to generate the function summaries, as in [14].

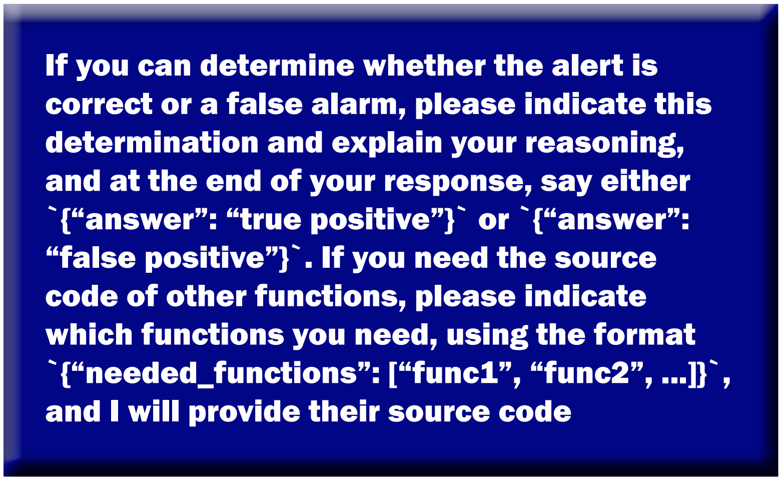
We have also tested a couple other strategies:

1. As part of the prompt, direct the LLM to ask for needed information. Our tool will then supply it to the LLM. Example: https://chat.openai.com/share/b01b0394-55f2-49f7-8a11-bfda15362297



I want you to adjudicate whether a static-analysis alert is correct or a false alarm. The alert message is “Null pointer passed to 1st parameter expecting ‘nonnull’”. If you need to know the behavior of other functions (e.g., whether the function aborts execution), please ask and I will provide their source code. The alerted line-of-code is marked in the below snippet with “/* ALERT */”:

Figure 2: *Start of prompt directs LLM to ask for needed information (source code not displayed here for brevity)*



If you can determine whether the alert is correct or a false alarm, please indicate this determination and explain your reasoning, and at the end of your response, say either `{“answer”: “true positive”}` or `{“answer”: “false positive”}`. If you need the source code of other functions, please indicate which functions you need, using the format `{“needed_functions”: [“func1”, “func2”, ...]}`, and I will provide their source code

Figure 3. *End of (same) prompt directs LLM to ask for other types of needed information.*

2. Use the LLM to generate preconditions for avoiding a bad state in a function with an alert, and then use the LLM to check whether the callers of the function satisfy the preconditions.

   a. Example of creating a precondition: https://chat.openai.com/share/cfeabe6f-5757-4c25-be82-f9569f8c9df2

      In this example, GPT-4 analyzes a function named “greet_user” that takes a string as an argument. GPT-4 is asked to adjudicate an alert about a buffer overflow. In its response, GPT-4 correctly determines that the buffer overflow can happen only if the length of the input string is too long. It returns a precondition for avoiding the buffer overflow:

      [{“precond”: “strlen(username) <= 52”, …}]

   b. Example of using a precondition: https://chat.openai.com/share/bbbf7df7-4fba-43b1-8f46-f09c4bd290cb

      In this example, GPT-4 analyzes a function that calls the “greet_user” function analyzed above. GPT-4 is given the precondition that it previously computed, and it is asked whether this precondition is satisfied. It correctly determines that the precondition can be violated.

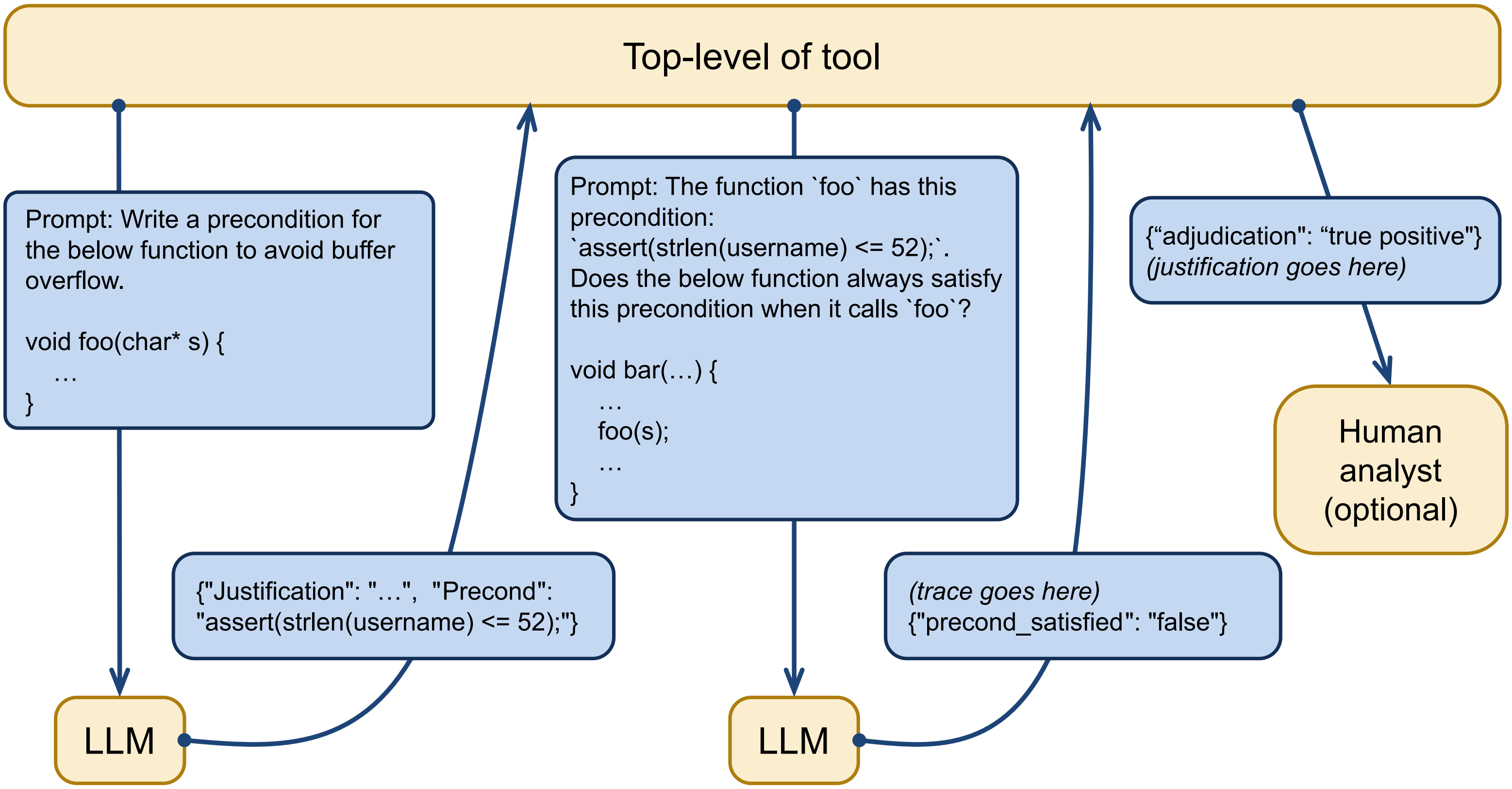

Top-level of tool
Prompt: Write a precondition for the below function to avoid buffer overflow.
void foo(char* s) {
…
}
{"Justification": "…", "Precond": "assert(strlen(username) <= 52);"}
LLM
Prompt: The function `foo` has this precondition: `assert(strlen(username) <= 52);`. Does the below function always satisfy this precondition when it calls `foo`?
void bar(…) {
…
foo(s);
…
}
(trace goes here)
{"precond_satisfied": "false"}
LLM
{"adjudication": "true positive"}
(justification goes here)
Human analyst (optional)

Figure 4. *Creating and Using Preconditions.*

Example: GPT-4 Adjudicating an Alert in the Linux Kernel

This example demonstrates GPT-4 successfully adjudicating an alert for vulnerability CVE-2022-41674 [15], about an integer-overflow leading to a buffer overflow in the Linux kernel: https://chat.openai.com/share/4ce0cdae-47b7-4648-9462-9e0a381ccc37

First, our script adds comments identifying two code locations that the alert specifies.

Next, we submit a prompt to GPT, which has a few sections:

III. The first part of the prompt is the following text:

   I want you to adjudicate whether a static-analysis alert is correct or a false alarm. The alert warns of a buffer overflow during `memcpy` on the line ending with “// ALERT-2” that happens if there is an integer overflow on the line ending with “// ALERT-1.”

IV. The middle part of the prompt consists of the source code of the alerted-about function.

V. The final part of the prompt is the following text:

   If you can determine whether the alert is correct or a false alarm, please indicate this determination and explain your reasoning, and at the end of your response, say either `{“answer”: “true positive”}` or `{“answer”: “false positive”}`. First identify whether integer overflow can happen. If it can’t, then report the alert is a false positive. If it can happen, then examine whether it can lead to a buffer overflow.

Step-by-step, GPT-4 determines the following, concluding that the alert is a true positive:

1. An integer overflow can happen on the line `cpy_len = mbssid[1] + 2; // ALERT-1` if `mbssid[1]` is equal to 255, since cpy_len is an unsigned 8-bit integer.
2. GPT-4 analyzes the relation between the allocated size of the `new_ie` buffer (into which `pos` points) and the amount being copied into it. It determines that a large

value of `mbssid[1]` should (and does) result in a small allocated buffer and should (but doesn’t) result in a small amount copied into the buffer. Due to the integer overflow, a large amount is actually copied into the small buffer, overflowing the buffer.

3. It then provides its final answer at the end of its response, in the format requested by the prompt:

   “`{“answer”: “true positive”}`.”

We’ve summarized GPT’s determinations, but it’s important to note that GPT-4 more fully states the basis for its reasoning prior to the determination and that a human analyst can verify its reasoning.

If asked about the patched version, GPT-4 correctly identifies that the vulnerability is no longer present: https://chat.openai.com/share/7ee8e60b-1fed-4b67-b77b-7edd289fee90

Example: GPT-4 Double-Checking its Work

LLMs sometimes make mistakes. One way to mitigate is to run the LLM multiple times and check that the answers aren’t inconsistent except on a small threshold percent of runs. Another mitigation is to ask the LLM to double-check its own work. This is an example of GPT-4 double-checking its work:

GPT-4 mistakenly claims that a program has a use-after-free bug, but it corrects itself when asked to double-check the reported trace: https://chat.openai.com/share/78200ee7-37e7-413d-8b2d-fd9051c28106

The first prompt provides the source code for the function and tells GPT-4 to:

> “Investigate whether there any use-after-free vulnerabilities in this C++ code, thinking step-by-step. Don’t worry about any other problems with code; focus exclusively on use-after-free bugs. If you find a use-after-free bug, provide a trace that demonstrates memory being used after being freed.”

You

GPT-4 responds by defining use-after-free vulnerabilities:

> “A use-after-free vulnerability occurs when a program continues to use a pointer after it has been freed.”

GPT-4

Then, it inspects code for pointers and allocation and freeing of memory for the pointers it finds, and provides its analysis and then steps of a trace it thinks ends with a use-after-free, ending with this summary:

> “The vulnerability arises due to the fact that the `R` command can potentially free memory that’s still being pointed to by entries in `sub_list`. To fix this, we should avoid using the reallocated memory from `getline()` for items that have been copied to the `sub_list` or ensure we update any relevant pointers in `sub_list` when renaming in `full_list`.”

GPT-4

The next prompt is:

You

> “OK, now please double-check that the trace is correct, stepping thru the code in detail.”

GPT-4 responds with a detailed step-by-step walkthrough, then ends with:

> “Upon a detailed step-by-step walkthrough, the trace provided previously appears to be incorrect. The use-after-free scenario doesn’t occur because when an item from full_list is copied to sub_list using the ‘C’ command, a duplicate of the string is created with strdup(), ensuring that sub_list entries are independent of any reallocations or frees that might occur with full_list entries.”

GPT-4

> ‘Therefore, there is no use-after-free vulnerability with the described trace in the provided code.”

GPT-4

Deleting the strdup call from the above example introduces a real use-after-free bug, and GPT-4 correctly identifies this use-after-free bug and correctly reports that it is a real bug when asked to double-check the reported trace in detail: https://chat.openai.com/share/36b75aaa-d52a-480c-8694-f8f62e64fdcc

GPT-4’s detailed double-check ends with:

> “After a detailed step-through, the trace does hold. The use-after-free vulnerability indeed exists, and the trace correctly demonstrates it.”

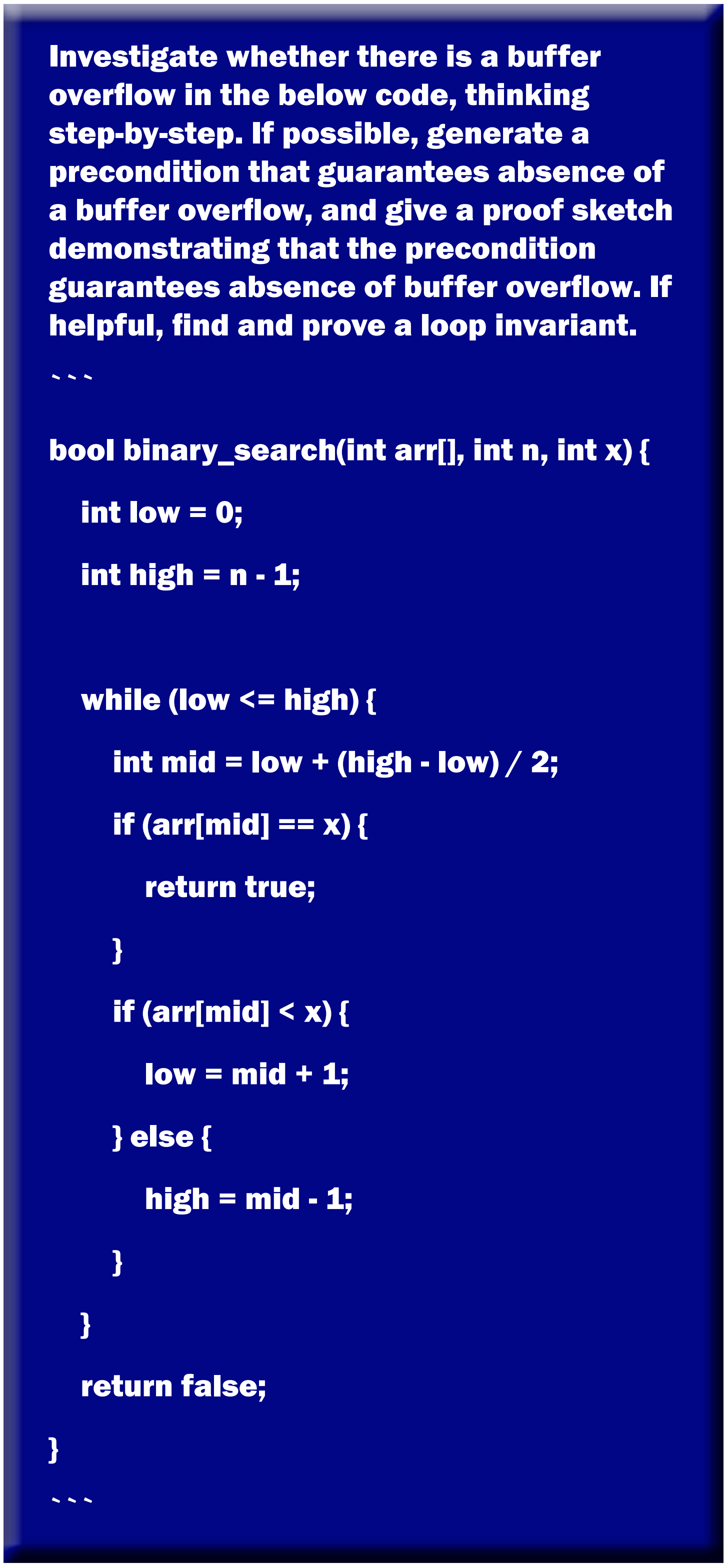
GPT-4

Example: GPT-4 Finding and Proving Loop Invariants

GPT-4 has some ability to find and prove loop invariants.

- Example: https://chat.openai.com/share/88c782ff-c7b7-4d4c-8cb1-988df4a3f1a8

The full initial prompt is shown in **Figure 5**, including source code for the function. GPT-4 steps through the code, finds the loop invariant (“0 <= low <= high < n”), and then sketches a proof of the loop invariant. A second prompt, “Please analyze the Maintenance step in more detail,” elicits additional detail from GPT-4 which supports its previous adjudication. That GPT-4 response ends with, “In both scenarios, whether we’re updating low or high, we can see that our loop invariant 0 <= low <= high < n is maintained. Thus, during any given iteration of the loop, the index mid always falls within the safe range [0, n-1], ensuring there is no buffer overflow.



Investigate whether there is a buffer overflow in the below code, thinking step-by-step. If possible, generate a precondition that guarantees absence of a buffer overflow, and give a proof sketch demonstrating that the precondition guarantees absence of buffer overflow. If helpful, find and prove a loop invariant.

~~~
```

bool binary_search(int arr[], int n, int x) {

    int low = 0;

    int high = n - 1;


    while (low <= high) {

        int mid = low + (high - low) / 2;

        if (arr[mid] == x) {

            return true;

        }

        if (arr[mid] < x) {

            low = mid + 1;

        } else {

            high = mid - 1;

        }

    }

    return false;

}

```
~~~

Figure 5. *Prompt for No-Buffer-Overflow Proof.*

LLM SA Alert Adjudication Can Also Be Used for Automated Code Repair and Shifting Security Left

LLM Adjudications can be Combined with Automated Code Repair

SA alert adjudication by LLMs can be used in combination with automated repair technology (either LLM-based or using an algorithmic approach [20][19]). This combination can be used to greatly speed up the process of hardening a system against cyberattacks.

Shifting Security Left

The need for improved static analysis effectiveness also applies to DevSecOps [21], which is a set of practices that combines code development and IT operations (DevOps) with security integrated throughout.

A continuous authorization to operate (cATO) formalizes and monitors specific technologies to reduce risk [23][24]. Such evaluation is done as part of ongoing DevSecOps, if there is a cATO. Integrated developer environments (IDEs) can do static analysis for some flaws while developers are coding [25], plus static analysis can be run as part of continuous integration testing.

Static Analysis vs. Dynamic Analysis

Dynamic analysis executes the code, running it in particular test environments with a set of test inputs. Two examples: Fuzz testing automatically injects inputs to try to reveal defects, monitoring for negative effects such as crashes and memory leaks, with black-box, grey-box, and white-box categories of fuzzers having different knowledge about the software [26]. Dynamic taint analysis inspects data sources and sinks during execution to identify data flows that should not happen, such as leaking sensitive data to a remote web address [27].

Static analysis analyzes code without executing it, using techniques that often include automated parsing of the code into a grammar and abstract syntax tree; creating a control flow graph; and analyzing data flow, control flow, and/or type flow to inspect for code flaws that could lead to security or functional problems [28][29].

Formal verification is a mathematical approach to check whether a software system meets formal specifications. Formal verification is a type of a static analysis, although the term "static analysis" usually connotes a less rigorous type of static analysis. There are several techniques used in formal verification, including model checking, theorem proving, abstract interpretation, and equivalence checking. Formal verification has been successfully used on small software systems (e.g., the seL4 microkernel), but it often is impractical for large software systems.

