



# Human-Agent Collaborative Decision-Making Framework for Naval Systems

Maria Olinda Rodas<sup>(✉)</sup>, Jeff Waters, and Cheryl Putnam

Space and Naval Warfare Systems Center Pacific, San Diego, CA 92103, USA  
maria.rodas@navy.mil

**Abstract.** This work provides an overview of a future human-agent collaborative decision-making framework to be developed for naval systems using an augmented reality platform. We present the basic concept behind the framework, key features of the application, and some details about a future proof of concept prototype that will demonstrate and evaluate the concept against a baseline design.

**Keywords:** Military · Virtual reality · Mixed reality · Augmented reality  
Decision optimization · Provenance · Data visualization · Design  
C2 · Innovation

## 1 Introduction

Increasingly, humans collaborate with intelligent and autonomous agents to meet military objectives [1]. The support of these agents is often used to acquire and analyze information, make decisions and plan future actions [2]. However, lack of trust is often mentioned as the main critical barrier to transition successful intelligent and automated agents into the military [3]. The concept of trust is often determined by three main factors: (1) the reliability of the agent; (2) the self-confidence of the human operator; and (3) the workload required for the task [4]. Often, intelligent agents are perceived as untruthful, not necessarily because they are unreliable, but because they are unable to adequately communicate their rationale to human operators [5].

Increasing trust in the intelligent agent via increasing reliability is crucial to avoid the “crying wolf effect”; however, helping the human operator to adequately calibrate trust in the agent is vital given that most intelligent agents will always be flawed (i.e., not designed for all requirements and therefore unreliable in some occasions) [6, 7]. This paper describes an effort to develop a human-agent decision-making collaboration framework to enable effective communication and trust calibration between intelligent agents and human’s operators.

## 2 Background

A key objective of the U.S. military is to achieve Information Superiority for enabling sound and timely decision-making [8]. As part of the planning process, military personnel must make informed decisions, solve complex problems, and ultimately

accomplish assigned missions. Decisions are critical as they drive actions which have impact that drive further decisions [2]. The consequences of a decision ripple through command structures driving more decisions of more organizations and entities which can be catastrophic if based on the wrong decision.

In the past these informed decisions were made without any support, while currently military personnel are making decisions with the support of intelligent and automated agents. While these technologies are meant to support humans in making faster and more efficient decisions; the current reality is that they are unable to effectively provide support because they fail to establish trustworthy relationships with the humans. Furthermore, they fail at providing humans with the supporting information in a way that will enhance their decision-making process. This involves providing information in a way that allows human operators to adequately weigh that information while making further decisions. In some cases, such as in the Automatic Target Recognition (ATR) technologies, human operators are forced to make blind decisions only considering yes or no answers on the part of the autonomous agent. In other such intelligent agents that make autonomous planning of optimal Unmanned Vehicle routes, the agent provides a course of action but again without any explanation on the algorithms rationale.

Most intelligent and automated agents still lack the adequate transparency critical to allow understanding of their rationale. There is the assumption that to make automation algorithms such as those of the ATR technologies transparent to human operators, the technologies must be able to communicate their purpose, process and performance [9]. Based on that assumption, researchers at the Army Research Laboratory developed the Situation Awareness-based Agent Transparency model (SAT) to explain aspects of SA that affect trust. The SAT model includes variables such as the agent's purpose (goal), process (intentions, progress), performance (constraints), reasoning process, projection of future state and limitations [10] to increase agent transparency and therefore overall trust in the agent. The same researchers showed the implementation of this SAT model during a human-agent teaming mission for multi-robot management demonstrating how transparency increased trust in agents without increasing the human operator's workload [11].

Understanding the decision-making process of intelligent and autonomous agents is not enough to ensure a successful collaboration between the two. Human operators must also be able to interact with the agents to allow further calibration of trust. Decision-making explanations by agents have shown to successfully calibrate trust in human-agent interactions allowing the agent to provide the right type of information to the humans [12, 13].

