

Model-Based Systems Engineering (MBSE) Approach to Develop an Artificial Intelligence Bill of Material (AIBOM) for Al System Compliance Verification

Dr. James Lee – DEVCOM C5ISR / DCI-Solutions Ahmad Alghamdi – University of Jeddah Prof. Abbas K. Zaidi – George Mason University



The AI BOOM vs Lack of Adoption

- Why haven't we seen AI implemented in more places across the Army/DoD?
 - Lack of explainability
 - Data privacy and security
 - Risk of errors and bias
- Can we trust AI / ML to make decisions for us?
 - Model stealing
 - Data poisoning
 - Al supply chain integrity



AI Layered Defense Framework (AI-LDF)

• A comprehensive library of AI-related risks and mitigations to inform and guide the development and implementation of AI models and software

AI LAYERED DEFENSE FRAMEWORK

System design engenders risk, necessitating risk mitigation strategies aligned to all component areas: Data (D), Models (M), Model Output (O), Infrastructure & Code (I), and Human Factors (H)

Risk Tolerant Models

- Al Models that are resilient due to low complexity and high transparency
- Easy to understand and recognize malicious actions

Risk Sensitive Models

- More complex models that make important mission decisions
- Models built around publicly know models or data

Risk Adverse Models

- Models with high adversarial value
- High complexity, opaque models for critical decisions

Al tools that offer broad value with minimal possibility that the occurrence of an event will adversely affect the achievement of the Army's objectives. The maximum benefit is achieved using the fewest controls.

> Al software that offers strategic value to a more limited number of users and has more significant mitigation strategies to **balance** value and security.

Al models that provide high value but cannot be compromised, employ state-of-theart defenses.

Low Control Mitigation Continuous Automated

Monitoring, Regular Cybersecurity Practices

Medium Control Mitigation

Robust Data Protection, Thorough Testing and Evaluation, Firewalling

High Control Mitigation

Comprehensive Risk Assessment, Strong Security Measures, Disaster Recovery Plan, Regular Audits



AI-LDF: AI Risks and Mitigations

A Lygerd Orbina Ria Hatis																					
Compo	nent Risk		Layer 1 Mitigation		Layer 2 Mighton									Layer 3 Mitigation							
	Note: Numbers in cells show correspondence between Risk, Mitigation and Layer (1, 2, 3). Color	Implement Establish Basic Alluber Pascess / Controls for Data Sources Pasce Alluber Pascess / Controls for Data Sources / Validation Engl Detervice	Implement Establish	Posteri Develop Perform Cross-Check Automatic Limit Padec Limit Padec	Create Maintain Ide	By and Implement Provide Enfo	rce Implement _{Waltebula and} Contro	4 Implement , Develop Trate	Check Enforce Entorce Code	Develop Inc. 2	and Data Implement Employ Act	wely Restrict VerBy and Isource	Employ Inpu	blandhu and Bastellar and Meeller	Employ Two Factor Authenticatio	Use Maintain Scorrettic Secure	Replicate Al Scan 1 Dutp	for Use Prohibi	E Report LLM Comprehe e Reven	siv Reduct Obscure Hyperparame Structure and and	Addet Utilize ring Clustering / d Anomaly Model Sensor
	LDF Figure 1.	Thaning Standards Dota and and Use and Centrol Process	Component	Risk											a internation	n Pacifies	Manpula	tion Methods	s Engreen Protocol	Weights Pecidina Adversari Training	na Detection Security
Det	Ingen Alleske Alle Manning Grangtion er Difk all Gas Egylandsten of Difk all Gas Desminist (all Cardinal State State Desminist (all Cardinal State States) Cardinal of Cardinal States (Dava Cardinal State Research and States) Dava Cardinal States			Note: Numbers in cells show correspondence between Risk, Mitigation and Layer (1, 2, 3). Color coding is for visual alignment to AI- LDF Figure 1.	Implement Basic Al User Training	Implement Basic MLOps Practices / Standards	Establish Access Controls for Data and Model	Document Data Sources and Use	Ensure Basic Input Validation and Control	Outline and Document ML Supply Chain	Create Detailed MLOps and Security Guides	Maintain Vetted and Monitored Model Libraries	Identify and Address Vulnerabilitie S	Implement Sophisticated Network Security							
	Lack of Data Trustworthiness Potential Privacy or Legal Issues from Da Dale, Ineccurate, or Intelevent Data			Loss of System, Model or Data Controls																	
	Insufficient Evaluation Data Data Assumptions Recursive Data		Data	Trojan Attacks																	
	Symbolic Data Model Asset Laak or Reverse Engineering Malicous Libraries / Code			Data Poisoning																	
Infrastru	Byzantine Attacks team insufficient input Resource Control Unsocured and Wulnerable Systems			Corruption or Drift of Data																	
	Label Flipping Attacks Test and Biobuston Data Poisoning Model Theft or Revense Engineering		Infrastructure	Model Asset Leak or Reverse Engineering																	
	Prompt Injection Attacks Unforeseen Model Manipulation Model Implementation Flaws			Malicious Libraries / Code																	
Models	ACC or Text Substitution Model Manipulation A Specific Denial of Service Model Theft or Reverse Engineering			Byzantine Attacks																	
	Mack-Hex Pealon Adversarial Text Extraction Prompt Injection Attacks			Insufficient Input Resource Control																	
	Model Collague Model Decay LMM Data Laskage			Unsecured and Vulnerable Systems																	
Outpu	All Disput Manipulation LLM Hallusinations and throns Looped Input Attacks			Label Flipping Attacks																	
	Lack of Al Technical Understanding Hadequate Oversight of Implementation-theopretation Prosmet Human in the Loop Involvement			Test and Evaluation Data Poisoning																	
Huma	en Black Itox and Transparency Issues Over Relace on Al System District of System Results District of System Results District Objections to Al																				

