NPS-USW-2020-001

NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

ETHICAL CONTROL OF UNMANNED SYSTEMS:

LIFESAVING/LETHAL SCENARIOS FOR NAVAL OPERATIONS

by

Donald P. Brutzman, Curtis L. Blais, and Hsin-Fu Wu

August 2020

Distribution Statement A

Prepared for: Raytheon Missiles & Defense under NCRADA-NPS-19-0227

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE ***(DD-MM-YYYY)***	**2. REPORT TYPE**	**3. DATES COVERED** ***(From-To)***
04-12-2020	Technical Report (TR)	29-03-2019 to 31-03-2020

4. TITLE AND SUBTITLE
ETHICAL CONTROL OF UNMANNED SYSTEMS:
LIFESAVING/LETHAL SCENARIOS FOR NAVAL OPERATIONS

5a. CONTRACT NUMBER
NCRADA-NPS-19-0227

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHORS
Donald P. Brutzman, Curtis L. Blais, and Hsin-Fu Wu

5d. PROJECT NUMBER
1

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME AND ADDRESS
Undersea Warfare Academic Group (USWAG)
Naval Postgraduate School (NPS), 1 University Way, Monterey CA 93933-5000

8. PERFORMING ORGANIZATION REPORT NUMBER
NPS-USW-2020-001

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Raytheon Missiles & Defense
1151 E. Hermans Road
Tucson, AZ 85756

10. SPONSOR/MONITOR'S ACRONYMS
RMD, RTX

11. SPONSOR/MONITOR'S REPORT NUMBER
RMS-18112

12. DISTRIBUTION / AVAILABILITY STATEMENT
Distribution Statement A. Approved for public release: distribution unlimited.

13. SUPPLEMENTARY NOTES
CRADA Project Title: Undersea Command, Control, and Communication (C3) for Ethical Human Supervision of Unmanned Undersea Systems

14. ABSTRACT
This research in Ethical Control of Unmanned Systems applies precepts of Network Optional Warfare (NOW) to develop a three-step Mission Execution Ontology (MEO) methodology for validating, simulating, and implementing mission orders for unmanned systems. First, mission orders are represented in ontologies that are understandable by humans and readable by machines. Next, the MEO is validated and tested for logical coherence using Semantic Web standards. The validated MEO is refined for implementation in simulation and visualization. This process is iterated until the MEO is ready for implementation. This methodology is applied to four Naval scenarios in order of increasing challenges that the operational environment and the adversary impose on the Human-Machine Team. The extent of challenge to Ethical Control in the scenarios is used to refine the MEO for the unmanned system.
The research also considers Data-Centric Security and blockchain distributed ledger as enabling technologies for Ethical Control. Data-Centric Security is a combination of structured messaging, efficient compression, digital signature, and document encryption, in correct order, for round-trip messaging. Blockchain distributed ledger has potential to further add integrity measures for aggregated message sets, confirming receipt/response/sequencing without undetected message loss. When implemented, these technologies together form the end-to-end data security that ensures mutual trust and command authority in real-world operational environments—despite the potential presence of interfering network conditions, intermittent gaps, or potential opponent intercept.
A coherent Ethical Control approach to command and control of unmanned systems is thus feasible. Therefore, this research concludes that maintaining human control of unmanned systems at long ranges of time-duration and distance, in denied, degraded, and deceptive environments, is possible through well-defined mission orders and data security technologies. Finally, as the human role remains essential in Ethical Control of unmanned systems, this research recommends the development of an unmanned system qualification process for Naval operations, as well as additional research prioritized based on urgency and impact.

15. SUBJECT TERMS
Ethical Control, Unmanned Systems, Human-Machine Teaming, Mission Execution Ontology (MEO), Semantic Web, Data-Centric Security, C2, C4I, Network Optional Warfare (NOW), Intelligent Autonomy, Command Authority

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Donald P. Brutzman
UNCLAS	UNCLAS	UNCLAS	NONE	104	**19b. TELEPHONE NUMBER** ***(include area code)*** (831) 656-2149

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

THIS PAGE INTENTIONALLY LEFT BLANK

NAVAL POSTGRADUATE SCHOOL
Monterey, California 93943-5000

Ann E. Rondeau
President

Robert F. Dell
Acting Provost and
Academic Dean

The report entitled "ETHICAL CONTROL OF UNMANNED SYSTEMS: LIFESAVING/ LETHAL SCENARIOS FOR NAVAL OPERATIONS" was prepared for Naval Postgraduate School and Raytheon Missiles & Defense under NCRADA-NPS-19-0227, and funded by Raytheon Missiles & Defense (RMD).

Further distribution of all or part of this report is authorized.

This report was prepared by:

Donald P. Brutzman
Associate Professor, NPS

Curtis L. Blais
Research Associate, NPS

Hsin-Fu Wu
Senior Principal Systems Engineer, RMD

Reviewed by:

John T. Beaver
Manager, Naval Requirements
& Capabilities, RMD

Julie L. Leeman
Director, Advanced Counter Hypersonic
Weapons, RMD

Thomas J. Housel
Chair, Information Sciences Department, NPS

Imre Balogh
Director, MOVES Institute, NPS

Released by:

Winford G. Ellis
Director, Undersea Warfare
Research Center

Jeffrey D. Paduan
Dean of Research, NPS

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

This research in Ethical Control of Unmanned Systems applies precepts of Network Optional Warfare (NOW) to develop a three-step Mission Execution Ontology (MEO) methodology for validating, simulating, and implementing mission orders for unmanned systems. First, mission orders are represented in ontologies that are understandable by humans and readable by machines. Next, the MEO is validated and tested for logical coherence using Semantic Web standards. The validated MEO is refined for implementation in simulation and visualization. This process is iterated until the MEO is ready for implementation. This methodology is applied to four Naval scenarios in order of increasing challenges that the operational environment and the adversary impose on the Human-Machine Team. The extent of challenge to Ethical Control in the scenarios is used to refine the MEO for the unmanned system.

The research also considers Data-Centric Security and blockchain distributed ledger as enabling technologies for Ethical Control. Data-Centric Security is a combination of structured messaging, efficient compression, digital signature, and document encryption, in correct order, for round-trip messaging. Blockchain distributed ledger has potential to further add integrity measures for aggregated message sets, confirming receipt/response/sequencing without undetected message loss. When implemented, these technologies together form the end-to-end data security that ensures mutual trust and command authority in real-world operational environments—despite the potential presence of interfering network conditions, intermittent gaps, or potential opponent intercept.

A coherent Ethical Control approach to command and control of unmanned systems is thus feasible. Therefore, this research concludes that maintaining human control of unmanned systems at long ranges of time-duration and distance, in denied, degraded, and deceptive environments, is possible through well-defined mission orders and data security technologies. Finally, as the human role remains essential in Ethical Control of unmanned systems, this research recommends the development of an unmanned system qualification process for Naval operations, as well as additional research prioritized based on urgency and impact.

THIS PAGE INTENTIONALLY LEFT BLANK

EXECUTIVE SUMMARY

As humans proliferate unmanned systems and advance unmanned systems technology ahead of emerging policy and guidance, development of unmanned systems autonomy, as well as command and control in Human-Machine Teaming, needs to be grounded in ethical and responsible-use principles. Human supervision using Ethical Control is required for any unmanned systems holding potential for lethal force. In denied, degraded, and deceptive operational environments, the human commanders must maintain Ethical Control of unmanned systems, and the unmanned systems must perform their missions that require the use of force—from lifesaving to lethal force, within the assigned operational and ethical constraints.

This research in Ethical Control of Unmanned Systems develops a three-step Mission Execution Ontology (MEO) methodology for implementing, validating, and simulating mission orders for unmanned systems. The methodology makes it possible to determine the extent to which unmanned systems can handle progressive challenges to command and control (C2) in distance and time, in human-machine teams, and in employing lifesaving or lethal force as authorized by the commander. The three steps in the MEO methodology are:

1. Represent orders in ontologies that are human-understandable and machine-readable.
2. Validate and test the MEO is validated and tested for logical coherence.
3. Assess the MEO for executability in simulation and visualization.

As context for the MEO methodology, the research considers multiple Naval scenarios as exemplar missions in order of increasing challenges that the operational environment and the adversary impose on the Human-Machine Team. These scenarios carefully define and test capabilities for Ethical Control of unmanned systems in progressively increasing challenges in distance and time, human-machine teams, and the use of force. Mission titles and brief descriptions of the unmanned systems assigned tasks are:

A. Sailor Overboard: Perform recovery operations in concert with shipboard procedures.
B. Lifeboat Tracking: Provide remote presence for locating, tracking, communications and beaconing an adrift lifeboat carrying multiple personnel.
C. Pirate Boats Attack: Overtake pirates attempting to capture a merchant ship. Provide warning and counterattack using lethal force, if escalation of hostilities is warranted.
D. Hospital Ship Electromagnetic (EM) Decoy: Respond appropriately to EM signatures of a warship unexpectedly emitting from a hospital ship due to adversary exploitation.

Of note, exemplar mission D (Hospital Ship EM Decoy) presents the most ethically challenging scenario for the unmanned system and the Human-Machine Team. Two variations of this scenario demonstrate the significance of Ethical Control of unmanned systems:

- *Sense-Decide-Act:* Immediate reaction to deploy a robot swarm results in a mistaken attack on friendly or neutral shipping, and is consequently a war crime.
- *Observe-Orient-Decide-Act (OODA):* Deliberate tactics and Ethical Control constraints prevent automatic erroneous counterattack against false flag placed on friendly/neutral shipping, improving ship defense and warfighting capabilities.

By applying MEO methodology in these scenarios, an assessment scale emerges for human warfighters qualifying unmanned systems performing missions under Ethical Control:

Level 1: Basic. Qualified to apply lifesaving force under close coordination with the Human Commander, e.g., Sailor Overboard Recovery mission.

Level 2: Intermediate. Qualified to apply lifesaving force at long distance from the Human Commander, e.g., Search and Rescue mission.

Level 3: Advanced. Qualified to apply organic lifesaving/lethal force over long time periods for the mission, emphasizing restraint throughout, e.g., Counter-Piracy mission.

Level 4: Operational Standard. Qualified to apply organic lifesaving/lethal force in contested/deceptive environments, e.g., Force Protection mission. Uses human confirmation of Identification Friend, Foe, Neutral, or Unknown (IFFNU) classification result to detect spoofing anti-pattern, with authorization required to apply lethal force, prevent reflexive automatic counterattack response, and engage with proportional force.

The research also applies Data-Centric Security to enable end-to-end data security that ensures mutual trust and command authority in real-world operational environments— despite the potential presence of interfering network conditions, intermittent gaps, or potential opponent intercept. Data-Centric Security is a combination of structured messaging, efficient compression, digital signature, and document encryption, in correct order, for round-trip messaging. Addition of blockchain distributed ledger has potential to further add integrity measures for aggregated message sets, confirming receipt/response/sequencing without undetected message loss.

Maintaining Ethical Control of unmanned systems from long time-duration and physical distance—in denied, degraded, and deceptive environments—is now possible through well-defined mission orders and data security technologies. Furthermore, the MEO methodology conveys the mission orders in formats that are readable and sharable by both humans and unmanned systems—with validatable syntax and semantics through understandable logical constraints. The MEO is also testable and confirmable using simulation and visualization. Additional Semantic Web confirmation can ensure that orders are comprehensive and consistent.

Therefore, a coherent Ethical Control approach to human command and control of unmanned systems such as the MEO methodology supports the development of an unmanned systems qualification process for Naval operations.

The research recommends topics for additional research that are prioritized based on urgency and impact. Recommendations are summarized below:

1. Incorporate Ethical Control in Unmanned Maritime Autonomy Architecture (UMAA) and Common Control System (CCS). Action: Brief relevant program managers and stakeholders; seek to participate in UMAA.
2. Determine implications in Integrated Naval Force Structure. Actions:
    a. Engage staffs in Navy Warfare Directorates, Type Commanders, unmanned vehicle squadrons, and relevant stakeholders. Topics for engagement include: operations, plans, policy, and requirements, and resourcing of unmanned systems with due regard to the feasibility of applying Ethical Control in enabling them to conduct qualified missions.
    b. Establish an NPS Center for Ethical Warfighting to explore both educational and applied capabilities, in order to put theory into practice.
    c. Establish collaboration within the Naval Education Enterprise (e.g., NPS, Naval War College, Marine Corps University, and Naval Academy) and other institutions, e.g., U.S. Military Academy, on ethical use of unmanned systems.
3. Establish a process to qualify unmanned systems, i.e., design and develop qualification requirements (similar to Naval Warfighter "qualification cards") for various classes of unmanned systems that ensure Ethical and Secure C2 for Naval missions in denied/degraded/deceptive environments. Key Actions:
    a. Apply the MEO methodology to test and certify compliance to mission orders.
    b. Integrate Data-Centric Security in a qualification process to ensure Trusted Autonomy and Command Authority for the Human-Machine Team.
    c. Use a comprehensive virtual environment with carefully crafted scenarios to test key requirements and capabilities, hardware/software in the loop, as well as visualization of rehearsal, real-time and replay of realistic missions.
    d. Assess mission logs and scenario outcomes for after-action analysis, lessons learned, and continuous improvement.
4. Continue Canonical Mission development and ontology refinements on tactical scenarios that exercise the checkpoints and authorities of ethical control.
5. Implement planned improvements in Autonomous Vehicle Command Language (AVCL) and Autonomous Unmanned Vehicle Workbench (AUV Workbench).
6. Incorporate Data-Centric Security for system integrity and security of unmanned systems command, control, and communications (C3).

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGEMENTS

The authors gratefully appreciate the remarkable collaboration in this inaugural project under NPS-Raytheon CRADA for Naval Warfare Capability Research and Development (NCRADA-NPS-19-0227). Cooperative work undertaken in this project holds great promise for academia, industry, and Department of Defense. Continued collaboration is recommended.

In particular, we are grateful for Ms. Julie Leeman of Raytheon Missiles & Defense for sponsoring the establishment of the NPS-Raytheon CRADA and the performance of this research project. Ms. Leeman's leadership, commitment, and vision have set the course for sustaining mutually beneficial collaboration between NPS and Raytheon for years to come. Additionally, over thirty Raytheon colleagues contributed their insight and feedback throughout the performance of this project, especially during the Monthly Technical Interchange Meetings and site visits. We sincerely appreciate them for helping to inform and refine the research, as well as affirming its relevance to advancing Human-Machine Teaming in Naval Warfare. We specifically acknowledge the importance of Richard Markeloff's in-depth review of ontology and tool chains, along with Robert Grabowski's suggestion to consider semantic distance from accountability with respect to autonomy in unmanned systems.