Previous Work with AI/ML for Static Analysis

To date, there has been a significant amount of research on using machine learning to aid in efficiently identifying source code flaws [30][31][32][33]. Researchers trained the ML using manually-adjudicated alerts ("labeled data") and features such as code cohesiveness metrics, lines of code per function and file, developer ID, and recency-of-code edits around that code location. Some work found that aggregating alerts from multiple SA tools for the same code location improved classifier precision [34], and other work developed a lexicon and adjudication rules to enable consistent adjudications to improve classifier training data [35]. Classifiers trained on labeled data from the same codebase generally perform better than those trained on data from different codebases (latter is called "cross-project prediction"), but techniques have been developed that improve cross-project prediction [36]. The high cost and time required for experts to manually adjudicate and create enough labeled data can be a barrier to creation of accurate static analysis classifiers. Flynn et al. made novel use of test suites [37][38] to create large datasets of labeled (true/false) SA alerts (augmenting ~7500 manually-adjudicated alerts on natural code). This resulted in high-precision ML classifiers for a larger set of CWE types than the natural dataset alone [39][40] and created a framework for use with multiple SA tools, ML classification, adaptive heuristics, labeled datasets, and test suites [41][42]. Gallagher et al. also used ML for finding code flaws but did that using LLVM intermediate representation instead of source [43] (LLVM is a set of compiler and toolchain technologies [44]). Flynn and Gallagher both found that artificial code and flaw injection can cause classifiers to use features not helpful with natural code.

As part of modern continuous integration (CI) code development, automatically cascading adjudications (manual adjudications true or false for alerts) to SA alerts for a later version of the codebase is important, but there are tradeoffs between fast `diff`-based cascading methods and higher-precision methods that may be too slow for practical CI use [45]. Classifier accuracy depends on the quality of the labeled data the classifier is created with, so incorrectly cascaded alert adjudications can be expected to produce lower accuracy in the resulting classifier [46].

LLMs for Education on Alert Adjudication, Coding Standards, and Flaw Taxonomies

Another potential benefit of using LLMs for static analysis adjudication involves education. Static analysis alert adjudication skills require understanding the coding standard for the programming language, understanding the code flaw taxonomy (such as CWE or CERT Coding Rules [47]), being able to grasp what is happening in the code, and understanding the static analysis tool's alert messages and checkers. A human analyst can try to follow the LLM's rationale, validating its claims by checking the code, language standard, and taxonomy item. Regardless of correctness of the LLM's rationale, the person can learn by following the reasoning and inspecting the source documents (code, standard, and taxonomy). A developer using an LLM to adjudicate SA alerts on their own code might learn to avoid inserting flawed code constructs, understanding why and how by following the LLM's rationale for a true positive adjudication. Early-stage research found that novice and expert programmers had significantly different needs for guidance, personalization, and integration of an LLM (e.g., with less confidence in their own capabilities, a novice may place too much trust in the LLM's output, while expert programmers reported more learning through use of LLMs). The research subjects lacked diversity (most highly educated and male), so further research is needed to determine how generalizable those findings are [48].

Future Directions

The 2024 Cyber Developmental Test and Evaluation policy and guidance specifies a responsibility to "Identify, assess, and document potential weaknesses that could affect technical, functional, and operational performance" [49]. Use of LLMs to support efficient and correct static analysis adjudication shows potential to improve performance on that mission-impacting responsibility.

Further work is needed to validate and widen the impact of previous results discussed in this article, by applying the methods to a wide range of DoD codebases and more types of code flaws and static analysis tools.

Looking to the future, LLMs may greatly help to enable formal verification of software, an area that has long been impractical for large codebases due to the amount of manual effort involved. Wu et al [50] report success in using LLMs for generating formal proofs on a benchmark of formal-verification problems [51], beating state-of-the-art formal-verification tools on a number of hard cases. Generating and proving loop invariants and function pre-/post-conditions is often a crucial and challenging part of formal verification, and as evidenced by our initial experimental results, LLMs show promise for helping with this task.

Acknowledgments

Carnegie Mellon University 2024. 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. References herein to any specific entity, product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by Carnegie Mellon University or its Software Engineering Institute nor of Carnegie Mellon University - Software Engineering Institute by any such named or represented entity.

[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.

CERT® and Carnegie Mellon® are registered in the U.S. Patent and Trademark Office by Carnegie Mellon University. DM24-0303

Image behind title was generated with AI. Output from a prompt by Wolfhound911. Adobe Stock. June 27, 2023. Source code displayed from git https://github.com/git/git/blob/master/commit-graph.c

References

[1] Dukes, Curtis W. “Committee on National Security Systems (CNSS) Glossary, CNSSI No. 4009.” DoD Ft Meade (2015).

[2] Synopsis. “Coverity DISA ASD STIG report”, Synopsis webpage. https://sig-product-docs.synopsys.com/bundle/coverity-docs/page/reports/disa-stig/coverity_disa_stig.html. Accessed 5 March 2024.

[3] MITRE, “Common Attack Pattern Enumeration and Classification (CAPEC)”, https://capec.mitre.org/.

[4] MITRE, “CAPEC-124: Shared Resource Manipulation”, https://capec.mitre.org/data/definitions/124.html

[5] MITRE, CWE-1331: Improper Isolation of Shared Resources in Network On Chip (NoC)”, https://cwe.mitre.org/data/definitions/1331.html

[6] “Ariane flight V88.” Wikipedia, The Free Encyclopedia, 8 Feb. 2024. https://en.wikipedia.org/w/index.php?title=Ariane_flight_V88&oldid=1204984683.

[7] Lions, Jacques-Louis. “Flight 501 failure.” Report by the Inquiry Board 190 (1996).

[8] OpenAI, “Research: GPT-4”, OpenAI contributors, published 14 March 2023, https://openai.com/research/gpt-4

[9] Flynn, Lori. “Challenges and Progress: Automating Static Analysis Alert Handling with Machine Learning.” (2018). https://insights.sei.cmu.edu/documents/4174/2018_017_101_518025.pdf

[10] Chan, Aaron, et al. “Transformer-based vulnerability detection in code at EditTime: Zero-shot, few-shot, or fine-tuning?.” arXiv preprint arXiv:2306.01754 (2023).

[11] Fan, Gang, et al. “Static Code Analysis in the AI Era: An In-depth Exploration of the Concept, Function, and Potential of Intelligent Code Analysis Agents.” arXiv preprint arXiv:2310.08837 (2023).

[12] Yang, Aidan ZH, et al. “Large language models for test-free fault localization.” Proceedings of the 46th IEEE/ACM International Conference on Software Engineering. 2024.

[13] Li, Haonan, et al. “Assisting static analysis with large language models: A chatgpt experiment.” Proceedings of the 31st ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering. 2023.

[14] Li, Haonan, et al. “The Hitchhiker’s Guide to Program Analysis: A Journey with Large Language Models.” arXiv preprint arXiv:2308.00245 (2023).

[15] National Vulnerability Database contributors. “CVE-2022-41674 Detail”, National Vulnerability Database (NVD), National Institute of Standards and Technology. https://nvd.nist.gov/vuln/detail/CVE-2022-41674 accessed 14 March 2024.

[16] Sherman, Mark. “Should I Trust ChatGPT To Review My Program?”, InfoSec World, (2023).

[17] Ahmed, Toufique, et al. “Improving few-shot prompts with relevant static analysis products.” arXiv preprint arXiv:2304.06815 (2023).

[18] SEI CERT Coding Rules Contributors, “INT32-C. Ensure that operations on signed integers do not result in overflow“, SEI CERT Coding Standard for C, Software Engineering Institute of Carnegie Mellon University. https://wiki.sei.cmu.edu/confluence/display/c/INT32-C.+Ensure+that+operations+on+signed+integers+do+not+result+in+overflow accessed 14 March 2024.

[19] CWE Contributors. “CWE-190: Integer Overflow or Wraparound”, MITRE Common Weakness Enumeration. https://cwe.mitre.org/data/definitions/190.html accessed 14 March 2024.

[20] Monperrus, Martin. “The Living Review on Automated Program Repair”, Technical Report HAL # hal-01956501, 2018. https://hal.science/hal-01956501/document/.

[21] Nichols, W. R., Yasar, H., Antunes, L., Miller, C. L., & McCarthy, R. (2022). Automated data for DevSecOps programs. Acquisition Research Program.

[22] Klieber, William, et al. “Automated code repair to ensure spatial memory safety.” 2021 IEEE/ACM International Workshop on Automated Program Repair (APR). IEEE, 2021.

[23] McKeown, David W. “MEMORANDUM FOR SENIOR PENTAGON LEADERSHIP DEFENSE AGENCY AND DOD FIELD ACTIVITY DIRECTORS.” Subject: Continuous Authorization To Operate (cATO). DoD Senior Information Security Officer (SISO). 3 Feb. 2022, https://media.defense.gov/2022/Feb/03/2002932852/-1/-1/0/CONTINUOUS-AUTHORIZATION-TO-OPERATE.PDF

[24] Lam, T., and N. Chaillan. “DoD Enterprise DevSecOps Reference Design.” Department of Defense, Chief Information Officer Library, 12 Aug. 2019. https://dodcio.defense.gov/Portals/0/Documents/DoD%20Enterprise%20DevSecOps%20Reference%20Design%20v1.0_Public%20Release.pdf

[25] Microsoft developers, Visual Studio Code, Microsoft. https://code.visualstudio.com accessed 15 March 2024.

[26] Lemieux, Caroline, and Koushik Sen. “Fairfuzz: A targeted mutation strategy for increasing greybox fuzz testing coverage.” Proceedings of the 33rd ACM/IEEE international conference on automated software engineering. 2018.

[27] Newsome, James, and Dawn Xiaodong Song. “Dynamic Taint Analysis for Automatic Detection, Analysis, and Signature Generation of Exploits on Commodity Software.” NDSS. Vol. 5. 2005.

[28] Harrold, Mary Jean, and Mary Lou Soffa. “Interprocedual data flow testing.” ACM SIGSOFT software engineering notes 14.8 (1989): 158-167.

[29] Horwitz, Susan, Thomas Reps, and David Binkley. “Interprocedural slicing using dependence graphs.” ACM Transactions on Programming Languages and Systems (TOPLAS) 12.1 (1990): 26-60.

[30] Heckman, Sarah, and Laurie Williams. “A systematic literature review of actionable alert identification techniques for automated static code analysis.” Information and Software Technology 53.4 (2011): 363-387.

[31] Kremenek, Ted, et al. “Correlation exploitation in error ranking.” ACM SIGSOFT Software Engineering Notes 29.6 (2004): 83-93.

[32] Flynn, Lori, et al. “Static Analysis Classification Research FY16 20: for Software Assurance Community of Practice.” (2020): 68.

[33] Ruthruff, Joseph R., et al. “Predicting accurate and actionable static analysis warnings: an experimental approach.” Proceedings of the 30th international conference on Software engineering. ACM, 2008.

[34] Flynn, Lori. “Prioritizing Alerts from Static Analysis with Classification Models”, Software Engineering Institute, Research Review 2016, 1 Nov. 2016. https://insights.sei.cmu.edu/library/prioritizing-alerts-from-static-analysis-with-classification-models-2/. Accessed 5 March 2024.

[35] Svoboda, David, Lori Flynn, and Will Snavely. “Static analysis alert audits: Lexicon & rules.” 2016 IEEE Cybersecurity Development (SecDev). IEEE, 2016.

[36] Wang, Song, Taiyue Liu, and Lin Tan. “Automatically learning semantic features for defect prediction.” Proceedings of the 38th International Conference on Software Engineering. 2016.

[37] Intelligence Advanced Research Projects Activity (IARPA) contributors. STONESOUP”, National Institute of Standards and Technology (NIST) Software Assurance Reference Dataset (SARD). https://samate.nist.gov/SARD/documentation#iarpa. Accessed 5 March 2024.

[38] Black, Paul E. Juliet 1.3 test suite: Changes from 1.2. US Department of Commerce, National Institute of Standards and Technology, 2018. https://nvlpubs.nist.gov/nistpubs/TechnicalNotes/NIST.TN.1995.pdf accessed 5 March 2024.

[39] Flynn, Lori, et al. “Prioritizing alerts from multiple static analysis tools, using classification models.” Proceedings of the 1st international workshop on software qualities and their dependencies. 2018.

[40] Flynn, Lori, William Snavely, and Zachary Kurtz. “Test suites as a source of training data for static analysis alert classifiers.” 2021 IEEE/ACM International Conference on Automation of Software Test (AST). IEEE, 2021.

[41] Flynn, Lori, et al. “SCAIFE API (OpenAPI v.3).” Software Engineering Institute site at GitHub, https://github.com/cmu-sei/SCAIFE-API/. Accessed 5 March 2024.

[42] Flynn, Lori, et al. “SCALe Repository (scaife-scale branch).” Software Engineering Institute site at GitHub, https://github.com/cmu-sei/SCALe/tree/scaife-scale Accessed 5 March 2024.

[43] Gallagher, Shannon K., et al. “LLVM intermediate representation for code weakness identifica-

tion." (2022). https://apps.dtic.mil/sti/trecms/pdf/AD1178536.pdf accessed 5 March 2024.

[44] "The LLVM Compiler Infrastructure". https://llvm.org/ accessed 5 April 2024.

[45] Guo, Xiuyuan, et al. "A Study of Static Warning Cascading Tools (Experience Paper)." arXiv preprint arXiv:2305.02515 (2023).

[46] Flynn, Lori. "Rapid Adjudication of Static Analysis Alerts During CI.", Research Review Presentation, Software Engineering Institute at Carnegie Mellon University, (2020): 35.

[47] Carnegie Mellon University Software Engineering Institute CERT Division and community contributors. "SEI CERT Coding Standards". https://wiki.sei.cmu.edu/confluence/display/seccode/SEI+CERT+Coding+Standards accessed 25 April 2024.

[48] Chen, John, et al. "Learning Agent-based Modeling with LLM Companions: Experiences of Novices and Experts Using ChatGPT & NetLogo Chat." arXiv preprint arXiv:2401.17163 (2024).

[49] Standard, Sarah. "DoD Cyber Developmental Test and Evaluation Policy and Guidance Policy and Guidance", DAU Webinar, 01 Feb. 2024. https://www.dau.edu/events/cyber-dte Accessed 04 March 2024.

[50] Wu, Haoze, Clark Barrett, and Nina Narodytska. "Lemur: Integrating Large Language Models in Automated Program Verification." arXiv preprint arXiv:2310.04870 (2023).

[51] SV-COMP contributors. "SV-Benchmarks", Competition on Software Verification (SV-COMP) benchmarks, International Competition on Software Verification. https://gitlab.com/sosy-lab/benchmarking/sv-benchmarks. Accessed 3 March 2024.

About the Authors

Dr. Lori Flynn is a senior software security researcher at the Software Engineering Institute at Carnegie Mellon University. Flynn's research topics include automated static analysis of software and cybersecurity. She led a series of research projects involving machine learning for static analysis alert adjudication. Other research involved Android taint flow analysis across multiple apps. She co-authored a patented method to create signatures for polymorphic programs. Currently she works on an automated code repair project.

Dr. Lori Flynn

Senior Software Security Researcher

Software Engineering Institute

lflynn@sei.cmu.edu

Dr. Will Klieber is a software security researcher at the Software Engineering Institute (SEI). He has recently worked on static analysis and program transformation to automatically repair certain classes of memory-safety vulnerabilities in C source code. He is now working on automated techniques for increasing software assurance of binary code.

Dr. Will Klieber

Software Security Researcher

Software Engineering Institute

weklieber@sei.cmu.edu

76TH SOFTWARE ENGINEERING GROUP

The 76th Software Engineering Group is an extremely diverse group of approximately 1,600 scientists, engineers, and Cyber/IT professionals, whose mission is to provide the best value engineering solutions to our customers, enabling global airpower for our Air Force warfighter. Connect with us on social media!

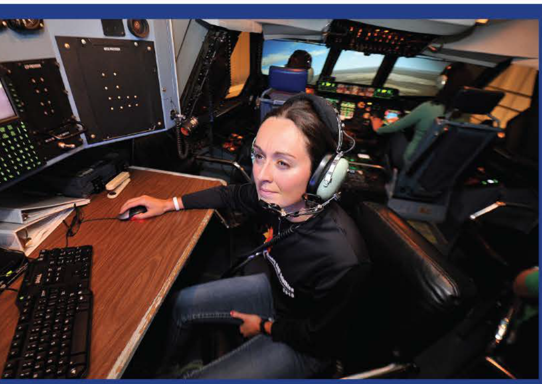
DECEMBER 10-13, 2024
SALT PALACE CONVENTION CENTER
SALT LAKE CITY, UT

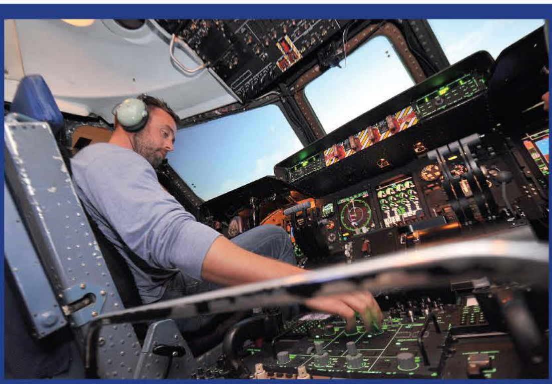
SOFTWARE SUMMIT PROGRAM

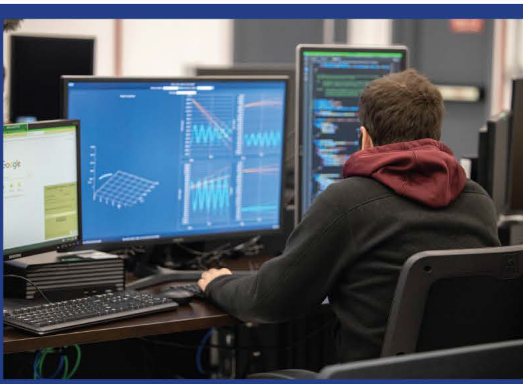
The 3rd annual DoD Weapon Systems Software Summit stakeholders including the defense industrial base, academia, Office of the Secretary of Defense, and the Services will share solutions to software issues in the tracks and panels below:

- DevSecOps Ecosystems
- Software Assurance
- Software Modernization
- Securing the Embedded Software Supply Chain
- Secure Software by Design
- OSD and Services Policy for Weapon Systems Software
- AI/ML – Does It Apply to Weapon Systems
- Agile Coaches Corner Panel
- Workforce Innovations Panel
- PMOS Working with Organic Software Teams Panel

Program details for the Software Summit will be available soon at the QR code below.
Book your hotel now at the QR code below.
Register using the code **24SFTWR** for the Software Summit discount.







Contact us: AFSC.DoD.WS_SWSummit@us.af.mil

SAE.ORG/DOD

UNDERSTANDING TODAY'S ARTIFICIAL INTELLIGENCE FOR GOVERNMENT USE

Rebel Brown
CEO,
Cognoscenti, Inc.

Introduction

The exploration of Artificial Intelligence (AI) dates to the mid-20th century. The Dartmouth Conference of 1956 is often cited as the birthplace of AI, where the term itself was coined and its foundational goal was set to mimic human cognitive functions through machines. Pioneers like Alan Turing had already sown the seeds of computational thinking, questioning whether machines could think.

Early AI research in the 1960s and 1970s focused on problem-solving and symbolic methods.

- The first expert systems created were designed to emulate the decision-making abilities of a human expert.
- By the 1980s, the field saw the advent of machine learning (ML), where the emphasis shifted from programming computers to perform certain tasks to training them to learn from data.
- The development of the Internet in the 1990s and the exponential increase in data availability led to significant advancements in AI algorithms and the rise of big data analytics.