Artificial Intelligence Bill of Material (AIBOM)

- Response to the growing need for managing AI supply chain risk
- Motivated by software bill of materials (SBOM), which is a nested inventory, a list of ingredients that make up software components
 - Key building block in software security and software supply chain risk
 - Builds stakeholder confidence in software systems through enhancing transparency, security, and compliance
- AIBOM can help organizations understand and manage the components that are part of their AI systems and reduce various risks by:
 - Managing vulnerabilities
 - Tracking compliance
 - Ensuring the integrity and security of components from the AI supply chain



D. Mor-Ofek, "It's Time to Talk About Al/ML BOM (Artificial Intelligence Bill of Materials) And Vulnerability Management," C2A Security - The Only Risk-Driven DevSecOps Platform. Accessed: Sep. 17, 2024. [Online]. Available: <u>https://c2a-sec.com/its-time-to-talk-about-ai-ml-bom-artificial-intelligence-bill-of-materials-and-vulnerability-management/</u>

AIBOM Application

- Malicious packages can be distributed through various means, including official channels
 - i.e. PyPI package manager
 - Some packages may have similar names to prominent Python packages, making it possible to install them by accident
- Once installed, bad actors could potentially steal sensitive information, disrupt processes, or gain control of other connected systems
- For organizations with complex software systems, code bases, and AI/ML applications, AIBOM can provide a compliance framework that enumerates and catalogs all components used in AI systems
- Automate checking against vulnerability reports (CVE's)



Created by Author using ChatGPT 40



Model-Based Systems Engineering (MBSE)

- A systems engineering methodology that focuses on creating and exploiting domain models as the primary means of information exchange between engineers rather than on document-based information exchange.
- More recently, the focus has also started to cover aspects related to model execution in computer simulation experiments, further overcoming the gap between the system model specification and the respective simulation software.
- As a reflection of this evolution, the term 'modeling and simulation-based systems engineering (M&SBSE)' has also come into use alongside 'MBSE ', underscoring the expanded scope and capabilities of the methodology.



Dr. Burak Gozluklu, "What is MBSE and why do industries start to use? - Model Based Systems Engineering (MBSE) on AWS: From Migration to Innovation." Accessed: Sep. 17, 2024. [Online]. Available: <u>https://docs.aws.amazon.com/whitepapers/latest/model-basedsystems-engineering/what-is-mbse-and-why-do-industries-start-to-use.html</u>



What is the Unified Architecture Framework?