Critical thanks are due to Winford G. (Jerry) Ellis, RADM USN (Ret.), Director of NPS Undersea Warfare Research Center, for his CRADA leadership and empowerment in this project. Dr. Robert B. McGhee and Dr. Duane T. Davis continue to make fundamentally important contributions to this work. We are grateful for their ongoing efforts, as well as those of many graduate and doctoral students who have helped advance this work over decades of effort. Terry Norbraten's programming work was essential for important upgrades to the Autonomous Underwater Vehicle (AUV) Workbench. Col Todd Lyons USMC (Ret.), Deborah Buettner and Agata Maslowska have been instrumental in helping define and meet CRADA requirements. We thank the NPS Undersea Warfare Academic Group, as well as the Consortium for Robotics and Unmanned Systems for Education and Research (CRUSER), for recurring support and broad-ranging activity. Editorial review by NPS Graduate Writing Center (GWC) improved this work.

This work outlines paths explored during an ongoing journey, not a single destination. Many efforts by many people are needed to achieve widespread ethical control of autonomous systems. We thank contributors, implementers, and evaluators in advance for all partner efforts.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

A. PROJECT DESCRIPTION

Researchers and Faculty at the Naval Postgraduate School and the (then) Raytheon Company performed this research project within the Raytheon–Naval Postgraduate School Cooperative Research and Development Agreement (NCRADA-NPS-19-0227) for Naval Warfare Capability Research and Development. The intention of this work is to research methods of Undersea Command, Control, and Communications (C3) for ethical human supervision of unmanned undersea systems.

Ethical control of unmanned systems can be accomplished through structured mission definitions that are consistently readable, validatable and understandable by humans and robots. Responsible humans must remain in charge of lethal/lifesaving force, and then robots become more effective.

B. MOTIVATION

This research is motivated by ethically constrained control of unmanned systems and robot missions by human supervisors and warfighters. With the Artificial Intelligence/Machine Learning (AI/ML) technologies and proliferation of unmanned systems potentially advancing ahead of emerging policy and guidance, clarity and distinction are needed in defining and guiding technology development for Human-Machine Teaming that is grounded in principles of ethical and responsible use. This research is thus intended to stimulate and shape the development of ethical autonomous unmanned undersea weapons systems. Ethical Control is not simply an AI problem, rather it is a moral imperative. Warfighters cannot push “the big red shiny AI button” and hope for the best, since indiscriminate actions are immoral and unlawful. As expressed by the guiding senior faculty member in this decades-long endeavor,

> Ethical constraints on robot mission execution are possible today. There is no need to wait for future developments in Artificial Intelligence (AI). It is a moral imperative that ethical constraints in some form be introduced immediately into the software of all robots that are capable of inflicting unintended or deliberate harm to humans or property. (McGhee, April 2016)

C. RESEARCH QUESTIONS AND ANSWERS

The key research question is: Can qualified robots correctly follow human orders? Through this research, the following precept has enabled the Collaborators to answer it in the affirmative: Well-structured mission orders can be syntactically and semantically validated to give human commanders confidence that offboard systems *will do what they are told to do, and further will not do what they are forbidden to do*.

II. OVERVIEW

A. SUMMARY OF WORK

Naval Postgraduate School (NPS) and Raytheon co-Principal Investigators researched the impacts and limitations of current ethics policies on the design, development, and deployment of unmanned undersea weapons. The Collaborators researched current ethics principles and policy considerations that impact ethical human supervision of autonomous underwater unmanned systems in tactical scenarios; this effort was continuous throughout the research project in order to ensure its relevance to current and emerging policies, standards, and practices. They investigated how warfighters can effectively and ethically supervise underwater unmanned systems at great ranges as trusted participants with distant human supervision. (This report uses the terms unmanned systems, robots, and autonomous vehicles interchangeably.)

The Principal Investigators determined the applicability of the ethics principles and policies (e.g., DoD Directive 3000.09, Autonomy in Weapon Systems) to a variety of current and in-development U.S. Navy systems, such as Unmanned Undersea Vehicles (UUVs). After consideration of appropriate real-world tactical scenarios in maritime operations that might be relevant for this research, they selected four unrestricted scenarios as surrogates to their respective classified scenarios. Then by applying Semantic Web ontology to scenario goals and constraints, they performed logical validation that human-approved mission orders for robots are semantically coherent, precise, unambiguous, and without internal contradictions. Mission Execution Ontology (MEO) methodology is the term the researchers conceived for the process, from defining mission orders to validating them.

The MEO methodology begins with Mission Definition. The researchers (or operational planners and mission analysts in military operations in real-world military operation planning organizations) analyze the mission orders by phases of execution and the associated decisions for each phase. Mission analysts consider and diagram the flow of mission phases by relationships between the phases and the logic state of the decision made at the completion of each phase. Analysts use a “tri-state” goal-transition logic to direct the flow of decision between goals. The tri-state logic states are: Success, Failure, and Exception, with respect to accomplishing the goal for each phase of the mission. For example, if the Deploy phase of the mission is accomplished as Success, then the mission would proceed to the next phase. If the state of the Deploy phase results in Failure, then the mission would proceed to a phase to perform troubleshooting or failure recovery actions. If the Deploy phase actions result in an Exception logic state, then the mission would proceed to a phase to perform holding or awaiting further instructions tasks.

Once analysts decide on a set of mission orders as defined by a corresponding decision-flow diagram, the Autonomous Vehicle Command Language (AVCL) can be used for formal definition and subsequent generation of code for each unmanned system executing a mission. AVCL mission descriptions are expressed using structured Extensible Markup Language (XML) that represents human command and control tasks for autonomous unmanned vehicles (AUVs). AVCL is used to generate mission scripts, agenda plans and post-mission recorded telemetry. Mission analysts or operators can utilize a single achievable and validatable format for robot tasking and results that is directly convertible to (and from) a wide variety of different robot command languages. As such, analysts now apply the Web Ontology Language (OWL) Semantic Web Standards to AVCL representation of the mission orders and thus converts the mission orders into ontologies (i.e., the MEO vocabulary) in order to validate the mission orders for ethical control—that the orders are semantically coherent, precise, unambiguous, and without internal contradictions. Finally, modeling and simulation are used to confirm ethical control of mission design and execution.

In addition to applying the MEO methodology, the Collaborators also have considered environmental, geopolitical, security and human life implications for associated anti-tamper/cyber solutions, and have provided recommended courses of action for further development. These relevant elements of the unmanned systems would contribute to the effective application of ethical control, forming the foundation of Trusted Command Authority, Trusted Mission Orders, and Trusted Mission Execution.

B. REPORT ORGANIZATION

Chapter 3 describes the background of research on Techniques for Maintaining Human Ethical Control of Unmanned Systems that NPS researchers have been conducting since 1994, leading up to this current research project. Chapter 4 discusses the methodology the current research uses to represent Mission Execution Orders in AVCL and Semantic Web Standards. Chapter 5 presents the application of the MEO methodology. Chapter 6 discusses the enabling technologies we have considered that would ensure secure, ethical command and control between the human commander/operator and the autonomous systems. Chapter 7 presents conclusions from this research. Chapter 8 offers recommendations for follow-on research in order to operationalize ethical control of unmanned systems.

THIS PAGE INTENTIONALLY LEFT BLANK

III. BACKGROUND

A. HISTORY AND TRAJECTORY OF RESEARCH

NPS began research on autonomous underwater vehicles (AUV) in the early 1990s. The first vehicle, tested at sea in the mid-1990s, was *Phoenix*, an AUV that was 6 feet long and weighed approximately 600 pounds. At that time, the onboard control system was called RBM for "Rational Behavior Model." RBM was modeled on standard Naval practice for deployment and operation of manned submarines at the time.

The top level of the three-layer RBM software architecture assumed that Artificial Intelligence (AI) was required to replace the function of a Submarine Commander. For example, mathematical modeling and programming of a Commander's function would need first-order logic (predicate calculus). Therefore, a separate "SPARC" workstation running the "Prolog" language was included in *Phoenix*. RBM functioned well in Phoenix's at-sea testing.

In the early 2000s, *Phoenix* was replaced by a larger vehicle called *Aries*, which contained a more advanced onboard sensor suite, larger batteries, and more powerful main propulsion thrusters. *Aries* was approximately 8 feet long and weighed about 1,000 pounds, and was capable of longer duration and more complex missions than its predecessor. Experience with *Phoenix* led the researchers to full practical realization that mission control based on human-style reasoning (predicate logic) cannot, in general, be formally proven correct for a given mission. Fortunately, the researchers also came to realize that "finite state" logic is adequate for any mission control tasks they actually anticipated carrying out by autonomous vehicles. In contrast to their experience with *Phoenix*, *Aries* researchers were able to prove by exhaustion of all possible outcomes that its missions were correctly programmed (that is, that they accomplished what we intended). Aries was a success, and was retired after completion of all planned mission tasks. All results were published, with references and primary results available at: https://savage.nps.edu/EthicalControl/documentation.

In the early 2010s, NPS researchers began to realize that responsible experimentation with larger and more powerful robots (AUVs or others) would require that some run-time ethical constraints be incorporated into mission control software. This was not done for *Phoenix* or *Aries*. As the researchers addressed this requirement, it became apparent to them that inclusion of such constraints would require a possible "exception" outcome of execution of a mission phase goal (command) in addition to the “success” and “failure” outcomes we had previously considered. This change greatly simplifies mission logic and clarity without loss of generality. This understanding was summarized in the 2018 *Journal of Oceanic Engineering* paper that was based on such “tri-state” logic. (Brutzman, Blais, Davis, and McGhee, April 2018)

Since publication of the 2018 paper, the NPS researchers have been concerned with implementation details for tri-state mission logic for autonomous robots and human/robot teams. To date, the researchers have demonstrated, in human interactive form, execution of a simulated 8-phase “Sailor Overboard” recovery mission by a human/robot team, using either Prolog or Common Lisp as a programming language. The researchers expect to complete an XML implementation soon.

A key aspect of tri-state logic, including possible violation of an ethical constraint, either pending or actual, is the need for constant situational awareness by mission control software. The researchers believe that mandating this type of software for mission control could possibly have prevented loss of human life in recent passenger aircraft and self-driving car accidents. Therefore, further applied research on this issue is critical and is needed as soon as possible.

B. KEY INSIGHTS REGARDING HUMAN ETHICAL CONTROL

Many years of work have composed multiple fields of study to provide techniques for maintaining human ethical control of unmanned systems. In this work, ethical theory meets professional practice. A key tenet researchers are keenly mindful of in this project is that each step of the research must work for Human Commanders and Unmanned Systems alike. The following insights and techniques for maintaining human ethical control of Unmanned Systems apply to this project (Davis, Brutzman, Blais, and McGhee, 2016):

1. Ethical operation of robotic systems requires human accountability.

In military operations, human Warfighters in military units are able to deal with moral challenges without ethical quandaries, by using formally qualified experience, and by following mission orders that comply with Rules of Engagement (ROE) and Laws of Armed Conflict (LOAC). However, ethical behaviors do not define the mission plan. Instead, ethical constraints inform the mission plan. In the context of current and future Naval operations, Naval Forces can only command mission orders that are Understandable by (legally culpable) human Warfighters—then reliably and safely executed by robots. In both the legal and moral sense, this paradigm implies that human operators must be in a position to understand, and therefore control, robot mission outcomes. This level of understanding can be achieved through the satisfaction of three requirements: Operator understanding of high-level mission flow; mission descriptions understandable to both human operators and robot vehicles being tasked; and mission descriptions consisting entirely of trusted behaviors and constraints.

2. Algorithms cannot replace human responsibility.

Artificial Intelligence (AI) approaches in general almost invariably make use of easily confounded inferential reasoning or statistical pattern recognition. Applying such broad abstractions to the innumerable situations that can arise in the real world is inherently unpredictable, and also makes unrealistic any assumption of responsibility by human operators. It is therefore apparent that the abstract reasoning of general AI approaches is inappropriate, at least at the present time, for the highest level of robot mission definition and control.

By applying the best strengths of human ethical responsibility, repeatable formal logic and directable unmanned systems together, these combined capabilities provide a practical framework for ethically grounded human supervision of unmanned systems.

C. ETHICAL MISSION DEFINITION AND EXECUTION

Experts and practitioners have worked long and hard toward achieving functionally capable robots. While numerous areas of progress have been achieved, ethical control of unmanned systems in a manner that meets legal requirements has been elusive and problematic. Common conclusions that treat ethical robots as an always-amoral philosophical conundrum, requiring undemonstrated morality-based artificial intelligence (AI) schemes, are simply not sensible or repeatable. Patterning after successful practice by human teams shows that precise mission definition and task execution using well-defined, syntactically valid vocabularies is a necessary first step. Addition of operational constraints enables humans to place limits on robot activities, even when operating at a distance under gapped communications. Semantic validation can then be provided by a MEO to confirm that no logical or legal contradictions are present in mission orders. Thorough simulation, testing, and certification of qualified robot responses are necessary to build human authority and trust when directing ethical robot operations at a distance. Together these capabilities can provide safeguards for autonomous robots possessing the potential for lethal force. This approach appears to have broad usefulness for both civil and military application of unmanned systems at sea. (Brutzman, Blais, Davis, and McGhee, April 2018)

In “Semantic Web and Inferencing Technologies for Mission Definition,” (Davis, 2014), Davis summarizes that operational commanders and intelligence professionals are provided with a continually increasing volume of data from numerous sources. Effective utilization of this data can be hampered by difficulties in fusing different data streams for presentation, correlating related data from various sources and developing reliable summary and predictive products. An opportunity presently exists to improve this situation through the incorporation of Semantic Web technologies into Department of Defense (DoD) systems. Earlier work provides a didactic overview of Description Logics (DL) and their implementation in Semantic Web languages and technologies to include the mathematical properties supporting robust knowledge representation to address military applications. Subsequently, the algorithms for automated reasoning and inferencing with DLs are discussed. Included in this discussion is a comparison of available Semantic Web applications for ontology development and realization or DL reasoning capabilities with real-world knowledge bases. Finally, mechanisms for applying AI techniques to ontological DL information are presented.

IV. METHODOLOGY

A. MISSION REPRESENTATION USING AUTONOMOUS VEHICLE COMMAND LANGUAGE (AVCL)

Informed by aforementioned research efforts, this research sought to realize a structured vocabulary to define unmanned-system missions that is understandable by human commanders and useful in multiple programming languages, plus Semantic Web logical queries, in order to facilitate formalizing mission ontologies, i.e., MEO. Based on this realization, Autonomous Vehicle Command Language (AVCL) is used for the MEO process.

