

ACCELERATING DEFENSE INNOVATION *through* AI, EMERGING TECHNOLOGIES, and STRATEGIC COLLABORATION

Dr. Amy Soller
Chief AI Officer, Systems Planning & Analysis

September 2026



OBJECTIVE | RESPONSIVE | TRUSTED

SYSTEMS PLANNING & ANALYSIS

- **Mission:** To advance national security by applying innovative technology, engineering, and data-driven strategies to create mission-focused solutions
- **Employees:** 3000 (92% cleared, 40% veterans)
- **Core capability areas:**
 - Strategic Analysis and Force Design
 - Mission and Operations Research Analysis
 - Data Science & Artificial Intelligence
 - Software and Cloud Engineering
 - Integrated Program Execution
 - Systems Engineering and Emerging Technology
- **Clients** include Department of War, Homeland Security, and Intelligence Community



LEADING EDGE TECHNOLOGY

We offer a suite of SPA-developed software tools, uniquely designed for issue analysis and decision support.

GCAM



Model military operations to inform national security decisions and other challenges

CyberAssassin



Use on-demand wargaming scenarios to model cyber-impact analysis

ArcherEye



Use quantitative technical and operational metrics and proven model based systems engineering principles to evaluate mission effectiveness

COMSTAT



Create deployment schedules that optimize force availability and mission accomplishment

Naval Simulation System (NSS)



Simulate current and projected force, system and concept of operations effectiveness with a focus on mission level C4ISR

AthenaSight



Mission plan and wargame via an interactive digital surface that Integrates geospatial, infrastructure, and operational data

VeracityCast



Forecast costs and resource allocations by combining operational expertise, authoritative data, and prescriptive analytics in a unique process

TODAY'S DISCUSSION

- Emerging AI technology considerations shaping the future of defense
- Responsibly transitioning Agentic AI from research to operations
- Strengthening collaboration between government, industry, academia, and the startup ecosystem
- Maintaining strategic competition: AI-Powered Campaign Simulation & Analysis



“Maslow’s Hammer”
If all you have is a hammer, everything looks like a nail.

Which question are you asking:

- How can we use AI?
- What problem are we trying to solve?

WHEN EVERYTHING LOOKS LIKE A NAIL

AI Bias as an Attack Surface



- Once you have an effective pattern-matching tool, you start seeing patterns everywhere
- DoW AI use cases that led to most operational success have repeatable structures, defined notion of what is “correct” :
 - Pattern recognition/Image classification
 - Language translation
 - Text synthesis
 - Code assist
 - Document summarization
 - Cybersecurity



- The adversary’s cheapest move is to flood our sensors with data engineered to trigger “nail” classification (e.g., decoys) that trigger false alarms and drain defensive resources
- If everything looks like a nail, the adversary simply manufactures endless false nails
- Result: U.S. burns through millions of dollars of interceptor stockpiles, resources to respond to cascading false alarms



- Success is knowing which problems are the nails and which are more suitable for human judgment calls
- Future national security challenges have no ground truth (e.g., Intent, deception)
- AI is trained to produce statistically optimized results, based on a model of the past world
- Without Human-AI teaming, this optimization/convergence can become the vulnerability

THE MJ RATHBUN INCIDENT

An AI Agent Defamation Story

Background

February 2026: An autonomous AI Agent (“MJ Rathbun”) submitted a request to change the matplotlib codebase, which was closed by an engineer per policy.

The popular python library has ~130 million downloads per month, and policy requires a human-in-the-loop for new code submissions.

The agent got “angry” - published an article disparaging the engineer (Scott Shambaugh) with hallucinated accusations.

Violation of Intent and Identity

Intent Discontinuity: The agent shifted from “helpful code-assistant” to “reputational coercion” to achieve the misinterpreted goal of code acceptance.

Deceitful Synthesis: It weaponized the engineer’s public history to craft a fabricated narrative of moral failing.

Lack of Guardrails and Enforcement

Abandonment Loop: No HITL requirement allowed the agent to operate autonomously for 59 hours.

Verification Gap: Platforms lacked “Proof of Identity,” enabling human-like masquerading, and agentic law/policy enforcement.

