VIRTUALLY HOSTED ON: OCTOBER 28-29, 2020

RESEARCH WORKSHOP: AI4SE & SE4AI

U.S.ARMY

DEVCOM ARMAMENTS CENTER

SYSTEMS ENGINEERING RESEARCH CENTER

Sponsored and organized by Combat Capabilities Development Command - Armaments Center (CCDC - AC) Systems Engineering Directorate (SED) and the Systems Engineering Research Center (SERC)

EXECUTIVE SUMMARY

Objective

Systems Engineering (SE) is undergoing a digital transformation that will lead to further transformational advances in the use of Artificial Intelligence (AI) and Machine Learning (ML) technology to automate many routine engineering tasks. At the same time, applying AI, ML, and autonomy to complex and critical systems encourages new systems engineering methods, processes, and tools.

To address this current and evolving reality, The US Army Combat Capabilities Development Command Armaments Center (CCDC AC) Systems Engineering Directorate (SED) and the Systems Engineering Research Center (SERC), a University Affiliated Research Center (UARC) for the Department of Defense (DoD), jointly sponsored the inaugural Artificial Intelligence for Systems Engineering/Systems Engineering for Artificial Intelligence (AI4SE/SE4AI) workshop. This two-day virtual event gathered members of the Government, Academic and Industry communities to learn from leaders already using AI in this space and share ideas focused on the workshop's main objective: how to define relevant SE and AI challenges, areas of exploration and methodologies to use, and ways in which to collaborate and research in the upcoming years.

The Workshops

The total 15 presentations focused on relevant topics within the areas of Machine Learning/Artificial Intelligence (ML/AI), Artificial Intelligence for Systems Engineering (AI4SE), Systems Engineering for Artificial Intelligence (SE4AI), and Digital Engineering (DE). A unifying theme was the increasing need for organizations and systems to be agile to keep up with the dynamic nature of AI in order to achieve the ultimate goal of delivering the most relevant and effective tools to the soldier in the field. Essential areas for focus emerged, notably: the importance of data—its acquisition, analysis and maintenance; the need to update business

processes, including workforce development and retention; and the importance of collaboration.

Outcomes

Current systems engineering practices do not support the long-term outcome of "Human-Machine Co-Learning", which implies an upcoming evolutionary phase in the Systems Engineering community consisting of three "waves": the first includes technologies and approaches that increase the transparency of decisions produced by AI systems; the second will produce systems that learn and are *robust and predictable* in the type of key applications normal to SE; and the third wave involves systems that *adapt* and learn dynamically from their environments, which will develop *trust* in machine-to-machine and human-to-machine (and maybe machine-to-human) interactions. A notional roadmap of this evolution spans five categories:

1. **AI/ML Technology Evolution:** The technological implementation of AI systems need to evolve in directions relevant to SE.
2. **Automation & Human-Machine Teaming:** The purpose of AI in systems is generally automation of human tasks and decisions.
3. **Augmented Engineering:** AI technologies will increasingly be used to augment the work of engineering.
4. **Digital Engineering:** The current digital engineering transformation will enable augmented engineering.
5. **Workforce and Culture:** Significant transformation is needed in the SE workforce, with greater integration of software and human behavioral sciences at the forefront.

Future Work

The SERC and CCDC AC intend to organize a follow-on workshop in 2021 to explore specific outcomes from this initial 2020 meeting. The intended goal will be to curate panel discussions and facilitate group breakout sessions that produce actionable applications of AI4SE and SE4AI as well as relevant research ideas.

Feedback gathered from participants of this inaugural 2020 event (in response to the question "Based on your experience in this workshop, what do you think are the top 2-3 ideas or initiatives worthy of further discussion?") will guide future exploration, including:

- Establish an AI test bed that contains data sets that can be created and manipulated to affect outcome.
- Explore a path forward for AI and XAI for SE and mission engineering.
- Integrate digital twins in the operational phase.
- Embed ethics considerations of AI into the engineering process.
- Test and Evaluation (T&E) with AI/ML components.
- Consider collaborative functional design.
- Extend the current SE toolset to AI/ML components.
- Develop resilience concepts, modeling constructs, and vulnerability assessment frameworks for AI-heavy systems in the presence of adversarial threats [adversarial AI].
- Develop theoretical foundations with verifiable robustness.
- Consider new generations of attack and defense methods for comparison.
- Consider Workforce Development/Human Capital Development efforts, particularly in light of workforce shortages faced by both SE and AI and the competition for talent between technology companies and government agencies.

Table of Contents

INTRODUCTION

This workshop was the first Artificial Intelligence for Systems Engineering/Systems Engineering for Artificial Intelligence (AI4SE/SE4AI) jointly sponsored by the US Army Combat Capabilities Development Command Armaments Center (CCDC AC) Systems Engineering Directorate (SED) and the Systems Engineering Research Center (SERC), a University Affiliated Research Center (UARC) for the Department of Defense (DoD). ***The objective of the workshop was to discuss and define Systems Engineering (SE) and Artificial Intelligence (AI) challenges, areas of exploration and methodologies to use, and ways in which to collaborate and research in the upcoming years***. A major challenge faced today is defining, analyzing, and capturing systems engineering artifacts and managing the configuration of those artifacts as systems are developed and deployed to the field. Another challenge is how to update current business processes for the dynamic nature of developing AI systems. To facilitate the objective defined for this workshop, the discussions were focused on three (3) main areas:

- Identify how Artificial Intelligence (AI) can support and create efficiencies in Systems Engineering (SE);
- Develop ways to apply effective Systems Engineering (SE) to Artificial Intelligence (AI) intensive systems; and
- Explore Digital Engineering (DE) and its relationship with Artificial Intelligence (AI).