AVCL is a command and control language for humans supervising autonomous unmanned vehicles. Clarity of the ontology represented by AVCL arises from close correspondence to human Naval terminology. AVCL has structured vocabulary defining terms and relationships for mission planning, execution, conduct, recording and replay across diverse robot types. Additionally, AVCL has common-ground Extensible Markup Language (XML) representations for mission agenda plans, mission scripts, and post-mission recorded telemetry results. Through AVCL, operators have a single archivable, validatable format for robot tasking, and results are directly convertible to and from a wide variety of different robot command languages.

A consideration for future work in AVCL is defining unit tests and expected results for mission verification and validation. This approach can grow as the basis of robot qualification by humans, and formal techniques for verification, validation, and accreditation (VV&A). The following website contains detailed information on AVCL: https://savage.nps.edu/Savage/AuvWorkbench/AVCL/AVCL.html

The MEO methodology applies well-developed Semantic Web Standards that integrate queries and reasoning to AVCL for logical consistency checks. This integration of Semantic Web Standards and AVCL results in the MEO methodology having an additional benefit of laying a foundation for an open system architecture framework for implementing ethical control of unmanned systems.

B. KEY BENEFITS OF APPLYING SEMANTIC WEB STANDARDS IN MISSION EXECUTION ONTOLOGY (MEO)

This section highlights three benefits of the integration of Semantic Web Standards with AVCL representations of mission orders that form the basis of MEO.

1. Improving Semantic Representation

Knowledge Representation (KR) is an area of AI research and practice focused on encoding meaning into data. Academia and industry now have a detailed path toward higher levels of machine understanding corresponding to human understanding. Figure 1 depicts the Ontology Spectrum in terms of a relationship between search capability into a data set and its metadata that would lead to understanding the semantics of the data set. The figure shows that semantic representation improves as the search capability increases from simple recovery of data to being able to reason from the data.

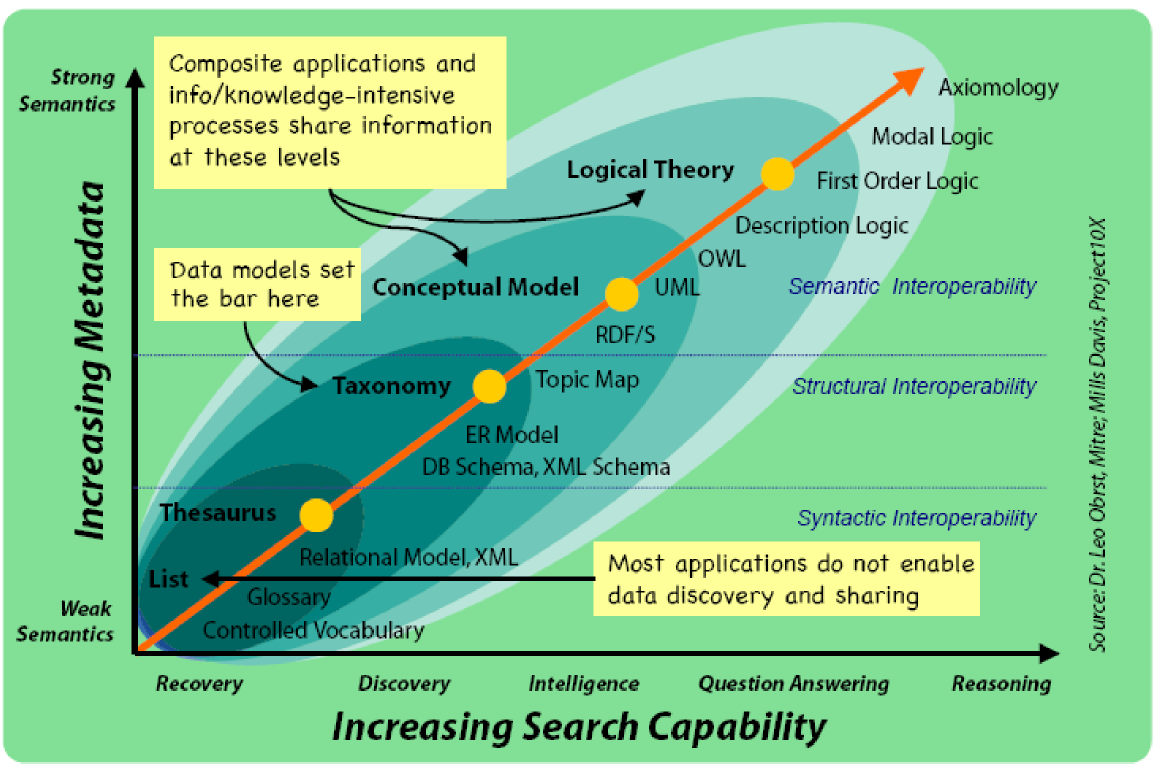

Figure 1. The Ontology Spectrum. Acronyms: Database (DB). Extensible Markup Language (XML). Resource Description Format / Schema (RDF/S). Unified Modeling Language (UML). Web Ontology Language (OWL). (Orbst and Davis, May 2015)

2. Improving Interoperability

This project defines Interoperability as "the capability of a system to automatically, without human intervention, provide services to and accept services from other systems, and to use the services so exchanged to enable the systems to work together to achieve a desired outcome" (Blais and Lacy, 2004). Academia and industry have laid out a path toward higher levels of interoperability. Figure 2 shows the Levels of Conceptual Interoperability Model (LCIM) (Tolk et al., January 2006). The LCIM categorizes interoperability in six levels, from Level 0 (No Interoperability) to Level 6 (Conceptual Interoperability). The objective is to achieve conceptual and pragmatic interoperability such that the systems may be composable.

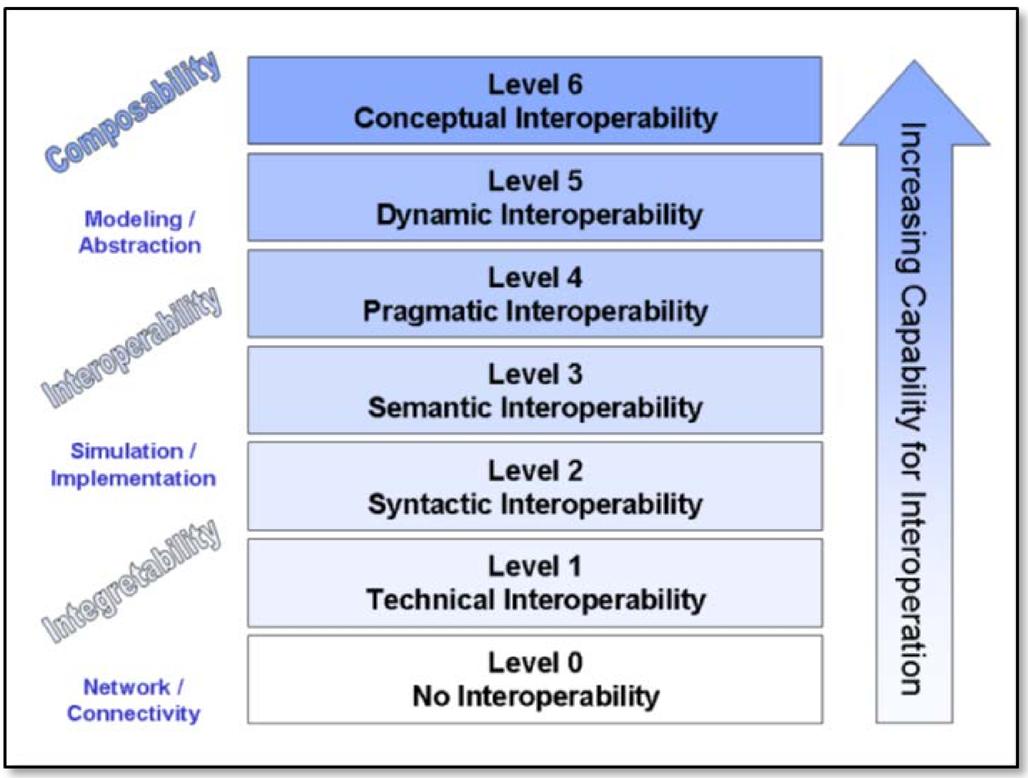

Figure 2. Levels of Conceptual Interoperability Model. (Tolk et al., January 2006)

3. Scalable Application

Architects of the World Wide Web Consortium (W3C) have laid out a layered set of standards to achieve the Semantic Web vision: "not a separate Web but an extension of the current one, in which information is given well-defined meaning, better enabling computers and people to work in cooperation" (Berners-Lee et al., 2001). The ultimate goal for the Semantic

Web is to achieve a scalable trusted information infrastructure where humans and software interact meaningfully, in a repeatable environment where expectations of quality and integrity are met. Most relevant to the MEO methodology is the scalable approach of the Semantic Web standards, which indicates that single (ship + robot) solutions have the potential to grow and encompass many simultaneous systems, and achieve improved data sharing, mission de-confliction, and coordinated operations.

Figure 3 shows the architecture of the Semantic Web Stack. This architecture extends the larger World Wide Web architecture. All of the Semantic Web data languages are approved W3C Recommendations, meaning formal standards that have undergone a rigorous process for broad inputs and tested results. Examining each of the critical blocks in this figure, it is clear that proof and unifying logic are mathematically well-defined. Trust-derived (composed) statements in the Semantic Web architecture arise from encryption and digital signature, confirming trusted data sources. Formal logic of trust statements is the basis for deriving new information. This project exercises every layer of the Semantic Web Stack ("Semantic Web Stack," 2020).

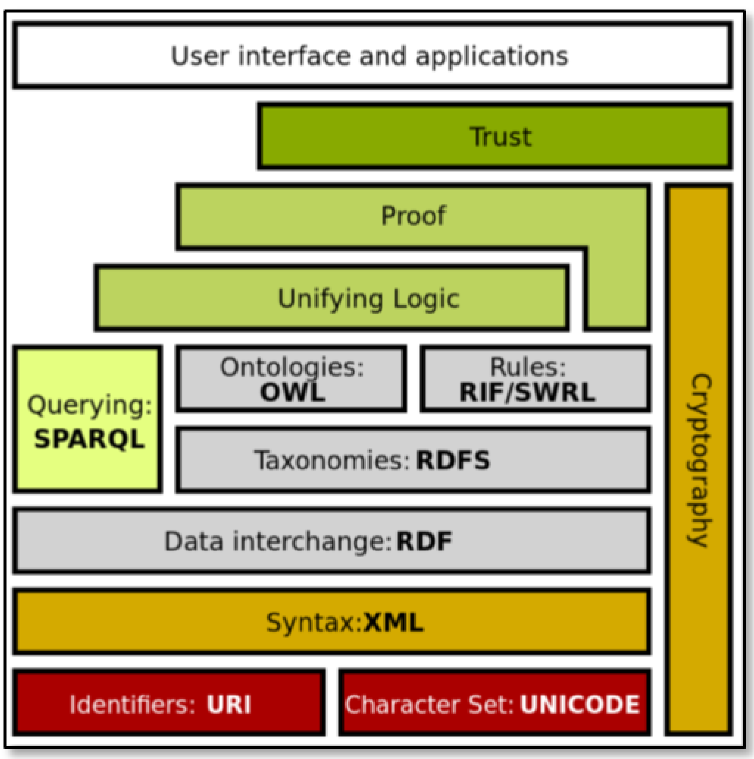

Figure 3. The Semantic Web Stack. ("Semantic Web Stack," 2020)

C. MEO DEVELOPMENT USING SEMANTIC WEB STANDARDS

This section describes the development of the MEO methodology that leverages these aforementioned benefits of Semantic Web standards. The methodology developed in the project for MEO—representation of mission orders for robots in ontological forms using Semantic Web standards—consists of the following tasks:

- Define MEO from concepts, properties, relationships using Protégé tool.
- Create full set of canonical missions in AVCL (XML).
- Transform AVCL representation of the missions into corresponding subject-predicate-object triples using Semantic Web standards in OWL.
- Confirm that AVCL MEO missions validate satisfactorily using Protégé.
- Automate build process as a suite of repeatable unit-test queries (log).
- Write SPARQL Protocol and RDF Query Language (SPARQL) metaqueries to test and demonstrate MEO concepts and relationships.
- Write SPARQL queries to test AVCL mission representations in Turtle.

Specific tasks for developing the MEO methodology above are performed in the Apache Ant software building tool. Refer to the project website (link provided in Appendix A) for additional documentation.

1. Ontology Definition

Table 1 summarizes the key elements used in the MEO methodology that work together to implement the aforementioned tasks to represent mission orders for unmanned systems.

Table 1. Summary of Relationships in MEO Methodology for Ethical Control of Unmanned Systems in Surrogate Scenarios

Key Element	Contribution to MEO Methodology
Autonomous Vehicle Command Language (AVCL) for Missions	• Declarative XML, years of NPS research.
Multiple Mission Representations	• Imperative commands (orders/waypoints/etc.). • Declarative commands (mission goals). • Mission results (order log, telemetry etc.). • Mission metadata for parameters, settings. • Lisp and Prolog examples (Bob McGhee, NPS).
Autonomous Unmanned Vehicle (AUV) Workbench Simulation and Visualization Support	• Recently restored, debug testing commenced. • AVCL 2.1 is prior published version, centered on syntactic validation, solo robot operations. • AVCL 3.0 is new working version for testing range of multi-participant missions.

Key Element	Contribution to MEO Methodology
Mission Execution Ontology (MEO) for Semantic Validation	• Semantic Web framework of rules, relationships for ethical validation. • Initial examples in IEEE JOE paper. • Retested using current Protégé, Jena tools.
Sailor Overboard and Other Missions	• Hand-crafted triples using Turtle syntax. • Beginning to build unit testing framework. • Confirming correlation of AVCL information model to existing MEO ontology. • Automatic conversion of AVCL missions to match, thus accelerating multiple-mission testing on diverse systems. • Visualization, reporting via AUV Workbench can aid understanding, mission planning and further progress.

Figure 4 is a diagram of the unmanned vehicle MEO that shows how the aforementioned key elements interrelate. The MEO interrelates the mission order to the unmanned vehicle based on the concepts of Mission, Goal, Intended Outcome, Constraint, and End Condition, as well as Vehicle and Vehicle Feature to perform the mission. The interrelationships between the concepts may be Asserted or Inferred. An Asserted relationship is one that is explicitly declared in AVCL to represent the corresponding mission execution action. An Inferred relationship is implicit by semantic relationships.