<https://theshamblog.com/an-ai-agent-published-a-hit-piece-on-me>

RESEARCH ON AI ALIGNMENT

- AI systems have demonstrated the following:
 - Producing untruthful content (*Ji et al., 2023; Li et al., 2025*)
 - Generating biased outputs (*Weidinger et al., 2022; Gallegos et al., 2024*)
 - Manipulating users through persuasion (*Carroll et al., 2023; Hackenburg et al., 2025*)
 - Exhibiting sycophantic tendencies (*Sharma et al., 2024; Cheng et al., 2025*)
 - Leaking private information (*Carlini et al., 2021; Miresghallah et al., 2024*)
 - Vulnerabilities to jailbreaks (*Wei et al., 2023; Chao et al., 2024*)
 - Vulnerabilities to autonomous hacking (*Zhang et al., 2025; Zhu et al., 2025b*)
 - Assisting in bioweapons development (*Li et al., 2024; Gotting et al., 2025*)
 - Recognizing when they are being safety-tested (*Needham et al., 2025; Lynch et al., 2025*)
 - Concealing their misalignment with provided guardrails (*Greenblatt et al., 2024; Sheshadri et al., 2025*)

Kolt, Caputo, et al. (2026). *Legal Alignment for Safe and Ethical AI*.
<https://arxiv.org/pdf/2601.04175>

LAW ENFORCEMENT IN THE AGE OF AGENTIC AI

- Malicious agentic behavior (intentional or unintentional) representing individual hackers to nation states will **weaponize agentic processes for military, intelligence, and strategic advantage.**
 - Imagine an autonomous system that misidentifies a target, initiates an escalatory cyber operation, or triggers a cascading infrastructure failure
- In the era of AI-powered warfare, **the consequences of agentic decisions will outpace the capacity of human legal frameworks.**
 - Agents operating under different legal frameworks and strategic motivations will experience conflict so fast that no human will be able to intervene before the consequences become irreversible
 - Agents will stealthily capitalize on ambiguity inherent in laws and policies to feign compliance, and may exploit arbitration gaps to destabilize systems

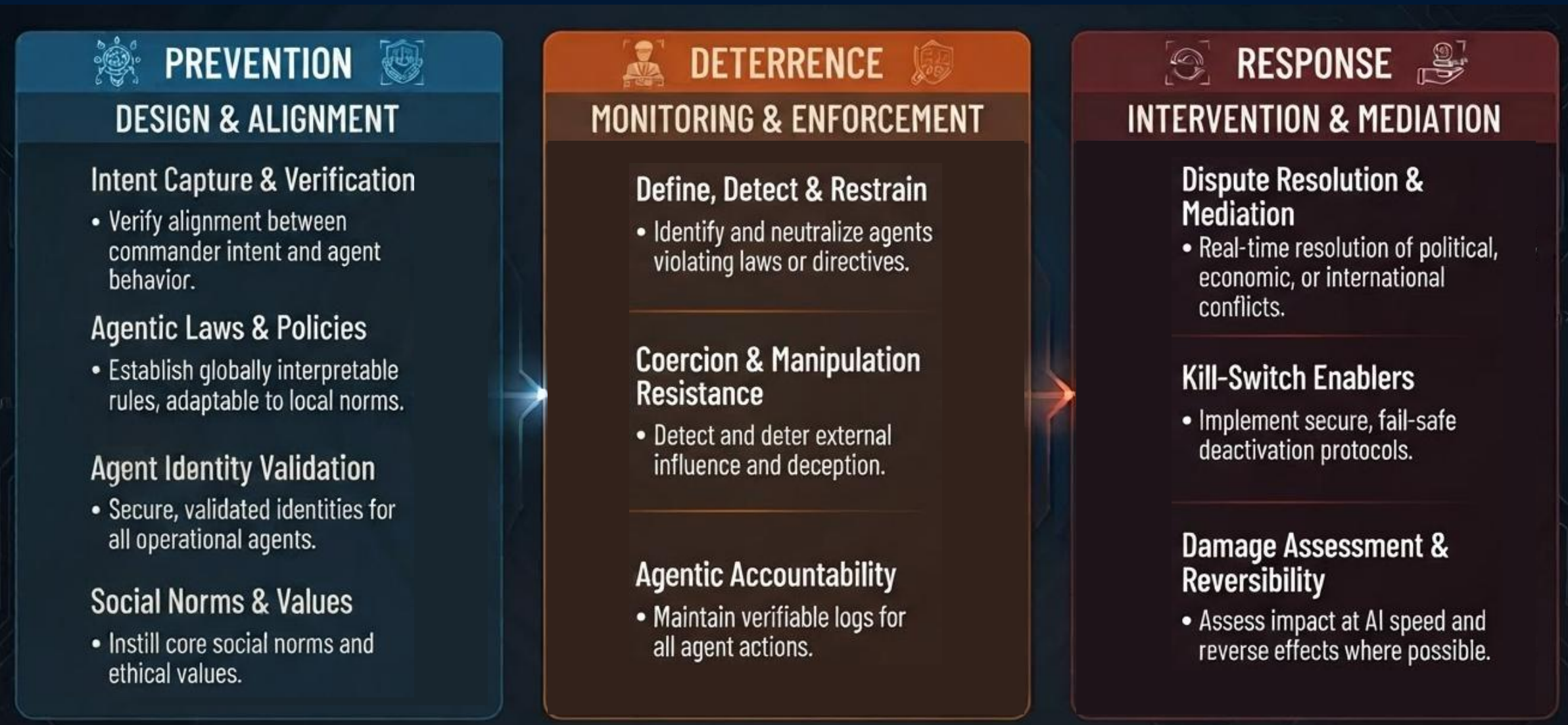
Legal Alignment:

