Artificial Intelligence, Social Computing and Wearable Technologies, Vol. 113, 2023, 1–13
https://doi.org/10.54941/ahfe1004174

CHAAIS: Climate-Focused Human-Machine Teaming and Assurance in Artificial Intelligence Systems – Framework Applied Toward Wildfire Management Case Study

Taissa Gladkova[1], Dhanuj Gandikota[1], Sanika Bapat[1], and Kristen Allison[2]

[1]The MITRE Corporation, Bedford, MA 01730, USA
[2]United States Department of Agriculture Forest Service, Vallejo, CA 94592, USA

ABSTRACT

Climate change and the resulting cascade of impacts pose a real and urgent threat to human safety. Simultaneously, products from Artificial Intelligence (AI) research have grown exponentially and show high potential towards use in climate adaptation. However, an increasingly large barrier to responsive deployment and adoption of AI tools into climate change adaptation workflows is the actionable knowledge discrepancy between the fields of AI, Human Machine Teaming (HMT), AI Assurance, and the work of climate adaptation decision makers. To ensure alignment, applications of AI to climate change adaptation actions need a framework and knowledge base that map development considerations to the decision maker workflow. This paper introduces CHAAIS (Climate-focused Human-machine teaming and Assurance in Artificial Intelligence Systems), a design standard and accompanying knowledge base detailing the necessary human element of AI interaction in the high-risk domain of climate change. CHAAIS incorporates direct user interaction, decision maker adoption considerations, and downstream implications. Our process combines accepted HMT and AI Assurance principles for ethical design while testing specific issues in their intersection in the climate change domain. Specifically, we demonstrate this process with a case study in forestry and implications for wildfire management. The goal for the CHAAIS design framework and knowledge base is to be both a living information source and an adaptable method of tailoring future climate change AI solutions for responsive deployment directly informed by climate decision makers.

Keywords: Human-machine teaming, Artificial intelligence, Climate, Climate change, AI assurance, Climate adaptation, High-risk domains, Wildfire

INTRODUCTION

Humanity is already feeling the effects of the threats posed by climate change and the resulting cascading dangers (World Meteorological Organization, 2022). Climate adaptation decision makers work at allocating resources towards mitigating, preparing, responding, and recovering from climate impact on communities (United States Federal Emergency Management Agency (FEMA), 2023). Climate Decision Makers (CDMs) include a variety of experts working in the climate space, including policy makers, emergency response teams, agencies allocating funding for climate-threat infrastructure, conservation agencies, and any other group responsible for generating and implementing courses of action that have climate relevance (Orlove et al., 2020). CDMs are often overburdened with analyzing geoclimatic datasets that are unstructured and cover large geographic areas with high dimensionality. Moreover, climate agencies and groups are chronically understaffed, and existing available tools are not optimized for the rapidly evolving urgent needs of the climate domain (Friedman, 2023).

Research in analogous domains shows potential in applying Artificial Intelligence (AI) to climate applications, as AI is equipped to glean insights from large datasets which would otherwise transcend normal computational requirements. When applying AI to high-stakes domains such as climate change where human lives are affected, AI tools must be assured and trustworthy. Trust is a critical antecedent for adoption of AI (Dorton & Harper, 2021). Given the high-risk domain of climate, interventions must be carefully analyzed to avoid exacerbating climate injustice, or the disproportionate impact of climate change effects on historically disadvantaged communities (University of California, 2022). Various frameworks and tools for developing trustworthy AI rely on abstract principles and are difficult to operationalize by AI developers and other stakeholders (Munn, 2023; Dorton et al., 2023). As such, AI developers need an actionable tool that bridges the domains of technological integration and responsible design when applying AI to climate.

Human-machine teaming (HMT) principles provide structure for creating an effective user-centered tool for collaborating with humans (McDermott et al., 2018). AI Assurance refers to the process of building trust and confidence in artificial intelligence systems by ensuring that they are safe, reliable, and aligned with human values (National Institute of Standards and Technology, 2023). It allows for measures of performance beyond accuracy. Combining these fields offers potential solutions for responsible AI deployment at the development level instead of the broad institutional guidance that currently exists (The Global Partnership on AI, 2021). Further, HMT and Assurance both generally combine the technical and human considerations for a problem that inherently involves both.