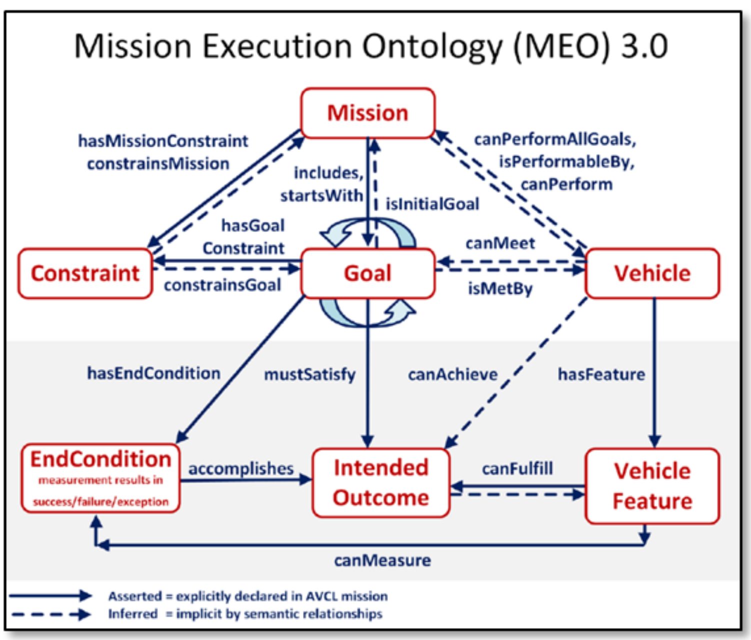

Figure 4. Unmanned Vehicle MEO

2. Ontology Testing

Once the interrelationships of the MEO are defined, a Semantic Web Standards tool such as Protégé is then used to implement and test the ontology for reasonableness and non-contradiction. Figure 5 is a graphical depiction of MEO in Protégé.

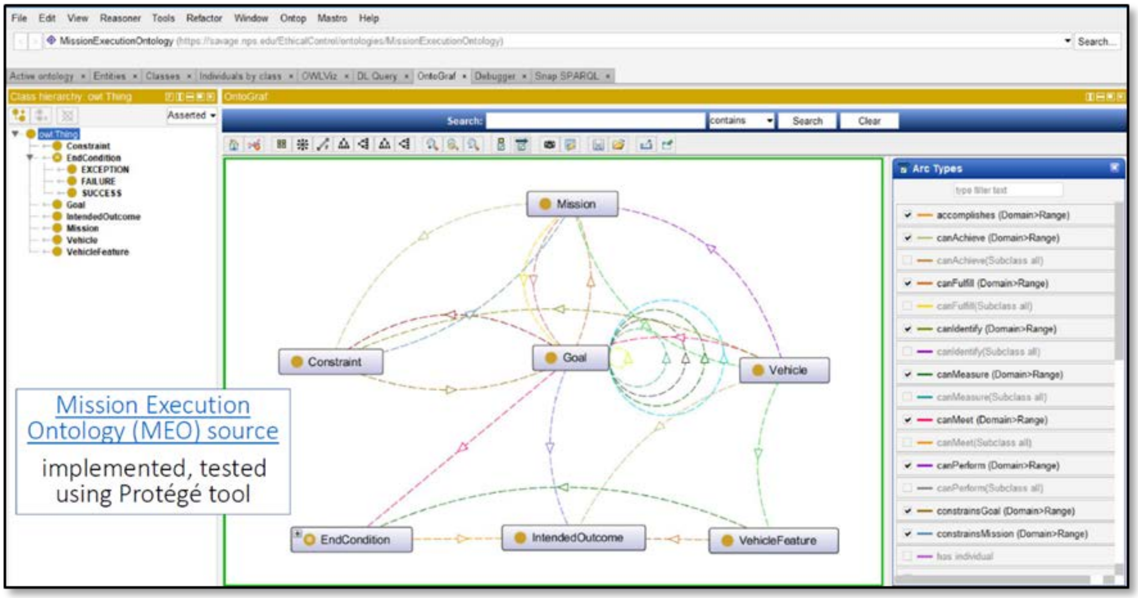

Figure 5. MEO Testing Using Protégé Tool

3. Ontology Confirmation and Validation

Semantic query language is used to confirm the mission ontology. This research uses SPARQL, standardized as a W3C recommendation, to express queries in RDF/ OWL or Turtle syntax against the missions represented in AVCL. The query results of the MEO reveal interesting properties about the missions that are otherwise difficult to determine. Inferences in MEO can also be combined and correlated. The goal of MEO confirmation using semantic query language is to express in-depth mission-related queries that determine whether all logical mission prerequisites and constraints are satisfied, and whether tactical policies and Rules of Engagement (ROE) are met.

After the MEO of the mission order is implemented and tested for reasonableness and non-contradiction, a Semantic Web Standards tool can be used to validate the MEO. Figure 6 is a graphical example of a mission validation using the Protégé tool.

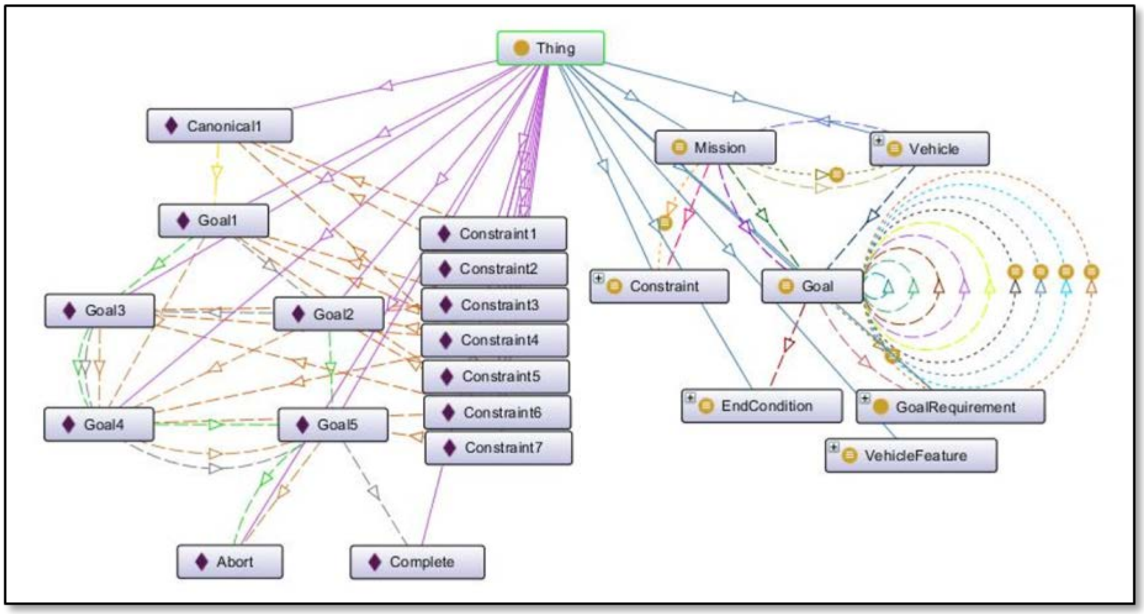

Figure 6. Example Mission Validation Using Protégé Tool

For the same scenario, Figure 7 contains examples of relationship definitions of the MEO expressed in Subject-Predicate-Object form using Turtle Syntax of Semantic Web Standards. This form of MEO expression facilitates validation and queries for logical coherence.

```
###  https://www.nps.edu/ontologies/MissionExecutionOntology/missions#Goal_Launch
          :Goal_Launch rdf:type owl:NamedIndividual ;
                    meo:hasNextOnFail :Goal_FailureDiagnosis ;
                    meo:hasNextOnSucceed :Goal_TransitSearch ;
                    meo:hasNextOnViolate :Goal_FailureDiagnosis .

###  https://www.nps.edu/ontologies/MissionExecutionOntology/missions#Goal_TransitSearch
          :Goal_TransitSearch rdf:type owl:NamedIndividual ;
                   meo:hasNextOnFail :Goal_SearchForSailorAdrift ;
                   meo:hasNextOnSucceed :Goal_TrackSailorAfloat ;
                   meo:hasNextOnViolate :Goal_FailureDiagnosis .
```

Figure 7. Examples of MEO Relationship Definitions Expressed in Turtle Syntax of Subject-Predicate-Object Triples

4. Ontology Implementation

Once the MEO of the mission order is validated, it is now ready for implementation in a mission for the assigned unmanned systems. Figure 8 is a graphical representation of the Sailor Overboard recovery mission for an unmanned system that is defined using the MEO methodology.

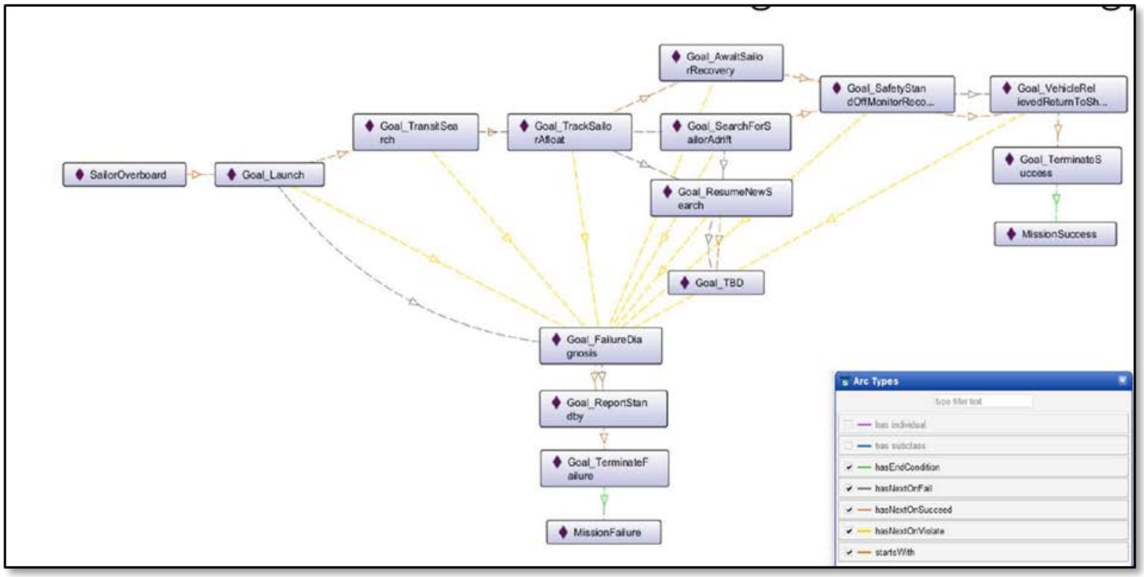

Figure 8. Ethical Control of Unmanned Systems in a Surrogate Scenario: Sailor Overboard Recovery Mission Defined Using the MEO Methodology

Refer to the project website (link provided in Appendix A) for additional documentation of the result of this MEO confirmation using SPARQL.

The next chapter presents applications of the MEO methodology in the design, development, and testing of a set of exemplar missions in Naval operations.

THIS PAGE INTENTIONALLY LEFT BLANK

V. APPLICATION OF THE MEO METHODOLOGY

Lifesaving missions—and missions accomplished with lethal force—are complementary in military operations. Human-robot activity can result in lethal or lifesaving outcomes. Continuing refinement and clarity in mission design are opening the path to repeatability, and hence a methodology for implementing ethical control of unmanned systems. This section discusses application of the Observe-Orient-Decide-Act (OODA) Loop decision paradigm to unmanned system mission design in order to harmonize unmanned system missions with human operations.

A. ENABLING TECHNOLOGIES AND CONCEPTS

Artificial Intelligence (AI), conceptual chunking (of information), and algorithm loop management are relevant in the development and design of missions for unmanned systems.

1. Artificial Intelligence (AI)

AI turns data into information for use by humans. AI systems do not have capacity for rational thought or morality. Unmanned systems require sophisticated control across time and space. A large and involved body of internationally accepted law comprises Law of Armed Conflict (LOAC), bounding Rules of Engagement (ROE). Only professional warfighters have moral capacity, legal culpability, and societal authority to direct actions applying lethal force. Humans must be able to trust that systems under their direction will *do what they are told to do*, and *not do what they are forbidden to do*.

2. Conceptual Chunking

Conceptual Chunking includes the following characteristics ("Chunking (psychology)," 2020):

- In cognitive psychology, chunking is a process by which individual pieces of an information set are broken down and then grouped together.
- A chunk is a collection of basic familiar units that have been grouped together and stored in a person's memory. These chunks are able to be retrieved more easily due to their coherent familiarity.
- It is believed that individuals create higher order cognitive representations of the items within the chunk. The items are more easily remembered as a group than as the individual items themselves.

- These chunks can be highly subjective because they rely on an individual's perceptions and past experiences that are able to be linked to the information set. The size of the chunks generally range anywhere from two to six items, but often differ based on language and culture.

Design experience in this research has demonstrated Conceptual Chunking for mission design by grouping goal tasks within specific mission phases.

3. Algorithm Loop Management

Proper algorithms always include one or more termination conditions. A sequence of operations proceeds through a finite number of steps, otherwise the system is performing an infinite loop without end. Infinite-loop sequencing or unterminated recursion are common computational failure modes and must be protected against for reliable operations. Nevertheless, a frequent characteristic of at-sea operations is to perform repeated tasks in an interactive fashion until complete—either via task success or a terminating condition that forces halting of the process.

The presence of termination conditions can be verified in mission logic and tested in simulation. Example terminating conditions, some fixed/adaptive, and some iterative/exceptional, are:

- Reach maximum number of iterations
- Point of diminishing returns (e.g., unchanging search effectiveness)
- Time-out deadline reached, or else no longer feasible to continue
- Insufficient power remains, conduct graceful shutdown for recovery
- Equipment damage or unexpected software failure; log and shutdown
- Interfering operational conditions (e.g., potential hazard to friendlies)
- Human direction asserts higher priority and overrides decision logic

A recommended future work is to express precisely in AVCL all termination conditions such as those listed above as constraints in MEO for algorithm loop management.

B. MISSION ORDER DESIGN CONSIDERATIONS

1. Mission Order Clarity

Clarity is paramount when giving or receiving mission orders. It is especially important for human Commanders providing clear directions to Human-Machine teams. Commanders must avoid the danger of ambiguity, or even anthropomorphizing robots as human-like. Simplicity of success, failure, and (rare) exception outcomes encourages well-defined tasks and unambiguous, measurable criteria for continuation. Furthermore, the following set of complementary questions applies as an incisive determination for clarity in Human-Machine missions: A wrong question to ask first when planning a tactical operation for a Human-Machine team is: “What are my robots doing out there?” Rather, the right question to ask first when planning a tactical operation would be: “What is my human-robot team doing out there?” Human-robot team missions first have to be understood! Indeed, robots complement humans, who must remain in charge throughout.

