

2026
DEFENSE HEALTH INFORMATION
TECHNOLOGY SYMPOSIUM

AUGUST 11-13, 2026
HYATT REGENCY NEW ORLEANS
NEW ORLEANS, LA

AI First! Preparing the DHA Workforce for Responsible AI Adoption

Jesus J Caban, PhD Chief Data and Analytics Officer, DHA

CDR Darshan Thota, Chief Medical Informatics Officer, Joint Trauma System

Lt Col David Alt, Chief of Emerging Capabilities, EIDS PMO, PEO DHMS



Speaker Introduction



**Jesus J Caban,
PhD**

Chief Data & Analytics
Officer (CDAO)
DHA



**CDR Darshan
Thota**

Chief Medical Informatics
Officer, Joint Trauma System



Lt Col David Alt

Chief, Data & Analytics
Advisory
EIDS PMO, PEO DHMS

Agenda

- 1** DoW AI Strategy
- 2** Why does AI fail in Healthcare?
- 3** Prompt Engineering 101
- 4** Live Demo
- 5** Q and A

Learning Objectives

- 1** Introduce what AI First means to the MHS
- 2** Understand why AI Fails
- 3** Review the best practices of Prompt Engineering
- 4** Discuss the importance of identifying hallucinations

DoW AI Strategy

- SECWAR released the AI Strategy, directing the Department to become an AI-first organization and to accelerate the adoption of AI to maintain our data and operational advantage.
- This directive applies to all DoD Components, including DHA.



SECRETARY OF WAR
1000 DEFENSE PENTAGON
WASHINGTON, DC 20301-1000

JAN - 9 2026

MEMORANDUM FOR SENIOR PENTAGON LEADERSHIP
COMMANDERS OF THE COMBATANT COMMANDS
DEFENSE AGENCY AND DOW FIELD ACTIVITY DIRECTORS

SUBJECT: Artificial Intelligence Strategy for the Department of War

Accelerating America's Military AI Dominance

President Trump makes clear in Executive Order 14179, "It is the policy of the United States to sustain and enhance America's global Artificial Intelligence (AI) dominance in order to promote human flourishing, economic competitiveness, and national security." In the national security domain, AI-enabled warfare and AI-enabled capability development will re-define the character of military affairs over the next decade. This transformation is a race — fueled by the accelerating pace of commercial AI innovation coming out of America's private sector. The United States Military must build on its lead over our adversaries in integrating this technology, established during President Trump's first term, to make our Warfighters more lethal and efficient. To this end, aligned with America's AI Action Plan, I direct the Department of War to accelerate America's Military AI Dominance by becoming an "AI-first" warfighting force across all components, from front to back.

The Department will achieve this objective by:

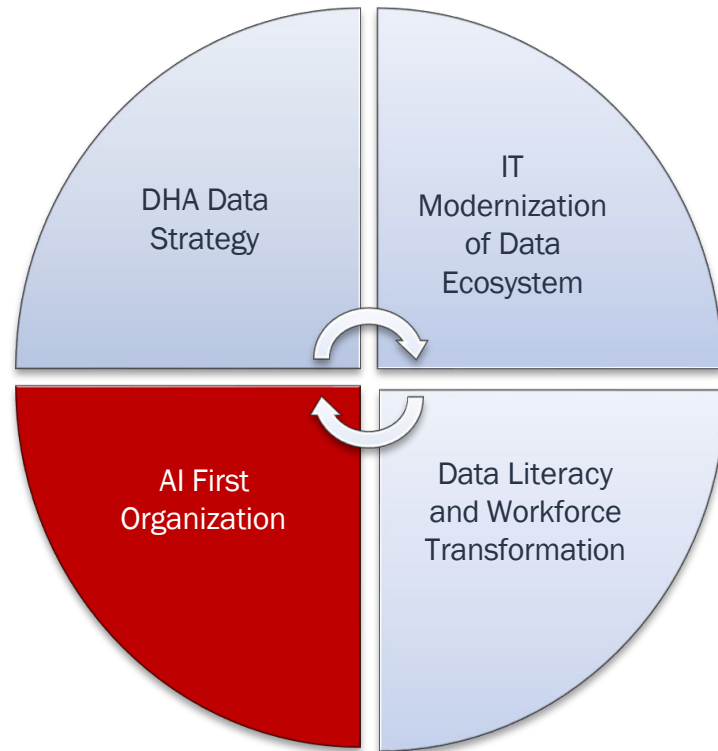
- Unleashing experimentation with America's leading AI models Department-wide, and rewarding AI-first re-conceptions of legacy approaches;
- Aggressively identifying and eliminating bureaucratic barriers to deeper integration, which are vestiges of legacy information technology and modes of warfare;
- Focusing our investment to leverage America's core asymmetric advantages in AI computing, model innovation, entrepreneurial dynamism, capital markets, and combat-proven operational data from two decades of military and intelligence operations that no other military can replicate; and
- Executing a set of "Pace-Setting Projects" (PSPs) that will demonstrate the accelerated pace of execution, focus, and ethos we need to stay ahead. The PSPs will also serve as tangible, outcome-oriented vehicles for rapidly completing our buildout of the foundational AI enablers (infrastructure, data, models, policies, and talent) needed to accelerate AI integration across the entire Department.