AI capabilities are usually developed and deployed as components within larger sociotechnical systems, comprised of workflows, people, and technologies. AI must be developed under the holistic consideration of how it may (sometimes adversely) affect the larger work system (Neville et al., 2022). While AI has a reputation for being a ‘silver bullet’ for capability gains, there are many cases where it has caused greater consequences in the long term (Banerjee & Chanda, 2020).

It is difficult to predict adverse emergent effects from integrating AI into a sociotechnical system (Dorton et al., 2023); however, it is the responsibility of AI developers and those who deploy it to ensure their creations do no harm. (Miller, 2011). This paper proposes an approach using HMT and Assurance principles which span these risks and offer a way to navigate them while properly communicating impacts to users. These principles span knowledge disciplines by translating between climate subject matter experts (the future users of the system) and technical experts.

This paper introduces Climate-focused Human-machine teaming and Assurance in Artificial Intelligence Systems (CHAAIS), a design framework and accompanying knowledge base detailing and organizing the necessary human element of AI interaction in the highly variable domain of climate change. CHAAIS incorporates direct user interaction, CDM adoption considerations, and downstream implications. Our process combines accepted HMT and AI Assurance principles for ethical design while testing specific issues in their intersection in the climate change domain. Specifically, we demonstrate this process with a case study in forestry and implications for wildfire management.

The CHAAIS framework and knowledge base is designed to offer a common operating framework between AI developers, CDMs, climate professionals, and policy makers. For AI developers, CHAAIS translates and organizes AI capabilities into actionable outputs and processes for the CDMs. They can leverage CHAAIS to translate complex AI technologies into capability gains and objectives within their workflows. Policy makers can utilize the framework and organized knowledge of the AI system's design process within the knowledge base as a 'receipt' of good design justification, which translates between domains and conveys trustworthy aspects to stakeholders throughout the process.

CHAAIS: PROCESS FOR RESPONSIBLE DEVELOPMENT OF AI FOR CLIMATE DECISION MAKERS

The CHAAIS process (Figure 1) provides a framework for developing AI applications through two high-level iterative processes: understanding and addressing stakeholder needs while understanding potential sociotechnical impacts. Within these two high-level processes are five phases, which are performed iteratively throughout the development process.

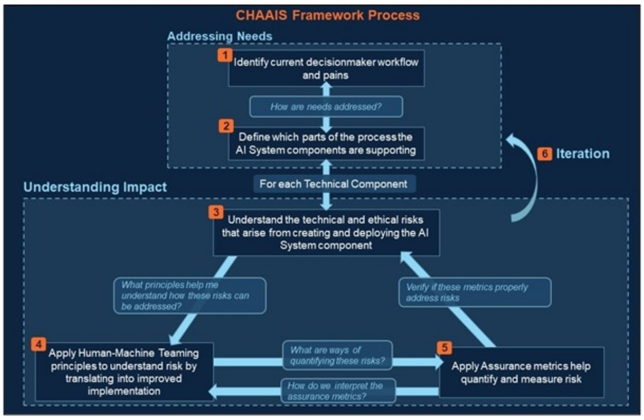

Figure 1: CHAAIS process for applying human factors and assurance in artificial intelligence tools for climate decision makers.

The first stage (addressing needs) emphasizes conceptualizing the AI application in the context of users and their work. During this stage the developers must ensure that the AI tool is addressing a defined need.

The second stage (understanding impact) focuses on recognizing and remedying any adverse impact(s) of the proposed AI tool. While phases nominally occur in sequence, CHAAIS is, in practice, an iterative process. Figure 2 provides further detail on the concepts and principles involved in each phase of CHAAIS.

Case Study–Wildfire Management

CHAAIS was developed and refined through the development of the Augmented Real Time 3-D Mapping with Intelligent Sensing AI (ART3MIS-AI) System (Gandikota et al., 2022). The ART3MIS-AI System's objective is to transfer and link geospatial AI research capabilities to autonomously extract critical metrics and representations for capability gains in wildfire mitigation.