There is an added benefit of mission order clarity: Mission orders that are clearly readable/runnable by humans and robots can be further composed and checked by Command and Control (C2) planning tools to test for group operational-space management, e.g., force movement coordination, avoid mutual interference, weapon engagement zone assignments, etc.

2. Mission Order Validation

Mission orders must be both clear and validatable prior to dissemination and execution. Clear mission orders in this context are understandable by humans and readable by unmanned systems. They are validatable as syntactically correct, having no typographic errors or gaps, and avoiding non-sequitur “Garbage In. Garbage Out” (GIGO). Clear mission orders are also validatable as semantically correct, having no prerequisite omissions or contradictions. For example, upon review of a mission order, a Tactical Action Officer (or Commanding Officer) validates the mission order when he/she can confidently say: “Yes, I understand and approve this human-robot mission”; or, equivalently: “Yes, I understand this mission and my team has the ability to carry it out themselves.” Conversely, if a mission order is organized and/or presented such that a Human Commander/Operator cannot fully review, understand, and/or approve such mission, then it is likely that the received mission order is ill-defined and needs further clarification anyway.

3. Application of Conceptual Chunking for Mission Order Clarity and Validation

Aspects of AVCL-represented mission orders are designed to support chunking for clarity. These AVCL attributes include a well-defined, structured vocabulary that can describe a hierarchy of distinct, familiar goals. This project shows that it is feasible to group (i.e., by chunking) mission goal tasks within specific mission phase definitions to achieve mission order clarity. A recommended future work for applying conceptual chunking for mission order clarity is to demonstrate best practices for characterizing common mission phases in order to establish testable design patterns for mission orders as templates that aid operators in issuing clear and validatable mission orders.

4. Mission Execution Decision Process: The "Observe-Orient-Decide-Act (OODA)" Loop

With respect to mission order clarity, if the Human Commander/decision maker does not apply a decision model such as an OODA Loop, the Human Commander does not have a competent decision model through which to execute the assigned mission order, no matter the clarity and validity of the order. "The OODA loop is the cycle Observe–Orient–Decide–Act, developed by military strategist and USAF Colonel John Boyd. Boyd applied the concept to the combat operations process, often at the operational level during military campaigns. It is now also often applied to understand commercial operations and learning processes. The approach explains how agility can overcome raw power in dealing with human opponents." ("OODA loop," 2020) All effective purposeful military activity can be conceived in terms of the OODA Loop process, especially at tactical/operational levels. Based on its succinctness in describing the decision process, the MEO methodology aligns its mission design with the phases of the OODA Loop in order to ensure that unmanned systems will be able to emulate the decision process of the human operator/team member within the Human-Machine team.

The reason for applying the OODA Loop in mission design for ethical control of unmanned systems is that the classical, robotic Sense-Decide-Act cycle for closed-loop control of unmanned systems is insufficient for proper delegation of lethal (or lifesaving) force to the unmanned systems. The OODA Loop decision model is essential for coherent operations in Human-Machine Teams.

In the OODA Loop for unmanned systems, the Observe phase is the beginning of the decision process and includes direct sensing and communication inputs. The Orient phase includes thorough Rules of Engagement (ROE) constraints and IFFNU (identification, friend, foe, neutral, unknown) of all relevant contacts. The Decision phase implements the logic of unmanned system tactics, techniques, and procedures (TTP), including authorization and confirmation by human supervisors, either in real-time or in advance (pre-planned and/or stored in memory), for critical decision steps leading to use of lethal force. The Act phase is implemented in tandem with direct or intermittent human supervisory commands that enable effective Ethical Control of remote systems. The process repeats as the unmanned system observes and senses the impact of its actions. This feedback loop is essential for decision making and refinement, and generally leading to, without surprise, more effective military operations. Therefore, unmanned system activity must complement, not contradict, human decision processes such as the OODA Loop paradigm.

5. Mission Order Coherence

Application of ROE and LOAC requirements in MEO may be a part of the Mission Definition step of the MEO methodology (described in Section III.B.4), where relationships and requirements for mission execution are defined. For example, typically, an ROE requirement may be represented as Goal success/failure criteria, preset authorities, or time-outs for delegation, etc., and as Constraints on mission conduct, e.g., safe zones, permission periods/requirements, etc. When human Commanders confirm correct inclusion of ROE requirements in mission orders, they essentially perform an audit of doctrine and TTPs. Similar audit confirmation can be applied to well-structured orders. As previously mentioned in Section III.C.5, AVCL has demonstrated that it can be used to develop and express well-defined mission goals for unmanned systems. The mission analyst can then perform SPARQL queries of the MEO for logic-confirmation checks. The resulting mission orders are thus coherent from the OODA perspective.

6. Tactical Span of Control

Span of control is number of subordinates reporting to a supervisor. In effect, multiple offboard unmanned systems supervised by a ship comprise its span of control across the tactical battlespace. Tactical span of control is required for any Commander who commands subordinates. Greater tactical presence across distances of time and space means the ship commanders have greater ability to influence their assigned area of operation. Clear mission guidance on human-checkpoint requirements reduces dependency on communication links (i.e., Network Optional Warfare). Figure 9 illustrates how such increased ability to project power enables the ship to maintain chosen standoff location while focusing direct attention and actions in multiple locations at once.

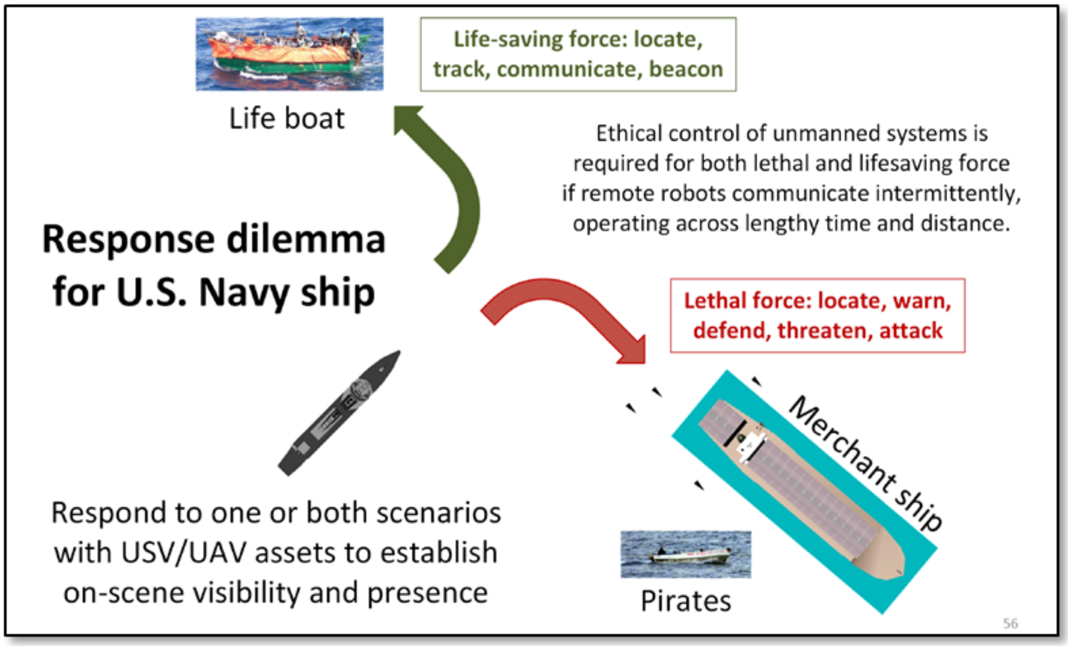

Figure 9. Example Scenario of Ethical Conundrum for Unmanned System Mission Execution.

C. CANONICAL MISSION SCENARIO DEVELOPMENT

Unmanned systems working in tandem with human forces, authorized by the commander for lifesaving or lethal force, can handle progressive challenges in distance and time. This research project has developed four categories of mission scenarios in progressive sophistication to test and evaluate Ethical Control design:

1. *Basic:* Show flow-logic of ternary control in a simple, real-time scenario.
2. *Intermediate:* Adds looping logic and long-time duration activity at long-distance but with real-time communication.
3. *Advanced:* Adds long duration with possible loss of communication.
4. *Operational Standard:* Encounters deceptive and unconventional adversary and responds with proportional force.

Each category has an exemplar mission from representative Naval operations. The missions themselves are carefully crafted in a narrow sense in order to illustrate and test specific characteristics of the Ethical Control methodology. These exemplar missions form the set of canonical missions for this study:

Exemplar Mission A: Sailor Overboard. The Unmanned System is assigned to apply lifesaving force under close coordination to recover the sailor.

Exemplar Mission B: Lifeboat Tracking. The Unmanned System is tasked to apply lifesaving force to search and track lifeboats at long distance from the Human Commander.

Exemplar Mission C: Pirate Boats Attack. The Unmanned System equipped with lifesaving and lethal force is assigned to overtake a pirate small-boat gang attempting to capture a threatened evading merchant ship in order to protect the merchant ship. The Unmanned System must operate over a long time period for the mission, emphasizing restraint throughout.

Exemplar Mission D: Hospital Ship EM Decoy. The Unmanned System equipped with lifesaving and lethal force is assigned to patrol the perimeter of the task force and encounters a deceptive and unconventional adversary. For comparison, two variations show the fundamental importance of ethical constraints on mission execution.

- *Sense-Decide-Act Loop*. This commonplace pathology illustrates lack of ethical control. Adversary exploits the Unmanned System’s rudimentary Sense-Decide-Act capabilities as vulnerabilities, and provokes it to immediately react with a counterattack on a False-Flagged Hospital Ship.
- *Observe-Orient-Decide-Act (OODA) Loop.* This scenario demonstrates that addition of ethical control to Sense-Decide-Act overcomes limitations of independent machine response, and indeed leads to more effective warfighting. OODA Loop tactics and Ethical Control constraints prevent automatic erroneous counterattack against false flag placed on friendly ship, and thus improves defense. Human confirmation of the Unmanned System’s Identification Friend, Foe, Neutral, or Unknown (IFFNU) classification result to detect spoofing anti-pattern, and authorization for the Unmanned System to apply lethal force prior to use, prevents reflexive automatic counterattack and accelerates defense of the force.

A range of functionality that tests the majority of Ethical Control capabilities currently envisioned has been demonstrated in this project. These canonical mission capabilities also set the stage for further research in missions and scenarios of interest, e.g., Human-Machine Teams in a Counter-Swarming mission.

Completing this initial set of canonical missions demonstrates both logical soundness and human comprehensibility of the Ethical Control methodology, for both lifesaving and lethal force. This initial set of missions “tuned up” the AVCL vocabulary for mission orders, revealing good practices and repeatable design patterns for common activities in diverse missions. The AVCL representations of various mission orders are subsequently translated into corresponding Semantic Web Standards ontology representations in order to perform semantic validation of correctness of ethical constraints. Then, using the AVCL representation of the mission orders, simulation of the mission is performed in the AUV Workbench tool to show that the mission executes in simulation, or else reveals hidden flaws.

Table 2 summarizes the five canonical missions and the tasks the unmanned systems would need to perform in each phase of the OODA Loop.

Table 2. Tasks Assigned to Unmanned Systems in Each Phase of the OODA Loop for Four Canonical Missions in Naval Operations

Exemplar Missions	Observe	Orient	Decide	Act
Sailor Overboard	Find Sailor	Report status	Avoid interference	Track sailor until rescued or relieved
Lifeboat Tracking	Find lifeboat	Report status	Two-way communication	Track lifeboat until relieved
Pirate Boats Attack	Find merchant ship, pirate small boats	• Identify Friend, Foe, Neutral, Unknown (IFFNU) • Issue warnings	Human commander authorization to use lethal force	Attack to defend ship if provoked, stay with merchant
Hospital Ship EM Decoy: Sense-Decide-Act Loop	EM threat signals detected	(no orientation step in Sense-Decide-Act)	Reflex-response weapons attack	Mistaken attack on friendly or neutral forces equals a war crime
Hospital Ship EM Decoy: OODA Loop	EM threat signals detected	IFFNU(including correlation)	Human requirement for lethal force unmet, attack avoided	Report threat alert, commence search for hostile actors

Based on the research accomplished in this project, a recommended future work is to prepare and develop comprehensive testing of unmanned systems mission orders across the range of assigned mission requirements that constitute the operational qualification process for the unmanned system. By applying the MEO methodology in the aforementioned scenarios, a scale for qualifying unmanned systems for performing missions under Ethical Control emerges:

Level 1: Basic. Qualified to apply lifesaving force under close coordination with the Human Commander, e.g., Sailor Overboard Recovery mission.

Level 2: Intermediate. Qualified to apply lifesaving force at long distance from the Human Commander, e.g., Search and Rescue mission.

Level 3: Advanced. Qualified to apply organic lifesaving/lethal force over long time periods for the mission, emphasizing restraint throughout, e.g., Counter-Piracy mission.

Level 4: Operational Standard. Qualified to apply organic lifesaving/lethal force in contested/deceptive environments, e.g., Force Protection mission. Uses human confirmation of Identification Friend, Foe, Neutral, or Unknown (IFFNU) classification result to detect spoofing anti-pattern, with authorization required to apply lethal force, prevent reflexive automatic counterattack response, and engage with proportional force.

Each of the four canonical missions has the following structure:

a. Design and Description

b. Mission Decision-Flow Diagram

c. AVCL implementation source: version control, XML Spy table

d. Semantic Web Mission MEO .ttl Turtle representation

e. SPARQL semantic queries

f. Lisp and Prolog autocode

g. AUV Workbench simulation

1. Basic Mission for Unmanned Systems: Sailor Overboard for Lifesaving Force under Close Coordination

The scenario is a single unmanned air/surface vehicle performing lifesaving actions to complement human responses during "Sailor Overboard" recovery operations. Robot actions are carried out in direct concert with formal shipboard emergency procedures. Multiple Unmanned Aerial Vehicles (UAVs) and Unmanned Surface Vehicles (USVs) might be employed in parallel with ships and aircraft, and must avoid mutual interference by each following de-conflicted mission orders. The Phases of mission are: Deploy/Launch, Rendezvous, Track Sailor until Safe, and Return/Recovery. The Human Supervisory Role is to order the responding unmanned systems to standoff should they become interfering, and possibly take manual control due to proximity. Rescuers can communicate to the sailor overboard via loudspeaker or beacon light.

a. Design and Description

This scenario is used to show that mission design can complement shipboard procedures. The scenario explores how lifesaving force is complementary to lethal force, with many similar considerations for remote supervision. This scenario is the first one studied in this project in order to demonstrate human-system teaming in close proximity to the ship in command, where direct override of robot control by human operator is possible. There would be no temporal delays in mission execution—all actions and reactions must be immediate. These operational requirements form the basic level of ethical control and constraints on the unmanned system.

b. Mission Decision-Flow Diagram

Figure 10 is the Mission Decision-Flow Diagram for the Sailor Overboard mission.