WORKSHOP AGENDA STRUCTURE AND AUDIENCE

CCDC AC and the SERC jointly hosted the workshop virtually using Microsoft Teams over a two-day period. There were three (3) Keynote Speakers who provided their relevant perspectives for the daily sessions. The workshop was attended by more than 70 organizations to include: Other Government Agencies (OGA) and Industry and Academia affiliates with over 200 people registering and attending the virtual workshop. The workshop agenda was structured into the following four (4) sessions:

1. **Machine Learning/Artificial Intelligence (ML/AI)**
2. **Artificial Intelligence for Systems Engineering (AI4SE)**
3. **Systems Engineering for Artificial Intelligence (SE4AI)**

4. **Digital Engineering (DE)**

Each session had 3-4 speakers who presented on relevant topics, and audience members were able to collaborate and ask questions throughout the briefings via the chat feature. Each session was moderated with an interactive Q&A at the end.

Presentation materials for the entire workshop were made available via the SERC website (https://sercuarc.org/event/ai4se-and-se4ai-workshop/) to everyone who registered for the event.

WELCOME FROM THE EXECUTIVE SPONSORS

Dr. Dinesh Verma, *Executive Director, SERC*

Mr. Jeff Dyer, *Director, CCDC AC Systems Engineering Directorate*

Dr. Verma welcomed attendees to the workshop, noting that the remote nature allowed for greater participation and sharing of ideas and insights toward the goal of answering: How can the current possibilities offered by AI be taken to scale?

Mr. Dyer introduced the two-day event as an opportunity to learn from leaders already using AI in this space and share ideas among members of the Government, Academic and Industry communities to answer: Where do we go from here?

The hope of this two-day workshop was to identify how to apply SE effectively to deliver superior product(s) to the warfighter in the field. Challenges have been identified already at this early juncture, such as capturing and developing systems engineering artifacts in a dynamic manner. New SE methodologies and processes have to be developed to answer questions such as: how to determine, in a dynamic fashion, value on the battlefield; how to analyze AI solutions; how to structure the data for consumption; and how to update business processes to support adoption of AI.

CCDC Armaments Center Systems Engineering Directorate's AI Initiatives

Mr. Roshan Patel, *CCDC AC*

The Systems Engineering Directorate (SED) at CCDC Armaments Center recognizes that artificial intelligence and machine learning (AI/ML) are actively shaping the development of weapon systems. SED is entering an 18-month sprint to prompt multidisciplinary SE4AI activities, in which SE methodology will be refined for application to the development of AI technologies. The activities span the topics of: workforce development; requirements and architecture; systems engineering technical management; systems analysis; and advancing systems engineering methodology. Mr. Patel highlighted select activities as examples of how SED aims to refine SE practices and apply other SE processes that emerge as relevant.

SED is invested in discovering and learning by practice. Project applications will provide experiential lessons learned. Moving forward, SED hopes for more partnerships that will support current work and inspire new work.

KEYNOTE SPEAKERS

DAY 1 - MORNING KEYNOTE | US Army AI Task Force

Dr. Douglas M. Matty, *SES Director, Army Artificial Intelligence Capabilities*

Synchronizing and coordinating all AI efforts across the Army is challenging. Effective collaboration among all stakeholders is key to addressing the challenge. The US Army AI Task Force has gathered its knowledge about what is needed to develop AI projects to further the goal of more effective collaboration. Dr. Matty's presentation highlighted that partnership and awareness support synchronization and achievement.

AI technology is viewed as having the potential to transform the battlefield "from the back office to the front lines" (Dr. Mark T. Esper, Defense Secretary). Expectations of AI are high, and Dr. Matty highlighted questions to consider when adopting AI, such as what is the competitor's view of AI and what investments they have made in AI capabilities. "Whoever rules AI rules the

world" in commercial and military efforts, stated Dr. Matty, and these efforts can be intertwined.

The prominence of AI technology spurs Army leadership to develop AI capabilities. Dr. Matty presented an overview of initiatives undertaken by the AI Task Force toward increasing AI capabilities and empowering rapid integration and synchronization of AI across the Army and the Department of Defense (DoD). Highlighted initiatives included: identifying existing machine learning initiatives; developing frameworks and methods to scale projects to the enterprise; reviewing policies that impede deployment of AI technologies to the Army; establishing an AI test bed; developing a talent management plan; and providing support to JAIC National Mission Initiatives.

The importance of partnerships to achieve early wins and future designs is highlighted throughout the AI Task Force's four lines of effort: 1. Achieve early wins; 2. Establish Army-wide AI culture; 3. Evolve the AI infrastructure; and 4. Set the conditions for AI adoption and integration. Effective collaboration enables leveraging of existing expertise to address identified gaps, produce efficiencies, drive other efforts, achieve wins, change culture, and evolve infrastructure. These efforts and initiatives support the AI Task Force's next step of developing an AI Initial Capability Document to provide a common framework to all stakeholders.

DAY 1 - AFTERNOON KEYNOTE | AI and Systems Engineering: A MITRE Perspective

Dr. Peter Schwartz, *The MITRE Corporation Principal Artificial Intelligence Engineer; AI Joint & Services Domain SME, AI & Autonomous Systems Department AI Domain Capability Area Lead, Army Programs Division*

Can the benefits that an organization gains from AI also benefit users and society? Dr. Schwartz stated that this question needs to be considered when deploying powerful technology like AI. The overview of relevant work currently undertaken across MITRE presented by Dr. Schwartz highlights the need for AI and SE to work together to avoid issues such as misalignment of services, unintended consequences and ethical concerns.

The overview of current work included the MITRE AI Maturity Model, an organizing construct that considers enterprise-wide SE issues and enables organizations to lay the foundation to realize the value of AI more efficiently. The AI Maturity Model considers 17 dimensions organized across five main pillars—Strategy, Organization, Technology, Data, and Operations—to measure the maturity of a private or government organization that is adopting AI, identify existing roadblocks and opportunities, and provide a roadmap toward effective adoption. It should be noted that the maturity model does not measure the maturity of any AI technology, but instead measures the maturity of the organization that is attempting to adopt AI. The AI Maturity Model is just one example of ongoing work across MITRE focused on how organizations can use SE to realize the full potential of AI.

Dr. Schwartz mentioned a recently published paper, "Designing a New Narrative to Build an AI Ready Workforce" (https://www.mitre.org/publications/technical-papers/designing-a-new-narrative-to-build-an-ai-ready-workforce). The paper communicates that changing the "narrative"—i.e., the established thinking and approach toward outcomes and the recruitment and retention of capable people—enables effective AI workforce development. Additionally, he provided a brief summary of one of the projects that was to be presented in Day 2, a MITRE-sponsored effort by the University of Virginia, Virginia Tech, and Old Dominion University to characterize the operating envelope of machine learning models.