- The Unified Architecture Framework[®] (UAF[®]) is a generic and commercially orientated architecture framework based on work in the defense domain by the Object Management Group (OMG)
- UAF defines ways of representing an enterprise architecture that enables stakeholders to focus on specific areas of interest in the enterprise while retaining sight of the big picture.
- UAF meets the specific business, operational, and systems-of-systems integration needs of commercial and industrial enterprises as well as the U.S. Department of Defense (DoD), the UK Ministry of Defense (MOD), the North Atlantic Treaty Organization (NATO), and other defense organizations.

WAF Order Water Parket Nork -	Motivation Mv	Taxonomy Tx	Structure Sr	Connectivity Cn	Processes Pr	States St	Sequences Sq	Information ^c If	Parameters ^d Pm	Constraints Ct	Roadmap Rm	Traceability Tr
Architecture Management ^a Am	Architecture Principles Am-Mv	Architecture Extensions Am-Tx ^e	Architecture Views Am-Sr	Architecture References Am-Cn	Architecture Development Method Am-Pr	Architecture Status Am-St		Dictionary Am-If	Architecture Parameters Am-Pm	Architecture Constraints Am-Ct	Architecture Roadmap Am-Rm	Architecture Traceability Am-Tr
Strategic St	Strategic Motivation St-Mv	Strategic Taxonomy St-Tx	Strategic Structure St-Sr	Strategic Connectivity St-Cn	Strategic Processes St-Pr	Strategic States St-St		Strategic Information St-If		Strategic Constraints St-Ct	Strategic Deployment, St-Rm-D Strategic Phasing St-Rm-P	Strategic Traceability St-Tr
Operational Op	Requirements Rq-Mv	Operational Taxonomy Op-Tx	Operational Structure Op-Sr	Operational Connectivity Op-Cn	Operational Processes Op-Pr	Operational States Op-St	Operational Sequences Op-Sq			Operational Constraints Op-Ct		Operational Traceability Op-Tr
Services Sv		Services Taxonomy Sv-Tx	Services Structure Sv-Sr	Services Connectivity Sv-Cn	Services Processes Sv-Pr	Services States Sv-St	Services Sequences Sv-Sq	Operational Information Op-If	Environment En-Pm and Measurements Me-Pm and Risks Rk-Pm	Services Constraints Sv-Ct	Services Roadmap Sv-Rm	Services Traceability Sv-Tr
Personnel Ps		Personnel Taxonomy Ps-Tx	Personnel Structure Ps-Sr	Personnel Connectivity Ps-Cn	Personnel Processes Ps-Pr	Personnel States Ps-St	Personnel Sequences Ps-Sq	Resources		Competence, Drivers, Performance Ps-Ct	Personnel Availability Ps-Rm-A Personnel Evolution PS-Rm-E Personnel Forecast Ps-Rm-F	Personnel Traceability Ps-Tr
Resources Rs		Resources Taxonomy Rs-Tx	Resources Structure Rs-Sr	Resources Connectivity Rs-Cn	Resources Processes Rs-Pr	Resources States Rs-St	Resources Sequences Rs-Sq	Information Rs-If		Resources Constraints Rs-Ct	Resources evolution, Resources forecast Rs-Rm	Resources Traceability Rs-Tr
Security Sc	Security Controls Sc-Mv	Security Taxonomy Sc-Tx	Security Structure Sc-Sr	Security Connectivity Sc-Cn	Security Processes Sc-Pr					Security Constraints Sc-Ct		Security Traceability Sc-Tr
Projects Pj		Project Taxonomy Pj-Tx	Project Structure Pj-Sr	Project Connectivity Pj-Cn	Project Processes Pj-Pr						Project Roadmap Pj-Rm	Project Traceability Pj-Tr
Standards Sd		Standards Taxonomy Sd-Tx	Standards Structure Sd-Sr								Standards Roadmap Sd-Rm	Standards Traceability Sd-Tr
Actual Resources Ar			Actual Resources Structure, Ar-Sr	Actual Resources Connectivity, Ar-Cn		Simulation ^b				Parametric Execution/ Evaluation ^b		

3DS CATiA NoMagic, "UAF 1.2 - UAF 1.2 Plugin 2021x Refresh2 - No Magic Documentation." Accessed: Sep. 17, 2024. [Online]. Available: https://docs.nomagic.com/display/UAF12P2021xR2/UAF+1.2