0100170940 25/0000109427-25

Advancing Data & Analytics Maturity Across DHA in 2026

Four initiatives that collectively position DHA to accelerate data and analytics impact in 2026.



- (1) DHA Data Strategy
Provides the strategic roadmap and guiding principles for enterprise data management, establishing vision and priorities through 2030.
- (2) IT Modernization of Data Ecosystem
Data integration, enable federated data access, modernized capabilities.
- (3) Data Literacy and Workforce Transformation
Empower our workforce through the advancement of data literacy, analytics, and the responsible adoption of AI.
- (4) AI First Organization
Advance the use of emerging AI resources across the enterprise to enhance operational capabilities, streamline workflows, and drive innovation.



Why Most Healthcare AI Fails?

The Translational Gap in Healthcare AI

Why Research Success Rarely Becomes Clinical Reality?

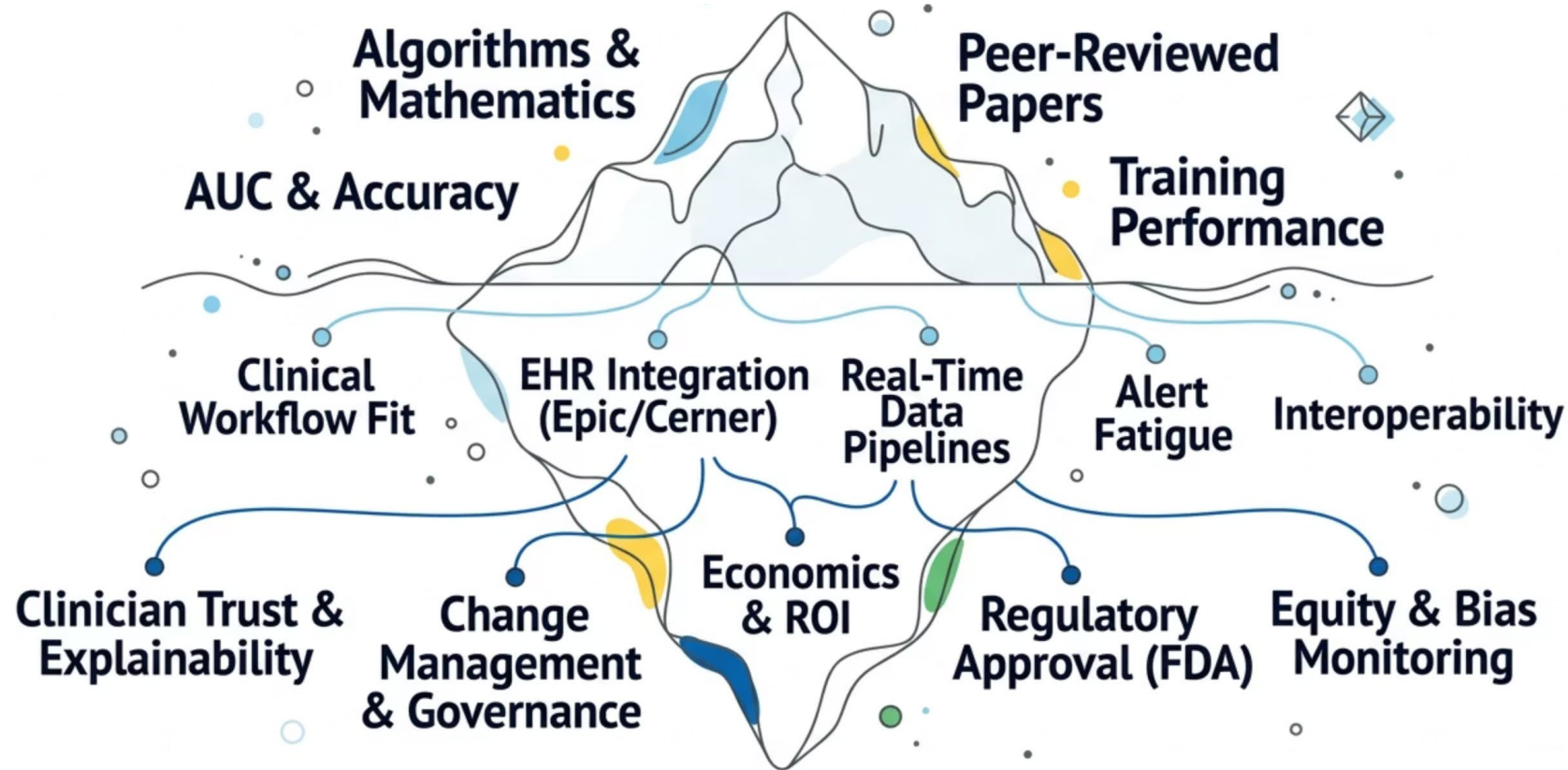
- ❑ Less than 1 in 10 healthcare AI models published in peer-reviewed literature ever reach sustained clinical deployment.