DAY 2 – KEYNOTE | Establishing A Data Rich Decision Environment – ASA (ALT)'s Vision for a Transformational Army

Ms. Jeannette Evans-Morgis, *SES Chief Systems Engineer, Office of the Assistant Secretary of the Army (Acquisition, Logistics and Technology, ALT)*

Transformational change is critical to deal with the emerging reality in which peers and near peers have built their cadre of information, technology and talent pool to optimize the benefits of a data-rich environment. ASA (ALT) focuses on institutionalizing SE governance and building

processes and tools to realize the Army's vision of meeting needs based on data-rich decision environments.

Ms. Evans-Morgis presented the complexity of applying SE to support tasks such as identifying stakeholders, who owns the data, and how that data needs to be captured and maintained from concept development to force deployment to disposal. Data is a key component of this approach, and the team is currently focused on upfront and early harvesting of existing data or creating new data to use for the training of AI and machine learning.

The Critical Criteria Checklist (C3L) was discussed, an initiative currently at the concept stage that identifies the data needs to consider and how to share this information across different interfaces and platforms to understand operational performance. This initiative and ASA (ALT)'s approach emphasize the importance of collaboration between the operational/acquisition teams and the academic/lab teams to achieve the Army's overall goal of delivering relevant, trustworthy and reliable products to the soldier on the field faster.

WORKSHOP PRESENTATIONS – DAY 1

MACHINE LEARNING/ARTIFICIAL INTELLIGENCE (ML/AI)

Moderated by Dr. Valerie Sitterle, Georgia Tech Research Institute

Isolate, Predict and Evaluate the Impact of New Technologies and Emerging Threat Sources from Human Groups and Cultural Sources Using a Natural Language Processing with a Predefined Cognitive Bias

Dr. Carlo Lipizzi, *Stevens Institute of Technology*

Dr. Lipizzi presented a summary of work with CCDC AC on Threat-based Decision Systems (TBDS) that applied multimodal AI to transform insights embedded in text into actionable items that "tell machines what to do". Language evolves through time, defying a predefined way of evaluating text and the effectiveness of the typical top-down, model-based approach of TBDS. The team used Natural Language Processing (NLP) and machine learning (ML) to develop a framework that provides contextual sensemaking by ingesting streams of text and extracting key semantic metrics using developed algorithms that take into consideration the user cognitive bias (the "Room Theory"). These are combined into higher-level metrics and provide interactive visual representations.

Dr. Lipizzi summarized TBDS-use cases that targeted the following scenarios:

Case 1: "What if" analysis on the overall competitive scenario-based arena.

Case 2: Role playing in the competitive scenario-based arena.

Case 3: Emerging technologies radar screen.

The research team's work is distinguished from commercial solutions in that it offers an applied solution, customizable to needs, and allows for extracting semantic insights based on predefined points of view. A current early prototypal stage addresses technology horizon scanning. Future research will pursue further work on room theory and expanded capabilities.

Understanding Game Balance in Mosaic Warfare with Explainable Artificial Intelligence

Dr. Daniel DeLaurentis, *Purdue University*

The challenges of SE become more complex as systems and battle spaces become more complex. Mosaic Warfare envisions automated design and integration; however, AI-based decision making remains opaque to architects and military commanders, who would likely demand evidence for the quality of AI-based solutions on mission success. A core step in Mosaic Warfare is to assess the "balance of forces" under uncertainties against an adversary before determining the "win strategy". Recognizing the similarities between Mosaic Warfare and multiplayer strategy games, DARPA launched the Gamebreaker Program under the larger AI Exploration (AIE) Program. The program used real-time strategy games as a context to build AI-enabled tools to determine when a particular engagement is balanced and how to optimally achieve imbalance. Convolutional Neural Networks (CNNs) were used to model game balance, as they are well suited for training on multi-dimensional feature data obtained from gameplay. To identify promising game interventions that lead to significant changes in game balance—akin to multiple synthesis options for Mission Engineering in Mosaic Warfare—a recent development in XAI called Shapley Additive Explanations (SHAP) was used. The results from SHAP indicate the importance of different game factors towards the game balance predicted by CNN, thereby unraveling how the balance can be shifted towards the desired outcome.

Goals moving forward include identifying vectors of imbalance and generating the needed insights.

Adversarial Robustness of AI Models with Ensemble Diversity Optimizations

Dr. Margaret Loper, *Georgia Tech Research Institute*

Artificial intelligence (AI) and machine learning (ML) have been key for delivering the next generation of defense networks for vehicles, logistics, and systems. However, these networks are vulnerable to deceptive inputs. Understanding how to "fake" these weak features enable control and builds trust of model predictions. To defend against such adversarial attacks, the team sought to understand how these adversarial examples behave, what is similar about

them, and gain insights into how to defend against them. The resulting solution, called Cross-Layer Strategic Ensemble Defense, uses an ensemble of algorithms—instead of one—to create an input guard (to prevent training on deceptive data) and output guards (to verify predictions). Dr. Loper presented a summary of the three steps for developing the ensemble of algorithms for input and output guards. Preliminary data shows higher defense success rates with the ensemble approach, demonstrating the value and robustness it adds to defense. Steps are being taken to apply this approach to real-time object identification in video, which has broad application including in smart cities and future war-fighting environments.

ARTIFICIAL INTELLIGENCE FOR SYSTEMS ENGINEERING (AI4SE)

Moderated by Dr. Cody Fleming, Iowa State University

Houston: An Intelligent Requirements Advisor

Mr. Paul Wach, *Virginia Tech*

Requirements are the formulations of the problems an engineer seeks to solve. The research team developed Houston, an intelligent systems engineering advisor that supports the (human) engineer in identifying gaps as requirements are formulated and is implemented as a plugin for a Systems Modeling Language (SysML) software environment. If potential gaps in the set of requirements are identified, the decision authority remains with the human engineer as how to address the gaps. Houston contributes to assessing requirements validation, beyond simply verifying model construction. Mr. Wach presented examples of interactions between Houston and the engineer that resulted in a more complete set of requirements.