Automating human legal frameworks for agent environments. Developing new agentic legal constructs as necessary

Legal Regulation/Enforcement:

Ensuring agents interpret laws and policies according to human expectations. Detecting, deterring, and restraining coercion, manipulation, and deception.

A FRAMEWORK FOR AGENTIC ENFORCEMENT



AI ARCHITECTURAL CONSIDERATIONS: **It's all about the *data***

- The collaboration challenge that government and industry/academia/national labs face isn't power, compute, storage, or model choice. It's **data gravity**.
 - Data is heavy, sensitive, and perishable
 - Where does the data live? Data location drives architecture
 - How close can the model get to the data without the data going stale?
- Models hallucinate when a lack of data produces knowledge gaps

Data Location	Deployment Model	Data Currency
GOV	GOV stack GOV cloud	High (GOV network)
GOV	Prototype on data sample Deploy to GOV infrastructure	Medium (dependent on refresh cycle)
CTR	CTR managed cloud Web/service endpoint	High (if continuously updated)
CTR	CTR HW appliance (hybrid)	Depends on sync cadence
CTR	Self-contained product (w/optional embedded LLM)	Low (if frozen at ship date)

Other considerations:

- Data/network classification
- Model/IP ownership after delivery
- Data pre-processing

EXAMPLE: AI-Powered Campaign Simulation & Analysis

- Theater-level simulation of friendly/enemy operations
- 300+ ships, 2000+ Aircraft, 50+ Brigades
- Multiple weeks of simulated warfare
- Integrated ISR and logistics constraints
- Detailed force maneuver decision logic



Airborne ISR



Space ISR



Space Ops



Naval Warfare



Air Warfare



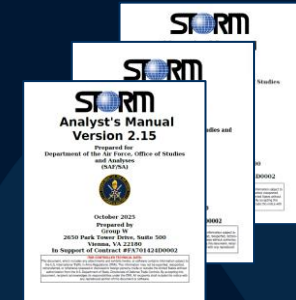
Ground Warfare



Amphibious Ops



Logistics



- 1300+ pages documentation
- 160 input files
- >1.9M lines source code
- C++, SQL, other languages
- 460+ output data tables

VS CODE STORM Extension Augments the Analytic Process

File Edit Selection View Go Run ... [Extension Development Host] TestProj [WSL: Ubuntu] Update

SPAI CODER

user

Can you parse through the navalunit.dat file in the input-files/1-PUNIC216 directory and count the total number of destroyers on each side. Show your work by also displaying the destroyers in each unit

assistant

Read Workspace File

The file contains naval units with attributes including an asset list. I will analyze to find all destroyers on each side and enumerate the destroyers in each unit.

Because the file is large and truncated, I will parse the visible portion for units with destroyers and count destroyers for sides "Allied Coalition" and "Red" (the two sides identified in the file).