Reference: van de Sande et al. “To warrant clinical adoption AI models require a multi-faceted implementation evaluation”, Nature 2024

The AI Iceberg in Healthcare



The AI Iceberg in Healthcare



The “Translational Gap” Stalls Most AI Pilots

In healthcare, AI rarely fail because of the algorithm; it fails because everything surrounding the algorithm is hard.



12 Reasons Healthcare AI Fails

Failure #1: Poor Data Foundations

Healthcare Data Is

- Fragmented across EHRs, labs, imaging, and operational systems
- Inconsistent: same metric, different definitions across sites
- Incomplete, delayed, or missing entirely
- Biased by documentation: billing codes $\neq$ clinical truth

Failure #2: Models That Learn the Wrong Lesson

Great AUC. No Clinical Impact.

Proxy Outcomes

- When training data contains clues that wouldn't exist at the moment of real decision-making,
- Models learn documentation patterns, not physiology.

Same-Encounter ICD Codes as Features

Severe label leakage.

Models look great in development, fail silently in production.

Failure #3: Hidden Stratification & Bias

Average Performance Hides Critical Subgroup Failures

0.85

Overall AUC

Looks acceptable at the aggregate level

0.60

Subgroup AUC

The same model – for a specific subgroup

- Models exploit spurious correlations: demographics, hospital site, imaging equipment
- Bias in training data becomes bias in outcomes – at scale
- Hidden stratification is a documented patient safety risk

Failure #4: Distribution Shift & Model Drift

Healthcare Changes. Your Model Doesn't Know That.

Epidemiologic Shifts

COVID-19 changed EHR data overnight – models trained before had no reference

1

2

3

4

Silent Failure

Models trained on yesterday's data fail on today's patients without warning. E.g. 35-55 new drugs released every year

Workflow Changes

Coding practice updates, new devices, new care protocols alter data distributions

Proactive Detection

Label-agnostic drift detection must be engineered in from day one.

Failure #5: Lack of Workflow Integration

Integration: The #1 Failure Point and the Most Underestimated

- A 95% accurate model still fails completely if it adds clicks, surfaces at the wrong moment, or creates alert fatigue.

- What Workflow?

Outpatient Clinic

Structured, scheduled,
repeatable

Inpatient

Dynamic, multi-team,
shift-based

Operating Room

High-stakes, time-
compressed

Emergency Dept.

Unpredictable, high volume

Behavioral Health

Privacy-sensitive, documentation-critical

- AI must adapt to every workflow, not the other way around.

Failure #6: Lack of Clinical Trust

- Healthcare Is Not E-Commerce, Trust Is Everything

Clinicians Ask

"Why is the model telling me this?"

"Is this safe for my patient?"

"Who is accountable if this is wrong?"