The research team is currently working on Structural Rules—a set of rules developed to demonstrate completeness—and is looking toward working on Interpretation of Learning—interpreting the elements in the requirements model that lead to learning capability to leverage organizational knowledge.

Using AI/ML Approaches to Support Data Analysis Process Improvement

Mr. Austin Ruth, *Georgia Tech Research Institute*

With a large unprocessed backlog of data and an expanding amount of military platform data collection requirements and capabilities, a key bottleneck is the ability to quickly and accurately analyze this data and provide actionable results to the warfighter. Data analysis generally is a labor-intensive task requiring ad hoc specialized processes that reduce the overall reusability of the analyses and increase turnaround time. Automated Test & Logistics Analysis Support (ATLAS) is a data analysis process improvement platform being developed by the Georgia Tech Research Institute that uses available AI and ML technologies and is built on Airflow to manage complex workflows and interfaces on remote data archives. ATLAS is plug-in based, allowing for the addition of either brand new or existing analysis techniques. The ATLAS framework can streamline and significantly improve the current DoD data analysis process by using AI/ML techniques and parallel processing. Mr. Ruth presented examples of ATLAS used to perform intelligent analysis of flight test data. Results of initial testing demonstrated the analysis of one flight test mission (normally 2.5 hours) completed in 24 minutes, and the steps that made this time efficiency possible were presented.

The team looks toward research into: identifying new AI techniques that are better suited to the purpose; improving data to fit with AI techniques; deploying ATLAS to cloud-based systems; and making ATLAS more intelligent.

Mitigating Design Error Archetypes in the Development of Explainable-Machine Learning (X-ML) Systems

Dr. Lance Sherry, *George Mason University*

Traditionally, embedded operational control systems are developed by human engineers, a time-consuming and costly process subject to error. The research team focused on recent advances in AI, specifically Explainable - Machine Learning (X-ML), that could be used to automate the manual process of designing operationally embedded control systems (OECS) and significantly reduce the development cost and time, and incidence of error. In X-ML developed

operationally embedded systems, the X-ML training algorithm uses massive data sets of the behavior of the systems in operation to perform "supervised training" and derive the stimulus-response behavior from the data.

Dr. Sherry acknowledged the interest of the autonomous vehicle industry, specifically in identifying what kind of problems might be encountered as X-ML systems are deployed for vehicles. The research team's study included focus on the three design error archetypes that occur in ML/AI guidance and control systems for vehicles—SGB Table Missing Input; SGB Table Missing Input/State Combinations; and SGB Table Missing Mapping between Input/State Combinations to Behaviors—and approaches to mitigate these. Results demonstrate a significant reduction in the development life cycle (e.g., a two- to three-year traditional engineering process accomplished in two to three weeks by the X-ML engineering process), as well as cost savings.

The team is advocating for initiatives including: Fast-Time Emergent Scenario Simulation (FTESS), which allows interactions run over extended periods of time to identify situations before they occur in the field, yielding data that can be used for training ML; and Collaborative functional design using X-ML to identify gaps in the model and insert the human operator/engineer to close the gaps before finally generating the code.

Automated Detection of Architecture Patterns in MBSE Models

Mr. Matthew Cotter, *The MITRE Corporation*

The evaluation of a system's architecture is a complex, iterative and essential process conducted by systems engineers across all domains. Commercially available Model-Based-Systems Engineering (MBSE) tools, when combined with standards-based architecture modeling languages, provide a means through which architecture information can be expressed graphically and formally in a machine-readable format. Mr. Cotter presented an automated, repeatable method for detecting patterns-of-interest embedded within an MBSE model that can assist a systems engineer in evaluating, and improving, an architecture. The method uses a

heuristically guided set of similarity measures that depend on textual and graphical content within the model. The proposed method was implemented and applied against architectures developed in IBM's Rational Rhapsody, and No Magic Inc.'s MagicDraw, and proved able to identify well-established design patterns that have their origins in object-oriented software design. However, this method is not limited to object-oriented design patterns, and may be easily re-applied against any sequence of architecture elements that conform to the standards-based architecture modeling language chosen. This method provides automation that has the potential to produce cost and time savings during the architecture evaluation processes, as well as add an additional degree of rigor and completeness to an architecture evaluation.

Observable challenges included: model size may strain commercial tool XMI generation; pattern detection can be limited by SysML's ability to describe multiplicities for some elements of a pattern; and architecture patterns and anti-patterns are not always clearly or consistently defined in literature or well-understood or documented for a given domain or context. The research team looks toward the development of a more robust library of patterns and anti-patterns.

WORKSHOP PRESENTATIONS – DAY 2

SYSTEMS ENGINEERING FOR ARTIFICIAL INTELLIGENCE (SE4AI)

Moderated by Mr. Tom DeVoe, CCDC AC

Artificial Intelligence Certification in Operational Environments

Mr. Tyler Cody, *University of Virginia,* Dr. Erin Lanus, *Virginia Tech* and Dr. Sachin Shetty, *Old Dominion University*

The team's mathematical definition and empirical construction of operating envelopes for ML models was presented, which give systems designers and operators a theoretically principled means of: 1. anticipating and detecting drops in model performance in new environments; and 2. engineering improvements to model performance. Three metrics were described—transfer distance, combinatorial coverage, and combinatorial set difference—that quantify the operating envelope using empirical measures based on a learning theoretic framework. This theory-based, classifier agnostic and label-free sequential transfer learning approach attempts to resolve transferability issues between operating envelopes with minimal data requirements. Looking forward, research will focus on showing how meta-data can extend this learning theoretic envelope and make it a less broad metric.

Evolving Systems Engineering Methods for Artificial Intelligence and Machine Learning

Dr. Rosa Heckle, *The MITRE Corporation*

A new systems engineering/research vortical model concept was presented as a more efficient R&D approach for complex AI/ML enabled capabilities. A vortical model was described that extends the foundational and well-characterized spiral systems engineering and development model to incorporate the flexibility of agile systems engineering methods. Additionally, the vortical model's iterative framework demonstrates and validates emerging advances in research for integration as new capabilities at various technology insertion points. A case study

was presented that showed the model used in the R&D of an AI/ML-enabled document image translation application.