### 3 Problem Formulation

Several types of autonomous agents, such as those used in ATR and automated mission planning, are currently utilized in the Navy to make decisions at lower levels of command such as the Unmanned Vehicle (UV) operator level. UV Operators make decisions which form the basis for critical planning decisions made at higher levels of command, posing a critical risk to mission success if early decisions are faulty.

Adequate trust calibration in intelligent and automated systems is crucial to effectively integrate the agent's support into the final decision while understanding their advantages and limitations.

The human decision-making process is mostly ignored by current agent designers since new technologies cannot offer the required human-decision making support in a user-friendly manner to enhance the information evaluation and the weighting process affecting the final decision. Adequate trust calibration in the agents is also missing, resulting in human operators ignoring the agents' recommendations and support. Finally, an effective collaboration enabler technology is also required to allow a trusted and efficient interaction between the agent and the human.

We propose that an effective human-agent decision-making collaboration framework can enhance human-agent communication by allowing the required transparency on the agent's rationale to enable a successful trust calibration. Furthermore, we propose that an effective framework can also enhance the human decision-making process making it easier for the operators to incorporate and use the agent's decisions. An effective interaction technology such as Augmented Reality (AR) will enable the interaction using interactive data displays.

## 4 Proposed Framework Details

This project will capitalize on existent resources across industry, academia and government to develop an effective framework for human-agent decision-making collaboration. Low-cost Augmented Reality (AR) technology such as the Microsoft HoloLens will be utilized as the main technology enabler for the human-agent interaction. Data visualization techniques and visual data analytic models will be employed to enhanced communication through effective user interface design techniques [14–16]. A well designed and elegant framework will improve the efficiency of creating software. It will improve developer productivity and importantly, improve the quality and reliability of the software.

The Agile Decision Computation Model (ADCM) and Agile SA Machine Model (ASAMM) previously developed at SSC Pacific will be adapted to this project. The ADCM specifies a standard decision format for agents to track the provenance, i.e. the who, what, when, where and why of the agent's decisions. The ASAMM will specify a standard SA format for agents to reflect and build rational explanations of why a decision was suggested. An Agent Communication Model (See Fig. 1) will be developed to integrate input from both the ADCM and ASAMM models and specifies how information from these models will be communicated back to the human operators. A Human Communication Model will also be developed to determine a standard format to allow humans to communicate with the agents at the different interaction modes. Finally, a Human-Agent Interaction Model will be developed to mediate the interaction between the human and the agent communication models. This model communicates to the AR capability to enable the interaction and visualizations. See Fig. 1.

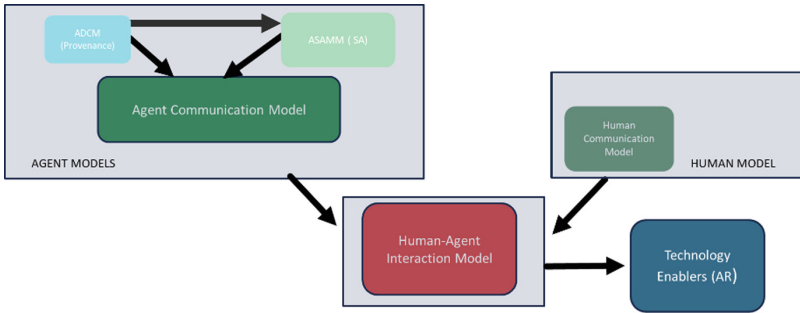
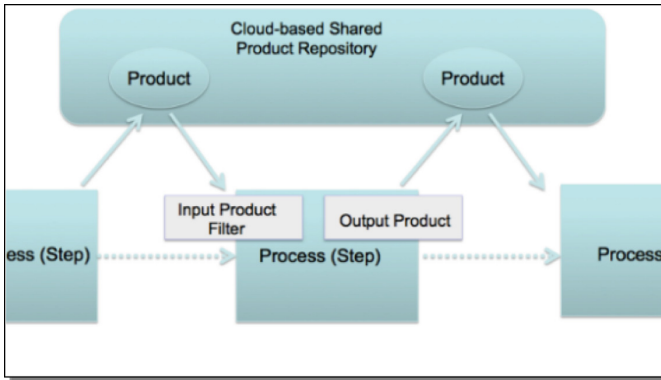