UAF for AIBOM

- Model overall system architecture, including relationships and dependencies between software components and other parts of the system
- Provides traceability features, allowing organizations to trace requirements through the system to components
 - Can help understand the impact of vulnerabilities as they are discovered
- Assist in managing changes to system components to maintain upto-date AIBOM

High-level Taxonomy of System



AI/ML Model Development Life Cycle



Example: AI Model Development

- Objective: build an AI model to predict the speed of sale on cover photo used on client facing real estate websites (i.e. redfin.com)
- Step1: scrape the website for cover photo and days on market value and store in database DATA ACQUISITION
 - Packages: beautiful soup, requests
- Step2: generate embeddings (numerical representation) of the images and create binary based on days on market (sold in less than a week of posting) DATA PROCESSING
 - Packages: pyspark, pandas
- Step3: partition the data and train an xgboost model for classification MODELING
 - Packages: scikit learn, xgboost





Benefits of UAF for AIBOM

- Help systematically define and organize the components of the AIBOM, ensuring that all relevant elements are covered and properly documented.
- Emphasizes detailed documentation of architectural elements. For AIBOM, this means having a clear and comprehensive record of AI components, dependencies, and interactions, which is essential for managing complex AI systems.
- Benefit from a common language and set of concepts. This facilitates better communication among different teams involved in the development and maintenance of the AIBOM
- Helps in identifying potential risks and vulnerabilities using the AIBOM. This can aid in proactive risk management and ensure that the AI system is robust and resilient.

References

[1] "Software Bill of Materials (SBOM) | CISA." Accessed: Jun. 14, 2024. [Online]. Available: https://www.cisa.gov/sbom

[2] J. Petersen, "Army Issues RFI for Project Linchpin AI Bill of Materials." Accessed: Jun. 14, 2024. [Online]. Available: https://executivegov.com/2023/11/armyissues-rfi-for-project-linchpin-ai-bill-of-materials/

[3] "PyPI · The Python Package Index," PyPI. Accessed: Jun. 14, 2024. [Online]. Available: https://pypi.org/

[4] "CVE - CVE." Accessed: Jun. 14, 2024. [Online]. Available: https://cve.mitre.org/

[5] J. J. López García and D. P. Pereira, "Analyzing System Security Architecture in Concept Phase Using UAF Domains," INSIGHT, vol. 25, no. 2, pp. 56–60, Jun. 2022, doi: 10.1002/inst.12388.

[6] M. Torkjazi, A. J. Davila-Andino, A. Alghamdi, and A. K. Zaidi, "UAF Strategic Planning for Enterprises," *IEEE Access*, pp. 1–1, 2022, doi: 10.1109/ACCESS.2022.3224456.

[7] A. Alghamdi, M. Torkjazi, A. J. Davila-Andino, and A. K. Zaidi, "Employing UAF Inter-Domain Traceability for Performance and Effectiveness Evaluation," in 2023 *IEEE International Systems Conference*(SysCon), Vancouver, BC, Canada: IEEE, Apr. 2023, pp. 1–8. doi:10.1109/SysCon53073.2023.10131056.

[8] D. Mor-Ofek, "It's Time to Talk About AI/ML BOM (Artificial Intelligence Bill of Materials) And Vulnerability Management," C2A Security - The Only Risk-Driven DevSecOps Platform. Accessed: Sep. 17, 2024. [Online]. Available: <u>https://c2a-sec.com/its-time-to-talk-about-ai-ml-bom-artificial-intelligence-bill-of-materials-and-vulnerability-management/</u>

[9] Dr. Burak Gozluklu, "What is MBSE and why do industries start to use? - Model Based Systems Engineering (MBSE) on AWS: From Migration to Innovation." Accessed: Sep. 17, 2024. [Online]. Available: <u>https://docs.aws.amazon.com/whitepapers/latest/model-based-systems-engineering/what-is-mbse-and-why-do-industries-start-to-use.html</u>

[10] 3DS CATIA NoMagic, "UAF 1.2 - UAF 1.2 Plugin 2021x Refresh2 - No Magic Documentation." Accessed: Sep. 17, 2024. [Online]. Available: https://docs.nomagic.com/display/UAF12P2021xR2/UAF+1.2