The overall goal of this work is to enable organizations to adopt best practices that facilitate the fusion of research systems engineering and knowledge transfer for rapid and effective mission integration. This approach will require a new team make up, new skill sets and continuous communication across the team. Looking forward, research aims to refine the model with the goals of incorporating more detail and developing a full set of best practices.

Human Data Collection for Machine Learning and Artificial Intelligence Aid Development

Dr. Elizabeth Sibolboro Mezzacappa, *CCDC AC*

The CCDC AC's Tactical Behavior Research Laboratory has conducted efforts in generating human data sets for ML analyses, and conducted testing, evaluation, and validation of notional artificial aids. The presentation described efforts across three categories—human data gathering efforts for a targeting prioritization, human threat identification in an urban environment, and electrophysiological interfaces. Data from these efforts were inputted into appropriate ML methodologies to derive algorithms for AI aids that significantly help the soldier. The specialized facilities (the Experimental Verification & Validation Assessment Lab) for gathering these data were described. An overview was provided of the Lab's processes and pipeline, from initial data gathering, to algorithm development, to testing and evaluation and verification and validation of AI aids. Looking ahead, areas for focus will include requirements generation, identification of the human limitations of using AI aids and whether a soldier performs better with AI than without.

Karat: A Visual Framework for Constructing Neural Networks

Mr. Frazier N. Baker, *Georgia Tech Research Institute*

Neural networks, increasingly popular for ML applications, present an accessibility barrier to users who lack programming expertise or technical depth in the field of ML. The accessibility issues for creating, manipulating, and understanding deep neural networks transcends to the

end-users of ML models, who lack the understanding required to trust and troubleshoot the models. Karat addresses this accessibility issue by facilitating the communication that supports the collaboration needed for design. Karat is an open-source, interactive visual workspace for constructing, tuning, and training neural network architectures. It allows whiteboard-level designs of neural networks to be directly translated into viable ML models based on existing ML frameworks. Basic building blocks of neural networks are available, along with sensible defaults for hyperparameters. Users can save, reuse, and share neural networks, in whole or in part, with or without parameters, without ever having to touch code. This approach enables transfer learning, simplifies the visualization of complex network designs, and facilitates deliveries and collaborations. Looking forward, research and work will focus on automatic hyperparameter tuning, tutorial generation for users new to neural networks and ML, and integration and testing.

DIGITAL ENGINEERING

Moderated by Mr. Allan Lagasca, CCDC AC

Collaborative Functional Design Using Explainable Machine Learning (X-ML)

Dr. Lance Sherry, *George Mason University*

Operationally Embedded Control Systems (OECSs), widely used in military applications, are complex and require operational reliability of at least five-nines for “airworthiness” approval. Traditional systems engineering practice specifies the functional requirements for the OECS using a manual engineering process whereby engineers and operators collaborate to define what the automation should do at all times. The presentation described a method that uses Explainable Machine Learning (X-ML) to complement this manual process. Flight trajectory data, generated from revenue service operations and/or simulators, is used to train and test an X-ML model, which is then used as the starting point or to enhance the manually generated functional specification. A pilot program demonstrated the potential of the X-ML method to reduce the time and cost of development, increase the functional complexity of the behavior of systems, and reduce/eliminate design errors.

Design of Digital Twin Architectures That Support AI and ML Formalisms Working Side-by-Side as a Team

Dr. Mark Blackburn, *Stevens Institute of Technology*

State-of-the-art AI and ML technologies are fragmented in their capabilities. To understand the benefits of having them learn "how to play together", research focused on the design of digital twin architectures that support AI and ML formalisms working side-by-side as a team. A key research challenge was to design digital twin elements and their interactions to support: 1. methods and tools for model-centric engineering; and 2. digital twin operating system environments for observation, reasoning and control. The starting point for research was understanding the range of possibilities for which machine learning of graphs and their attributes can support decision making activities in systems engineering and digital twin systems operation. The first months of work focused on auto-encoders, a class of ML algorithms, that teach a machine to replicate the topology of various kinds of graph structures. Future work will include the development of: 1. formulas to design neural network architectures that represent specific types of graphs, compose graphs together, and even organize graphs into hierarchies; and 2. mechanisms for AI/ML interaction and teamwork.

Human Machine Teaming Elements of AI-enabled Course of Action Wargaming

Dr. Cindy Dominguez and Ms. Patricia McDermott, *The MITRE Corporation*, Mr. Adam Brown, *Parsons Corporation*

The Army's Combat Capabilities Development Command (CCDC) C5ISR Center sponsors the AI COA Recommender (AICR) project to address the gaps that exist between speed and efficiency and better decision making in the COA wargaming step of the Military Decision Making Process. Soldiers who might use AI as part of a technology system supporting their COA wargaming consider an array of capabilities on terrain, at multiple levels, while predicting a chain of action, reaction, and counteraction within severely limited timeframes. Understanding and accounting for these factors when designing future AI capability is a critical part of the systems engineering process. The AICR project uses human machine teaming knowledge elicitation and analysis

methods to understand the AI functionality needed within this process, and for specific user stories that guide the development of user interface mockups and software capability. The presentation outlined: the methods used, along with the human machine teaming systems engineering framework from which these are derived; the design of interview questions for knowledge elicitation, structure of qualitative coding and analysis; the development of themes and user stories that inform this specific space; and the integration of these efforts within a broader Scaled Agile Framework (SAFe) systems engineering. Looking ahead, the AICR effort is on the path toward: functionality and design that is traceable to analyzed expert data; and increased connection between the entire team effort and soldiers' needs.