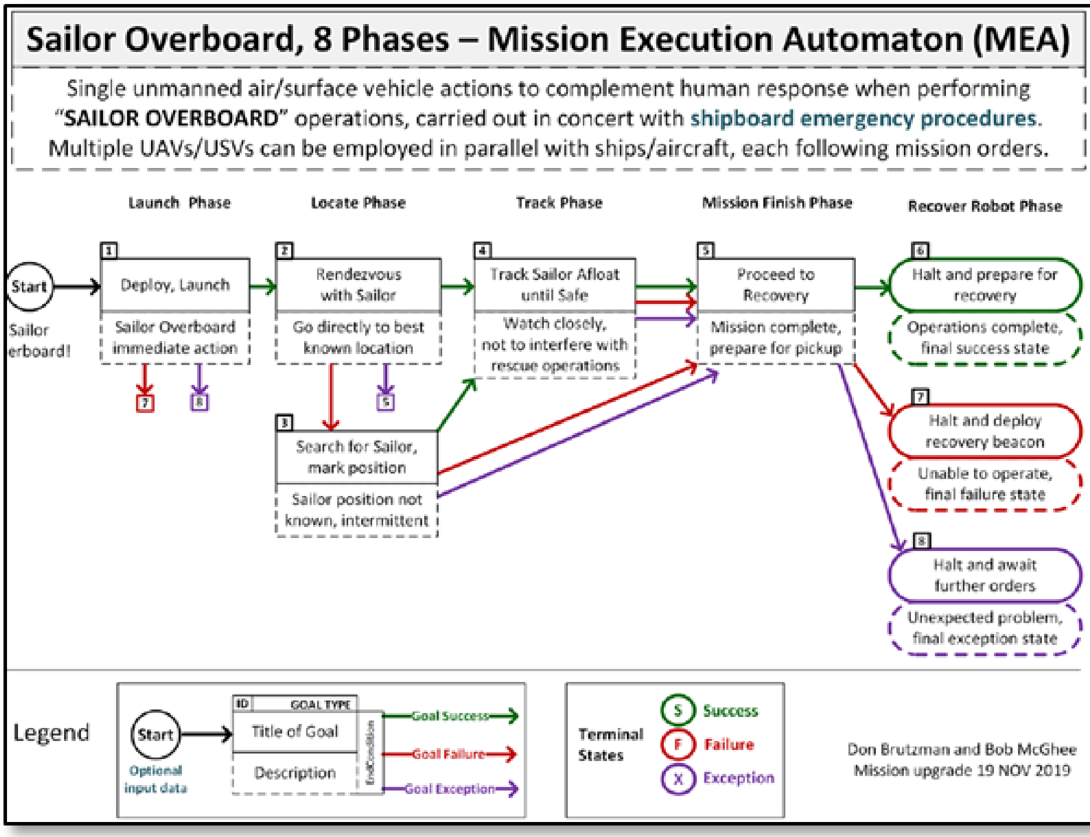

Figure 10. Sailor Overboard Mission Decision-Flow Diagram

A lesson learned from implementing the MEO methodology for this scenario is the need for good mission design patterns that visually represent the AVCL mission satisfactorily, and diagramming missions into decision phases using a temporal flow representation (left to right, e.g., Gantt chart). Hence, the Mission Decision-Flow diagram is an important step in the analysis of the mission in a particular scenario.

c. AVCL Implementation Source: Version Control, XML Spy Table

The AVCL representation of the Unmanned Systems' mission order for this scenario may be reviewed here:

https://gitlab.nps.edu/Savage/EthicalControl/blob/master/missions/avcl/SailorOverboard.xml

d. *Semantic Web Mission MEO .ttl Turtle representation*

The Semantic Web Standards in Terse RDF Triple (.ttl) format of the MEO for the scenario may be reviewed here: https://gitlab.nps.edu/Savage/EthicalControl/-/blob/master/ontologies/MissionExecutionOntology3.0.ttl

e. *SPARQL Semantic Queries*

For confirmation and validation, the SPARQL queries of the MEO may be reviewed here: https://gitlab.nps.edu/Savage/EthicalControl/blob/master/queries/MissionQuery_01_GoalBranches.rq

f. *Lisp and Prolog Autocode*

The query results in Lisp and Prolog form may be reviewed here: https://gitlab.nps.edu/Savage/EthicalControl/-/blob/master/queries/SailorOverboardConverted.MissionQuery_01_GoalBranches.rq.txt

g. *AUV Workbench Simulation*

Figure 11 is a screenshot of the graphical layout of the simulation of the scenario in the simulation environment in the AUV Workbench tool.

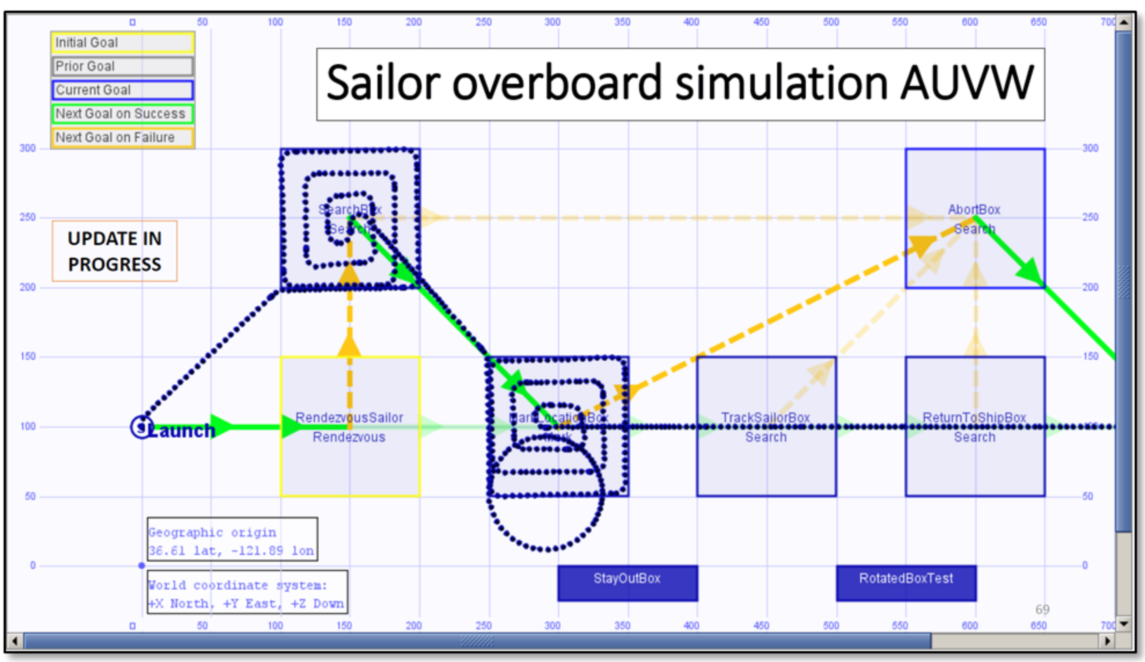

Figure 11. Sailor Overboard Mission Scenario as Simulated Using AUV Workbench (AUVW)

For brevity, this report presents the Mission Design and Description, as well as the Mission Decision-Flow Diagram for each of the subsequent scenarios. The NPS project website contains the MEO in AVCL, the Semantic Web SPARQL queries for confirmation and validation and the respective results, as well as the AUV Workbench simulation. These results for all missions are posted here: https://gitlab.nps.edu/Savage/EthicalControl/-/tree/master

2. Intermediate Mission: Lifeboat Tracking and Lifesaving Force Application under Remote Condition

This mission is similar to Sailor Overboard in demonstrating use of lifesaving force, but with far greater distances, over the horizon, and increases the ship Commander's tactical span of control. Potential for intermittent or lost communications in real time requires advance guidance for default behaviors desired by the human controller. Consideration of possible transfer of supervisory control mid-mission to another cooperating vessel appears feasible. The lessons learned from this scenario include Vertical grouping of related subtasks in a mission phase helps in structuring mission goal sets, without requiring a change to the ternary logic of AVCL mission goals. Also, coexistence of multiple constraints is possible, but requires careful thought to design and implement.

a. Design and Description

The mission for the unmanned system in this scenario is to provide remote presence for locating, tracking, communicating, and beaconing. The mission Phases are: Deploy/Launch, Rendezvous, Track Lifeboat, Beacon/Communicate, and Return/Recovery. The Human Supervisory Role and Constraints are to monitor, communicate, respond or coordinate rescue effort. Considerations for a low fuel condition and graceful-degradation response are included in the MEO. The mission presents an intermediate level of ethical challenge for the unmanned system and the Human operator due to the increased physical distance and tactical span of control.

b. Mission Decision-Flow Diagram

Figure 12 is the Mission Decision-Flow Diagram for this mission.

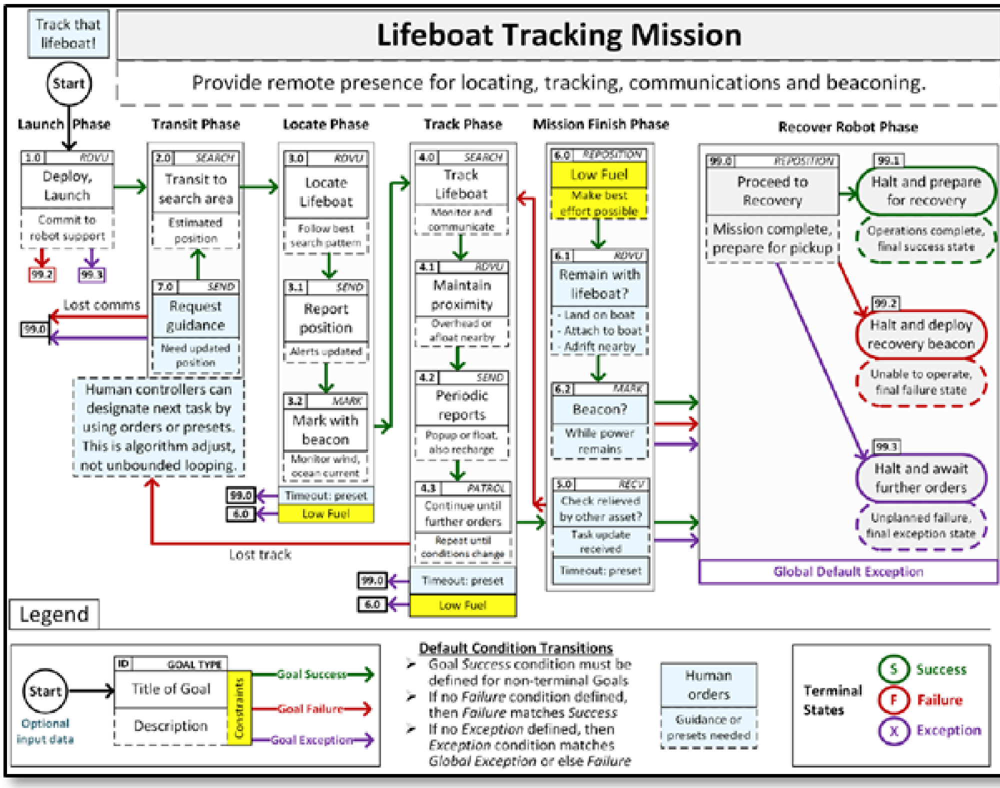

Figure 12. Decision-Flow Diagram for Lifeboat Tracking Mission

3. Advanced Mission: Deter Pirate Boats Seizing Merchant Ship with Ability for Steady Escalation to Lethal Force

The motivations for the design of this mission are the necessity to apply lethal force against pirates distant from ownship, and corresponding lifesaving force potential for the hostage merchant crew. In the scenario, unmanned systems must operate over a long time period, emphasizing restraint throughout. Additionally, supervisory checkpoints for human operator control of the unmanned systems are soft and strict. The lessons learned from this scenario include the concept of phases helps organize overall mission structure sensibly (e.g., approach, warning, attack, recovery). Additionally, decision looping is necessary, with human control as checkpoints to help avoid deadlock (i.e., algorithm loop management).

a. Design and Description

The unmanned systems' mission in this scenario is to overtake a pirate small-boat gang attempting to capture a threatened evading merchant ship. The mission has the following phases: Deploy/Launch, Search, Approach and Track, Warning, and Attack. For Human Supervisory Role and Constraints, the Human Commander controls the pace of engagement, and the role of careful, deliberate escalation in the use of force. The Human Commander/Operator must also confirm IFFNU classification, and must order lethal force prior to use. The scenario includes a low-fuel condition for the unmanned system and consequent graceful-degradation response for realism. Additionally, the scenario presents an ethical conundrum for the Human-Machine Team with low ammunition condition in the Unmanned Systems, with consideration for the team to assign the Unmanned Systems to fight to the finish, or stand in reserve.

b. Mission Decision-Flow Diagrams

Figures 13 through 15 are the Mission Decision-Flow Diagrams for the counter-piracy scenario covering the Approach, Escalation, and Counterattack phases of the mission.