Fig. 1. Human-agent decision-making collaboration framework

### 4.1 Capturing Agents Decision Structure

The ASAMM will allow a situational understanding of an agent’s current decision in terms of past context and the future impact of their decisions on mission goals through the machine-understandable representation of the intersection of processes. Situational Awareness has been defined as “...the *perception* of the elements in the environment within a volume of time and space, the *comprehension* of their meaning, and the *projection* of their status in the near future” [17]. Based on this definition, what is needed for machine situational awareness of an agent’s decision is the perception of the decision (e.g. a machine representation) and its components, along with the decision’s provenance providing the comprehension of the meaning of the decision, and the impact of the decision on relevant ongoing processes. The ASAMM model is designed to capture these components. For instance, the model captures processes that can be thought of as a series of “steps” for accomplishing a goal (decision goals) [18]. A “step” in this setting is a “task”, or an “action”, something the machine needs to do. See Fig. 2. For a task to be completed, something needs to have been accomplished, and that accomplishment can be considered an output “product” of the task. Each task may have one or more input or output products. We all execute processes in our daily lives and work; however, the processes are not always obvious or well defined. A well-defined machine task will have clearly defined tangible input and output products. A “process” can be represented as a sequence of steps where each “step” represents an action that has an input and an output “product”. The products represent interim or final results and are instances of generic product types relevant for command and control such as agent’s Observation, Course of Action, Request, Approval, Decision, and Metric. Utilizing this type of machine SA model will allow agents to communicate with humans on why and how a decision was made. To do this, a process needs to specify by the human operator in a machine understandable way. Once that done that information will be used to check on the agent status and provide feedback to the user that will allow adequate calibration of trust.



**Fig. 2.** The process step building block

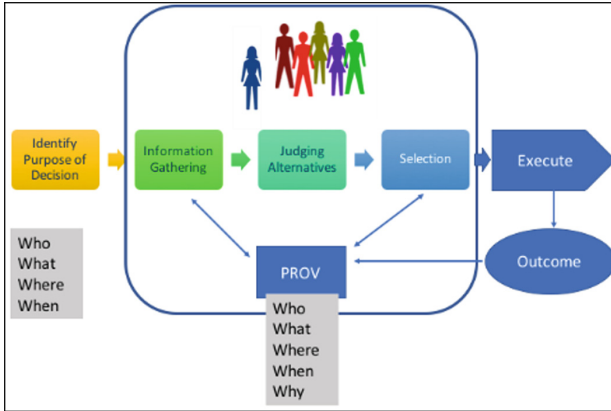
## 4.2 Tracking Agent's Decision Provenance

Often, when information is passing from intelligent agents to human operators there is no information about the agent's confidence in the decision and the overall decision criteria used by the agent. This type of information is critical for military decision makers since decision criteria requirements can rapidly change in real time missions. Understanding information reliability, especially decision confidence and criteria, is key to effective decision performance [19].

To understand the agents' decision-making a standardized, machine-understandable representation of their decisions is required. The Agile Decision Computation Model (ADCM) being developed at SSC Pacific will allow us to capture agent's decision provenance [20]. It consists of 5 simple stages: (1) identify the need to make a decision; (2) gather information; (3) judge alternatives; (4) select an alternative; and (5) add the outcome of the decision following execution of the decision. Using the PROV representation of decisions, the ADCF can consider managing the agent's decision. The framework will provide an Application Programming Interface (API) and reusable software library in different programming languages enabling the management (e.g. creation, retrieval, update, deletion and query) of decisions. The framework will also provide a validator, a graphical visualization, a cloud repository and other common services to ease the burden on developers and users. The ADCF will allow us to capture the agent's decision rationale in detail (Fig. 3).

## 4.3 Enabling Human-Agent Interaction Technology

Choosing a technology that will effectively support the human-agent collaborative decision-making is key to the success of any framework. Through the process of considering interactive technologies some requirements were identified as critical to success in the type of military settings being considered. Those requirements include: (1) a user-friendly platform that allows different type of interaction and effective data visualization; (2) single or multiple decision-makers; (3) little or no modification of existing



**Fig. 3.** Agile decision computation framework

military systems (separate platform); (4) lightweight capability for expeditionary types of missions; and (5) easy integration of information coming from different systems.