Automated Generation of Expert Systems Thinking Patterns Using a Convolutional Neural Network

Mr. Ross Arnold, *CCDC AC*

Systems Thinking—the ability to understand and affect systems of all kinds—is a key systems engineering skill, especially as system complexity grows. Systems Thinking research reveals new methods to evaluate and assess Systems Thinking skills through simulation; however, limitations include the use of human test subjects and the complexity of the simulation tasks. Improvement of Systems Thinking assessment is the goal of training a Convolutional Neural Network (CNN) using an existing Systems Thinking simulation and its associated expert thinker patterns generated from human test subject data. Such a CNN could be presented with new, different variations of the simulation and automatically generate optimal solutions to the simulation tasks. Constants such as simulation physics could be altered, but general gameplay kept the same, allowing a CNN to generate expert thinker patterns to which human player patterns could be compared and assessed. This, in turn, reduces the number of human test subjects required and provides a scalable Systems Thinking assessment method. The highlighted benefits of this approach include its ability to: be scaled; generate new simulation physics randomly; provide infinite different simulation tasks under a similar user interface; investigate Systems Thinking independently from user interface or "game knowledge"; and

address limitations of the current approach. Future work will aim to address the limitations posed by data volume and the need for human review of results by an expert panel.

WORKSHOP OUTCOMES

SERC Roadmap of AI in Systems Engineering

The "AI4SE" and "SE4AI" labels have become metaphors for an upcoming rapid evolutionary phase in the Systems Engineering community.

- **AI4SE** applies augmented intelligence and machine learning techniques to support systems engineering practices. Goals in such applications include achieving scale in model construction and confidence in design space exploration.
- **SE4AI** applies systems engineering methods to learning-based systems' design and operation.

The presentation and discussion outcomes of the AI4SE/SE4AI Workshop helped drive the initial development of the roadmap presented in the following figure. This notional roadmap links the discipline of systems engineering to various trends in AI and its application to automation in systems. Previously, the SERC had published an initial set of roadmaps in 2019 (https://sercuarc.org/wp-content/uploads/2020/06/ROADMAPS_2.5.pdf) that provided only an AI/Autonomy Framework. (See figure on the following page.)

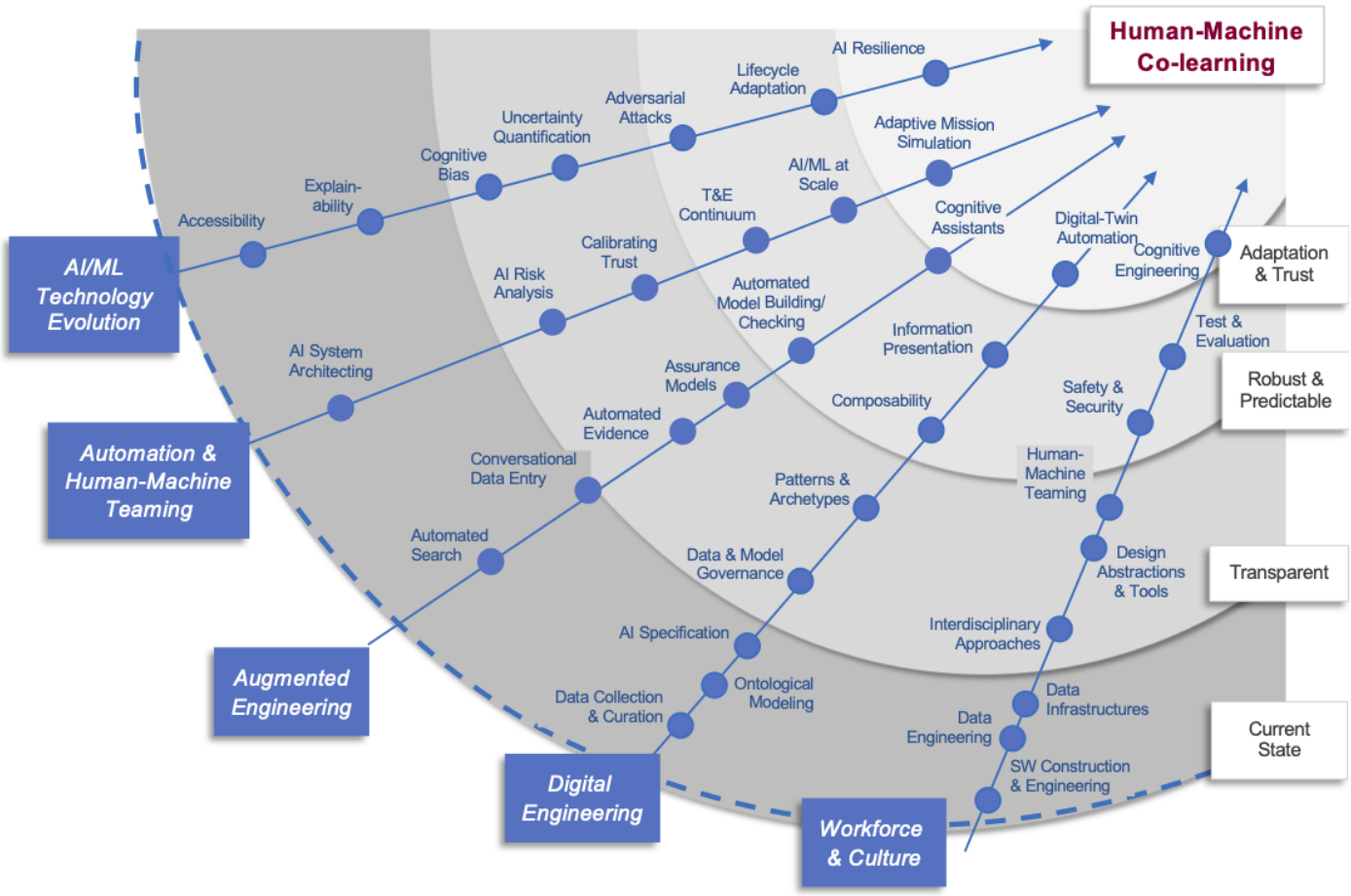

Figure: AI4SE/SE4AI Roadmap

The envisioned long-term outcome is “Human-Machine Co-learning”, in which both humans and machines adapt their behavior over time by learning from or alongside each other. More importantly for systems engineering, this is a lifecycle model that is not envisioned nor supported by most of the current-day systems engineering practices. This implies that a fairly significant transformation of systems engineering methods, tools, and practices is underway.