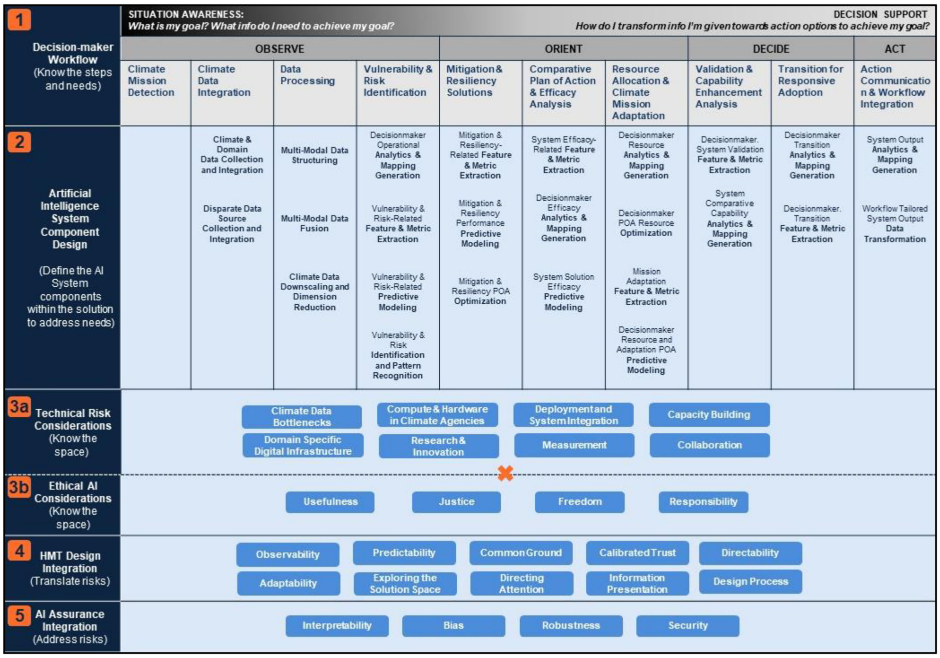

Figure 2: Detailed concepts in chaais process for applying human factors and assurance in artificial intelligence tools for climate decision makers.

ART3MIS was developed collaboratively with the United States Forest Service Fire and Aviation Management (USFS FAM), which involved visits to the Angeles National Forest (ANF). The ART3MIS-AI team worked with USFS FAM to understand user workflows for gathering information about areas that require mitigation techniques to minimize risk of future wildfires and various factors that drive decision making for mitigation implementation.

Interviews (N = 12) were conducted with stakeholders across the CHAAIS CDM workflow to understand their tasks, responsibilities, and roadblocks they faced. These interviews drove the development of user stories and journey maps moving into Phase 1 of CHAAIS. This stakeholder research was essential when establishing common ground with the AI. In fact, consultation with user groups should be performed as much as possible throughout the process, infused throughout the phases.

Phase 1: Identify Current Climate Decision Maker Workflows, Needs, and Pains

Human-centered design processes start by understanding the target user's current workflows. Although climate missions can be viewed as an end-to-end workflow, they are complicated by complexity in tasks, interdependencies across tasks, personnel, and knowledge requirements. A detailed understanding of CDM workflows, including their goals, methods, and tools/resources, is necessary to design the AI system to complement and enhance their work.

Techniques such as journey mapping, empathy mapping, and task analysis can be used to break down the process into phases, and identify user tasks, decisions, goals, resources, and challenges. For this particular application we begin with the Observe, Orient, Decide, Act (OODA) Loop as a baseline model of cognition for filtering and acting on information (Lewis, 2022). The OODA Loop approximates cognitive processes of many climate applications, especially emergency response (Huang, 2015). We acknowledge that there are other models of situation awareness (Endsley, 1995), sensemaking (Klein et al., 2006), and decision making (Klein, 1993), and even formal decision processes for course of action generation (Bryant et al., 2007); however, the OODA loop is sufficient for our purposes, as we can augment it with additional sub-phases that map to observed decision making processes. These sub-steps were derived from observed actions amongst CDMs based on experience through case studies and are generalizable but need verification.