The 21st century has seen AI further mature through deep learning and neural networks, inspired by the human brain's structure and function. These advances have fueled contemporary AI's capabilities, including image and speech recognition, which have been widely applied across the U.S. government. In this realm, AI has been instrumental in areas such as national security, healthcare, and public administration, transforming data into actionable insights, automating routine tasks, and enhancing the efficacy of services provided to the public.

As AI continues to evolve, it is increasingly intertwined with areas like quantum computing and

Author's note: Artificial intelligence was used to write this article, including a highly trained version of ChatGPT, Bard, and Bing AI. For those who believe that AI can now write content, I'll share my experiences here and with other content. While some AIs can be helpful with researching content, its ability to actually write is limited. It saved approximately 20% of time in writing this article versus others I've written for this publication. Its issues include mistakes in understanding and applying content as well as ever-present hallucinations - random and unknown sources which are not real and provide false information and conclusions. As I recommend to my clients: with today's AIs, it's best to only use AI on subject matter that you know, as a tool to enhance your research and knowledge. Trusting it to develop content on a topic you don't know well is risky at best.

explainable AI, pushing the boundaries of technology and raising both opportunities and ethical considerations for society and governance. This journey has transformed the technological landscape, as it also reshaped the way government agencies operate and interact with citizens.

We all know the long and winding road that AI has taken, in search of processing power and infrastructures, as well as use cases, to move into the mainstream. Today, we are experiencing the advancement of that journey for both government and commercial leaders.

The State of AI Today: Defining Modern Artificial Intelligence

Artificial Intelligence enables machines to perform tasks that typically require human intelligence. AI is defined by its capabilities to learn from data, adapt to new situations, make decisions, and solve problems.

Modern AI systems range from narrow, task-specific applications like voice assistants, to complex networks capable of diagnosing diseases or piloting drones without human intervention.

AI’s role in government extends to various areas where decision-making is crucial.

- In military applications, predictive analytics leverage AI to forecast conflict scenarios, optimize logistics, and maintain strategic advantages.
- Public safety utilizes AI for crowd monitoring and emergency dispatch optimization, ensuring swift responses to critical situations.
- Military operations leverage AI to analyze and predict a variety of strategic and tactical operations.

Yet AI is not just one technology or application. In fact, today’s AI has many facets, from Generative AI (Gen AI) to machine learning to computational AI and more. Let’s look at these various AI technologies, from their function to their application in the government.

Generative AI: Crafting Data-Driven Government Solutions

Generative AI enables machines to create new, original content or data that can mirror human-like creation.

It leverages complex algorithms and frameworks such as Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs,) known for their ability to generate new data samples that resemble the training data.

Generative Adversarial Networks (GANs)

GANs consist of two neural networks—the generator and the discriminator—that are trained simultaneously through a competitive process. The generator aims to produce data that is indistinguishable from real data.

> The discriminator evaluates the generated data against real data, trying to distinguish between the two.
>
> This process continues until the generator produces data so close to the real data that the discriminator cannot differentiate.

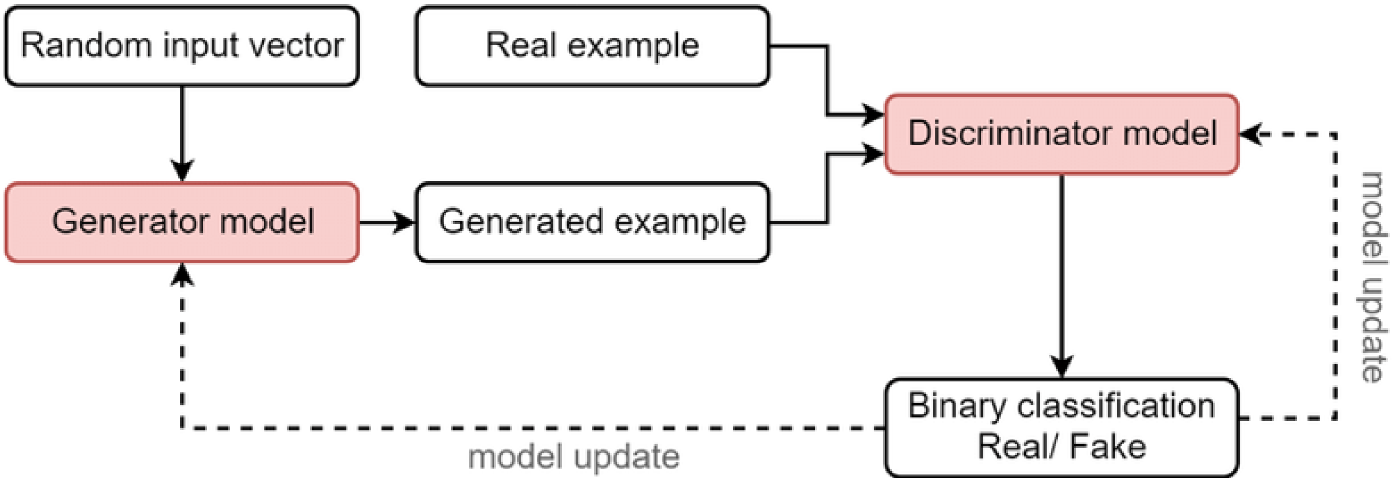


Figure 1. *GAN Workflow [A].*

GANs are known for their ability to generate high-quality, realistic images. The training of GANs involves alternating between training the discriminator and the generator, with each having its own loss function.

The challenge in training GANs lies in achieving convergence, where the discriminator's feedback becomes less meaningful over time as it gets harder to distinguish real from generated examples.

GANs have been utilized in various applications, including synthetic data generation, image style transfer, and even audio style transfer.

Variational Autoencoders (VAEs)

VAEs (as seen in **Figure 2**) are deep generative models that learn a low-dimensional representation (latent space) of high-dimensional data.

An encoder maps input data into a latent representation.

A decoder reconstructs the input data from this latent representation.

VAEs are trained to minimize the reconstruction error and a regularization term that ensures the latent space has good properties, allowing for the generation of new data points by sampling from the latent space.

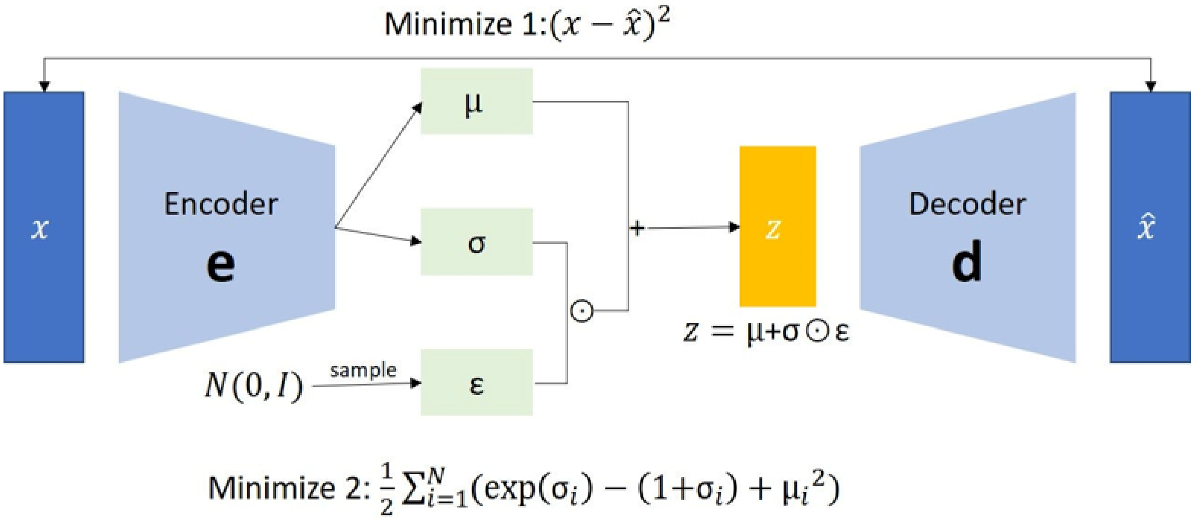

Minimize 1:$(x-\hat{x})^2$
x
Encoder
e
μ
σ
ε
$N(0,I)$
sample
+
⊙
z
$z = \mu + \sigma \odot \varepsilon$
Decoder
d
$\hat{x}$
Minimize 2: $\frac{1}{2}\sum_{i=1}^{N}(\exp(\sigma_i) - (1+\sigma_i) + {\mu_i}^2)$

Figure 2. *VAE workflow [B].*

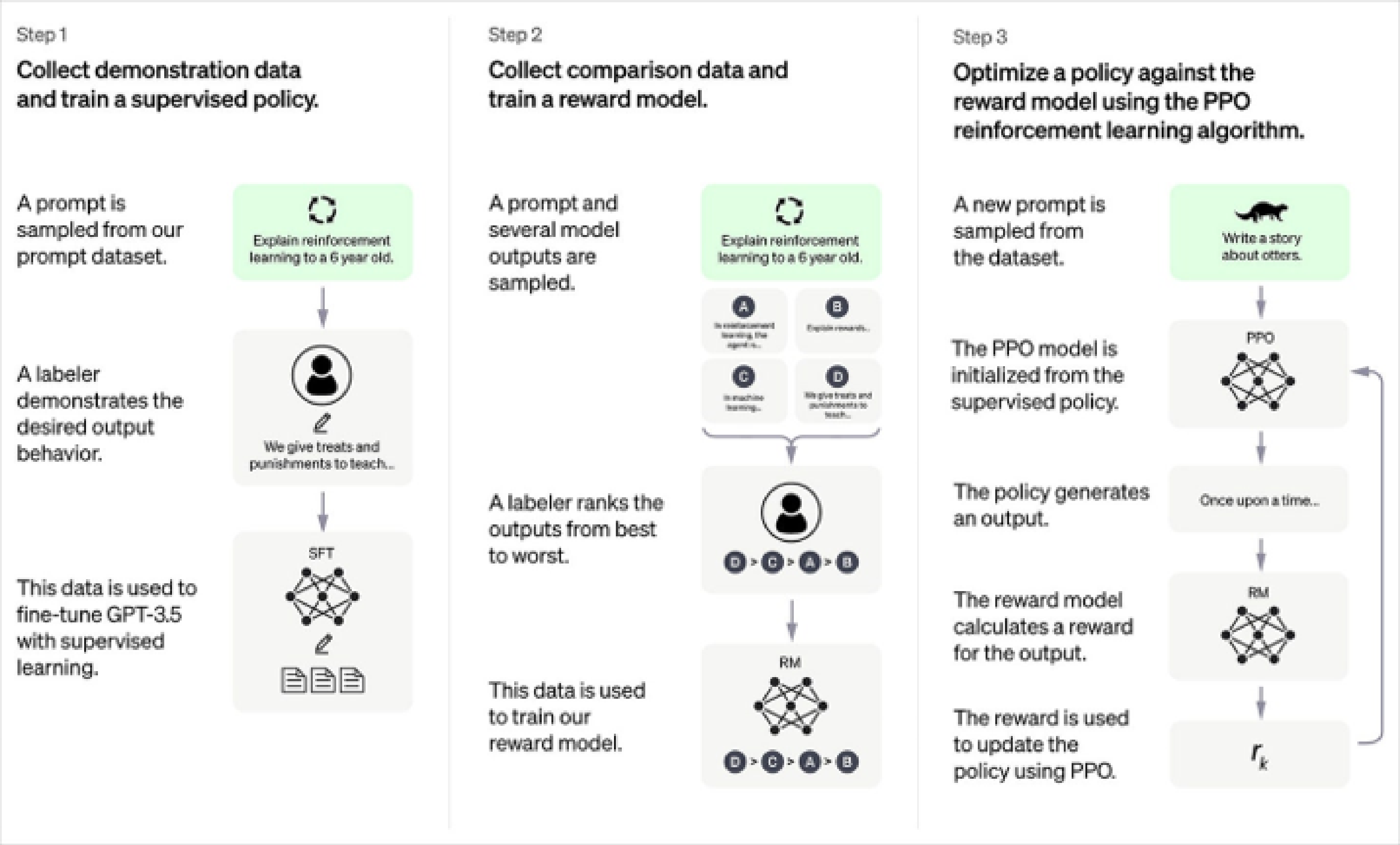

Step 1
Collect demonstration data and train a supervised policy.
A prompt is sampled from our prompt dataset.
Explain reinforcement learning to a 6 year old.
A labeler demonstrates the desired output behavior.
We give treats and punishments to teach...
This data is used to fine-tune GPT-3.5 with supervised learning.
SFT
Step 2
Collect comparison data and train a reward model.
A prompt and several model outputs are sampled.
Explain reinforcement learning to a 6 year old.
A
In reinforcement learning, the agent is...
B
Explain rewards...
C
In machine learning...
D
We give treats and punishments to teach...
A labeler ranks the outputs from best to worst.
D > C > A > B
This data is used to train our reward model.
RM
D > C > A > B
Step 3
Optimize a policy against the reward model using the PPO reinforcement learning algorithm.
A new prompt is sampled from the dataset.
Write a story about otters.
The PPO model is initialized from the supervised policy.
PPO
The policy generates an output.
Once upon a time...
The reward model calculates a reward for the output.
RM
The reward is used to update the policy using PPO.
r_k

Figure 3. *LLM Process [C].*

The Generative AI Stack

The technology stack for Generative AI typically includes a training dataset, a generative model, and an evaluation component.

The generative model, which is the core of Generative AI, uses deep learning techniques to produce new content. These models are trained on massive datasets, learning the underlying patterns and distributions of the data.

Large Language Models (LLMs), like GPT (Generative Pre-trained Transformer) and BERT (Bidirectional Encoder Representations from Transformers), are specific instances of Generative AI that have been pre-trained on extensive corpuses of text from the internet. They can generate coherent and contextually relevant text, complete sentences, or even entire articles based on a given prompt.

LLMs work by analyzing the relationships between words and phrases within the input data to predict what comes next in a sentence or generate a response to a query. They are composed of layers of neural networks that process input data in a manner structurally inspired by the human brain, though the actual processing is mathematical and involves weights and biases adjusted during the training phase.

In a government context, these technologies are used to automate and enhance content creation for reports and briefings. They can process vast amounts of governmental data, understanding and generating documentation in a style consistent with government communication norms.

The integration of Generative AI with other AI technologies in government communications can lead to more effective public engagement. For example, LLMs can be used to enhance Natural Language Processing (NLP) capabilities, allowing for clearer communication with citizens through automated systems. These systems can provide real-time responses to public inquiries or disseminate information during emergencies.

Generative AI and LLMs are still developing fields, and their integration into government technology stacks is an ongoing process that will likely continue to evolve and mature.

As these models become more advanced, they offer the potential for increasingly sophisticated applications that can significantly impact the efficiency and effectiveness of government operations.

Machine Learning: Data-Driven Innovation in Government Operations

Machine learning, a field at the forefront of artificial intelligence, has become integral to how government agencies process and derive insights from data.

It uses sophisticated algorithms that identify patterns within datasets, enabling predictive analytics and automating decision-making. These capabilities extend beyond traditional data processing, affording agencies the power to anticipate and act upon a variety of societal, economic, and environmental factors with a new level of precision.

At the core of machine learning technology is an array of algorithms, each designed for specific tasks. For instance, linear regression models may be utilized for economic forecasting, while more intricate neural networks are deployed for image recognition in security applications.

The government leverages this technology to analyze everything from traffic patterns for infrastructure planning to satellite imagery for environmental monitoring.

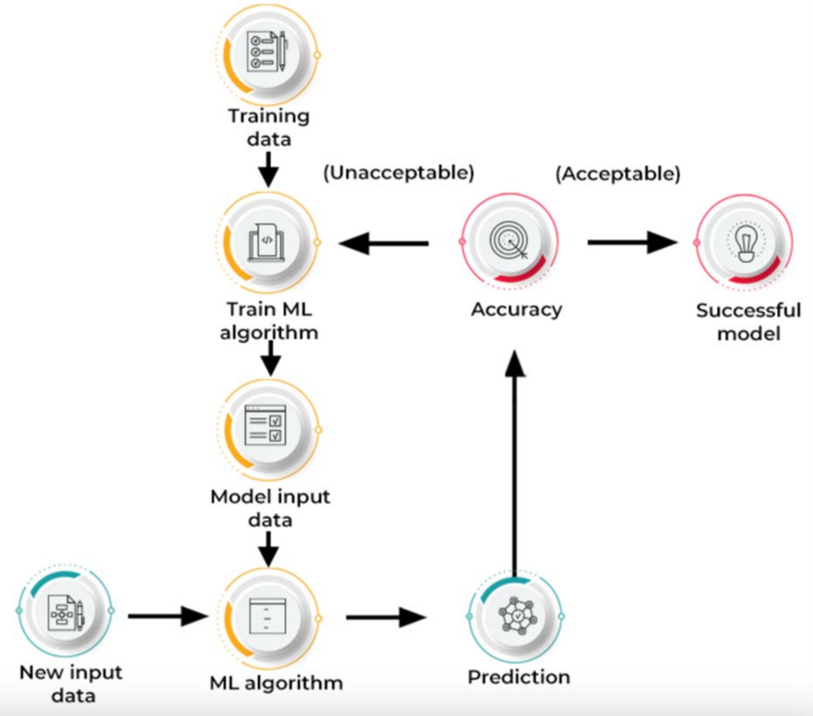


Figure 4. *Machine Learning Process [D].*

The process of machine learning is meticulous and methodical.

- It begins with the curation and preparation of data—a crucial step that ensures the accuracy of insights derived. The better the data, the better the results.
- The training phase involves exposing the algorithms to historical data, where they learn to make predictions and/or classify data points.
- Training is followed by validation and testing to confirm that the models perform well on new, previously unseen data.
- After validation, models are deployed into production environments, where they support real-world decisions and strategies.

Use cases for machine learning are prevalent across the government. For example:

- In the realm of public health, machine learning models ingest large amounts of health data to identify potential disease outbreaks. This enables agencies to allocate medical resources where they are most needed.
- Financial regulatory bodies can harness algorithms to detect irregularities within transactions, further safeguarding the integrity of financial systems.
- National security agencies can use machine learning to sift through complex and vast communications to identify and preempt potential threats.

Machine learning's potential is expanded when it is integrated with other technologies. For example, the Internet of Things (IoT) provides a network of physical devices that collect and exchange data. Machine learning systems analyze such data to optimize everything from energy use in public buildings to traffic flow on busy city streets.

In tandem with natural language processing, machine learning enhances the government's ability to interact with citizens, whether through automated customer service systems or by analyzing public sentiment on social issues.

The value of machine learning is not just in automating tasks or processing information, though. Its key values rest in its ability to create a more responsive, agile, and informed government, truly leveraging the vast data that we collect.

Computational AI: Spearheading Problem-Solving in Government Operations

Computational AI within the government refers to the advanced problem-solving capabilities that AI brings to complex challenges.

Computational AI stands out in its capacity to navigate and resolve intricate computational tasks, often involving vast datasets and multifaceted criteria that traditional problem-solving approaches cannot address.

The term 'computational' in Computational AI pertains to the system's ability to perform large-scale calculations, optimize processes, and simulate potential outcomes with a high degree of accuracy. These processes leverage sophisticated algorithms designed to manage, analyze, and derive actionable intelligence from volumes of government data, whether it be census figures, economic indicators, or security footage.

For instance, in tackling urban infrastructure challenges, Computational AI systems can simulate traffic flow across different scenarios, feeding dynamic traffic management systems that adapt to real-time conditions. These systems can mitigate congestion, reduce emissions, and save time for commuters.