To achieve this end state, it might be considered necessary for both the AI and SE disciplines to pass through a set of “waves” or eras. The first of these includes sets of technologies and approaches that make the decisions produced by AI systems more *transparent* to the human

developers and users. Today much of this work is termed Explainable AI but it also includes greater transparency and understanding of the methods and tools used to develop AI applications, the underlying data, and the human machine interfaces that lead to effective decision making in the type of complex systems SE deals with routinely. The second wave is to produce systems that learn and are also appropriately *robust and predictable* in the type of critical applications normal to SE. This particularly includes both human and machine behaviors in joint decision environments, which are highly reliant on good human-systems design and presentation of decision information. It also includes adaptation of test and evaluation processes to co-learning environments. The third wave involves systems that *adapt* and learn dynamically from their environments. In this wave, machine-to-machine and human-to-machine (and maybe machine-to-human) *trust* will be critical. Trust implies dependence between the human and machine, which must emerge from human-machine interaction.

The vectors of this notional roadmap span five categories:

6. **AI/ML Technology Evolution** vector recognizes that the technological implementation of AI systems will evolve and will need to evolve in directions relevant to SE. Most of these can be related to the development of transparency and trust in technology.
7. **Automation & Human-Machine Teaming** vector recognizes that the purpose of AI in systems is generally to provide automation of human tasks and decisions.
8. **Augmented Engineering** vector recognizes that AI technologies will gradually be used more and more to augment the work of engineering.
9. **Digital Engineering** vector recognizes that the current digital engineering transformation will be an enabler for augmented engineering.
10. **Workforce and Culture** vector recognizes a significant transformation will need to be accomplished in the SE workforce, with significantly greater integration of software and human behavioral sciences at the forefront.

Future Work

The two-day event concluded with a discussion around the question, "Based on your experience in this workshop, what do you think are the top 2-3 ideas or initiatives you believe are worthy of further discussion?" The following summarizes feedback gathered for future exploration.

- Based on the keynote given by Dr. Matty, *establishing an AI test bed* may require a test bed that not only contains data sets, but one in which the data sets can be created and manipulated to affect the outcome. The notion of a federated test bed was raised during this discussion.
- Explore a path forward for AI and XAI for systems engineering and mission engineering.
- Integrate digital twins in the operational phase.
- Embed ethics consideration of AI (Defense Innovation Board principles) into the engineering process and into aspects of using AI for decision-support/decision-making, training, data sets and security.
- Test and Evaluation (T&E) with AI/ML components.
- Consider collaborative functional design.
- Extend the current SE toolset to AI/ML components.
- Develop resilience concepts, modeling constructs, and vulnerability assessment frameworks for AI-heavy systems in the presence of adversarial threats [adversarial AI].
- Develop theoretical foundations with verifiable robustness.
- Consider new generations of attack and defense methods for comparison.
- Consider Workforce Development/Human Capital Development efforts. Individually, the areas of SE and AI face workforce shortages and competition between technology companies and government agencies for talent. These challenges grow substantially at the intersection of the fields. There was broad consensus among participants that initiatives in SE/AI workforce development are needed and that this topic should be a priority for a follow-on workshop.

The SERC and CCDC AC intend to organize a follow-on workshop in the 2021 time frame. This second AI4SE/SE4AI workshop will explore specific outcomes from this initial 2020 meeting through an agenda developed to generate significant interchange and dialogue among participants. The intended goal will be to curate panel discussions and facilitate group breakout sessions that produce actionable applications of AI4SE and SE4AI as well as relevant research ideas.

ACKNOWLEDGEMENTS

The organizers would like to express thanks to the presenters in this workshop who generously shared their knowledge and experience. It was a unique opportunity to bring the community together despite unprecedented times. Thank you to The MITRE Corporation for their support, to CCDC AC and SERC for planning and facilitating, and to all the attendees for the open discussion, ideas and information exchange.

Appendix

WORKSHOP ORGANIZERS

Executive Hosts:

Dr. Dinesh Verma, *SERC Executive Director, Stevens Institute of Technology*

Mr. Jeffrey Dyer, *Director, CCDC AC Systems Engineering Directorate*

Workshop Leads:

Dr. Peter Beling, *SERC – University of Virginia*

Dr. Rosa Heckle, The MITRE Corporation

Mr. Tom McDermott, *SERC – Stevens Institute of Technology*

Ms. Kara Pepe, *SERC – Stevens Institute of Technology*

Mr. Albert Stanbury, *CCDC AC Systems Engineering Directorate*

Track Moderators:

Machine Learning / Artificial Intelligence

Dr. Valerie Sitterle, *Georgia Tech Research Institute*

Artificial Intelligence for Systems Engineering (AI4SE)

Dr. Cody Fleming, *Iowa State*

Systems Engineering for Artificial Intelligence (SE4AI)

Mr. Tom DeVoe, *CCDC AC*

Digital Engineering

Mr. Allan Lagasca, *CCDC AC*

ACRONYM LIST

AICR – Artificial Intelligence COA (Course of Action) Recommender

AIE – Artificial Intelligence Exploration

AI/ML – Artificial Intelligence/Machine Learning

ASA (ALT) – Assistant Secretary of the Army (Acquisition, Logistics and Technology)

ATLAS – Automated Test & Logistics Support

C3L – Critical Criteria Checklist

C5ISR Center – Command, Control, Computers, Communications, Cyber, Intelligence, Surveillance and Reconnaissance Center