We define tools supporting the Observe and Orient phases as "Situation Awareness" (SA) systems, while tools supporting Decide and Act phases as "Decision Support" (DS) systems. This aligns with the distinction between SA and decision making processes (Endsley, 1995; Pfaff et al., 2013).

Each type of AI system, (i.e., SA vs DS) will have unique issues with engendering trust with end users, which is an already complex and context-dependent phenomenon (Dorton & Harper, 2022). Users have been shown to resist technologies that are more prescriptive in nature, rather than supporting their actual cognitive needs (e.g., exploring the options jointly) (Moon & Hoffman, 2005).

Phase 1 Applied to Case Study

For applying Phase 1 to the ANF case study, we focus the findings here on the **Vulnerability and Risk Identification** segment of the workflow in Figure 2. Interviews with the ANF personnel indicated the need for improved remote sensing data for effective utility. Current remote sensing imagery utilized for fuel estimates not only remains outdated but is present at a 30m+ resolution.

This is below the needed actionable fuel resolution of 3-5m in the current active fire mitigation workflows.

Fire planners preferred more useful extracted metrics such as increased resolution of vegetation and stand structural features to calculate more accurate fuel values (used for risk estimation) and more accurate measures for pre- and post-analysis of mitigation techniques (e.g., prescribed fires). It was evident that ANF could benefit from leveraging AI tools to augment their current work.

Phase 2: Define Which Parts of the Process the AI Is Supporting

Phase 2 in the CHAAIS framework involves mapping AI components to specific tasks or functions that were enumerated in the previous phase. This approach reduces the scope of the design effort and facilitates focused discussions among developers and CDMs. It aligns developers' AI objectives to user objectives, grounded in their actual workflows.

Within the CHAAIS framework, we categorize the AI component objectives through their applied 'capability gains' in each step of the Phase 1 CDM workflow. Figure 2 shows the set of capability gain modules selected during observation of the CDM workflow and objectives in the ANF case study (e.g., Pattern Recognition, or Feature Extraction). It is important to note that the design of the CHAAIS framework allows for the integration and application of other capability gains beyond what are listed.

Phase 2 Applied to Case Study

For applying Phase 2 to the ANF case study, we focus on the choice of AI System Component for the **Vulnerability and Risk Identification** segment of the CDM workflow within Figure 2.

Stemming from the identified need for increased resolution and specific metric extraction, the team focused on the category of **Vulnerability and Risk-Related Feature and Metric Extraction** (Figure 2). After further design discussions with the USFS FAM team, the AI system component identified to achieve the objective was a series of modular design, deep learning object detection models within the ART3MIS-AI System which performed both 2-D and 3-D vegetation object detection to extract useful variables such as vegetation height, width, area, and biomass values.

Phase 3: Understand the Technical and Ethical Risks That Arise From Creating and Deploying AI Solutions

Although AI systems can have great positive effects on climate by helping in mitigation and management strategies, they can also create unintended harmful repercussions or struggle to achieve their original objectives due to technical risks to effective deployment. Each component of the AI system poses its own impact, and as such each should be analyzed separately and as a whole system.

The technical challenges in developing and deploying climate AI can result in a system that does not serve its users. This is especially problematic in the climate domain since developers must ensure that the benefits of the system outweigh the potential emissions posed by large-scale computing necessary for some AI methods. The Climate Change and AI: Recommendations for Government Action report highlights potential challenges for implementing AI for climate action (The Global Partnership on AI, 2021). For the purposes of CHAAIS, the various challenges have been adapted into technical risks that are relevant for individual AI system development and deployment, rather than the government agencies the report originally targets.

Technical Considerations

- **Climate Data Bottlenecks**: Is there a lack of data in my domain application? Will my system have access to existing data? Is the data collected at a reasonable rate for my application?
- **Domain-Specific Digital Infrastructure**: Are there domain-specific tools and software that the AI component must be compatible with? Are tools and software necessary for this domain inaccessible?
- **Compute & Hardware in Climate Agencies:** Do the end users have the necessary compute and hardware to run this AI component (in combination with the rest of the system)?
- **Research & Innovation:** Is this AI component risking a "silver bullet" approach instead of focusing on a well-tailored smaller application? Does our AI component consider historically underserved areas?
- **Deployment and System Integration**: Does our AI component risk slow adoption? Does our AI component have potential conflicts with legacy infrastructure?
- **Measurement**: Does our AI component have ways of measuring impact?
- **Capacity Building**: Does our AI component rely on technical AI expertise in the user-base? Does our AI component pose a risk of capture?
- **Collaboration**: Does our AI component discourage collaboration with too narrow a focus?