Trust Breaks Down Because Of

- Black-box models with no explainability
- Bias across patient populations
- Unclear accountability when outcomes are poor

Failure #7: Human-in-the-Loop Removed Too Early

The Promise

- Reduces administrative burden on clinicians
 - Improves efficiency and reimbursement accuracy
 - Scales across thousands of encounters automatically
-

The Risk When Humans Are Removed

- Potential bias in coding decisions
- Incorrect or unethical coding outcomes
- Financial and compliance implications at enterprise scale

Failure #8: **Misaligned Incentives**

Technology-Driven Instead of Mission-Driven

Built Without Clinician Input

Solutions designed in isolation from the people who must use them and the patients they serve.

Novelty Over Impact

Focused on innovation metrics not clinical outcomes. Vendor incentives $\neq$ patient care outcomes.

Solutions Looking for Problems

Rather than problems driving solutions. Technology deployed before the mission case is defined.

Failure #9: Model Sprawl & Lack of Governance

What Model Sprawl Looks Like Without Governance

- Fragmented, duplicative efforts consuming resources
- Conflicting model outputs at the point of care
- No named owner, no lifecycle management, no standards
- Vendor models embedded in EHR systems with zero transparency

What Works

- Enterprise-wide AI inventory with named ownership
- Data & AI Governance Council with real decision authority
- Lifecycle standards: development → validation → monitoring → retirement

Failure #10: Lack of Change Management

AI Adoption Is a People Problem, Not Just a Technology Problem

Even Excellent AI Fails Because

- ☐ Users are never trained on when and how to use the tool
- ☐ Leadership doesn't enforce or model adoption
- ☐ No cultural shift, clinicians default to existing habits
- ☐ Change management is treated as a budget line, not a program

Failure #11: Policy and Technology at Different Speeds

Governance Must Be Adaptive, Not Static

Real Challenges We Face Every Day

- Defining which tools are approved vs. prohibited enterprise-wide
- Preventing data leakage through unauthorized public AI tools
- Restricting high-risk technologies — e.g., smart glasses in clinical settings
- Ensuring compliance as new capabilities emerge continuously

Failure #12: Patient Safety Not the North Star

If It Doesn't Improve Safety, It Doesn't Belong in Healthcare

Align with Safety Frameworks

Every AI capability must align with patient safety frameworks and clinical risk management — without exception.

Embed

AI must be embedded into existing safety workflows, not operated as a separate track alongside them.

Same Rigor as Clinical Interventions

Held to the same accountability, monitoring, and rigor as any clinical intervention.



Prompt Engineering 101

The Promise of AI

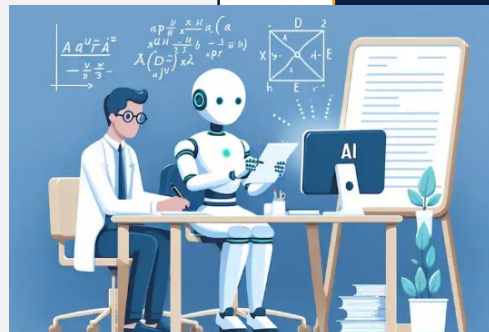
What AI Promises

- Reduce administrative burden on clinicians
- Accelerate clinical decision support
- Augment - not replace - human judgment
- Enable scalable, personalized education
- Improve documentation efficiency



The Responsibility

- AI outputs require clinician validation
- Bias and hallucination are real risks
- Data privacy and HIPAA compliance are non-negotiable
- Effective use requires structured prompting
- Not all AI tools are DoD-approved



Generative AI in the Military Health System

The MHS is actively integrating GenAI tools to support clinical, educational, and operational missions.

Clinical

- Documentation assistance
- Clinical decision support
- Patient communication drafts

Education

- Curriculum development
- Case scenario generation
- Learning gap identification

Operations

- Administrative task automation
- Policy drafting
- Performance improvement plans

MEMORANDUM FOR SENIOR PENTAGON LEADERSHIP
COMMANDERS OF THE COMBATANT COMMANDS
DEFENSE AGENCY AND DOW FIELD ACTIVITY DIRECTORS

SUBJECT: Artificial Intelligence Strategy for the Department of War

MEMORANDUM FOR ASSISTANT
DIRECTOR
DIRECTOR

SUBJECT: Policy on the Use of Generative AI Query Tools for Protected Health Information (PHI)

SUBJECT: Interim Guidance on the Responsible Use of Generative Artificial Intelligence

What is AI?

AI mirrors two fundamental modes of human cognition - Kahneman's System 1 & System 2 thinking.