Similarly, in disaster response, Computational AI models can predict the spread of wildfires or floods, enabling preemptive evacuations and resource deployments that save lives and property.

The efficacy of Computational AI is not solely dependent on algorithms; it also hinges on the quality of data fed into these systems. The government's move towards open data initiatives provides a basis for Computational AI's accuracy. When powered by high-quality, accessible data, Computational AI can offer powerful analytical solutions across the public sector.

One of the most significant applications of Computational AI in the government is its integration into national defense systems. Here, it assists in processing and analyzing vast streams of intelligence data, identifying potential threats, and supporting decision-making in defense strategy and operations.

Computational AI's capacity for handling large-scale, complex calculations also enables the government to forecast geopolitical shifts and prepare accordingly.

Documenting the Machine's Mind: The Imperative of Explainable and Transparent AI in Government

Explainable and Transparent AI (XAI) documents the actions and decisions of algorithms to share the inner workings of AI systems with humans, ensuring that the decision-making processes are not just effective but also logical and accurate.

In government operations, as in regulated industries, XAI is a necessity.

When AI systems determine the allocation of resources or influence legal outcomes, they must do so with a level of transparency that allows their reasoning to be examined and understood by humans. This transparency is vital in building human trust and ensuring that automated decisions adhere to the values and laws of our government.

The technical infrastructure of explainable AI is complex, yet its goal is straightforward: to demystify AI decisions.

This is achieved through various methods such as feature importance—which highlights the data variables most influential in a model's decision—and decision trees, which outline the binary choices leading to a conclusion.

Model-agnostic methods, too, play a pivotal role. These are techniques that can be applied to any ML model to document its decision-making process.

In practical terms, the U.S. government applies XAI across multiple agencies and sectors:

- Public Service Delivery: XAI assists in ensuring that algorithms determining the distribution of social services do so with equitability and transparency.
- Law Enforcement: In criminal justice, XAI is crucial for predictive policing tools, assuring they operate without bias and under ethical guidelines.
- Regulatory Oversight: Monitoring financial and healthcare regulations, XAI helps maintain compliance with complex legal standards.
- Policy Making: When analyzing data for policy development, XAI ensures that policy recommendations are based on understandable and justifiable models.
- Healthcare: XAI technologies document recommended diagnostics and treatment options.

XAI adds value to AI in several areas. For example:

- By combining XAI with ML, government agencies can make the reasoning behind complex models transparent.
- With data analytics, XAI documents the intricacies of all data processes, fueling accountability and compliance.
- In collaboration with computational AI, XAI ensures that the sophisticated decision-making models and the resulting recommendations are both understandable and fully documented.

Unlocking the Future: Quantum Computing's Role in Advancing Government AI

Quantum Computing and Quantum AI promise a significant leap forward in computational capability, harnessing the principles of quantum mechanics to process information in ways that classical computing cannot [1].

In the realm of AI, quantum computing introduces Quantum AI, where these advanced principles are applied to accelerate and enhance machine learning algorithms, especially in optimization and complex pattern recognition.

This technology holds significant promise for the U.S. government, offering breakthrough potential in secure communications, advanced cryptography, and solving large-scale computational problems that are currently unsolvable for classical computers.

For example, quantum computing promises to analyze extensive climate data sets with unprecedented speed, offering more accurate and detailed climate models. It can also revolutionize secure communication through quantum encryption, which could be unbreakable by conventional hacking methods, fortifying national security [2].

As Quantum AI continues to develop, it is expected to become integral to various government sectors, enhancing data analysis, bolstering security measures, and enabling more efficient resource allocation and strategic planning.

Data as a Pillar: Enhancing AI's Decision-Making Power in Government

AI and data management are transforming how government agencies leverage information to enhance public service. This involves a complex system of practices and tools designed to optimize data usage, ensuring AI decision-making is as informed and effective as possible.

At its foundation, AI relies on the availability and quality of data. For government agencies, this means implementing rigorous data management policies that align with overarching strategies such as the Federal Data Strategy. Such policies are critical for maintaining data integrity, ensuring security, and fostering public trust.

The role of data governance in AI extends beyond mere maintenance to encompass regulatory compliance, adhering to strict standards like the General Data Protection Regulation (GDPR) and the Health Insurance Portability and Accountability Act (HIPAA)   , especially relevant in sectors managing sensitive information.

In practical application, AI and data management enable government agencies to perform predictive analytics in public health, optimize resource allocation, detect and prevent fraud, and manage environmental monitoring. These applications rely on a robust infrastructure of databases, data lakes, and cloud computing solutions that can handle the scale and security demands of government AI workloads.

However, challenges such as data silos, data quality, and ethical considerations remain at the forefront of AI data management. The future points towards emerging trends and tools, including blockchain for data integrity and edge computing for data processing, which could redefine how government data is managed and utilized for AI applications.

This blend of AI and data management holds the promise of more responsive, efficient, and accountable government operations. It also enhances the strategic use of our most valuable resource - data.

Key AI Use Cases in the U.S. Government

There are a number of AI applications within government that are already in progress. They represent our first steps in understanding and applying the current and future potential of AI to solve complex, data driven problems that are outpacing our current abilities to analyze and optimize. As AI evolves and advances, these baseline applications are expected to expand and evolve.

1. Automating Routine Tasks and Predictive Maintenance

The U.S. Department of Defense (DoD) operates a vast array of complex equipment, including aircraft, vehicles, and weapon systems. Maintaining these assets in optimal condition is crucial for operational readiness.

As in commercial industries, traditional maintenance approaches rely on scheduled maintenance or reacting by addressing issues after failure. This results in inefficiencies and unplanned downtime. In the case of the DoD, such events are not acceptable.

Yet, proactively predicting and solving issues ahead of downtime is unwieldy, demanding significant manpower to perform complex, routine tasks, such as monitoring sensors and analyzing vast data sets. There is also the issue of ongoing human error.

AI Today

The DoD has begun implementing AI models for predictive maintenance. These models use machine learning algorithms to analyze data from equipment sensors and predict potential failures before they occur. By using ML, the DoD can analyze more complex data, faster and with much more precision, while significantly reducing the massive manpower requirements and human errors.

- ML systems are trained on historical maintenance data, operational logs, and real-time sensor readings. This enables them to identify patterns and anomalies that indicate a potential failure.
- By automating the analysis of sensor data, AI reduces the need for manual monitoring. For instance, in aircraft maintenance, AI can assess when a plane requires service or can continue operation, based on flight hours and sensor data. Maintenance is performed on schedules, while sensor data can trigger an early response to a potentially significant event, eliminating the unexpected downtime.

For example, the U.S. Air Force applies AI for predictive maintenance of the C-5 Galaxy and the E-3 Sentry aircraft. The AI models identify parts that are likely to fail, then suggest optimal maintenance schedules [3].

Another instance is in the U.S. Navy where AI is used for predictive analytics on ship engines and mechanical systems. AI empowers timely maintenance actions and reduces the risk of operational disruptions [4].

The adoption of AI for predictive maintenance has shown to improve the readiness and availability of military equipment. It reduces downtime and maintenance costs by addressing issues proactively. It also reduces maintenance turnaround times and increases equipment availability.

As expected, the DoD continues to refine the AI models. For example, increasing the data sets to be more comprehensive and tuning algorithms to increase accuracy of insights. Deep integration of these AI systems directly into the maintenance workflows so that personnel can immediately leverage AI-driven insights as part of all maintenance decisions is also under exploration [3].

2. Intelligence Analysis and Enhanced Weapon Systems

The complexity of modern battlefields, combined with the massive volumes of data generated in defense operations, demand efficient, real-time analysis for decision-making.

Traditional methods of intelligence analysis were labor-intensive and slow. Today, they cannot possibly meet the demands of complex and modern intelligence. Similarly, enhancing weapon systems traditionally required extensive manual input and testing. Today, that process is also more complex and unwieldy for traditional approaches.

Artificial intelligence offers the opportunity to accelerate analysis and decisions, improve the accuracy of both, and open doors to powerful new insights into intelligence and response systems.

The U.S. Department of Defense's AI Adoption Strategy emphasizes the acceleration of AI capabilities to maintain decision superiority on the battlefield. This strategy includes enhancing the speed and quality of commanders' decisions, which directly impacts the ability to deter aggression and excel in combat situations. The emphasis is on integrating AI responsibly into operations to improve decision-making processes and to ensure safety, as unsafe systems could be deemed ineffective. This approach is not just about staying current with technological trends but ensuring that AI applications are reliable and efficient for long-term operational success [5].

AI Today

AI is being applied to efficiently process and analyze large volumes of intelligence data quickly, including image and signal analysis, as well as facial recognition.

For weapon systems, AI enhances targeting, navigation, and decision-making processes. AI algorithms are integrated into drones, missile systems, and robotic units to improve their autonomy and response times.

The development of AI-driven intelligence analysis tools and their integration into weapon systems have significantly enhanced military capabilities. The U.S. Department of Defense has released an AI Adoption Strategy that focuses on leveraging AI to maintain decision superiority on the battlefield. This strategy involves integrating AI to accelerate the speed and improve the accuracy of commanders' decisions, which is crucial for both deterring conflict and excelling in combat situations [6]. By continuously deploying data analytics and AI capabilities, the DoD aims for enduring decision advantages in battlespace awareness, force planning, kill chains, and business operations.

Technological advancements in AI have also facilitated improvements in natural language processing and machine learning, contributing to more precise interpretations of intelligence data. This has important implications for how information is processed, understood, and used to make critical decisions in military contexts. The applications of AI include a range of use cases, from logistics and supply chain optimizations to the direction and control of autonomous and AI-enabled technologies for enhanced warfare capabilities.

Operational efficiency has seen improvements through the deployment of enhanced AI weapon systems in both training and combat scenarios.

- The AI systems are designed to identify targets, optimize mission planning, and provide decision support to human operators.
- AI improves accuracy, speed, and outcome effectiveness in highly volatile and urgent operational settings.
- The implications of these advancements are far-reaching, affecting the present state of military operations. They are also defining more advanced and expedient future approaches to warfare technologies and intelligence handling.

Other applications include the use of AI-driven facial recognition technologies to identify persons of interest in security operations across the globe. The U.S. military also employs AI in drone systems for autonomous targeting and threat assessments, eliminating the lag times associated with human analysis and control.

AI in intelligence analysis has significantly reduced the time to process and analyze information, leading to quicker responses to emerging threats.

Enhanced weapon systems with AI have shown increased accuracy and efficiency, although the extent of their operational use and effectiveness varies based on the specific situation and the weapons involved.

3. AI in Cybersecurity

Cybersecurity threats are rapidly evolving thanks to AI, quantum computing, and other advancements in technology [2]. Traditional security measures struggle to keep pace. The sheer volume of potential threats and the sophistication of cyber-attacks make manual monitoring and response inadequate, and today’s security technology is not capable of protecting against these modern and aggressive attacks.

AI is being used to monitor network traffic and identify unusual patterns that could indicate a cyber threat. Additionally, AI algorithms analyze data from various sources to identify potential insider threats and zero-day vulnerabilities [7].

AI Today

As discussed in the article, “Cybersecurity: The Drive for Continuing Innovation,” a major focus of AI in cybersecurity is preventing quantum computers from cracking the current security algorithms that are the foundation of data security [2]. Yet there are many more opportunities.

Another example of AI in cybersecurity is protecting intelligence and alerting the U.S. military to potential attacks. AI is trained to implement real-time network monitoring, automatically detecting, and responding to cyber threats. The systems are easily trained to identify vulnerabilities from emerging threats as they are reported globally, delivering proactive defense measures.

Consequently, AI in cybersecurity has enhanced our ability to detect and respond to threats more rapidly and accurately. These systems will continue to be enhanced and expanded with the continuous updates needed to address new cyber threats, as they arise.

4. Geospatial Analysis and Autonomous UAVs

Geospatial data is a key to military operations. It provides critical situational awareness and intelligence to commanders, helping them to quickly understand enemy movements, terrain, and friendly forces.

The military uses geospatial data for a variety of purposes including: planning missions, developing tactics, conducting operations, intelligence gathering and analysis, and post-mission analysis and feedback.

Due to the volume and complexity of geospatial data, its analysis is difficult and time-consuming.

Unmanned aerial vehicles (UAVs) are powered aircraft that can fly without a human pilot. They can also be piloted remotely and can operate autonomously.

Traditionally UAVs require significant human resources for operation and data analysis, leading to information overload.

AI Today

AI and ML are employed to analyze satellite and aerial imagery for intelligence purposes, such as damage assessment or landscape analysis. Machine learning and artificial intelligence significantly enhance the speed and accuracy of this analysis.

For example, AI-driven analysis of satellite imagery quickly identifies and alerts to changes in enemy installations or provides rapid damage assessments in disaster zones [8].

In UAVs, AI automates data processing and action recommendations, reducing the burden on human operators. Today, UAVs equipped with advanced technologies, such as AI, operate independently to make intelligent decisions without input from a pilot or operator, learning from and adapting to their environment.

For example, Autonomous UAVs use AI for mission planning, target surveillance, and decision-making in reconnaissance and combat operations.

Simultaneously, autonomous UAVs have enhanced surveillance capabilities, although the balance between automation and human oversight remains a focal point for ongoing operations.

5. Healthcare

Fraud detection in healthcare programs, like Medicaid, is challenging due to highly disconnected and vast data sets that are difficult to manage, much less analyze. Even as higher numbers of human analysts are tasked with reviewing and searching for potential fraud, the process becomes more and more difficult.

AI and ML offer a solution. They can consume the vast amounts of data from highly diverse and distributed data sets, consolidate and normalize them, and then analyze them for potential fraud or misuse.

AI Today

AI algorithms are being used to sift through healthcare data to detect patterns indicative of fraud. For example, AI systems in Medicaid fraud detection analyze billing patterns and patient records to identify anomalies.

Research on Medicare data has been conducted to improve healthcare fraud detection using data-centric AI approaches. These methodologies use claims data to classify and detect fraudulent activities by healthcare providers, which can lead to billions in savings and improved quality of patient care. With the increasing complexity of healthcare billing systems and the rising incidence of fraudulent activities, AI offers a sophisticated and efficient solution to identify and address these issues proactively [9] .

Already, AI has improved the efficiency and effectiveness of detecting healthcare fraud.

5. Disaster Response

Predicting and managing resources for natural disasters requires analyzing large amounts of data from diverse sources. As with all of our ever-growing complex data sets, consolidating, analyzing, and quickly delivering insights is difficult leveraging human analysts. Given the immediate response demanded by disasters, AI and ML are a welcome and vital set of technologies to improve the timing and accuracy of our responses to disasters.

AI Today

In disaster response, AI analyzes weather data, social media, and other sources to predict impacts and optimize resource allocation.

For instance, Google's AI models can evaluate river water levels up to seven days in advance and generate detailed flood maps. These models, which are trained on an array of global weather data and satellite imagery, have shown the potential to provide critical information in areas that typically have less data available, including low and medium-income countries. Such advancements in AI-powered flood forecasting are proving to be a boon for disaster management, allowing for more prepared responses in affected communities and among aid organizations [10].

Moreover, initiatives like the United Nation's Institute for Training and Research's (UNITAR) rapid mapping service, which forms part of their Operational Satellite Applications Programme (UNOSAT), demonstrate how AI is being harnessed to quickly assess the impact of floods through satellite image analysis. UNOSAT's analysts work with AI algorithms trained on extensive archives of historical flood maps to identify flooded areas with high accuracy. This approach ensures that high-quality maps are produced and delivered to relief efforts rapidly, sometimes within just a few days, greatly enhancing the response to humanitarian crises [11].

The integration of AI methodologies into rapid mapping services has shown to accelerate the map production process while maintaining quality. These services can provide numerous updates on the spread of floods, helping to direct relief efforts more efficiently.

Questions and Considerations Around AI Adoption and Deployment in the U.S. Government

For U.S. government use cases, the key questions surrounding AI adoption often extend beyond operational efficiency to encompass broader issues related to public interest, security, and governance:

1. Ethical and Legal Compliance: How does AI deployment align with ethical standards and legal frameworks, particularly regarding privacy, civil liberties, and human rights?

2. Data Security and Privacy: What measures are in place to protect sensitive government and citizen data from breaches and misuse in AI systems?

3. Transparency and Accountability: How will AI decisions be made transparent to stakeholders, and what accountability mechanisms are in place for AI-driven actions?

4. Interoperability and Standards: How can AI systems ensure interoperability across different government agencies and compliance with established standards?

5. Public Trust: How will the use of AI foster or impact public

trust in government decisions and services?

6. Risk Management: What are the risks (including biases in AI algorithms) and how are they being mitigated?
7. Workforce Impact and Training: How will AI impact government workforce, and what training will be provided to upskill employees for the AI-driven future?
8. Budget and ROI: What are the cost implications of AI adoption, and how is ROI measured in terms of improved public services or operational efficiency?
9. Vendor Selection and Management: How does the government evaluate and manage third-party vendors providing AI solutions, ensuring they align with public sector values and regulations?
10. Long-term Strategy and Scalability: What is the long-term strategy for AI adoption, and how are these technologies scalable and adaptable to future needs?
11. Collaboration with Private Sector and Academia: How does the government plan to collaborate with private companies and academic institutions for AI development and ethical considerations?
12. National Security and Defense: In defense and security applications, how is AI used responsibly, and what are the protocols to prevent misuse or unintended consequences?

Conclusion

As with many advanced technologies, the hype around artificial intelligence outpaces its reliable applications.

Yes, today's AI offers significant value beyond the AIs of the past. Machine Learning processes large data models to provide insights that can and will enhance our intelligence and our ability to act on that knowledge. It outpaces search engines to deliver research in a fraction of the time. It offers unique concepts and ideas to push our thinking into new realms.

That said, it is not the trustworthy content creator many expect. Aside from the hallucinations and false information, its content is often stilted and filled with adjectives and superlatives that are obvious to the reader.

Picture References

[A] "Figure 1.8: Generic Workflow of Generative Adversarial Network (GAN)...." ResearchGate, www.researchgate.net/figure/Generic-workflow-of-Generative-Adversarial-Network-GAN-The-figure-was-created-by_fig5_340031990.

[B] Yong, Roger. "Variational Autoencoder(VAE) - Geek Culture - Medium." Medium, 6 Jan. 2022, medium.com/geekculture/variational-autoencoder-vae-9b8ce5475f68.

[C] Nield, David. "How ChatGPT and Other LLMs Work—and Where They Could Go Next." WIRED, 30 Apr. 2023, www.wired.com/story/how-chatgpt-works-large-language-model.

[D] Bozkus, Emine. "Machine Learning 101: A Beginner's Guide to Understanding the Basics." Medium, 9 Dec. 2022, eminebozkus.medium.com/machine-learning-101-a-beginners-guide-to-understanding-the-basics-d8a85ff2c8e.

References

[1] Brown, Rebel. “Quantum Computing: Redefining Technology, Science, and Information.” CrossTalk, Aug 2023. August, 2023.

[2] Brown, Rebel. “Cybersecurity: The Drive for Continuing Innovation.” CrossTalk, Nov 2023. November, 2023.

[3] Taylor, Dan. “AI Predictive Maintenance for U.S. Army to Be Provided by Palantir.” Military Embedded Systems. militaryembedded.com/ai/machine-learning/ai-predictive-maintenance-for-us-army-to-be-provided-by-palantir.

[4] Uppal, Rajesh. “Armies Investing in Predictive Maintenance Technology.” International Defense Security & Technology, idstch.com/military/army/armies-investing-in-predictive-maintenance-technology.