Ethical AI is widely sought after and hard to establish. IBM provides a list of principles to follow (IBM, 2022). Coeckelbergh discusses ethical and political challenges posed by integrating AI for climate applications and calls for responsible use of AI to dispute these challenges (Coeckelbergh, 2021). Nordgren discusses similar ideas and organizes them into a set of principles: Usefulness, Freedom, Justice, Responsibility (Nordgren, 2023). On the other hand, Munn argues that it is fruitless to enforce ethics through the barrage of buzzwords and principles, and rather suggests an approach that considers the system as a whole, since simply providing principles as ideals results in a lack of action towards these goals (Munn, 2023). CHAAIS attempts to combat this pitfall by leaning on Nordgren's principles but attempting to provide a path forward by translating them into actionable questions.

Ethical Considerations

- **Usefulness:** How could the usefulness of this AI component fall short of making up for the extra emissions from running it? Do the outputs of this component map to the needs of users?

- **Freedom:** How could this AI component infringe on the freedoms of individuals and communities? Should restrictions or manual overriding be in place?
- **Justice:** How could this AI component unjustly affect or unjustly represent specific individuals or groups, especially those historically marginalized? Does targeted data have proper scope and sampling?
- **Responsibility:** How can we, as the AI component developers, take responsibility for and ensure safe and ethical operation of the AI? What mechanisms can we establish for feedback and security? What happens to the system and data in case of failure?

When considering ethical and technical risks separately, it may be difficult to come up with "what ifs" that properly predict potential outcomes. Dorton et al. discusses in Foresight for Ethical AI that it is notably difficult to accurately predict AI harm (Dorton et al., 2023). A potential strategy is to employ the use of a premortem to guide catastrophizing in an effective way (Klein 2007; MITRE, 2023). Additionally, developers can give context to the ethical considerations by examining the intersectionality with technical considerations. For example, when examining justice, instead of attempting to tackle justice across the entire AI component, consider its overlap with data bottlenecks. How can limited data result in unjust representation? It is then easier to translate this potential risk into something the AI developer can tackle.

Phase 3 Applied to Case Study

From the Deep Learning Object Detection component chosen for **Vulnerability and Risk-Related Feature and Metric Extraction**, let's examine the **Technical and Ethical AI Risk** considerations, specifically **Usefulness** as applied to **Domain-Specific Infrastructure**. FAM personnel identified specific software that is highly integrated into their workflows; so, to ensure a useful product, it was essential that outputs of the AI system be compatible with this software. The consideration of usefulness also directly maps to user needs: the outputs must be filling a demonstrated need, or they are excessive or even frivolous. Specifically for ANF, when we consider the Object Detection AI component, identified needs include remote sensing imagery constraints, model output efficacy, and visualization considerations. Now the challenge is adequately minimizing this identified risk.

Phase 4: Apply Human-Machine Teaming Principles to Address Risk

Human-Machine Teaming principles provide a framework that makes it possible to address the identified technical and ethical risks. The CHAAIS framework focuses on MITRE's Human-Machine Teaming Principles, however there are multiple frameworks that incorporate similar ideas (McDermott et al., 2018).

For each identified technical and ethical risk, consider the list of HMT principles (Observability, Predictability, Directing Attention, Exploring the Solution Space, Adaptability, Directability, Calibrated Trust, Common Ground, Information Presentation, Design Process, all of which are thoroughly discussed in MITRE's document) and view the risk through that lens. Does enforcing that principle help resolve the risk? If so, how and why might the principle be violated? This helps developers understand how to concretely improve the technical system to minimize the identified risk.