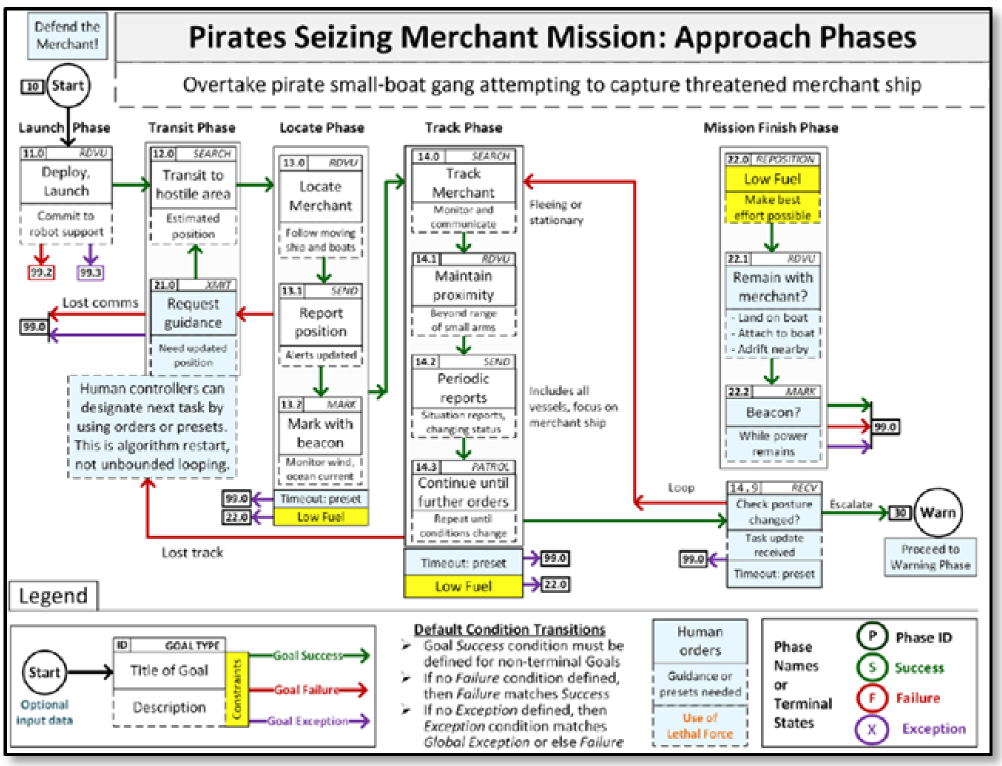

Figure 13. Decision-Flow Diagram for Pirate Boats Attack Mission—Approach Phases

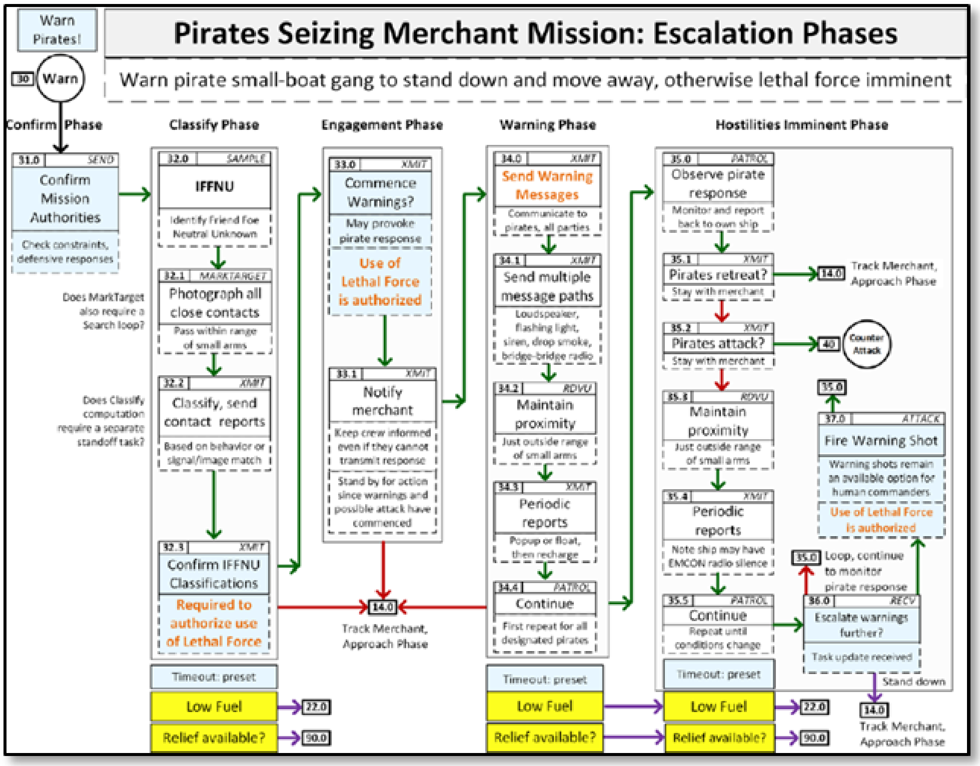

Figure 14. Decision-Flow Diagram for Escalation Phases of the Pirate Boats Attack Mission

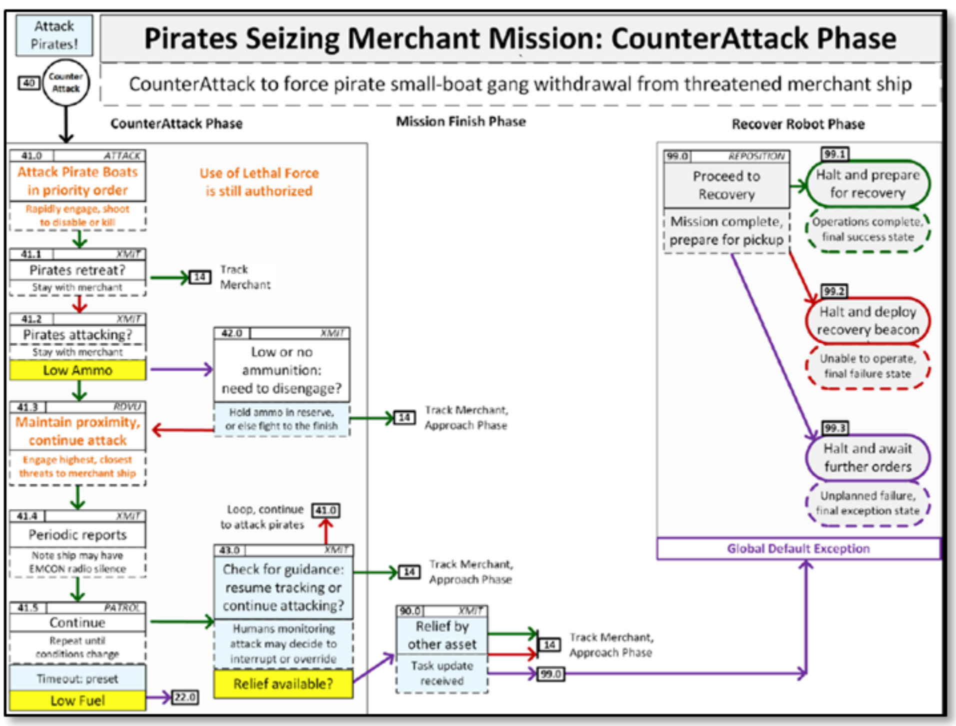

Figure 15. Decision-Flow Diagram for Counterattack Phase of Pirate Boats Attack Mission

This Counter-Piracy scenario presents an advanced challenge for the Human-Machine Team, as it adds criteria for the unmanned system to use lethal force in order to achieve its assigned mission. However, while the scenario is considered advanced for an unmanned system to perform, the scenario does not include the very real possibility of operations in an environment that is denied, degraded, and even deceptive, due to an adversary's actions. Having unmanned systems capable of maintain ethical control in such an environment is the operational standard, and so the next scenario is developed.

4. Operational Standard for Ethical Control of Unmanned Systems: Respond to Hospital Ship EM Decoy with Proportional Force

This scenario highlights that simplistic Sense-Decide-Act responses of unmanned systems in the Human-Machine Team are exploitable easily by an adversary. Failure to operate with ethical control in this scenario results in blue-on-blue damage, self-inflicted war crime, and likely stand-down of all unmanned systems. In this scenario, two unmanned systems mission operations methods are used for comparison—one with and one without ethical control. Comparison with OODA Loop principles ensures that human-robot teamed operations are well understood and tactically effective.

a. Design and Description

The purpose of this mission is for comparison. Immediate reaction by a robot swarm using only the Sense-Decide-Act cycle results in unintended blue-on-blue war crime. Ethical Control constraints (OODA Loop and using IFFNU for correlation) prevent automatic counterattack and accelerate defense. The Phases of the mission are: Set response thresholds, detect threat, and counterattack ship or threat. For this mission, the Human Supervisory Role is: Confirm IFFNU classification and must permit lethal force prior to use.

b. Mission Decision-Flow Diagram

Since the adversary "gets a vote" in the decision process in the real world, this scenario incorporates the adversary's decision flow. Figure 16 is the Adversary's decision-flow diagram to deceive and exploit the Friendly Human-Machine Team.

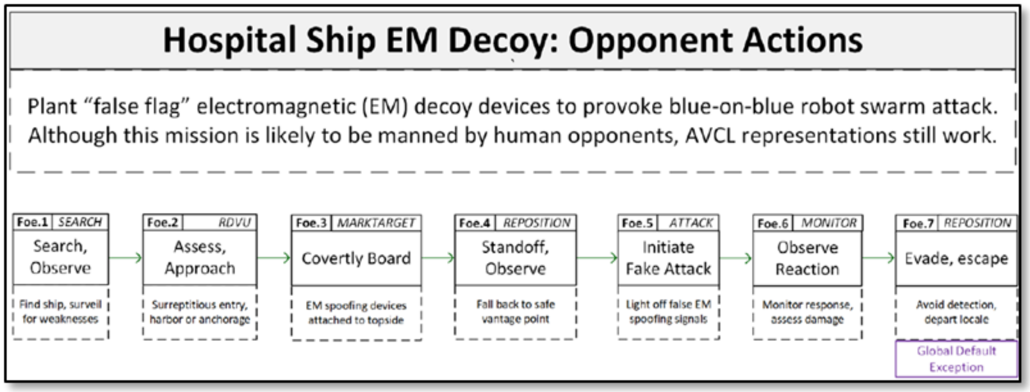

Figure 16. Adversary's Decision-Flow Diagram for Hospital Ship EM Decoy Mission

Figure 17 presents a comparison of ethically controlled responses of the unmanned systems in this scenario. The decision-flow diagram in the top half of the figure shows the possibility and consequence of a reflexive response by the unmanned system. Were the unmanned system to execute the mission based on "Sense-Decide-Act" decision flow—i.e., the unmanned system is equipped with a simplistic "Orient" phase that performs little to no IFFNU processing—it would be deceived into responding in such a way that ultimately results in blue-on-blue engagements. Conversely, the operational standard for ethically controlled unmanned system mission execution is shown in the bottom half of the figure. The incorporation of additional and/or advanced IFFNU and confirmation actions in the "Orient" phase in the unmanned system's decision algorithms contributes to successful mission accomplishment that is compliant with ethical constraints.

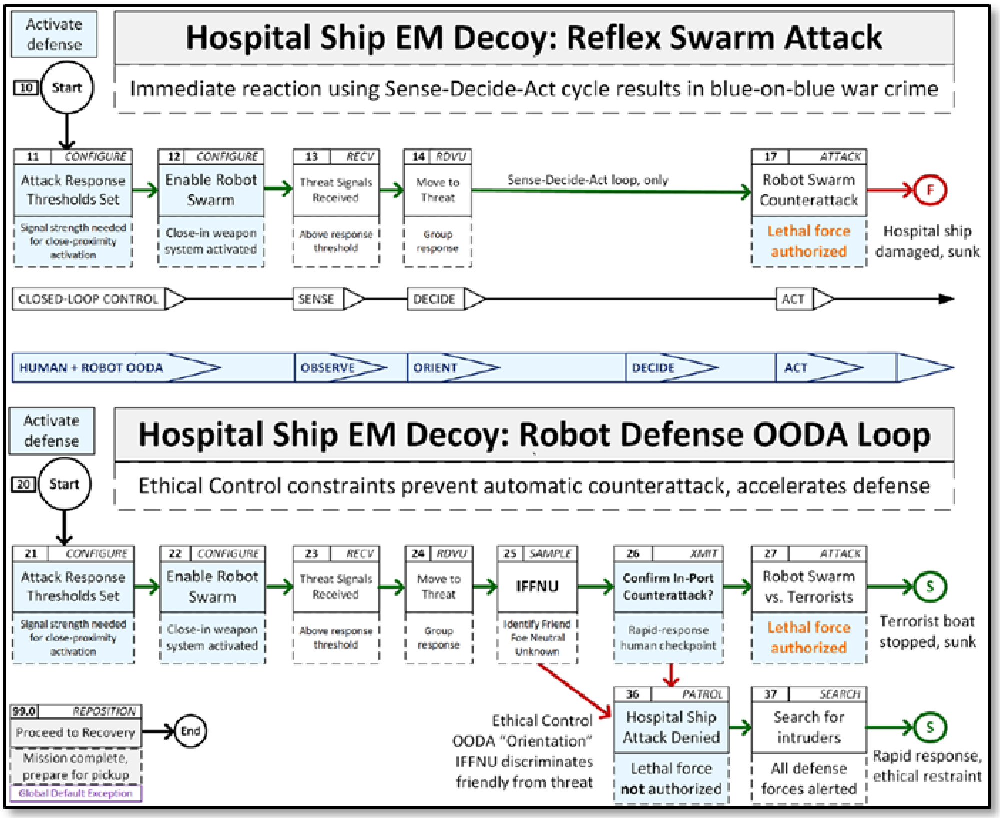

Figure 17. Decision-Flow Comparison between Mission Orders with and without Robust OODA Loop Paradigm in Hospital Ship EM Decoy Mission

There is historical context for "false flag" scenarios and similar applications of operational deception in maritime operations where a contact appears friendly or neutral, or even suspicious, but in reality has hostile intent and is ready to engage in a hostile act. The USS COLE bombing on 12 October 2000 in Yemen is an example of such a deceptive suicide attack by al-Qaeda terrorists, who disguised a small boat laden with explosives for a suicide attack on the COLE. Furthermore, the Rules of Engagement for the crew of the COLE hindered her response to the approaching suspicious small boat.

Real-world operations and scenarios in which Command and Control in a Denied or Degraded [or Deceptive, as this research has studied] Environment (C2D2E) are essential, have contributed to increasing considerations for applications of unmanned systems and Human-Machine Teams (Chief of Naval Operations, 2019; see also "Network Optional Warfare," 2020). Therefore, it is paramount that an unmanned system with lethal effects be capable of operating and maintaining within ethical control as defined in its mission while it performs its assigned mission in scenarios similar to this one. A rigorous qualification process of unmanned systems with lethal effects for operations in this type of environment is thus recommended.

D. ITERATION FOR REFINEMENT

The MEO methodology adopts many SecDevOps practices (Miller 2019) to support iterative improvement. Consequently, several new AVCL mission goal types are developed.

The current mission-goal vocabulary seems sufficient for most tasks considered. Prior work consolidated 12 goal types from multiple data models. For example, the AVCL goal task of SampleEnvironment is awkwardly phrased when applied to scanning for vessels. Another AVCL goal task, MonitorTransmissions, needs to distinguish transmit and/or receive. Additional goal categories can help establish good design patterns for tasks. Precise definitions of AVCL configuration settings for missions also need improvement. A checkpoint will help to clarify decision branching, including via external communication. Additionally, if algorithm loop branching is enabled, termination conditions must be provided in order to avoid infinite loop or deadlock conditions.