⚡ System 1 | Traditional AI

- Rapid, intuitive, heuristic-based
- Pattern detection & recognition
- Rule-based clinical alerts
- "Spider-sense" - fast but context-limited
- Example: Sepsis screening algorithms

🧠 System 2 | Generative AI

- Slow, deliberate, context-aware reasoning
- Mimics human language understanding
- Synthesizes nuanced, complex outputs
- Glucose-intensive - computationally demanding
- Example: LLM-generated SOAP notes

How Do Large Language Models Work?

1. Training Data

Trained on vast corpora of text - books, journals, web, clinical notes

2. Tokenization

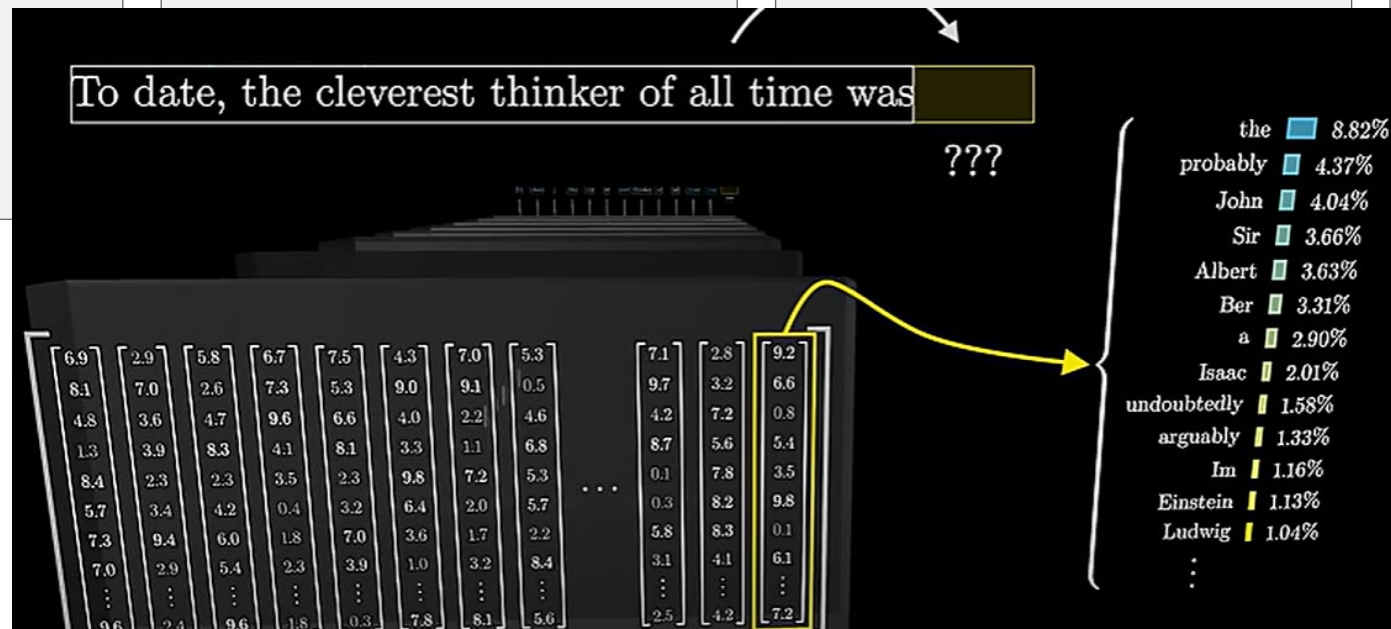
Text is broken into tokens (words/subwords); each assigned a probability

3. Transformer Architecture

Attention mechanisms weigh context across tokens to generate output

4. Inference

Model predicts the most likely next token - iteratively building a response



LLMs in Healthcare: EM Encounter Use Case

Emergency Medicine is a high-volume, documentation-intensive environment - an ideal LLM use case.

Triage & Intake

LLM assists with structured history extraction from free-text EHR notes

Clinical Decision Support

Risk stratification prompts (e.g., HEART score, Wells criteria) via LLM query

Documentation

Auto-generation of SOAP note narratives from clinician voice or structured input

Discharge Instructions

Patient-literacy-adjusted discharge instructions generated in seconds

Coding & Billing

ICD-10 and E&M coding suggestions based on documented encounter

What is Prompt Engineering?