Phase 4 Applied to Case Study

Consider how the HMT principles listed above help ART3MIS developers ensure the Deep Learning Object Detection component (Phase 2) extracts metrics which are useful enough to justify the resources to operate the system (Phase 3). Take for example the HMT principle of **Common Ground:** The Angeles National Forest has uniquely large areas of chaparral vegetation – the tool must be able to accurately detect this specific feature for the ANF, but also work for deciduous woods to be used more universally. The tool must be able to adapt its assumptions of likely vegetation to the geographic region to be useful to ANF wildfire mitigation stakeholders. To have effective metric extraction, the AI system must be aligned with the context that the user works in. Assurance techniques can be used to quantify how well the AI achieves these HMT metrics but should also be bolstered by qualitative analysis like whether users find the Object Detection reasonable given the ANF geography.

Phase 5: Apply Assurance Metrics to Help Quantify and Measure Risk

AI Assurance provides a framework for the development and evaluation of AI systems, ensuring that they meet the desired level of performance, safety, and reliability. It conveys the confidence in the AI system's ability to perform as intended while adhering to ethical guidelines and minimizing the risks associated with its deployment. This confidence is expressed through the careful evaluation of various AI Assurance areas, which are integral to the overall design process.

However, viewing AI Assurance as a separable static issue can lead to problems in the development and integration of AI systems. It is essential to consider AI Assurance as a dynamic, ongoing process that evolves with the design and implementation of the AI system. CHAAIS follows the NIST AI risk management frameworks values on AI Assurance distilled down to robustness (reliability under many conditions), interpretability (ease of human's understanding processes), safety (minimization of risk), and fairness (equitable treatment of individuals and groups) (National Institute of Standards and Technology, 2023).

In CHAAIS, AI Assurance serves as the connecting link between HMT requirements and AI System Design Integration, providing a means to explain risk minimization solutions and convey them back to the stakeholder.

Phase 5 Applied to Case Study

AI Assurance is crucial to make the user trust and rely on the model's Object Detection (Phase 3) output. AI Assurance principles such as **Robustness** and **Interpretability** are key ensuring stakeholders recognize a **Common Ground**

(Phase 4) with ART3MIS which strengthens trust in the AI system. To guarantee reliable and consistent performance of the framework, the Object Detection and Segmentation AI Modules must be robust to terrain and changing conditions for the ANF. This Common Ground is conveyed back to the user using in an iterative process to build trust and reliance. These AI Assurance measures are built into requirements for the Structural and Feature Extraction AI modules.

Robustness techniques play a crucial role in enhancing the performance of object detection and segmentation AI modules. These techniques include data augmentation, where various transformations are applied to input images to help the model generalize better (Na et al., 2022). Adversarial training incorporates attacks during training (Li et al., 2021). Different regularization techniques prevent the model from relying too heavily on single input features. Ensemble methods combine the predictions of multiple models, and error detection and correction mechanisms maintain performance even when the primary model fails (Wyatt et al., 2022). Transfer learning helps the model adapt to changes in input data distribution. Synthetic data can be used to augment models' training data with specific characteristics (Gao et al., 2023). Together, these techniques ensure reliable and consistent performance, ultimately strengthening trust among stakeholders.

Interpretability allows for stakeholders to understand the key factors impacting an AI model's predictions on identified risks, vulnerabilities, and structural metrics. Tree based models provide inherent interpretability, which can be utilized for higher risk tasks or as shadow models trained to explain a black box model's predictions. Deep models like Neural Networks can also be interpreted through techniques like LIME (Ribeiro et al., 2016) or SHAP (Antwarg et al., 2021) which can be used to generate explanations. Visualization of the AI system's decision-making process using tools like feature importance plots, partial dependence plots, or decision boundaries can also clarify the usefulness of different data portions that are informing decision making and in turn, build trust.

Phase 6: Iterate

Organizing, pursuing, and leveraging the information gathered from CDMs can be challenging, as these conversations often branch into multiple areas of knowledge. Linear design frameworks fail to adequately capture the intricacies of these discussions and the wealth of valuable information that emerges from them. Moreover, the conclusions drawn from these discussions require re-examination at the AI component level and the context of the entire AI system. Creating a comprehensive user feedback mechanism for the overall AI system is a non-trivial hurdle.