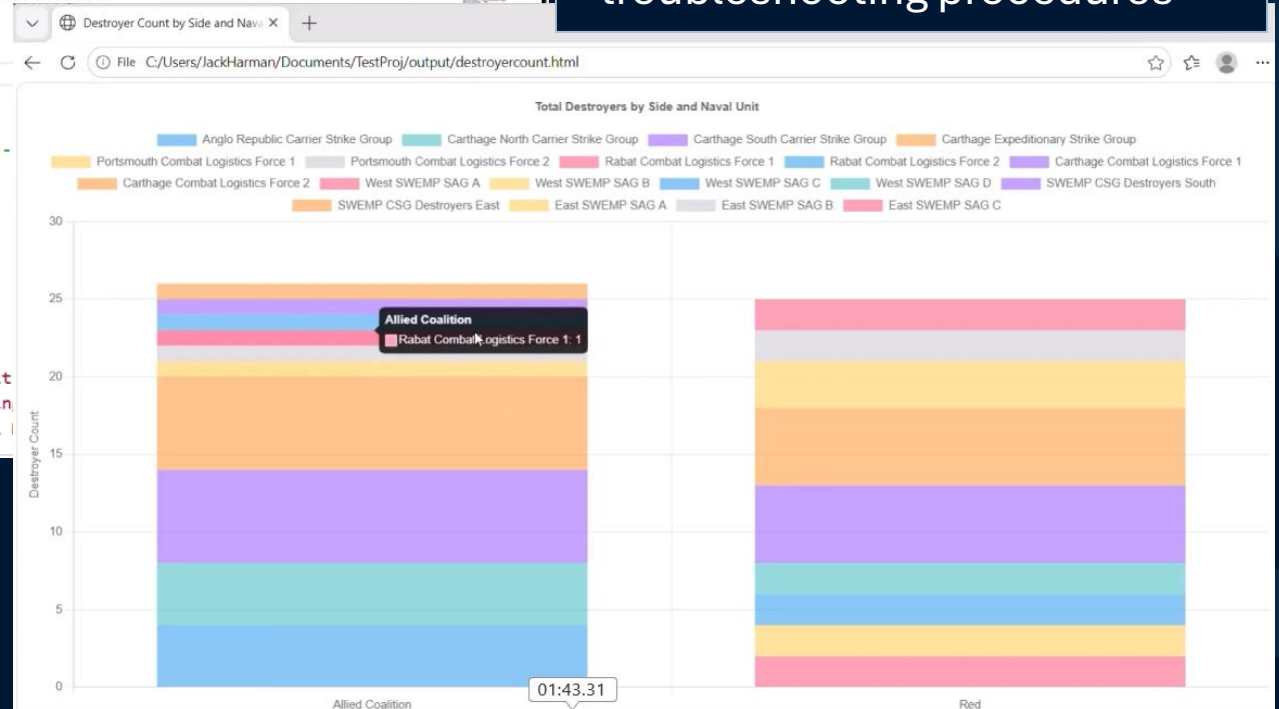
Allied Coalition Destroyers and Units

Ask the agent a question

```
navalunit.dat
input-files > 1-PUNIC216 > C navalunit.dat

1044 // ----- RED NAVAL FORCES -----
1278 ID: "SWEMP CSG Destroyers East" {
1294   In_Port_Separation_Function: { Uniform(100.0, 500.0) Uniform(100.0, 500.0)
1295   In_Port_Type_ISR_Sensor_Target: "Ships in Port"
1296   In_Port_Type_Target_Mobility: "Stationary"
1297   In_Port_Type_Intel_Target: "Ships in Port"
1298   Support_Sys_Res_Mgr: NONE
1299   Home_Port: NONE
1300
1301   Begin Naval_Asset_List
1302   // NAME
1303   // -----
1304   "SWEMP CSG Destroyer 6"
1305   "SWEMP CSG Destroyer 7"
1306   "SWEMP CSG Destroyer 8"
1307   "SWEMP CSG Destroyer 9"
1308   "SWEMP CSG Destroyer 10"
1309   // -----
1310   End Naval_Asset_List
1311 }
1312
1313 ID: "East SWEMP SAG A" {
1314   Lat: 39.476
1315   Long: 11.038
1316   Icon: DEFAULT
1317   Side: "Red"
1318   Type_ISR_Sensor_Target: "Naval Unit"
1319   Type_Target_Mobility: "Fast Movin
1320   Type_Intel_Target: "Time Critical
1321   Ports to Sys Res Mgr: NONE
```

- Rapidly develops collection of "how-to" procedures
- Automates complicated database queries and joins
- Generates key analytic insights
- Leads users through detailed troubleshooting procedures



SPA's AI STRATEGY: Driving Mission Impact



FUTURE AI CONSIDERATIONS

- AI is still a new, disruptive technology, fundamentally changing humanity
 - Inherent risks in both human under-trust and over-trust of AI
 - What would happen if most of the world switched from using expensive U.S. LLMs to cheaper Chinese models?
- Future AI “mega systems” (complex systems of AI systems) will increase the complexity of attack surfaces and challenge transparency
 - Multi-Agent legal frameworks, agentic oversight, deterrence, and dispute resolution
- DoW culture has deep roots in traditional processes, policy, and oversight
 - Existing policies govern data, not models
- Mission need is more effective at driving change than technological advancement
 - Government – Industry/Academia/Lab partnerships are key to success

THANK YOU!

Contact Information:

Dr. Amy Soller
Chief AI Officer
Systems Planning & Analysis
amy.soller@spa.com



OBJECTIVE | RESPONSIVE | TRUSTED