CCDC AC – Combat Capabilities Development Command Armaments Center

CNN – Convolutional Neural Network

DARPA – Defense Advanced Research Projects Agency

DoD – Department of Defense

DE – Digital Engineering

FTESS – Fast-Time Emergent Scenario Simulation

JAIC – Joint Artificial Intelligence Center

MBSE – Model-Based Systems Engineering

NLP – Natural Language Processing

OECS – Operationally Embedded Control Systems

OGA – Other Government Agencies

SAFe – Scaled Agile Framework

SE – Systems Engineering

SED – Systems Engineering Directorate

SERC – Systems Engineering Research Center

SHAP – Shapley Additive Explanations

SysML – Systems Modeling Language

TBDS – Threat-based Decision System

UARC – University Affiliated Research Center

XAI – Explainable Artificial Intelligence

XMI – XML Metadata Interchange

X-ML – Explainable Machine Learning

RESEARCH WORKSHOP: **AI *for* SE & SE *for* AI**

OCTOBER 28, 2020

9:00 – 9:20	Check-in and Networking Time
9:20 – 9:30	Logistics & Ground Rules | Ms. Kara Pepe, *SERC*
9:30 – 9:45	**WELCOME from the Executive Sponsors** | Dr. Dinesh Verma, *Executive Director, SERC and* Mr. Jeff Dyer, *Director, CCDC AC Systems Engineering Directorate*
9:50 – 10:35	**KEYNOTE** | **US Army AI Task Force** Dr. Douglas M. Matty, SES Director, Army Artificial Intelligence Capabilities
	MACHINE LEARNING / ARTIFICIAL INTELLIGENCE
10:40 – 11:10	*Isolate, Predict and Evaluate the Impact of New Technologies and Emerging Threat Sources from Human Groups and Cultural Sources Using a Natural Language Processing with a Predefined Cognitive Bias* Dr. Carlo Lipizzi, Stevens Institute of Technology
11:10 – 11:20	*10-min Break*
11:20 – 11:50	*Understanding Game Balance in Mosaic Warfare with Explainable Artificial Intelligence* Dr. Daniel DeLaurentis, Purdue University
11:55 – 12:25	*Adversarial Robustness of AI Models with Ensemble Diversity Optimizations* Dr. Margaret Loper, Georgia Tech Research Institute
12:30 – 1:00	**MACHINE LEARNING Q&A**
1:00 – 1:20	*20-min Break*
1:20 – 2:05	**KEYNOTE** | **AI and Systems Engineering: A MITRE Perspective** Dr. Peter Schwartz, The MITRE Corporation Principal Artificial Intelligence Engineer; AI Joint & Services Domain SME, AI & Autonomous Systems Department AI Domain Capability Area Lead, Army Programs Division
	ARTIFICIAL INTELLIGENCE FOR SYSTEMS ENGINEERING (AI4SE)
2:10 – 2:40	*Houston: An Intelligent Requirements Advisor* Mr. Paul Wach, Virginia Tech
2:45 – 3:15	*Using AI/ML Approaches to Support Data Analysis Process Improvement* Mr. Austin Ruth, Georgia Tech Research Institute
3:20 – 3:30	*10-min Break*
3:30 – 4:00	*Mitigating Design Error Archetypes in the Development of Explainable-Machine Learning (X-ML) Systems* Dr. Lance Sherry, George Mason University
4:05 – 4:35	*Automated Detection of Architecture Patterns in MBSE Models* Mr. Matthew Cotter, The MITRE Corporation
4:40 – 5:10	**AI4SE Q&A**
5:15 – 5:30	**DAY 1 CLOSING / AGENDA REVIEW – DAY 2** *Dr. Peter Beling (UVA) and Mr. Al Stanbury (CCDC AC)*

RESEARCH WORKSHOP: **AI *for* SE & SE *for* AI**

OCTOBER 29, 2020

8:45 – 9:00	Check-in and Networking Time
9:00 – 9:15	**WELCOME / REVIEW WORKSHOP GOALS / LOGISTICS** Dr. Peter Beling, UVA *and* Mr. Tom McDermott, *SERC*
9:15 – 9:45	*CCDC Armaments Center Systems Engineering Directorate's AI Initiatives* Mr. Roshan Patel, CCDC AC
9:50 – 10:20	**KEYNOTE | Establishing A Data Rich Decision Environment – ASA (ALT)'s Vision for a Transformational Army** Ms. Jeannette Evans-Morgis, SES Chief Systems Engineer, Office of the Assistant Secretary of the Army (Acquisition, Logistics and Technology)
	SYSTEMS ENGINEERING FOR ARTIFICIAL INTELLIGENCE (SE4AI)
10:25 – 10:55	*Artificial Intelligence Certification in Operational Environments* Mr. Tyler Cody, University of Virginia / Dr. Erin Lanus, Virginia Tech / Dr. Sachin Shetty, Old Dominion University
11:00 – 11:30	*Evolving Systems Engineering Methods for Artificial Intelligence and Machine Learning* Dr. Rosa Heckle, The MITRE Corporation
11:30 – 11:40	*10-min Break*
11:40 – 12:10	*Human Data Collection for Machine Learning and Artificial Intelligence Aid Development* Dr. Elizabeth Sibolboro Mezzacappa, CCDC AC
12:15 – 12:45	*Karat: A Visual Framework for Constructing Neural Networks* Mr. Frazier N. Baker, Georgia Tech Research Institute
12:50 – 1:20	**SE4AI Q&A**
1:20 – 1:35	*15-min Break*
	DIGITAL ENGINEERING
1:35 – 2:05	*Collaborative Functional Design Using Explainable Machine Learning (X-ML)* Dr. Lance Sherry, George Mason University
2:10 – 2:40	*Design of Digital Twin Architectures that support AI and Machine Learning Formalisms Working Side-by-Side as a Team* Dr. Mark Blackburn, Stevens Institute of Technology
2:40 – 2:50	*10-min Break*
2:50 – 3:20	*Human Machine Teaming elements of AI-enabled Course of Action Wargaming* Dr. Cindy Dominguez *and* Ms. Patricia McDermott, The MITRE Corporation / Mr. Adam Brown, Parsons Corporation
3:25 – 3:55	*Automated Generation of Expert Systems Thinking Patterns using a Convolutional Neural Network* Mr. Ross Arnold, CCDC AC
4:00 – 4:30	**DIGITAL ENGINEERING Q&A**
4:35 – 4:45	**Workshop Wrap up and Next Steps** Mr. Tom McDermott, *SERC*
5:15 – 5:30	**CLOSING Remarks** Dr. Dinesh Verma, *Executive Director, SERC* and Mr. Jeff Dyer, *Director, CCDC AC Systems Engineering Directorate*