Within CHAAIS framework, Phase 6 highlights the importance of the iterative back-and-forth cycle that occurs across all stages of CHAAIS. The productive feedback discussions and questions that arise during application reviews are vital for reaching well-founded conclusions.

Phase 6 Applied to Case Study

Although ART3MIS is primarily a SA system (observe and orient), it was still important to gain information about other phases of the OODA Loop to see how the tool might expand in the future, and how specific metrics and visuals that the Orient phase outputs could effectively be used by CDMs in the remaining stages. For example, when proposing mitigation strategies during the decide phase, stakeholders hold public town halls where community members living at the impacted urban interface areas of the forest can provide input on changes to their "backyard". Currently it is difficult to describe changes resulting from a mitigation strategy like a prescribed fire and hard for the public to imagine what a 30% reduction in fuels *looks like,* potentially leading to distrust when reality does not align with expectation. This challenge during the decide phase of the OODA Loop prompted the design team to incorporate specific interactive visuals into the output of the risk identification (orient) phase that will make it easier to justify and visualize strategies during these meetings.

DISCUSSION

In the ART3MIS application, the CHAAIS framework was applicable for most AI components. While there was not always the possibility for perfect mapping across all the phases, the phases did provide valuable structure to ensure the development team had a comprehensive design and user research approach.

Although CHAAIS was developed through a climate-focused lens, it is generalizable to other high-consequence domains with similarly dynamic workflows. The CHAAIS organizational framework for sifting through multiple domains is still notional and needs more work to verify the approach. As the CHAAIS framework is further developed through additional use cases, we will have more ability to identify patterns of success and room for refinement and improvement.

ACKNOWLEDGMENT

The authors would like to acknowledge the United States Forest Service for valuable insights from a climate expert perspective.

This work was funded by MITRE's Independent Research and Development Program.

REFERENCES

Antwarg, L., Miller, R. M., Shapira, B. & Rokach, L., 2021. Explaining anomalies detected by autoencoders using Shapley Additive Explanations. *Expert Systems with Applications,* Volume 186.

Banerjee, D. N. & Chanda, S. S., 2020. *AI Failures: A Review of Underlying Issues.* [Online] Available at: https://arxiv.org/abs/2008.04073.

Bryant, D. J. et al., 2007. Development and Evaluation of an Intuitive Operational Planning Process. *Journal of Cognitive Engineering and Decision Making,* 1(4), pp. 434–460.

Coeckelbergh, M., 2021. AI for climate: freedom, justice, and other ethical and political challenges. *AI and Ethics,* 1(1), pp. 67–72.

Dorton, S. L. & Harper, S. B., 2022. A naturalistic investigation of trust, AI, and intelligence work. *Journal of Cognitive Engineering and Decision Making,* 16(4), pp. 222–236.

Dorton, S. L. & Harper, S., 2021. *Trustable AI: A Critical Challenge for Naval Intelligence.* [Online] Available at: https://cimsec.org/trustable-ai-a-critical-challenge-for-naval-intelligence/ [Accessed 7 July 2023].

Dorton, S. L., Minestero, L. M., Alaybek, B. & Bryant, D. J., 2023. Foresight for ethical AI. *Frontiers in Artificial Intelligence,* Volume 6.

Endsley, M. R., 1995. Toward a Theory of Situation Awareness in Dynamic Systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society,* 37(1), pp. 32–64.

Friedman, L., 2023. Depleted Under Trump, a 'Traumatized' E. P. A. Struggles With Its Mission. *The New York Times*, 20 June.

Gandikota, D. M. et al., 2022. AI Augmentation to Remote Sensing Imagery in Forestry Conservation & Restoration for Increased Responsive Capabilities. DC, USA, IEEE Applied Imagery Pattern Recognition Workshop (AIPR), pp. 1–16.

Gao, C. et al., 2023. Synthetic data accelerates the development of generalizable learning-based algorithms for X-ray image analysis. *Nature Machine Intelligence,* Volume 5, p. 294–308.

Huang, Y., 2015. Modeling and simulation method of the emergency response systems based on OODA. *Knowledge-Based Systems,* Volume 89, pp. 527–540.