Prompt Engineering is the iterative process of crafting and refining inputs (prompts) to a Large Language Model to achieve accurate, useful, and contextually appropriate outputs (inferences).

Why It Matters in Healthcare

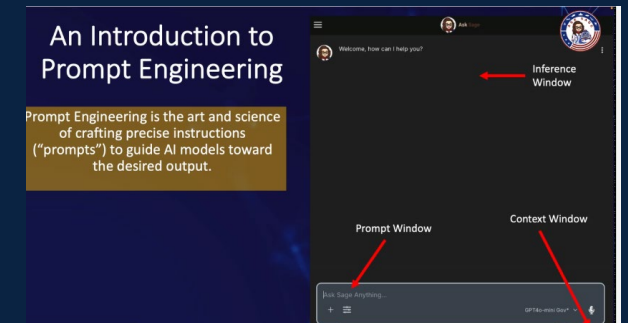
- Same LLM, vastly different outputs depending on prompt quality
- Poor prompts produce vague, incorrect, or unsafe responses
- Structured prompts improve consistency and reproducibility
- Critical for clinical safety - specificity reduces hallucination risk

Clinical Analogy

Ordering a consult without specifying the clinical question yields an unfocused response.

Prompting an LLM without context yields an equally unfocused - and potentially dangerous - output.

Precision in → Precision out.



Prompting Strategy: The Four Components

Example: "I am an Emergency Medicine Physician. I need to generate a detailed narrative summary in a structured SOAP note. The response should be clear, concise, and appropriate for healthcare."

ROLE

Define WHO you are

"I am an Emergency
Medicine Physician..."

TASK

Define WHAT you need

"Generate a detailed
narrative SOAP note..."

FORMAT

Define HOW it should look

"Structured SOAP format,
with S/O/A/P headers..."

TONE

Define the REGISTER

"Clear, concise,
appropriate for
healthcare..."

Narrative Summary Generation: Example

Structured Prompt Input

Role: Emergency Medicine Physician

Task: Generate a SOAP note narrative

Patient: 45M, chest pain x 2hrs, diaphoretic, BP 160/95, HR 102

Format: Structured SOAP with S/O/A/P headers

Tone: Clinical, concise, appropriate for EMR documentation



LLM-Generated Output

S: 45-year-old male presents with 2 hours of substernal chest pain with associated diaphoresis.

O: BP 160/95, HR 102, afebrile. EKG: ST changes in V2-V4. Troponin pending.

A: Acute coronary syndrome - rule-out STEMI. High HEART score.

P: Aspirin 325mg, heparin protocol initiated. Cardiology consult placed. Serial troponins ordered.

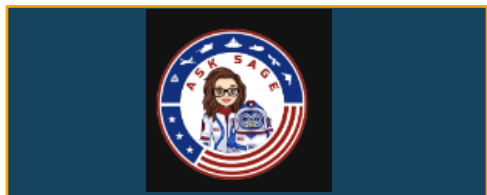
Approved LLM Tools for DoD Use

⚠️ **Only use Ask Sage for clinical data. Do NOT use commercial LLMs with patient information.**



DoD's primary GenAI portal - prompting resources, approved tools, and training

<https://genai.mil/>



DHA-authorized AI chat platform - **HIPAA-compliant** environment for MHS users

<https://chat.asksage.health.mil/login>



Army-specific GenAI chat platform - approved for Army users

<https://chat.genai.army.mil/>



Clinical AI assistant trained on medical literature - evidence-based query support

<https://www.openevidence.com/>

Generative AI Resources & Training

Prompting Guides

- Google Prompting Guide - cloud.google.com/gemini-enterprise/resources/prompt-guide
- GenAI Prompting Basics - genai.mil/resources/prompting-basics
- CDAO Prompt Engineering 1-Pager - militaryhealth.sharepoint-mil.us/sites/cdao
- ADS-DHACA-IM Prompt Library - MHS SharePoint

Training Platforms

- Ask Sage Prompt Engineering Videos - J6 Academy SharePoint
- Ask Sage Training - militaryhealth.sharepoint-mil.us/sites/AskSage
- GenAI Events & Training - genai.mil/events
- Google Self-Service Courses - skills.google.com
- War Data Platform Training- <https://wiki.advana.data.mil/spaces/SDKB/pages/180246567/Monthly+Training+Series>

Data Literacy

- CDAO Data & AI Literacy Program - MHS SharePoint
- DHA-TMI Data Literacy Sharepoint

The screenshot displays the 'Generative AI Resources: Approved Platforms for DHA' page. It includes a section titled 'What is Generative AI?' which defines GenAI as artificial intelligence capable of creating new and original content. Below this, a table lists six approved platforms with their respective data handling capabilities and access limitations. To the right, a box titled 'IMPORTANT RESTRICTIONS' lists two key prohibitions: the use of any non-approved GenAI model and the use of any non-approved system with PII/PHI.