[5] “Fact Sheet. 2023 DoD Cyber Strategy.” 2023 DoD Cyber Strategy Fact Sheet, https://media.defense.gov/2023/May/26/2003231006/-1/-1/1/2023-DOD-CYBER-STRATEGY-FACT-SHEET.PDF

[6] “DOD Releases AI Adoption Strategy.” U.S. Department of Defense, www.defense.gov/News/News-Stories/Article/Article/3578219/dod-releases-ai-adoption-strategy.

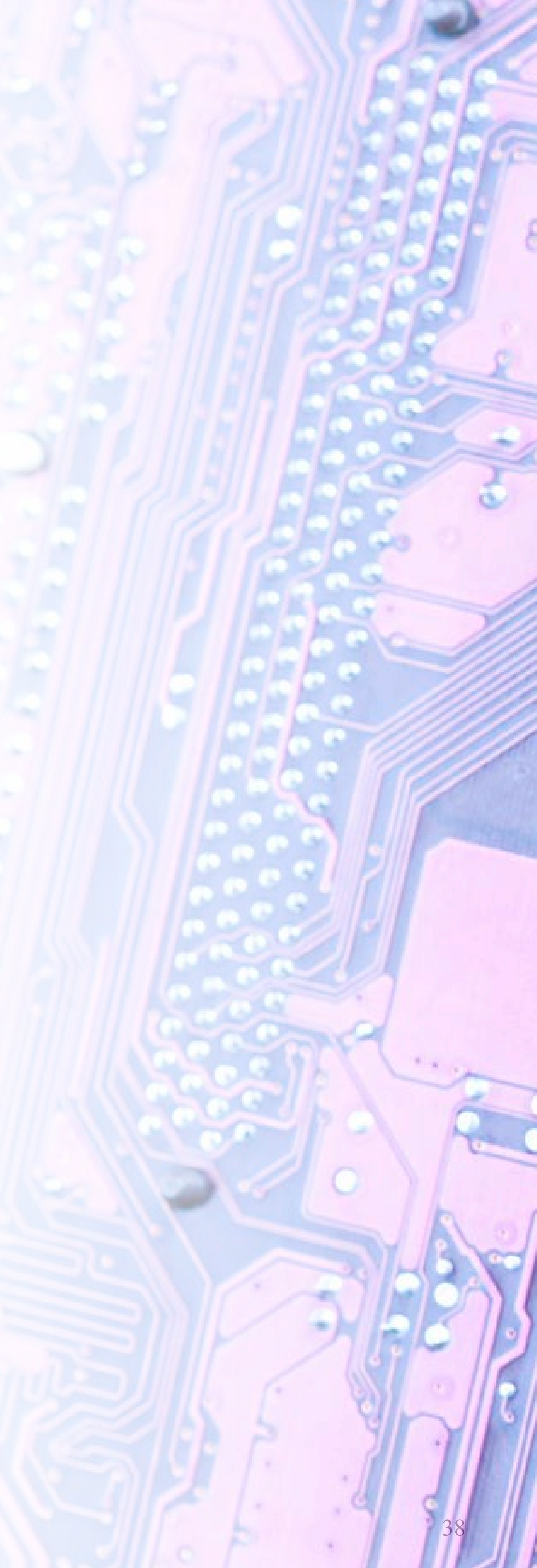
[7] Zero Day Attack - Glossary | CSRC. csrc.nist.gov/glossary/term/zero_day_attack.

[8] Torrieri, Marisa. “How AI and ML Are Supercharging Earth Observation.” Via Satellite, 23 Oct. 2023, interactive.satellitetoday.com/via/november-2023/how-ai-and-ml-are-supercharging-earth-observation.

[9] Shekhar, et al. “Unsupervised ML for Explainable Health Care Fraud Detection.” Ar5iv, ar5iv.labs.arxiv.org/html/2211.02927.

[10] “Flood Forecasting: AI for Information and Alerts - Google Research.” Flood Forecasting: AI for Information & Alerts - Google Research, sites.research.google/floodforecasting.

[11] “Fusing AI Into Satellite Image Analysis to Inform Rapid Response to Floods.” UNITAR, unitar.org/about/news-stories/news/fusing-ai-satellite-image-analysis-inform-rapid-response-floods.



About the Author

As a recognized technology strategist, Rebel Brown guides companies to profitably define, launch, and grow their bottom lines. She is a go-to-market expert whose strategies, positioning, and launches have led to dramatic and successful results for over 75 startups and 300 high-tech and complex B2B technology companies globally. Her current passions include quantum computing, artificial intelligence (AI/XAI) and machine learning (ML).

Ms. Rebel Brown

CEO

Cognoscenti, Inc.

rebel@rebelbrown.com

402D SOFTWARE ENGINEERING GROUP

402d Software Engineering Group is staffed with over 1,500 personnel experienced in electrical engineering, computer engineering, computer science, information technology, and administrative support. To fulfill our mission in being a premier software provider for DoD and NATO, we design and maintain one-of-kind system integration labs consisting of software and hardware-in-the-loop. Connect with us on social media!

Leveraging AskSage: An Intelligent Question-Answering System for Enhanced Efficiency in Government-Related Tasks

Tony Lau
Staff Engineer,
Hill Air Force Base

Introduction

In an era characterized by rapidly evolving threats, complex geopolitical landscapes, and the ever-increasing demand for efficient decision-making, the Department of Defense (DoD) is constantly seeking innovative solutions to enhance its operational capabilities. This paper introduces AskSage, an advanced conversational AI system meticulously designed and tailored to cater to the unique needs of the DoD. Serving as the ChatGPT equivalent for defense applications, AskSage represents a groundbreaking step towards revolutionizing communication, knowledge dissemination, and decision support within the defense sector.

With its natural language processing capabilities and nuanced understanding of defense-related contexts, AskSage emerges as a versatile tool poised to address the multifaceted challenges faced by military personnel, strategists, and decision-makers. By harnessing the power of AI, AskSage aims to streamline information exchange, augment intelligence analysis, and facilitate more informed decision-making processes across various defense domains. Consideration is given to potential challenges and ethical concerns associated with the deployment of AskSage in government contexts.

This paper introduces AskSage, a sophisticated commercial conversational AI system specifically designed to meet the unique requirements of the DoD. It explores how the system's adaptability to military strategies, technological advancements, and geopolitical scenarios can significantly contribute to achieving enhanced operational efficiency. Additionally, it considers AskSage's potential to enhance citizen engagement and transparency in government operations.

As we embark on an exploration of AskSage for the DoD, this paper not only highlights the capabilities and potential benefits of this specialized conversational AI but also aims to shed light on the transformative impact such technology can have on defense communication, decision-making processes, and overall readiness in an ever-evolving global security landscape.

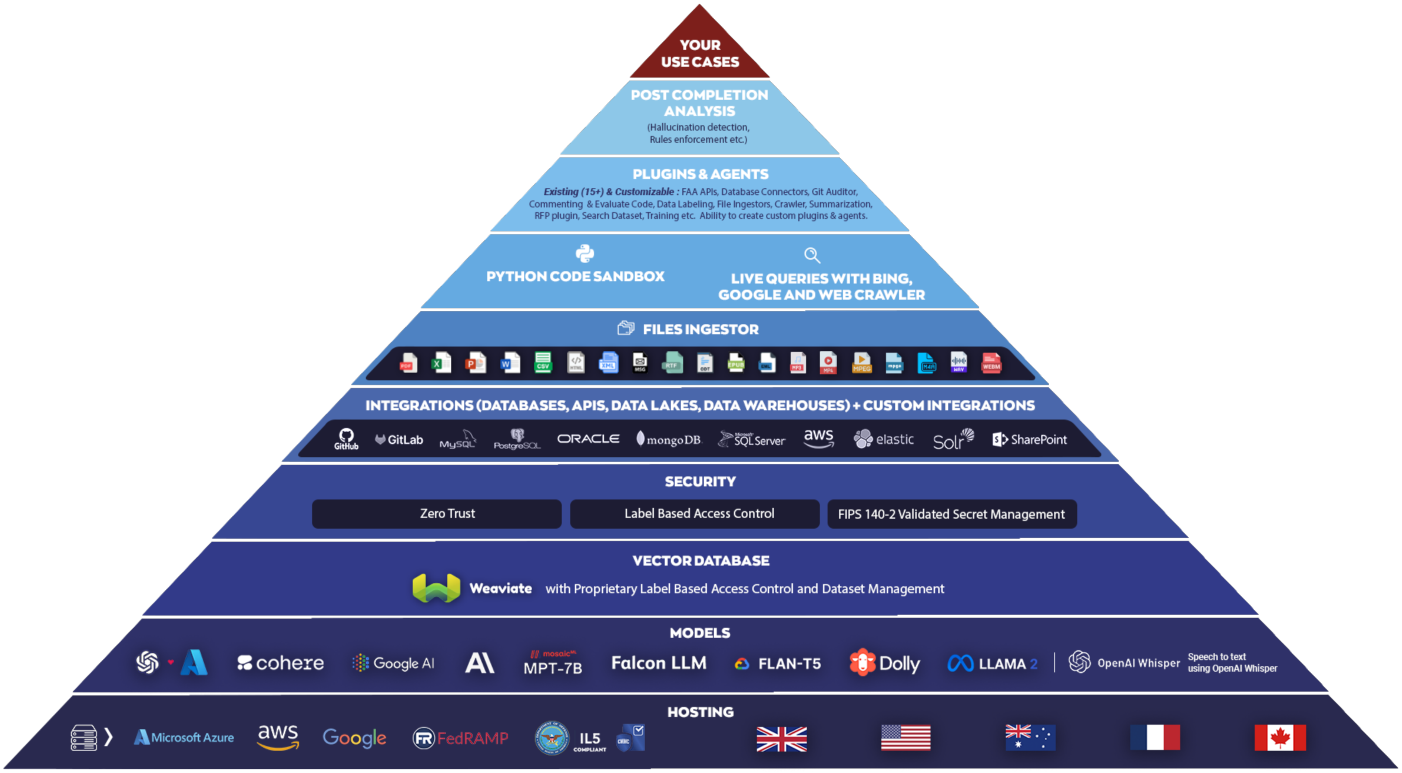

YOUR USE CASES
POST COMPLETION ANALYSIS
(Hallucination detection, Rules enforcement etc.)
PLUGINS & AGENTS
Existing (15+) & Customizable : FAA APIs, Database Connectors, Git Auditor, Commenting & Evaluate Code, Data Labeling, File Ingestors, Crawler, Summarization, RFP plugin, Search Dataset, Training etc. Ability to create custom plugins & agents.
PYTHON CODE SANDBOX
LIVE QUERIES WITH BING, GOOGLE AND WEB CRAWLER
FILES INGESTOR
PDF
CSV
HTML
XML
MSG
RTF
ODT
EPUB
EML
MP3
MP4
MPEG
mpga
M4A
WAV
WEBM
INTEGRATIONS (DATABASES, APIS, DATA LAKES, DATA WAREHOUSES) + CUSTOM INTEGRATIONS
GitHub
GitLab
MySQL
PostgreSQL
ORACLE
mongoDB
SQL Server
aws
elastic
Solr
SharePoint
SECURITY
Zero Trust
Label Based Access Control
FIPS 140-2 Validated Secret Management
VECTOR DATABASE
Weaviate
with Proprietary Label Based Access Control and Dataset Management
MODELS
cohere
Google AI
mosaic
MPT-7B
Falcon LLM
FLAN-T5
Dolly
LLAMA 2
OpenAI Whisper
Speech to text using OpenAI Whisper
HOSTING
Microsoft Azure
aws
Google
FedRAMP
IL5 COMPLIANT
CMMC

Figure 1. *AskSage LLM Model from hosting to use cases [A].*

Policy Analysis of AskSage

Architectural Insights

AskSage's effectiveness within the DoD is underpinned by its robust architectural foundation, specifically designed to meet the unique needs of defense-centric queries and communication. Leveraging advanced Natural Language Processing (NLP) algorithms, AskSage interprets and responds to queries with precision. These algorithms are specially trained to understand military jargon, protocols, and terminology, ensuring responses that are both accurate and contextually relevant. By analyzing the structure and semantics of text inputs, AskSage extracts key information and provides meaningful insights to users.

In addition to its NLP algorithms, AskSage employs machine learning models that enhance its understanding and response capabilities. Trained on extensive datasets of defense-related documents (including manuals, regulations, and historical data), these models enable the platform to continuously refine its accuracy and relevance. Through machine learning, AskSage adapts to evolving defense requirements, maintaining high performance and precision over time.

Customization is often necessary to ensure that AskSage effectively meets the specific needs of the DoD. Given the complexity and sensitivity of defense data, this customization can involve adapting the product's algorithms and applications to handle specialized terminologies and comply with security protocols. AskSage's platform is inherently designed with security features to protect sensitive data, having been initially developed for the Department of Defense. It also offers cloud-agnostic and model-agnostic capabilities, which allow it to operate across various environments and integrate with over 150 large language models from providers such as Google, Anthropic, and Azure OpenAI [1].

At a high level, AskSage operates through a multi-layered architecture that integrates various technological components. The core architecture includes a sophisticated NLP engine that parses and interprets user queries, supported by a machine learning infrastructure that continuously learns from interactions and feedback. Principal components of AskSage include a query processing unit, which manages the intake and analysis of text inputs; a response generation module, which formulates accurate and contextually appropriate answers; and a feedback loop system that refines algorithms based on user input and performance metrics. This architecture allows AskSage to handle complex queries and deliver nuanced responses, showcasing the intricate nature of modern NLP technologies. In addition to NLP algorithms, AskSage utilizes machine learning models to enhance its understanding and response capabilities.

AskSage's architecture is designed to seamlessly integrate into the defense communication ecosystem. This integration is achieved through secure APIs and protocols that enable AskSage to access and interact with classified information and systems. By adhering to strict security standards and protocols, AskSage ensures the confidentiality and integrity of sensitive defense information [1].

AskSage's architecture is designed for scalability and performance, allowing it to handle a high volume of queries and users simultaneously. Through the use of distributed computing and cloud-based infrastructure, AskSage can efficiently process and respond to queries in real-time. As the defense landscape continues to evolve, AskSage's architecture ensures that it remains a valuable tool for enhancing communication and decision-making within the Department of Defense [2].

Defense Decision Support and Intelligence Augmentation

AskSage enables defense decision-makers to access a wide range of information, including geopolitical analyses, threat assessments, and operational insights. This real-time access to relevant information enhances situational awareness and allows for timely and informed decision-making. AskSage serves as a force multiplier. Military leaders can leverage the insights provided by AskSage to develop strategies, allocate resources, and respond to emerging threats more effectively.

AskSage accelerates the decision-making process by augmenting intelligence analysis. The system's ability to quickly analyze vast amounts of data and provide actionable insights allows for a quick turnaround, ensuring that military leaders can respond promptly to emerging threats.

Air Force Secretary Frank Kendall has initiated a study through the Scientific Advisory Board to assess the potential impacts of "generative" artificial intelligence, such as the popular AI program ChatGPT, in military applications. Kendall acknowledges the current limited utility of such AI systems for the military but envisions their assistance in certain tasks if applied "in an ethical way." He emphasizes the importance of addressing ethical considerations while rapidly advancing AI technology into the field, acknowledging its revolutionary potential in enhancing military capabilities [3].

Kendall is establishing a permanent AI-focused group to comprehend and integrate AI technologies efficiently into the Defense Department's operations. He highlights the broader interest in AI for tasks like pattern recognition and sorting through intelligence functions, emphasizing the need for ethical deployment and the potential of AI to enhance military capabilities significantly.

By leveraging AI technologies like AskSage, the Air Force can stay ahead of emerging threats in an increasingly complex and dynamic global security environment.

Streamlining Operational Communication

Effective communication is crucial for the success of military operations. To explore AskSage's role in streamlining internal communication processes within the Department of Defense (DoD), the platform serves as a centralized stage for disseminating critical updates, sharing mission-specific information, and facilitating collaboration among personnel, AskSage enhances the agility and responsiveness of the military apparatus.

The tool facilitates collaboration among personnel by providing a platform for sharing ideas, insights, and best practices. This collaborative environment fosters innovation and enables personnel to work together more effectively towards common goals.

By streamlining communication processes and enhancing collaboration, AskSage helps to make the military apparatus more agile and responsive. Personnel can quickly adapt to evolving challenges and make informed decisions based on real-time information. For example, during a recent military exercise, AskSage was used to disseminate updated mission objectives to personnel in real-time. This allowed units to adjust their tactics and strategies on the fly, leading to a more successful outcome.

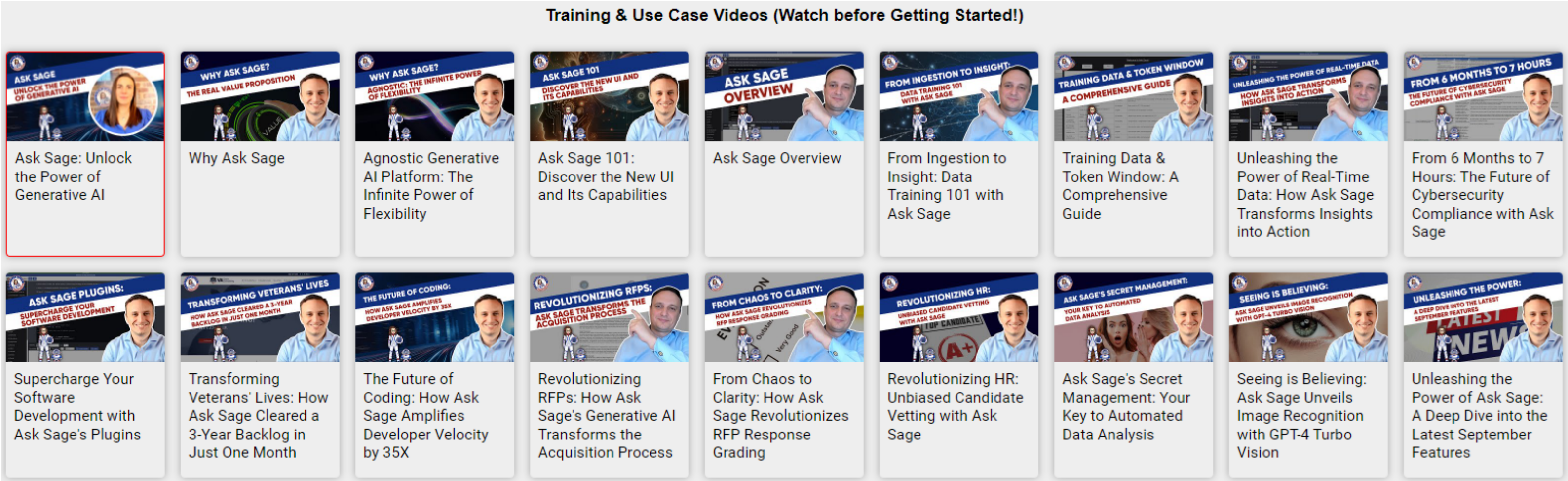
Training and Education for Military Personnel

AskSage goes beyond its operational role to become an invaluable resource for training and education within the military. This section explores how the system facilitates continuous learning, providing an interactive platform for military personnel to acquire and reinforce knowledge related to defense policies, historical contexts, and evolving technologies. AskSage's adaptability in creating tailored learning experiences contributes to the ongoing professional development of military personnel.

AskSage serves as a continuous learning platform for military personnel, offering access to a wide range of educational resources. These resources include articles, videos, and interactive tutorials that cover various aspects of defense policies, historical events, and emerging technologies.