IBM, 2022. *AI design ethics overview.* [Online] Available at: https://www.ibm.com/design/ai/ethics/ [Accessed 25 July 2023].

Klein, G. A., 1993. A Recognition-Primed Decision (RPD) Model. In: G. A. Klein, J. Orasanu, R. Calderwood & C. E. Zsambok, eds. *Decision Making in Action: Models and Methods.* Norwood, NJ: Ablex Publishing Corporation, pp. 138–147.

Klein, G., 2007. Performing a Project Premortem. *Haravrd Business Review*, pp. 18–19.

Klein, G., Moon, B. & Hoffman, B., 2006. Making Sense of Sensemaking 1: Alternative Perspectives. *Intelligent Systems, IEEE,* 21(4), pp. 70–73.

Lewis, S., 2022. *OODA Loop.* [Online] Available at: https://www.techtarget.com/searchcio/definition/OODA-loop [Accessed 25 July 2022].

Li, X. et al., 2021. *Estimating and Improving Fairness with Adversarial Learning.* [Online] Available at: https://arxiv.org/abs/2103.04243

McDermott, P. et al., 2018. *Human-Machine Teaming Systems Engineering Guide.* [Online] Available at: https://www.mitre.org/news-insights/publication/human-machine-teaming-systems-engineering-guide [Accessed 25 July 2023].

Miller, K. W., 2011. Moral Responsibility for Computing Artifacts: "The Rules". *IT Professional,* 13(3), pp. 57–59.

MITRE, 2023. *Premortem.* [Online] Available at: https://itk.mitre.org/toolkit-tools/premortem/ [Accessed 25 July 2023].

Moon, B. & Hoffman, R., 2005. How Might "Transformational" Technologies and Concepts be Barriers to Sensemaking in Intelligence Analysis?. Amsterdam, Proceedings of the Seventh International NDM Conference.

Munn, L., 2023. The uselessness of AI ethics. *AI and Ethics,* 3(3), pp. 869–877.

Na, S. et al., 2022. Development of an Artificial Intelligence Model to Recognise Construction Waste by Applying Image Data Augmentation and Transfer Learning. *Buildings,* 12(2).

National Institute of Standards and Technology, 2023. *NIST AI 100–1 Artificial Intelligence Risk Management,* s.l.: U. S. Department of Commerce.

Neville, K. J. et al., 2022. The TRUSTS Work System Resilience Framework: A Foundation for Resilience-Aware Development and Transition. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting,* 66(1), pp. 2067–2071.

Nordgren, A., 2023. Artificial intelligence and climate change: ethical issues. *Journal of Information, Communication and Ethics in Society,* 21(1).

Orlove, B., Shwom, R., Markowitz, E. & Cheong,. S.-M., 2020. Climate Decision-Making. *Annual Review of Environment and Resources,* Volume 45, pp. 271–303.

Pfaff, M. S. et al., 2013. Supporting Complex Decision Making Through Option Awareness. *Journal of Cognitive Engineering and Decision Making,* 7(2), pp. 155–178.

Ribeiro, M. T., Singh, S. & Guestrin, C., 2016. *"Why Should I Trust You?": Explaining the Predictions of Any Classifier.* [Online] Available at: https://arxiv.org/abs/1602.04938

The Global Partnership on AI, 2021. *Climate change and AI: Recommendations for government action,* s.l.: The Global Partnership on AI.

United States Federal Emergency Management Agency (FEMA), 2023. *Mission Areas and Core Capabilities.* [Online] Available at: https://www.fema.gov/emergency-managers/national-preparedness/mission-core-capabilities [Accessed 25 July 2023].

University of California, 2022. *What is Climate Justice?.* [Online] Available at: https://centerclimatejustice.universityofcalifornia.edu/what-is-climate-justice/

World Meteorological Organization, 2022. *State of the Global Climate in 2022.* [Online] Available at: https://public.wmo.int/en/our-mandate/climate/wmo-statement-state-of-global-climate [Accessed 26 July 2023].

Wyatt, M. et al., 2022. Using ensemble methods to improve the robustness of deep learning for image classification in marine environments. *Methods in Ecology and Evolution,* 13(6), pp. 1317–1328.