AR enables natural communication by displaying the visual cues necessary for a human and intelligent agent to collaborate in a decision-making situation. Augmented Reality (AR) technologies combine real and virtual objects while the real objects are still perceived in the real world. Furthermore, AR can enable task collaboration with multiple users or decision-makers at different locations. AR allows displays of information without modification of existing systems, and introduction of new displays and capabilities in otherwise constrained and limited environments. AR platforms such as Microsoft HoloLens provide a lightweight and easy to integrate capability. Finally, AR can easily integrate data found in different systems into one single display that will provide the required information for collaborative decision-making without modification of existing systems.

## 5 Proof of Concept with AR Capability Implementation

Once the models are adapted to this project, a proof of concept prototype implementation will be developed for a mine-hunting naval mission scenario. In this scenario, an ATR technology is used to support the human decision-making in finding underwater mines. With AR, a sonar operator could use AR glasses, such as the HoloLens, and see overlaid on top of the ATR target on the sonar display, as needed, a 3D rendering (derived from the ATR system) hovering in mid-air of the specific type of target detected, intelligence about that target or type of target, and then to further drill-down, the operator could see a pop-up graph of the metrics used by the ATR algorithm and the scores and confidence. Moreover, AR technology could support human decision-making by providing a step-by-step explanation of how the target was chosen by the agent and allowing a more informative decision while also learning about the agent’s limitations and advantages. This type of interaction is expected to allow an effective trust calibration.

To the operator, the interaction will appear as if the operator is interacting with one system, a very helpful system which provides not simply an alert about a target, but a realistic visualization, intelligence, drill-down information and agent's confidence levels. The operator can "communicate" by indicating interest by gaze, or by touch, or by voice, or by input device, and with the system responding with the desired information. So, for this purpose, AR assists the operator by integration of disparate capabilities. AR assists system developers by not requiring separate, individual interfaces and not requiring onerous early integration limitations. For this project, AR allows significant flexibility in providing a lightweight but effective integration capability so that we can integrate capability directly on existing systems without impacting those systems directly. AR allows us to do our experiments and measures directly on top of the operator's existing systems.

## 6 Conclusions

This paper briefly discusses the details of a human-agent collaborative decision-making framework to be developed at SSC Pacific. Tracking an agent's decision-making information through a provenance model will help us uncover critical information that is required for effective collaborative decision-making and allow the user to better understand and trust the information. The framework will avoid a ripple effect on decisions caused by using unreliable information. Additionally, an agent SA model will allow the agent to track its decisions and offer feedback in terms of their past and future performance. The goal is to provide greater transparency and decision-support to the human operator's decision-making process. Performance data along with details on the decision-making used by agents can help human operators to effectively understand the reliability of informational sources provided by the agents. Moreover, AR data integration and visualizations will support a user-friendly interaction while supporting drill-down examination of the agent rationale. A proof of concept prototype will be developed for the mine countermeasures Navy scenario to demonstrate and evaluate the human-agent decision-making collaboration framework.

## References

1. Barnes, M.J., Evans, A.W.: Soldier-robot teams in future battlefields: an overview. In: Barnes, M., Jentsch, F. (eds.) *Human-Robot Interactions in Future Military Operations*, pp. 9–29. Ashgate, Farnham (2010)
2. Parasuraman, R., Sheridan, T.B., Wickens, C.D.: A model for types and level of human interaction and automation. *IEEE Trans. Syst. Man Cybern. Part A Syst. Hum.* **30**(3), 286–297 (2000)
3. Office of the Under Secretary of Defense for Acquisition, Technology and Logistics. Washington, DC. 20301-3140. Department of Defense: Defense Science Board. Task Force Report: The Role of Autonomy in DoD Systems, July 2012. <https://fas.org/irp/agency/dod/dsb/autonomy.pdf>. Accessed 12 Jan 2018
4. Parasuraman, R., Riley, V.: Humans and automation: use, misuse, disuse, abuse. *Hum. Factors J.* **39**(2), 230–253 (1997)