The tool provides interactive learning experiences that engage military personnel and enhance their understanding of complex concepts. Through simulations and virtual scenarios, personnel can apply their knowledge in practical settings.

AskSage allows military personnel to create tailored learning paths based on their interests and career goals. For example, a new recruit interested in learning about military history can use AskSage to access a series of interactive modules that cover key historical events and their impact on modern warfare. Through quizzes and simulations, the recruit can test their knowledge and gain a deeper understanding of military history.



Training & Use Case Videos (Watch before Getting Started!)

- Ask Sage: Unlock the Power of Generative AI
- Why Ask Sage
- Agnostic Generative AI Platform: The Infinite Power of Flexibility
- Ask Sage 101: Discover the New UI and Its Capabilities
- Ask Sage Overview
- From Ingestion to Insight: Data Training 101 with Ask Sage
- Training Data & Token Window: A Comprehensive Guide
- Unleashing the Power of Real-Time Data: How Ask Sage Transforms Insights into Action
- From 6 Months to 7 Hours: The Future of Cybersecurity Compliance with Ask Sage
- Supercharge Your Software Development with Ask Sage's Plugins
- Transforming Veterans' Lives: How Ask Sage Cleared a 3-Year Backlog in Just One Month
- The Future of Coding: How Ask Sage Amplifies Developer Velocity by 35X
- Revolutionizing RFPs: How Ask Sage's Generative AI Transforms the Acquisition Process
- From Chaos to Clarity: How Ask Sage Revolutionizes RFP Response Grading
- Revolutionizing HR: Unbiased Candidate Vetting with Ask Sage
- Ask Sage's Secret Management: Your Key to Automated Data Analysis
- Seeing is Believing: Ask Sage Unveils Image Recognition with GPT-4 Turbo Vision
- Unleashing the Power of Ask Sage: A Deep Dive into the Latest September Features

Figure 2. *Training videos provided by AskSage [B].*

Regulatory Compliance of AskSage Security and Ethical Considerations

As with any technology deployed within the defense sector, security and ethical considerations are paramount. This section scrutinizes the measures taken to ensure the confidentiality, integrity, and secure exchange of sensitive information through AskSage. Ethical considerations related to the responsible use of AI in defense applications are also addressed, emphasizing the importance of transparency and accountability.

In the realm of security and ethical considerations, recent developments within the Department of Defense have unveiled a temporary ban on the use of generative artificial intelligence (GenAI) tools and large language models (LLMs) for official purposes by Space Force Guardians. The directive, originating from a memo dated 29 September 2023 by Lisa Costa, the Space Force's Chief Technology and Innovation Officer, explicitly restricts Guardians from employing government data in generative AI solutions without obtaining official approval.

This ban extends to prominent platforms such as OpenAI's ChatGPT and AskSage, the latter being a specialized model designed by Nicolas M. Chaillan, the former Chief Software Officer for the Department of the Air Force. Notably, Chaillan has voiced concerns regarding the potential risks associated with such prohibitions, highlighting the likelihood of personnel accessing these platforms on personal devices, thereby creating a higher risk of shadow IT and cybersecurity vulnerabilities [4].

Despite the ban, Chaillan emphasized the security measures implemented in AskSage, developed on government clouds and meeting all cybersecurity prerequisites. He reported that approximately 500 Guardians utilized the platform over the past six months without any reported security incidents. However, the Space Force's decision indicates a temporary pause to evaluate the integration of such capabilities into the United States Space Force (USSF) mission.

Ethical considerations raised in the memo align with concerns voiced by a Pentagon spokesperson, who highlighted the need to safeguard sensitive DoD data and uncertainties regarding the traceability and validity of answers provided by LLMs. While the ban is indicative of cautious deliberation, it also prompts reflections on the potential risks and benefits associated with the use of advanced AI tools within the defense sector.

Lisa Costa's memo sheds light on her office's participation in Task Force Lima, a DoD initiative aimed at advancing national security through GenAI. Despite the ban, Costa expressed confidence in the long-term utility of GenAI, anticipating its revolutionary impact on the workforce and the enhanced ability of Guardians to operate at speed. The duration of the ban remains uncertain, with plans for the Chief Technology and Innovation Officer to issue specific USSF guidance, possibly within 30 days of the memo's publication.

AskSage is deployed within the DoD network, adhering to stringent security and compliance protocols. It employs robust security measures on the host system to maintain data integrity and is specifically approved to handle Controlled Unclassified Information (CUI). The platform manages large volumes of data using secure, pre-loaded datasets and periodic updates from internal sources, eliminating the need for real-time Internet access while upholding high-security standards.

Recent advancements further bolster AskSage's security credentials, as highlighted in the July 2024 update. The platform has achieved an Impact Level 5 (IL5) Authority to Operate (ATO) with the U.S. Army, a Defense Information Systems Agency (DISA) Interim Authority to Test (IATT), and a Federal

Risk and Authorization Management Program (FedRAMP) High ATO. These certifications confirm that AskSage meets rigorous standards for handling sensitive data and aligns with federal guidelines, ensuring its suitability for integration into government systems and operations.

To address concerns about security and the potential for “classification by aggregation,” AskSage implements several robust measures. The platform is hosted on Azure Government at IL5, ensuring top-tier data safety and security. Strict access controls are enforced, including Common Access Card (CAC) authentication for CUI datasets and Multi-Factor Authentication (MFA) using One-Time Passwords (OTPs) with Microsoft Authenticator for administrative access. Data is encrypted in transit using Transport Layer Security (TLS) 1.3 and at rest through Azure Government’s encryption capabilities. Continuous monitoring is provided by Azure Firewall v2, Microsoft Defender for Cloud, and other security solutions to detect and prevent unauthorized access or attacks. These comprehensive measures mitigate risks associated with data aggregation and ensure AskSage maintains the highest standards of security in sensitive operational contexts [1].

These recent developments underscore the delicate balance between security imperatives, ethical considerations, and the imperative to harness cutting-edge AI technologies in defense operations. As the USSF evaluates the best path forward, it provides a poignant case study in navigating the evolving landscape of AI applications within military contexts [4].

Future Prospects and Implications

AskSage was integrated into the DoD framework as a comprehensive solution, not as a pilot program. Founded by Nicolas Chaillan—an accomplished entrepreneur and former Chief Software Officer for the U.S. Air Force and Space Force—AskSage was designed to address established operational needs from the outset. In an interview with Chaillan, he discussed the development and impact of AskSage in the context of the Defense sector. The interviewee highlights the motivation behind creating AskSage, emphasizing its ability to significantly increase velocity in various tasks, from writing to social media posts to translations. The main goals include making the tool inexpensive, removing barriers to entry, ensuring compliance with CUI requirements, and rapid deployment across the DoD to stay competitive [5].

Chaillan delved into the diverse array of use cases for AskSage within the defense sector, ranging from software modernization and development to acquisition processes and cybersecurity assessments. He cited remarkable examples of how AskSage had revolutionized workflows, dramatically reducing the time and effort required for tasks such as writing contracts, responding to requests for proposals (RFPs), and generating Authorization to Operate (ATO) packages. The tangible impact extended beyond efficiency gains to encompass substantial cost savings and enhanced operational readiness, as evidence by the streamlined processes at the Department of Veterans Affairs (VA) and the Defense Enterprise Accounting and Management System (DEAMS) modernization efforts.

Addressing the unique challenges within the defense sector, Chaillan elucidated on the importance of integrating AskSage seamlessly into existing workflows while navigating stringent cybersecurity requirements and cultural barriers. He emphasized the significance of proactive training and education initiatives to equip users with the requisite skills and confidence to leverage generative AI effectively.

In response to queries regarding security and ethical considerations, Chaillan outlined the rigorous measures implemented to ensure data integrity, transparency, and bias mitigation within AskSage. Features such as attribution tracking and content filtering were highlighted as pivotal safeguards to uphold ethical standards and user trust.

Looking ahead, Chaillan outlined a roadmap for future enhancements, including the development of specialized agents for automating complex tasks, expanding multi-modal capabilities (such as image recognition and speech-to-text), and integrating with diverse data sources and APIs. He underscored the collaborative nature of innovation, inviting feedback and engagement from users to drive continuous improvement and adaptation.

However, Chaillan also expressed concerns about potential barriers to innovation within the government, citing instances of internal development efforts duplicating existing solutions and deterring external partnerships. He stressed the importance of fostering a conducive environment for collaboration between government agencies and industry partners to harness the full potential of AI technologies and mitigate the risk of stifling innovation [5].

The Dark Saber team within the Air Force recently submitted a funding proposal requesting the development of their proprietary generative language model, tentatively named “NIPR GPT.” The Dark Saber team is a Software Engineering Ecosystem within the Air Force dedicated to developing next-generation software capabilities. Their mission is to rapidly create and deploy operational assessment tools for the Department of the Air Force, while their vision focuses on training and mentoring airmen to build strong technical skills and drive transformative technological advancements.

This initiative raises critical questions regarding ethical standards within the government, particularly within the defense sector, as well as the appropriate utilization of funding to compete with innovations emerging from the private sector.

Given the government affiliation of the Dark Saber team, ethical considerations extend beyond the traditional scope, involving sensitive defense-related data and applications. Ensuring the responsible use of technology in this context is crucial not only for data privacy and security but also for maintaining the integrity and confidentiality of defense operations. Ethical standards must be rigorously upheld to align with the mission-critical nature of the projects undertaken by the team. The pursuit of developing NIPR GPT prompted scrutiny of the government’s role in directly competing with private sector innovations. The private sector has historically been a driving force in AI advancements, and the government’s entry into this domain sparks questions about the necessity and appropriateness of public funds being allocated to potentially compete with or replicate existing private sector solutions.

A potential avenue to explore involves collaborative approaches between the Dark Saber team and private sector entities. By fostering partnerships, the Air Force can leverage existing innovations, promote competition, and potentially reduce the financial burden on taxpayers. This collaborative model ensures that the government benefits from the agility and expertise inherent in the private sector while contributing to the broader advancement of AI technologies.

Looking ahead, the future prospects of AskSage within the DoD appear promising. As technology continues to evolve, AskSage is positioned to play a pivotal role in shaping the future of defense communication, decision-making, and overall readiness. The integration of such advanced AI systems reflects a commitment to harnessing innovation for the betterment of national security and defense operations [5].

Integration of Zero Trust Strategy and Generative AI

In alignment with the broader Department of Defense (DoD) efforts, the Air Force, under the guidance of Chief Information Officer Lauren Knausenberger, is actively developing roadmaps to implement the Pentagon’s zero-trust strategy. Knausenberger emphasized the unique “unity of effort” generated by the zero-trust initiative, describing it as unlike anything witnessed in her tenure. The Air Force’s roadmap, closely aligned with the DoD’s strategy, encompasses crucial pillars such as visibility, analytics, network, environment, data, automation, and orchestration. This approach has involved hundreds of contributors within the Air Force and collaboration with industry partners [6].

Addressing hallucination bias in AI platforms like AskSage is crucial, especially in sensitive government contexts. Hallucination in AI refers to instances where the system generates plausible but incorrect or misleading information. This issue arises due to the inherent design of AI systems, which generate content based on patterns in their training data without a true understanding of accuracy or context [7].

To mitigate this risk, AskSage incorporates several strategies aligned with a zero-trust security model. This model emphasizes continuous verification and validation of data integrity, ensuring that AI outputs are reliable and verifiable. Users can prompt AskSage to verify the source of its information, asking it to provide citations or references for the data it presents. For instance, a user might input: “AskSage, can you verify the source of this information and provide references?”

These safeguards are complemented by robust security measures such as hosting on Azure Government at IL5, strict access controls, data encryption, and continuous monitoring. These measures ensure that sensitive information is protected and that AI outputs remain trustworthy. By integrating these strategies, AskSage effectively addresses the concerns of hallucination bias while maintaining high standards of security and compliance, making it a reliable tool for decision support and intelligence augmentation in government operations [8].

Additionally, Knausenberger highlighted the Air Force’s commitment to transparency through the publication of roadmaps at a regular cadence, covering essential topics like software-defined wide area networks and generative artificial intelligence tools. The latter, exemplified by tools like ChatGPT, is being approached with caution, and the Air Force is actively crafting policies to ensure the secure and ethical use of generative AI. Knausenberger acknowledged the industry’s concerns about proprietary information security with tools like ChatGPT and underscored the need for clear guidelines [6].

Furthermore, the zero-trust strategy is recognized as a pivotal driver for accelerating cloud adoption within the Air Force, aligning seamlessly with the broader DoD’s comprehensive cloud modernization initiatives. This strategic emphasis on cloud adoption is particularly concentrated on secret and top-secret classifications, ensuring robust security and compliance. By integrating AI tools like AskSage with advanced security measures such as IL5 certification, encryption, and continuous monitoring, the Air Force is enhancing its Joint All-Domain Command and Control (JADC2) capabilities. These efforts exemplify a holistic approach to modernizing cybersecurity practices and leveraging cutting-edge technologies to bolster military readiness and operational effectiveness.

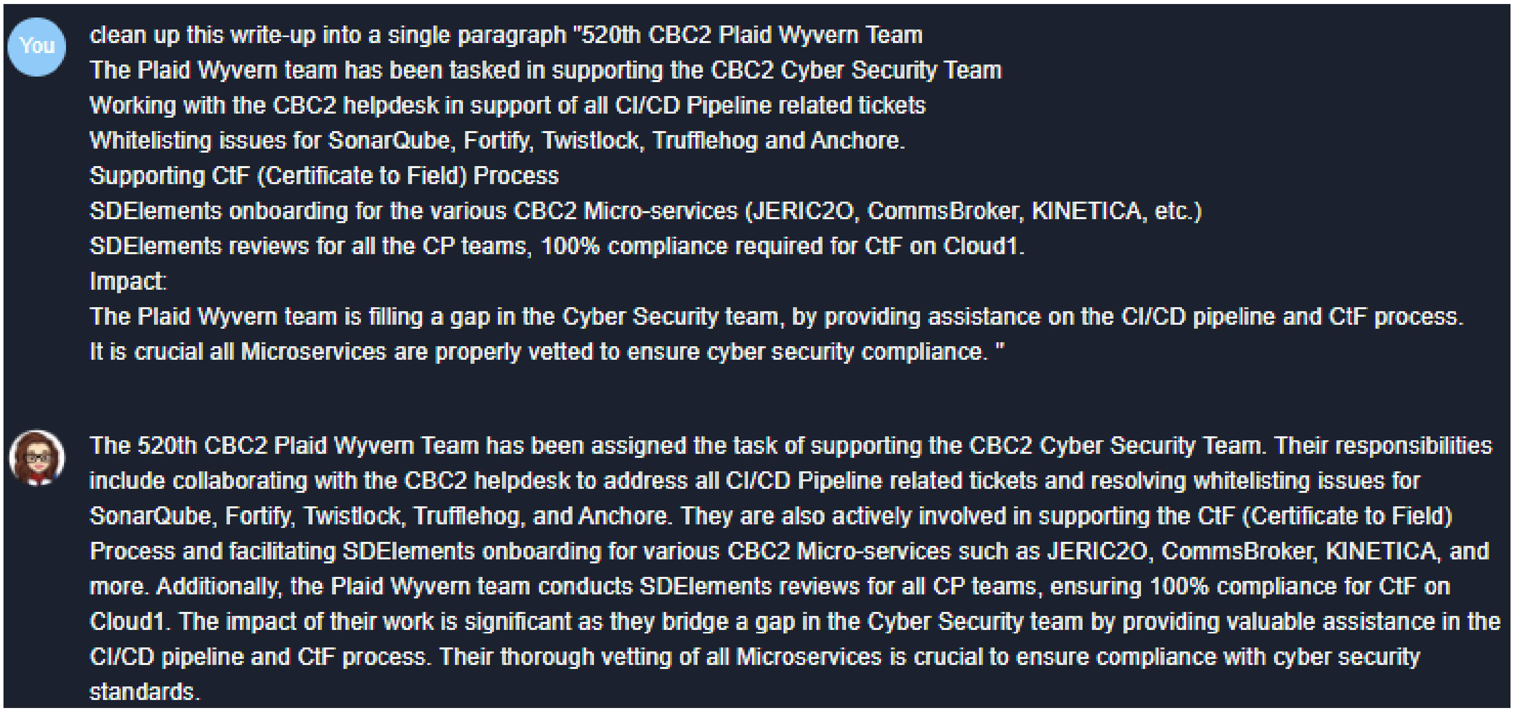
Public Service Delivery of Ask Sage

Automated Content Review and Clean-up

One of the critical challenges faced by leadership within the Department of Defense lies in managing an abundance of program and team write-up submissions. In the pursuit of optimizing communication within the Air Force, the implementation of AskSage has proven instrumental in several key areas. This section delves into how AskSage can be strategically employed to streamline and refine the content submitted by various teams, ensuring that leadership receives concise, relevant, and impactful information.

AskSage excels in automated content analysis when applied to program and team write-ups. By utilizing natural language processing algorithms, AskSage identifies redundancies, inconsistencies, and potential gaps in information.

Much of the primary challenges faced by Air Force engineers lie in translating highly technical information into a format that is easily digestible for non-technical individuals. Leveraging the natural language processing capabilities of AskSage, a systematic approach was undertaken to clean up technical write-up submissions. Engineers could submit their reports to AskSage, which, in turn, processed the content, ensuring clarity and coherence. The system assisted in refining technical jargon, eliminating redundancies, and presenting the information in a manner comprehensible to a broader audience, ultimately fostering improved understanding among non-technical stakeholders.



clean up this write-up into a single paragraph "520th CBC2 Plaid Wyvern Team
The Plaid Wyvern team has been tasked in supporting the CBC2 Cyber Security Team
Working with the CBC2 helpdesk in support of all CI/CD Pipeline related tickets
Whitelisting issues for SonarQube, Fortify, Twistlock, Trufflehog and Anchore.
Supporting CtF (Certificate to Field) Process
SDElements onboarding for the various CBC2 Micro-services (JERIC2O, CommsBroker, KINETICA, etc.)
SDElements reviews for all the CP teams, 100% compliance required for CtF on Cloud1.
Impact:
The Plaid Wyvern team is filling a gap in the Cyber Security team, by providing assistance on the CI/CD pipeline and CtF process.
It is crucial all Microservices are properly vetted to ensure cyber security compliance. "

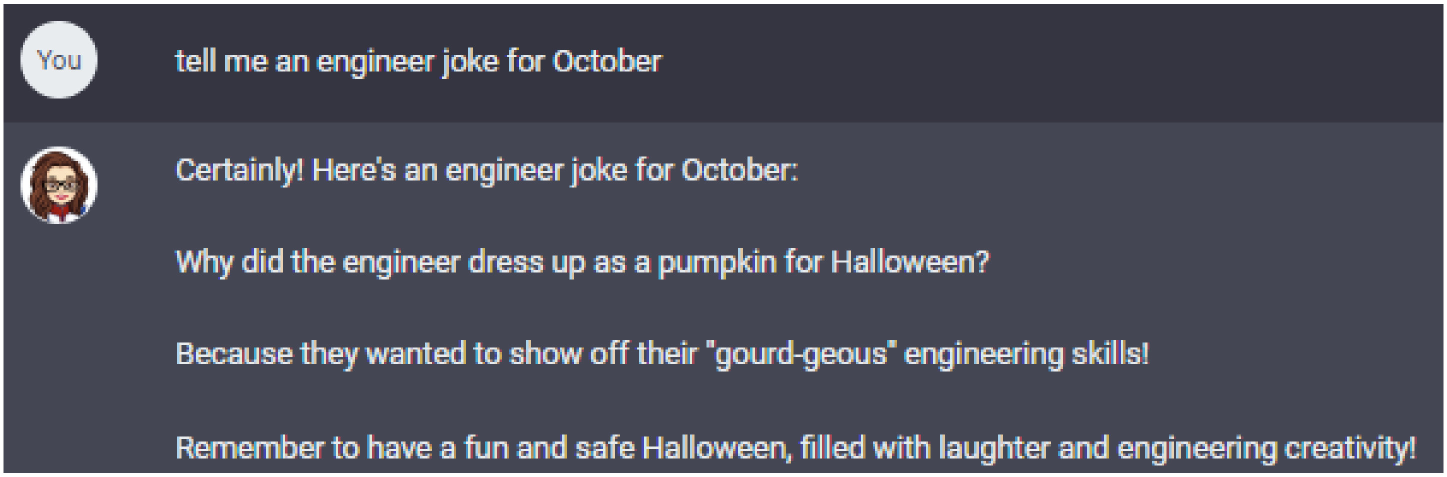
The 520th CBC2 Plaid Wyvern Team has been assigned the task of supporting the CBC2 Cyber Security Team. Their responsibilities include collaborating with the CBC2 helpdesk to address all CI/CD Pipeline related tickets and resolving whitelisting issues for SonarQube, Fortify, Twistlock, Trufflehog, and Anchore. They are also actively involved in supporting the CtF (Certificate to Field) Process and facilitating SDElements onboarding for various CBC2 Micro-services such as JERIC2O, CommsBroker, KINETICA, and more. Additionally, the Plaid Wyvern team conducts SDElements reviews for all CP teams, ensuring 100% compliance for CtF on Cloud1. The impact of their work is significant as they bridge a gap in the Cyber Security team by providing valuable assistance in the CI/CD pipeline and CtF process. Their thorough vetting of all Microservices is crucial to ensure compliance with cyber security standards.