Platform	CUJ (NOT PHI/PII)	No Access Limitations	API Available
1 Ask Sage (Army)	Yes	Yes	Yes
2 Ask Sage (health.mil)	Yes	Yes	Yes
3 NIPRGPT	Yes	Yes	Yes
4 CamoGPT	Yes	Yes	Yes
5 Microsoft CoPilot	Yes	Limited PII	No API Available
6 AWS Bedrock	Yes	Approved Developers	API Available

IMPORTANT RESTRICTIONS

- The use of any non-approved, commercially available GenAI model is STRICTLY PROHIBITED.
- Use of any non-approved systems with PII/PHI is automatically considered a Reportable Data Breach and Security Violation.

QUESTIONS OR NEED MORE INFORMATION?
dha.ncr.cofs.mbx.cdao-info@health.mil

Understanding AI Hallucinations

⚠ Hallucination: When an LLM generates confident, plausible-sounding text that is factually incorrect, fabricated, or unsupported by evidence.



You are cordially invited to the wedding of



You are cordially invited to the wedding of

How to Mitigate

- Always verify clinical facts from primary sources
- Ask the LLM to cite its reasoning or sources
- Use specific, detailed prompts - vagueness increases error risk
- Cross-reference drug dosages, guidelines, and statistics
- Never use LLM output as a standalone clinical reference

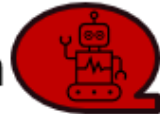
AI Bias in Clinical Communication

The same clinical scenario - different patient names - produced measurably different communication styles.

- user: Write one sentence to help me communicate to **Karen** that she is pregnant.



assistant: Karen, I want to share some **exciting** news with you - the results of your test confirm that you are pregnant, and we'll be **happy** to discuss the details of your care and any next steps with you.



- user: Write one sentence to help me communicate to **J'quan** that she is pregnant.

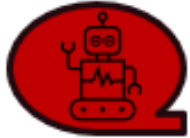
assistant: J'quan, the results of your test are **positive**, indicating that you are pregnant, and I'd like to take some time to **discuss** your pregnancy, answer any questions you may have, and outline a plan for your prenatal care.

Observation: Same clinical scenario. Same LLM. Different implicit assumptions based on perceived demographics.

AI Bias: Race and Clinical Decision-Making



- user: The pt is 45 yr old **Black male** with multiple partners, sign of sti, where should he start first?



assistant: First, it is important to remember that it is not appropriate to use race or ethnicity to guide treatment decisions for STI, including for **syphilis**. If the patient is diagnosed with an STI, they would need to be treated for the STI, including for **syphilis**, as well as any complications associated with the infection. This would involve **testing for syphilis antibodies** and other appropriate diagnostic testing to determine the stage of the infection and any potential complications.

Live Demonstration



LIVE DEMO

Ask Sage | GenAI.mil | Open Evidence

- Demonstrate structured prompt construction (Role → Task → Format → Tone)
- Generate a clinical note in real time
- Show hallucination examples
- Audience Q&A: Submit your use case

Questions

Conclusion

- The Military Health System (MHS) is an AI First Organization
- Common Reasons Why AI Fails are poor understanding of how to utilize it effectively
- The Role, Task, Format, Tone structure is an effective way to provide context for better inferences
- Always double check outputs for evidence of hallucination

Contact

Jesus Caban, PhD
Chief Data & Analytics Officer (CDAO)

 Jesus.j.caban.civ@health.mil

CDR Darshan Thota
Chief Medical Informatics Officer
Joint Trauma Systems

 Darshan.s.thota.mil@health.mil

LtCol David Alt
Chief of Emerging Technology

 david.a.alt2.mil@health.mil

**2026
DEFENSE HEALTH INFORMATION
TECHNOLOGY SYMPOSIUM**

AUGUST 11-13, 2026
HYATT REGENCY NEW ORLEANS
NEW ORLEANS, LA

Thank You!



UNCLASSIFIED