5. Hancock, P.A., Billing, D.R., Shaefer, K.E., Chen, Y.C., DeVisser, E.J., Parasuraman, R.: A meta-analysis of factors affecting trust in human-robot interaction. *Hum. Factors J.* **53**(5), 517–527 (2016)
6. Hoff, K.A., Bashir, M.: Trust in automation: integrating empirical evidence on factors that influence trust. *Hum. Factors J.* **57**(3), 407–434 (2015)
7. Navy War College. Navy Planning, December 2013. [https://www.usnwc.edu/getattachment/17afb3-a1e2-46b3-ble9-d1fa4b0fec5a/5-01\\_\(Dec\\_2013\)-\(Promulgated\).aspx](https://www.usnwc.edu/getattachment/17afb3-a1e2-46b3-ble9-d1fa4b0fec5a/5-01_(Dec_2013)-(Promulgated).aspx). Accessed 20 Jan 2018
8. Wagenhals, L.W., Levis, A.H.: Course of action development and evaluation. In: DTIC, Fairfax (2000)
9. Lee, J.D.: Trust, trustworthiness, and trustability. In: The Workshop on Human Machine Trust for Robust Autonomous Systems, Ocala, Florida (2012)
10. Chen, J.Y.C., Barnes, M.J.: Human-agent teaming for multirobot control: a review of human factors issues. *IEEE Trans. Hum.-Mach. Syst.* **44**(1), 13–29 (2014)
11. Mercado, J.E., Rupp, M.A., Barnes, M.J., Barber, D., Procci, K.: Intelligent agent transparency in human-agent teaming for multi-UxV management. *Hum. Factors J.* **58**(3), 401–415 (2016)
12. Wang, N., Pynadath, D.V., Hill, S.G., Merchant, C.: The dynamics of human-agent trust with POMDP-generated explanations. *Intelligent Virtual Agents. LNCS (LNAI)*, vol. 10498, pp. 459–462. Springer, Cham (2017). [https://doi.org/10.1007/978-3-319-67401-8\\_58](https://doi.org/10.1007/978-3-319-67401-8_58)
13. Wang, N., Pynadath, D.V., Hill, S.G.: The impact of POMDP-generated explanations on trust and performance in human-robot teams. In: Proceeding of the 15th International Conference on Autonomous Agents and Multiagent Systems, Singapore, pp. 997–1005 (2016)
14. Lim, K., Suresh, P., Schulze, J.P.: Oculus rift with stereo camera for augmented reality medical intubation training. In: Proceedings of IS&T the Engineering Reality of Virtual Reality, San Francisco, CA, 1–2 February 2017
15. Wiley, B., Schulze, J.P.: archAR: an archaeological augmented reality experience. In: Proceedings of IS&T/SPIE Electronic Imaging, the Engineering Reality of Virtual Reality, San Francisco, CA, 9–10 February 2015
16. McCarthy, D., Schulze, J.P.: Distributed VR rendering using NVIDIA OptiX. In: Proceedings of IS&T the Engineering Reality of Virtual Reality, San Francisco, CA, 1–2 February 2017
17. Endsley, M.R.: Toward a theory of situation awareness in dynamic systems. *Hum. Factors: J. Hum. Factors Ergon. Soc.* **37**(1), 32–64 (1995)
18. Jeff, W., Bruce, P., Pilcher, J., Odland, A., Jones, D.: A dynamic agile process model for situational awareness: a machine-understandable, fractal-based, data-driven approach. Presented at the Cognitive Methods in Situational Awareness and Decision Support (CogSIMA) Conference, San Diego, California (2016)
19. Langloz, C.: Fundamental measures of diagnostic examination performance: usefulness for clinical decision making and research. *Radiol. J.* **228**(1), 3–9 (2003)
20. Putman, C., Waters, J., Rodas, O.: A standard decision format using provenance. In: Proceedings of IEEE International Symposium on Signal Processing and Information Technology, Bilbao, Spain (2017)