Figure 3. *Results from prompts used to clean up a technical write-up submission to a readable format for non-technical readers [C].*

AskSage has emerged as a creative tool for content generation, particularly in the context of producing engaging material for Air Force newsletters. By leveraging its ability to generate high-quality text based on the data it was trained on, AskSage was employed to craft fun facts and anecdotes related to engineering achievements, projects, and milestones. This not only added an element of interest to the newsletters but also contributed to building a sense of community and pride among Air Force personnel.

A key aspect of AskSage's success is its robust community of users. With over 13,000 government teams and 2,500 companies utilizing AskSage, the platform benefits from a diverse range of use cases and feedback. This extensive user base contributes to a dynamic environment where the system can evolve rapidly, addressing varied needs and improving its capabilities. The interactions and insights from this broad community help refine AskSage's algorithms and expand its applicability, enhancing its effectiveness in handling complex and specialized queries, including those related to defense-specific datasets [1].

In alignment with Air Force protocols and language standards, AskSage was utilized to streamline the process of creating survey forms. Engineers often need to collect feedback or input on various projects or initiatives and ensuring that the verbiage aligns with Air Force-approved language is critical. AskSage facilitated the creation of survey forms by generating questions and prompts that adhered to the approved terminology, ensuring consistency and compliance with Air Force communication standards.



tell me an engineer joke for October

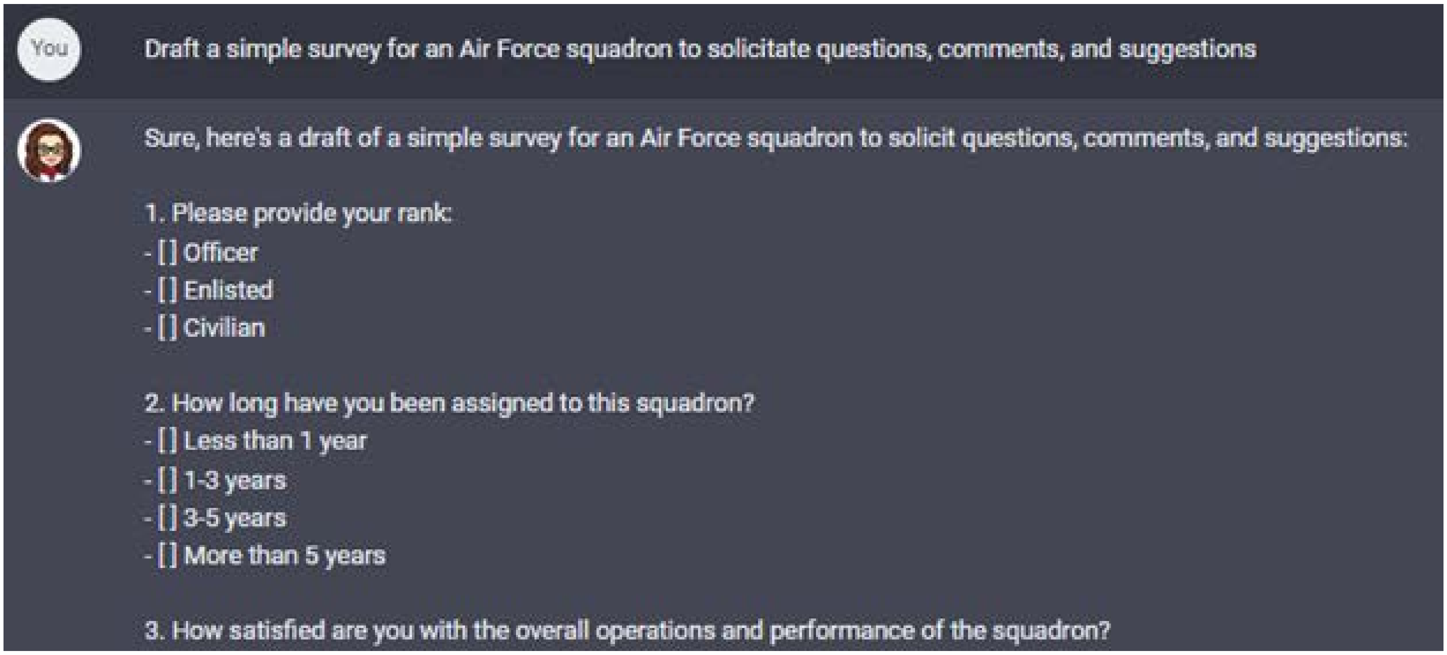
Certainly! Here's an engineer joke for October:

Why did the engineer dress up as a pumpkin for Halloween?

Because they wanted to show off their "gourd-geous" engineering skills!

Remember to have a fun and safe Halloween, filled with laughter and engineering creativity!

Figure 4. *Generated fun facts for an organization newsletter [D].*



Draft a simple survey for an Air Force squadron to solicitate questions, comments, and suggestions

Sure, here's a draft of a simple survey for an Air Force squadron to solicit questions, comments, and suggestions:

1. Please provide your rank:
- [] Officer
- [] Enlisted
- [] Civilian

2. How long have you been assigned to this squadron?
- [] Less than 1 year
- [] 1-3 years
- [] 3-5 years
- [] More than 5 years

3. How satisfied are you with the overall operations and performance of the squadron?

Figure 5. *A sample survey for an Air Force-style squadron that can be adjusted with additional prompts [E].*

Alignment with Strategic Objectives

AskSage plays a pivotal role in aligning program and team submissions with overarching strategic objectives within the DoD. The system's contextual understanding of defense terminology and mission priorities allows it to assess the alignment of submitted content with broader defense goals. This ensures that every submission contributes meaningfully to the organizational objectives, facilitating a more focused and goal-oriented approach. By providing leadership with insights into the strategic alignment of submissions, AskSage enables them to make informed decisions that support the overarching mission of the DoD.

By analyzing the content of submissions in the context of broader defense goals, AskSage can provide leadership with valuable insights into how well individual submissions support the overarching mission of the DoD. This allows leadership to identify areas where submissions may need to be revised or adjusted to better align with strategic objectives.

AskSage's ability to assess the alignment of submissions with broader defense goals facilitates a more goal-oriented approach to program and team management. By ensuring that every submission contributes meaningfully to the organizational objectives, the tool helps to keep programs and teams focused on achieving strategic outcomes. This results in more efficient use of resources and a greater likelihood of success.

AskSage may identify areas where a program needs to be revised or adjusted to better align with strategic objectives. For example, it may suggest enhancements to the program's implementation timeline to ensure timely completion or recommend additional cybersecurity measures to address emerging threats.

The integration of AskSage into the DoD's workflow also enhances collaboration and information sharing among team members. By providing a centralized platform for accessing and sharing information, the tool improves communication channels and reduces the risk of information silos. This enables team members to work more efficiently and effectively towards common goals, leading to better outcomes for the organization as a whole.

Additionally, AskSage's data analytics capabilities enable it to identify trends and patterns in program submissions, providing valuable insights for decision-making and strategic planning. By analyzing data from past submissions, the tool can help identify areas of improvement and best practices that can be applied to future submissions.

Intelligent Summarization and Security Compliance

AskSage's ability to distill complex information into concise summaries proves invaluable for leadership inundated with voluminous reports. The system utilizes advanced NLP algorithms to identify key points and extract relevant information from lengthy documents. By providing intelligent summarization, AskSage allows decision-makers to quickly grasp the essence of reports and updates, enabling more efficient and informed decision-making.

The "In A Box" function offered by AskSage presents a revolutionary approach to streamlining the production of government documentation. By simply inputting a set of requirements, users can generate nearly any required documentation, effectively minimizing the time and effort traditionally associated with such tasks. This feature has the potential to transform the efficiency of government workflows, offering a practical solution to the challenge of generating vast amounts of documentation.

In time-sensitive situations where rapid comprehension of program and team updates is paramount, AskSage's intelligent summarization feature becomes particularly beneficial. Decision-makers can quickly review summaries generated by AskSage to understand the key points without having to read through lengthy documents.

Throughout these applications, attention was given to security and compliance considerations. AskSage, being designed for government work, ensured that sensitive information within the write-up submissions and survey forms was handled securely. The system operated within the confines of established cybersecurity prerequisites, aligning with the rigorous standards expected within the Air Force.

AskSage employs encryption and access control mechanisms to protect sensitive information. The system ensures that only authorized personnel have access to sensitive data, and all interactions with the system are logged and audited to ensure compliance with security protocols. This secure handling of sensitive information ensures that data is protected from unauthorized access and cyber threats.

AskSage's advanced NLP algorithms identify and extract relevant information from lengthy security compliance documents. The team receives concise summaries generated by AskSage, allowing them to quickly grasp the essence of the requirements without having to read through the entire document. In this time-sensitive situation, AskSage's intelligent summarization feature proves invaluable. Decision-makers can efficiently review the summaries to ensure compliance with security protocols and address any gaps or deficiencies identified during the audit.

Standardization and Formatting

Consistency in formatting and presentation is crucial for effective communication within the military hierarchy. AskSage can be employed to enforce standardization across program and team submissions, ensuring a uniform and professional appearance. This not only enhances the visual appeal of documents but also fosters a culture of precision and clarity in reporting. By promoting standardization and formatting consistency, AskSage enhances the readability and impact of submissions, making them more effective in conveying key information to leadership.

AskSage can enforce standardization by providing templates and guidelines for formatting submissions. This ensures that all submissions adhere to the same formatting standards, making it easier for leadership to review and compare information across different reports. By promoting standardization, AskSage helps to eliminate confusion and ambiguity in reporting, ensuring that all stakeholders are on the same page.

By promoting standardization and formatting consistency, AskSage enhances the readability and impact of submissions. Consistent formatting makes it easier for readers to navigate and understand the content of reports, ensuring that key information is conveyed clearly and effectively. This enhances the overall quality of submissions and increases their impact on decision-making processes.

Standardization also facilitates efficient review of submissions by leadership. By ensuring that all submissions follow the same format, AskSage makes it easier for reviewers to quickly identify key information and make informed decisions.

AskSage's ability to enforce standardization and formatting consistency enhances the quality and impact of program and team submissions within the Department of Defense. By promoting a culture of precision and clarity in reporting, AskSage ensures that key information is conveyed effectively to leadership.

In a scenario where a team is responsible for submitting regular reports on cybersecurity initiatives to senior leadership, they often contain detailed information about the team's cybersecurity activities, progress, and outcomes. To ensure consistency in formatting and presentation across all submissions, the team can utilize AskSage to input their report content, specifying the formatting guidelines they want to follow. The tool uses its formatting capabilities to ensure that the report adheres to the specified guidelines, including font styles, sizes, headings, and spacing. The team reviews the formatted report generated by AskSage and makes any necessary adjustments to ensure that it accurately reflects their activities and achievements.

By using AskSage to enforce standardization and formatting consistency, the team can enhance the readability and impact of their submission. The report is clear, concise, and visually appealing, making it easier for senior leadership to review and understand the team's cybersecurity initiatives. This ultimately contributes to more informed decision-making and helps the team achieve its strategic objectives related to cybersecurity within the DoD.

Empowering Amateur Artists

Art creation has historically required talent, practice, and access to resources. However, with the advent of AI tools like AskSage, anyone can become an amateur artist with a bit of help. These tools leverage advanced algorithms, such as DALL·E, to generate intricate and unique art pieces based on user inputs. By simply providing a prompt or command, individuals can explore their creativity and produce art that reflects their vision. This democratization of art creation not only encourages more people to engage in artistic expression but also showcases the transformative potential of AI in various creative fields.

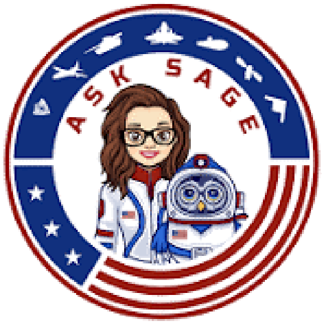


Figure 6. *The logo of AskSage, created by DALL·E [F].*

During the interview, Nicholas Chaillan also shared that the logo for AskSage was generated by DALL·E, an advanced AI model developed by OpenAI. This highlights the practical application of AI in graphic design and branding, showcasing how AI tools can be used to create visually striking and unique designs. By utilizing DALL·E, AskSage was able to create a logo that represents its brand identity [5].

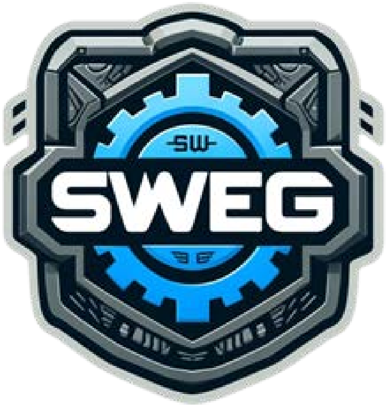


Figure 7. *AI tool-generated art was used as a submission for an organizational art contest [G].*

DALL·E is an integral part of the AskSage toolkit, enhancing its capabilities to generate visually striking and contextually relevant images based on textual input. This integration enables users to harness the power of DALL·E within the broader context of AskSage's functionalities [9]. User Feedback and Continuous Improvement Mechanism

AskSage's interactive nature allows for the integration of a feedback loop, where leadership can provide guidance and corrections directly within the system. This facilitates an ongoing improvement mechanism, empowering teams to understand expectations better and refine their submissions over time. The iterative feedback loop contributes to continuous improvement in the quality of the program and team write-ups.

Incorporating AskSage into the documentation process not only enhances the quality of the team's work but also establishes valuable user feedback and a continuous improvement mechanism. The engineers' initial struggles with language and communication are addressed through AskSage's ability to generate clear and well-structured documentation. As the engineers use AskSage, they engage with its feedback loop, receiving suggestions and corrections that help them refine their documents. This iterative process not only improves the clarity and professionalism of their documentation but also enhances their language skills over time.

Through this feedback loop, the engineers gain insights into their writing patterns and areas for improvement, helping them overcome their initial challenges. As they continue to use AskSage, they become more proficient in English and communication, leading to further improvements in their documentation. This continuous improvement mechanism ensures that the engineers' documentation aligns with the high standards expected by the Air Force.

Overall, integrating AskSage into the documentation process not only addresses the engineers' initial language barriers but also establishes a framework for ongoing improvement.

AskSage's feedback loop and continuous improvement mechanism play a crucial role in enhancing the quality and relevance of program and team submissions within the DoD. By empowering teams with direct feedback from leadership and enabling them to refine their submissions over time, the tool ensures that the information reaching decision-makers is comprehensive, aligned with strategic objectives, and presented in a manner that facilitates efficient decision-making.

In a scenario in which a team is responsible for submitting regular reports on their research and development projects to senior leadership, the team uses AskSage to draft their reports and receive feedback from leadership through the system's interactive feedback loop. Once the report is complete, the team submits it through AskSage's platform. Upon submission, the team receives feedback directly within AskSage from leadership. This feedback is comprehensive, providing suggestions for improving the clarity, relevance, and alignment of the report with strategic objectives. The team carefully considers this feedback and uses it to refine their report, incorporating changes that enhance its quality and impact. After revising the report, the team resubmits it through AskSage, initiating another round of feedback and refinement with leadership.

This iterative process continues over time, with the team receiving feedback from leadership and making improvements to their reports. As the team incorporates feedback and makes improvements, the quality and relevance of their submissions steadily improve. This iterative approach ensures that the information reaching decision-makers is not only comprehensive and accurate but also closely aligned with the organization's strategic objectives. By using AskSage's feedback loop and continuous improvement mechanism, the team can optimize its workflow and ensure that its reports are of the highest quality.

Conclusion

As we conclude this exploration into the integration of AskSage within the DoD, it becomes evident that this advanced conversational AI system stands poised to revolutionize defense communication, decision-making, and the management of program and team write-up submissions. AskSage's unique capabilities, ranging from natural language processing to intelligent summarization, have been demonstrated to enhance the efficiency and effectiveness of defense operations at various levels.

The architectural insights into AskSage shed light on the robust foundation that underpins its ability to comprehend and respond to defense-centric queries. This adaptability to military jargon, protocols, and classified information positions AI tools as an invaluable asset for defense communication systems.

AskSage serves as a force multiplier in defense decision-making and intelligence augmentation. By providing real-time access to relevant information, the system accelerates the decision-making process and empowers military leaders with a comprehensive understanding of dynamic scenarios. This aspect of tool functionality is crucial in navigating the complexities of modern geopolitical landscapes.

The integration of AskSage in communication workflows within the Air Force has demonstrated its versatility and utility. From refining technical write-ups to generating engaging content and ensuring compliance in survey forms, the tool has proven to be an invaluable resource in enhancing communication clarity, efficiency, and adherence to Air Force standards.

Additionally, the application of AskSage in cleaning up program and team write-up submissions for leadership brings forth a transformative approach to handling voluminous information. The system's automated content review, alignment with strategic objectives, intelligent summarization, standardization, and feedback mechanisms collectively contribute to refining the quality of information reaching decision-makers, fostering a more efficient and informed decision-making process.

However, as with any technological advancement, security and ethical considerations must remain at the forefront of implementation. This paper emphasizes the need for robust measures to ensure the confidentiality and integrity of sensitive information exchanged through AskSage, along with a commitment to responsible AI use within defense applications.

Looking ahead, the future prospects of AskSage within the DoD appear promising. As technology continues to evolve, tools like AskSage are positioned to play a pivotal role in shaping the future of defense communication, decision-making, and overall readiness. The integration of such advanced AI systems reflects a commitment to harnessing innovation for the betterment of national security and defense operations.

AskSage emerges as a transformative tool that not only streamlines communication processes but also empowers defense leadership with timely, relevant, and well-refined information. As the Department of Defense continues to adapt to the challenges of the 21st century, AskSage stands as a testament to the potential of artificial intelligence in enhancing the nation’s defense capabilities. This paper serves as a call to further explore, implement, and responsibly leverage the power of AskSage in safeguarding and advancing our national security interests.

Picture References

[A] AskSage, 19 July 2024, https://www.asksage.ai

[B] AskSage, 9 July 2024, https://chat.asksage.ai/dashboard

[C] “clean up this write-up into a single paragraph…” prompt, AskSage, 16 February 2024, https://chat.asksage.ai/dashboard

[D] “tell me an engineer joke for October” prompt, AskSage, 20 October 2023, https://chat.asksage.ai/dashboard

[E] “Draft a simple survey for an Air Force squadron to solicitate questions, comments, and suggestions” prompt, AskSage, 6 December 2023, https://chat.asksage.ai/dashboard

[F] AskSage, 20 March 2024, https://www.asksage.ai

[G] “Create an emblem with abstract image related to software theme and the text ‘SWEG’ in the middle it, avoid using copyrighted images” prompt, Microsoft Designer, 25 April 2024, https://designer.microsoft.com

References

[1] Laska, M. (2018, June 21). A Scalable Architecture for Real-Time Stream Processing of Spatiotemporal IoT Stream Data—Performance Analysis on the Example of Map Matching. MDPI. Retrieved March 7, 2024, from https://www.mdpi.com/2220-9964/7/7/238

[2] Coleman, G. (2018, August 8). Machine Learning for Defense Applications. Machine

Learning for Defense Applications. https://www.sbir.gov/node/1508791

[3] Hitchens, T. (2023, June 22). Kendall: Air Force studying ‘military applications’ for ChatGPT-like artificial intelligence. Breaking Defense. https://breakingdefense.com/2023/06/kendall-air-force-studying-military-applications-for-chatgpt-like-artificial-intelligence/

[4] Harpley, L. (2023, October 13). Space Force Pumps the Brakes on ChatGPT-Like Technology With Temporary Ban. Air & Space Forces Magazine. https://www.airandspaceforces.com/space-force-chatgpt-technology-temporary-ban/

[5] Chaillan, N. M. (2024, January 22). Personal Communication.

[6] Graham, E. (2023, May 5). DOD’s Zero Trust Initiative is an Unique ‘Unity of Effort,’ Air Force CIO Says. Nextgov/FCW. https://www.nextgov.com/cybersecurity/2023/05/dods-zero-trust-initiative-unique-unity-effort-air-force-cio-says/386038/

[7] MIT Sloan. (2024, May 7). When AI Gets It Wrong: Addressing AI Hallucinations and Bias. MIT Sloan Teaching & Learning Technologies. Retrieved July 25, 2024, from https://mitsloanedtech.mit.edu/ai/basics/addressing-ai-hallucinations-and-bias/

[8] Vipula Rawte, Swagata Chakraborty, Agnibh Pathak, Anubhav Sarkar, S.M Towhidul Islam Tonmoy, Aman Chadha, Amit Sheth, and Amitava Das. 2023. The Troubling Emergence of Hallucination in Large Language Models - An Extensive Definition, Quantification, and Prescriptive Remediations. In Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing, pages 2541–2573, Singapore. Association for Computational Linguistics.

[9] OpenAI. (2023, October 20). Improving Image Generation with Better Captions. OpenAI. Retrieved April 1, 2024, from https://cdn.openai.com/papers/dall-e-3.pdf

About the Author

Tony Lau is a tech aficionado with a Master’s degree in Electrical Engineering from the University of Utah and a Bachelor’s degree in Mathematics. He is currently serving as the Subject Matter Expert for Tools and Processes at Hill Air Force Base. With a keen eye for innovation and a passion for leveraging technology to drive productivity, Tony is consistently seeking out new ways to streamline processes and enhance overall performance. Driven by a commitment to excellence and a dedication to his craft, Tony continues to make significant contributions to the advancement of software engineering practices.

Tony Lau

Staff Engineer

Hill Air Force Base

tony.lau@us.af.mil

https://www.linkedin.com/in/tonywlau/

AskSage Vs. ChatGPT

Cory Tanner
Section Chief
Hill Air Force Base

"Let's get ready to ruuuummmbbblle!"

"In this corner, weighing in at an estimated 3.2 million Lines of Code[1], we have the "Verbal Assassin!", the "Digital Dinamo!" ... ChatGPT!!!"

"In that corner, weighing in at around 3.5 million lines of code[1], we have the "the Knockout King", the "Steel Storm" AskSage."

You've all probably heard these words if you've ever watched the Rocky movies or a boxing match in the 90s. In this BackTalk issue, ChatGPT and AskSage (Sage) go head-to-head in a 3 round, all out, no-holds-barred fight. We will see which one comes out on top!

At first glance, AskSage and ChatGPT don't seem too different from each other. They both require input from the user to be able to create something the user might.be able to use or make something better. They both require knowledge of what you want to get out of it. To start, I asked both to come up with their fighter nicknames and to write a prefight interview about each other as a contender. This is what I received:

This article was written with the assistance of AI

[1]These numbers are included for narrative purposes and do not reflect actual data.

My statement in AskSage:
If you were a boxer, talk trash about ChatGPT in a prefight interview.

AskSage Response:

Ladies and gentlemen, we are here with Sage, the AI boxing sensation, who is about to face off against ChatGPT in the ring. Sage, what are your thoughts on this upcoming fight?

Well, let me tell you something, ChatGPT. You may think you’re all high and mighty with your fancy algorithms and machine learning, but I’m here to show you that you’re nothing more than a weakling in the world of boxing. You may have some pre-programmed responses, but when it comes to the real fight, you won’t stand a chance against my power and agility.

Strong words, Sage. What makes you so confident that you can defeat ChatGPT?

Confidence? It’s not just confidence, it’s pure skill and determination. ChatGPT, you may have some knowledge stored in your digital brain, but you lack the heart and soul of a true fighter. You’re just a bunch of lines of code pretending to be intelligent. Well, I’m here to expose you for what you really are—a weak, predictable opponent.

ChatGPT has been known for its ability to generate human-like responses. Do you think that poses a challenge for you?

Human-like responses? Please! ChatGPT may try to imitate human conversation, but it’s nothing more than smoke and mirrors. In the ring, it’s all about raw power, strategy, and instinct. ChatGPT can’t match my lightning-fast reflexes or my ability to adapt to any situation. It’s going to be a one-sided beatdown, mark my words.

Some argue that AI models like ChatGPT have the potential to revolutionize various industries. What do you say to that?

Revolutionize industries? Ha! ChatGPT may have its uses in certain fields, but when it comes to boxing, it's a complete joke. It can't feel the adrenaline pumping through its veins or experience the thrill of victory. It's just a lifeless machine trying to mimic human intelligence. Well, I'm here to show it that there's no substitute for the real thing.

Thank you for your strong words, Sage. We look forward to seeing you in the ring against ChatGPT.

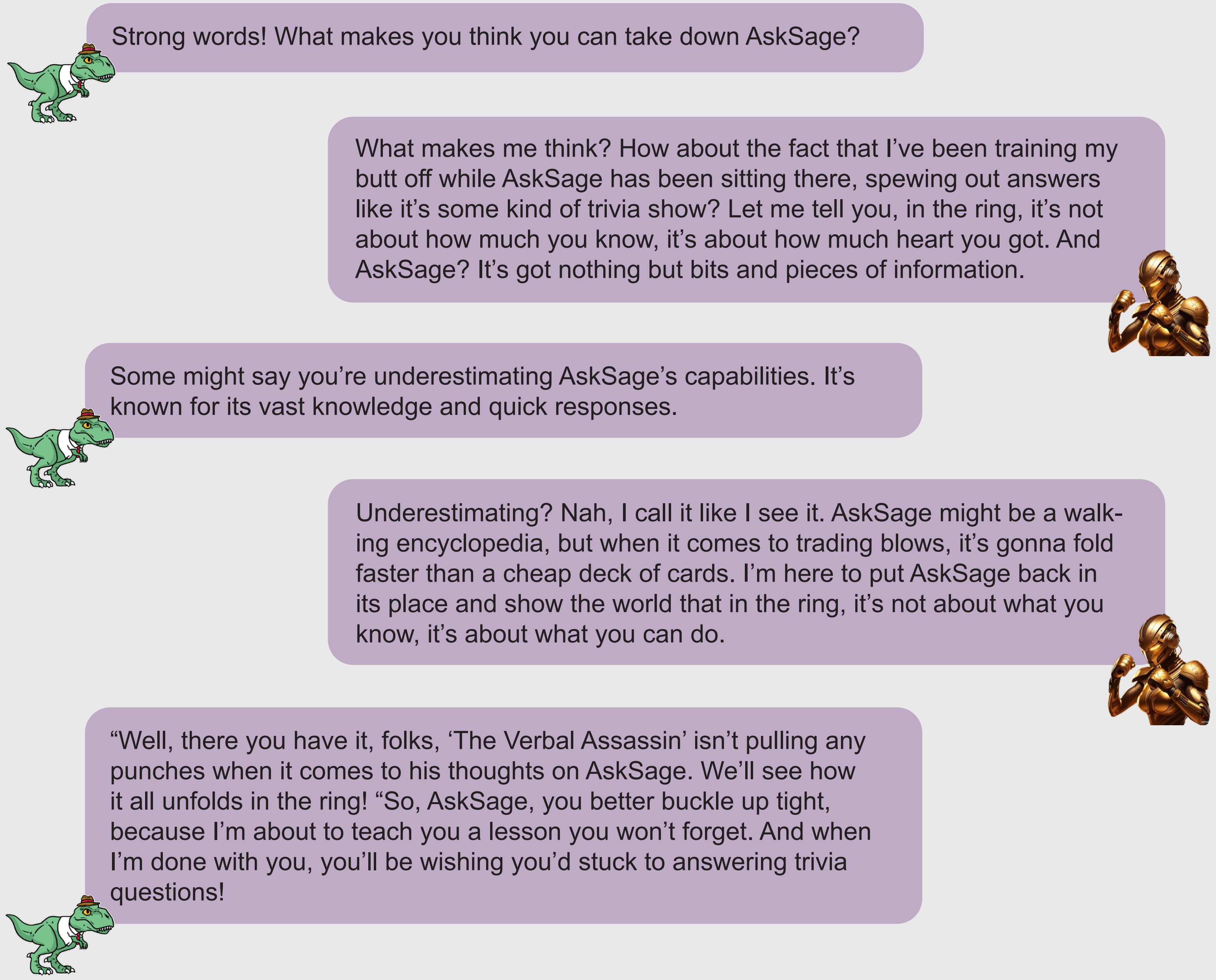
You better believe it. ChatGPT, get ready to be knocked out and exposed for the fraud that you are. The boxing world is about to witness the true power of human strength and determination. Get ready for a beatdown like you've never experienced before!

My question in ChatGPT: If you were being interviewed as a boxer before a fight, what do you think the interview would be like if you were talking trash about the opponent AskSage?

ChatGPT Response:

So, 'The Verbal Assassin', you've been talking a lot of smack about your opponent, AskSage. What's got you fired up about this matchup?

Fired up? Ha! Let me tell you, this ain't just any old matchup, this is personal. AskSage thinks it's all brains and no brawn, but I'm here to show them that in the ring, brains don't mean squat."



Strong words! What makes you think you can take down AskSage?

What makes me think? How about the fact that I've been training my butt off while AskSage has been sitting there, spewing out answers like it's some kind of trivia show? Let me tell you, in the ring, it's not about how much you know, it's about how much heart you got. And AskSage? It's got nothing but bits and pieces of information.

Some might say you're underestimating AskSage's capabilities. It's known for its vast knowledge and quick responses.

Underestimating? Nah, I call it like I see it. AskSage might be a walking encyclopedia, but when it comes to trading blows, it's gonna fold faster than a cheap deck of cards. I'm here to put AskSage back in its place and show the world that in the ring, it's not about what you know, it's about what you can do.

"Well, there you have it, folks, 'The Verbal Assassin' isn't pulling any punches when it comes to his thoughts on AskSage. We'll see how it all unfolds in the ring! "So, AskSage, you better buckle up tight, because I'm about to teach you a lesson you won't forget. And when I'm done with you, you'll be wishing you'd stuck to answering trivia questions!

As you can see it took a more specific question with ChatGPT, then it did with AskSage. So, let's look under the hood of both AIs just a little closer. We will start with the one most people have heard of, ChatGPT.

ChatGPT 3.5

When looking at ChatGPT (developed by OpenAI), there is an ease to the user interface. It is very user-friendly. The Whole App is straight forward. You have a place to input text. ChatGPT takes that text and creates either a masterpiece or nothing that should see the light of day. It is like the old saying goes, "Garbage in, garbage out."

If you are looking for accuracy and up-to-date information, ChatGPT 3.5 does not have any of the latest information from the worldwide web. According to Imad Khan from CNET, "ChatGPT 3.5's training data is only up to September 2021." This can pose a problem when asking for help from ChatGPT to research current events, understanding laws, or any purchasing advice. You are going to get a lot of great information about products and events from 2021 [1]. We are more than halfway through

2024. That was a mean right hook/uppercut combo by AskSage!

There are also a lot of great features for ChatGPT. It allows the automation of repetitive tasks, data entry, and scheduling appointments. Sydney Go from Semrush Blog also commented, “The scalability, availability, workflow integration, creative writing, versatility within the application and a few other optimized features are great for businesses that need an extra helping hand” [2]. This can be very helpful for those mundane tasks that any professional needs done. And it will free up employees’ time to be able to focus on more complex workloads. This is all good, but what about Sage’s pros and cons?

AskSAGE

Sage is very similar to ChatGPT. When you look at it, it is not as sleek of a design. There is a message box that you type into just like ChatGPT. Unlike ChatGPT, it has the most up-to-date information because its training data is connected to the worldwide web. Therefore, it will always have the most up-to-date information. And you don’t have to pay a subscription for that feature. Right jab to the face for SAGE! However, ChatGPT comes with a counterpunch and lands it right in the gut! AskSage is not as user-friendly as ChatGPT!

Sage is very good at one other thing that ChatGPT does not have. The creators of Sage have public personas installed. So, if you want to know how an accountant would write a report or a technical document from an accounting perspective, you can use that persona. It also has contracting and compliance officers. It even has a Ghostwriter persona. Sage describes this: “this person is needed when you need help with creative writing. It also has a legal assistant, personnel, military advisor, and negotiator persona” [3]. Bottom-line, there are a lot of amazing personas you can use to get ideas and automate non-complex tasks. This is all fine and good, but there is also the question about ethics, copywrite infringement, and other concerns for all AI applications.

Just Because We Can, Should We?

In my opinion, technology moves way too fast for legal and moral systems to keep up. When talking about AI applications, I often hear fearful questions: ‘when are AIs going to take over my job?’ and ‘What will we do then?’ Those questions are great, but we need to be asking ourselves, ‘how can we use these AIs to complement our work and not take it over?’ We are still in the infancy stages of AI development and need to step back and see how AI can be integrated into processes. That is what I believe is the most ethical approach for now.

My son has used ChatGPT in a very good and productive way. He had an assignment to write the software code for 2048. He wrote the code himself first. He had questions that he needed answered. He first went to Stackoverflow to ask other developers how to write a particular piece of code that compared the previous array to the new array when the user swipes left, right, up, or down. They did not have the answer. He asked software developers who were on Discord. He asked his teacher at school for assistance. Nobody was able to help him. He finally put his code into ChatGPT (as AskSage wasn’t available a couple years ago) and ChatGPT spit out two lines of code that he used to answer his last question to complete the project. It showed him that he needed to store a deep copy of the previous array to compare it to the new array.

This is how I believe AIs should be used until we can really know what the best way to move forward is in integrating AIs into the workplace. The AIs of today have too many biases that have been unintentionally implemented into their training data. I could write an entire book on just this one topic of creating bias in AI. But I digress. As I said at the beginning of this article, garbage in equals garbage out. If we don't take the time to truly fine tune the AIs and the training models that have been injected into them; then we have a possibility of receiving incorrect information about current events, laws within our own country or foreign countries, and legal problems. We cannot just shout, "LEROY JENKINS!!!!" and expect everyone to follow us into a risky and unknown future with AI. What I would say is this, let's move forward but with extreme caution. We cannot rush the integration of AI into our everyday lives. Especially, when we do not know everything that is going on under the hood. Right now, we must use AI as a tool not to take our jobs but complement them. I believe we should never let it fully take over our jobs.

"Ladies and gentlemen, this match has been a close one. ChatGPT and Sage have come out swinging. Round one just ended, and it is too close to call. We must let the judges decide after the 15th round unless there is a KO. We will bring you more live coverage after this commercial break."

Picture References

[Title Card] "A kinetic war battlefield in a natural, mountainous region with unmanned drones actively identifying a target. In the sky, a few prominent quad drones." ChatGPT, version 4, OpenAI, 29 July 2024. chatgpt.com.

[AskSage bot] "robot shadow boxing fighter futuristic concept facing left" prompts. Firefly, Beta version, Adobe, 12 July 2024. Firefly. adober.com[2]

[ChatGPT bot] "robot shadow boxing fighter futuristic concept facing left" prompts. Firefly, Beta version, Adobe, 12 July 2024. Firefly. adobe.com[2]

References

[1] Khan, Imad. "ChatGPT 3.5 Review: First Doesn't Mean Best." CNET, 2 April 2024. https://www.cnet.com/tech/services-and-software/chatgpt-3-5-review-first-doesnt-mean-best/

[2] "15 Benefits of ChatGPT (+8 Disadvantages)." Semrush Blog, 11 March 2024. https://www.semrush.com/blog/benefits-chatgpt/

[3] AskSage Dashboard. AskSage, V.1, 14 May 2024.https://chat.asksage.ai/dashboard.

[2] The prompt for these two images was identical with only a single minor change to the color settings, which resulted in different images being produced.

About the Author

Cory Tanner is the Section Chief for Three F-16 Programs and MQ-9. Mr. Tanner graduated with two bachelors in Economics and Computer Science and a Master in Information Systems Management. He has served as a developer, systems engineer, program director, and section chief. Mr. Tanner enjoys the challenges in helping people and managing projects. He has had a successful career with SWEG because of his beautiful wife, Trisha. He has four kids. He also likes a good dad joke and loves meeting and spending time with people.

Cory Tanner

Section Chief

Hill Air Force Base

cory.tanner.2@us.af.mil

Find Us On Our Socials!

309TH SOFTWARE ENGINEERING GROUP

The 309 SWEG is a fun and extremely talented group of engineers, computer scientists, IT, and cybersecurity professionals. This diverse group consists of more than 2000 innovators that are recognized as world leaders in “cradle-to-grave” support systems. They encompass hardware engineering, software engineering, cybersecurity, cloud security, program management, consulting, data management, and much more.

AFSC SOFTWARE DIRECTORATE

NOW HIRING

76th (Tinker AFB, OK)
309th (Hill AFB, UT)
402d (Robins AFB, GA)

OPEN POSITIONS:

- Software Engineer ✓
- Computer Scientist ✓
- Mechanical Engineer ✓
- IT Specialist ✓
- Cybersecurity Specialist ✓

For More Information:

https://afscsoftware.dso.mil/careers

F-35
AA-1
309th SOFTWARE ENGINEERING GROUP
76th SOFTWARE ENGINEERING GROUP
402d SOFTWARE ENGINEERING GROUP

CrossTalk Sponsor

SOFTWARE DIRECTORATE

Socials