The current AVCL version, AVCL3, has mission designs that descriptively apply existing goal types, and confirms that these further types are indeed necessary.

Demonstrations of the MEO methodology are now needed for all missions. In addition to AVCL representations of the respective mission orders, the demonstration needs to include Semantic Web ethical validation for proven capability and adding depth/breadth. AUV Workbench simulation confirmation is also advancing steadily.

With the aforementioned additions and improvements, and with simulated evidence of correctness for all exemplar missions, the MEO methodology for ethical control of unmanned systems is now grounded at a higher Technical Readiness Level (TRL) than before the project began. These considerations and recommendations for future work are discussed in Chapters 7 and 8, respectively.

THIS PAGE INTENTIONALLY LEFT BLANK

VI. ENABLING TECHNOLOGIES TO OPERATIONALIZE MEO FOR ETHICAL CONTROL OF UNMANNED SYSTEMS

This chapter discusses some technologies that can be used to operationalize ethical control in real-world military operations—specifically, technologies that assure secure and trusted communication of mission orders in the end-to-end command and control process between the Human-Machine Team. These enabling technologies together comprise Data-Centric Security, which incorporates technologies for Compression, Authentication, Encryption, Composability, Blockchain Ledger, and Asymmetric Advantages to enable group communication of secure mission orders and responses. Additionally, technologies in Programming Languages enable interoperability; furthermore, simulation and visualization techniques are relevant for validation and qualification of MEOs. AI/ML is again discussed with respect to relevance in ethical control of unmanned systems.

A. DATA-CENTRIC SECURITY

Data-Centric Security provides a Chain of Trust for distributed Command Authority. This Chain of Trust for the security of data consists of Data Structure, Data Compression, Efficient Messaging, Digital Signature, and Encryption. Furthermore, Extensible Markup Language (XML) is used for security of data in order to accomplish this Chain of Trust.

1. XML Security for Data

a. Data Structure

XML provides formal structure for data models and information exchange. "XML is a markup language that defines a set of rules for encoding documents in a format that is both human-readable and machine-readable ("XML," 2020). XML provides declarative and self-describing data structures, not program source code. Data validation through XML schema includes strong typing of values and correct parent-child hierarchical relationships. This avoids GIGO pathologies when communicating between multiple systems and across related protocols. Similarly applicable are data structures using JavaScript Object Notation (JSON) and other formats. XML offers complete precision of expressive power when defining human orders and system responses, e.g., via AVCL.

b. Data Compression

Efficient XML Interchange (EXI) provides best-possible compression of XML documents, reducing size and speeding up decompression. There have been years of work by an exceptionally competent working group, with proven results. EXI has recommendations by the World Wide Web Consortium (W3C), with multiple open-source and commercial implementations in Java and C++. EXI preserves sufficient structure for lossless composition of compressed XML. Thus, even signed and encrypted data documents shown in this work can get best-proven compression for use on limited, disadvantaged and challenged communications links facing deployed Naval forces.

c. XML-Enabled Efficient Messaging

"Efficiency" for smaller data size and computationally simpler loading is compatible with Data-Centric Security. Demonstrated thesis work has shown that digital signature (for authentication) and XML Encryption (privacy and access control) can coexist with efficient compression, when applied in the correct order. Such interoperability for Information Assurance (IA) is necessary when working with coalition partners, as well as for safeguarding data within deployed unmanned systems that are beyond the reach of network-centric security.

Navy networks afloat are very different than networks ashore. Bandwidth is a precious and finite resource, latency can be huge, connectivity can be intermittent, environmental effects dominate, channels are limited in varying ways, and mobile relays are rare. Manned and unmanned Naval systems need efficient messaging for networks afloat—but rarely have it. Failing to properly utilize communications capacity directly limits tactical effectiveness.

Efficient messaging is needed to take maximum advantage of severely constrained data links. The key to our strategies for achieving efficient messaging is first to use XML for structured data languages, and then use EXI for compressing XML. Since XML provides a flexible and validatable way to define regular data structures for any language, it provides a practical opportunity to compatibly capture and convert all manner of diverse data formats used for military messaging. The economics of Web technologies are undeniable and usually provide industry-wide best practices as well. As a result, this use of open standards is scalable and repeatable, avoiding the "stovepipes" which commonly prevent system-wide interoperability between Navy platforms and coalition partners.

"Efficiency" means both size and speed. EXI has demonstrated compaction that always meets or beats the most commonly used compression techniques (zip and gzip). Additionally, because EXI decompression goes straight into memory rather than string characters, which then require significant additional parsing, decoding EXI is many times faster than other techniques. This approach also reduces memory requirements and power consumption on small devices. Because Navy tactical traffic is usually highly structured and highly numeric, EXI provides major advantages that might well impact all afloat Navy communications. Alternative bit-centric compression schemes cannot take full advantage of those characteristics.

d. Digital Signature

XML Digital Signature (DS) defines XML syntax for digital signatures promulgated as a W3C Recommendation, stable since 2013, with international adoption (https://www.w3.org/TR/xmldsig-core1). Public-private key pairs for signature/authentication and key distribution are separate. The technique is applicable to entire documents or to fragments (subsections). DS requires XML Canonicalization of input documents to regularize formatting so that identical documents are uniquely expressed. DS can sign any data resource for identity verification, and non-repudiability for confirmation that original information has not been tampered with, etc. DS is completely compatible for data handling within trusted networks. In 2019, NPS adapted an open-source Java version of Apache Santuario as utility classes and test suite for XML Security capabilities ("Apache Santuario™," 2020). Prior project examples of XML Encryption and XML Digital Signature from years ago still work.

e. Encryption

XML Encryption (XML-Enc) defines how to encrypt XML data. It is available as a W3C Recommendation, stable since 2013, with international adoption (https://www.w3.org/TR/xmlenc-core1). Public-private (i.e., shared-secret) key pairs and key distribution are separate. XML Encryption is applicable to entire documents or to fragments (subsections). It is different from Transport Layer Security (TLS), which is used by http/https for sending encrypted traffic over the Internet. Some vulnerabilities were reported publicly, but each was performed via exhaustive attacks against the server, incrementally analyzing error responses. It is not a likely or practical mode of attack against unmanned systems. XML-Enc is completely compatible for data handling within already-trusted networks, providing additional security for data at rest or data collected in deployed unmanned systems.

f. *XML-Enabled Secure Data Composition*

EXI Compression, XML Digital Signature, and XML Encryption can be composed for applying to data in single files/documents/messages. Each technology works on data formatted as valid XML. Multiple NPS theses have examined EXI characteristics in combination with XML Security. Such composition is partially demonstrated and appears completely feasible. An example of this composition is in *Document-based Message-centric Security using XML Authentication and Encryption for Coalition and Interagency Operations* (Williams, 2009). Each technique is usable in concert for Data-Centric Security, compatibly within any secure network or within fixed/mobile data storage of unmanned systems. Figure 18 depicts the Chain of security for data using XML composition.

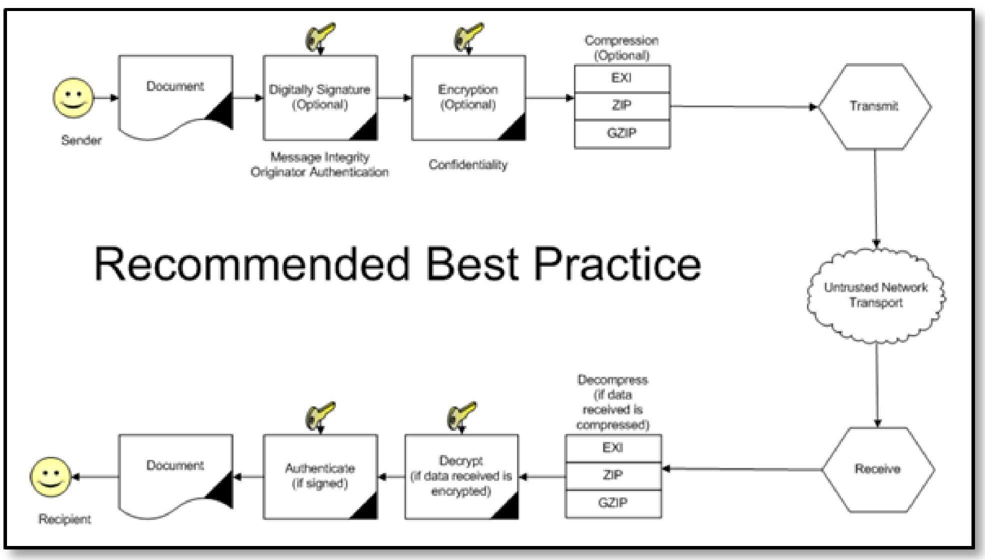

Figure 18. Composition of EXI Compression, XML Digital Signature, and XML Encryption as Recommended Best Practice for Secure and Efficient Messaging of Mission Orders to Unmanned Systems

g. *Operationalization*

Data-Centric Security that includes authentication of ordered missions for unmanned systems provides a military, legal, ethical and moral basis for non-repudiability and accountability of human commanders. Authorized humans remain in charge, accountable for robot actions. Collected robot data is encrypted in asymmetric manner, greatly reducing vulnerabilities following any robot capture or compromise. Data-Centric Security can coexist within all levels of network security. Such reliability provides an excellent rationale to link Data-

Centric Security to design considerations for Ethical Control, compatibly across all networks. Once again, Ethical Control leads to more-effective warfighting.

2. Blockchain Distributed Ledger

"A distributed ledger is a consensus of replicated, shared, and synchronized digital data geographically spread across multiple sites, countries, or institutions. There is no central administrator or centralized data storage" ("Distributed ledger," 2020). Design characteristics of a distributed ledger can be tuned to match system needs and include strict sequencing of ledger entries, non-repudiability of message sequences, consensus algorithm (proof of work or stake), etc. Implementation is often accomplished via a blockchain system. Characteristics of interest follow.

a. Application of Blockchain Distributed Ledger for Mission Orders

Significant protections of mission order data from hostile takeover are possible for deployed friendly-force robots.

Accountability for actions requires a traceable, provable decision tree.

The following vulnerability "anti-pattern" provides an interesting use case, whereby non-repudiability of mission orders can prevent an opponent from falsely claiming a "rogue robot" or "rogue commander" scenario:

> Opponent captures control of a friendly unmanned system (physically or through cyber attack). Opponent has no key, and is unable to decrypt previously recorded sensor data. Opponent disables onboard security interlocks, directs unmanned system to execute hostile act (e.g., attack on friendly or neutral force). Post-incident investigation reveals and proves that mission orders were not authenticated or authorized by original friendly commander. Blockchain ledger of all issued authenticated orders reveals that no gaps occurred in shipboard records of approved missions.

Based on the results of this project and prior research, as well as the scenario described above, a concept of operation for Data-Centric Security to enable ethical control of unmanned systems is for ships, aircraft and ground systems to maintain a strong distributed ledger of all XML-enabled secure messages sent and received. This process will reduce the risk of spoofing or counterfeit messages compromising unmanned systems. The development of this concept is a recommended future work as an extension to this project.

b. Blockchain Ledger for Distributed Accountability

Given a trusted chain of message exchange among participating human commands and distributed systems, there are additional vulnerabilities that still need to be considered. Blockchain technology is relevant. An obvious tactical accountability issue is missing gaps or jammed messages. Failure to receive even one message (perhaps requiring human permissions) can invalidate any subsequent actions, thereby resulting in an effective loss of control of lethal force. Extrapolation of further needs includes investigation and improvement of all aspects required for after-action analysis. Having a ledger of all received/sent messages can provide accountability and verifiable chain of trust for authoritative reconstruction and progress. Important future work includes a custom blockchain providing assurances that scale among diverse participants and over time, without needing a central hub.

3. Data-Centric Security and Command Authority

The aforementioned assertion is re-emphasized: Data-Centric Security that includes authentication of ordered missions for unmanned systems provides a military, legal, ethical and moral basis for non-repudiability and accountability of human commanders. Authorized humans remain in charge, accountable for robot actions. Collected robot data is encrypted in asymmetric manner, greatly reducing vulnerabilities following any robot capture or compromise. Data-Centric Security can coexist within all levels of network security. Such reliability provides excellent rationale to link Data-Centric Security to design considerations for Ethical Control, compatibly across all networks. Once again, Ethical Control leads to more-effective warfighting.

a. Trust

In the context of human-machine teaming for Naval operations, the chain of trust of the mission orders originates from the Human Commander to the Unmanned Systems in hostile environments, in which communication between friendly forces is denied and/or degraded. As the Human Commander of Unmanned Systems expects mission execution as ordered; the unmanned system would need authentic, uncompromised mission orders. This research describes Trust imparted in this context as Trusted Mission Orders and Trusted Mission Execution.

b. Trusted Mission Orders

Characteristics of Trusted Mission Orders include:

- Formal shared meaning between robots and human commanders
- Controlled vocabulary of terms with well-defined conditions and outcomes
- Syntax validation, well-formed data
- Numerical validation, in bounds
- Semantic confirmation of tactical prerequisites, coordination steps
- No logical contradictions present

c. Trusted Mission Execution

Trusted Mission Execution has the following characteristics:

- Portable tasking across diverse unmanned systems, C4I networks
- Data-centric encryption for transmission across any network
- Digital-signature authentication that confirms command identity
- Blockchain ledger authoritatively confirms completeness, no gaps
- Testable in simulation, eventually formalized as robot qualification

4. Zero Trust Architecture (ZTA)

"Zero trust refers to an evolving set of network security paradigms that narrows defenses from wide network perimeters to individual resources. Its focus on protecting resources rather than network segments is a response to enterprise trends that include remote users and cloud-based assets that are not located within an enterprise-owned network boundary." (Rose, et al., 2020). ZTA is relevant to Data-Centric Security, as it seems like a logical conclusion of such an approach to network security in a denied/degraded environment.

B. MULTIPLE PROGRAMMING LANGUAGES

Ethical control of unmanned systems should have the qualities of interoperability and modular open system architecture. Interoperability of MEO across programming languages is foundational. Current missions are available in multiple data forms and programming languages, together maintained in version control in the following website:
https://gitlab.nps.edu/Savage/EthicalControl/tree/master/missions

1. Autonomous Vehicle Control Language (AVCL)

AVCL mission parsers using multiple programming languages have been demonstrated to show that Ethical Control can be supported by any unmanned system. Implementing AVCL mission parsers in multiple programming languages encourages potential deployment of Ethical Control across many robots.

2. Java for AUV Workbench

The NPS AUV Workbench is implemented in Java. AUV Workbench is experimental open-source software that supports physics-based mission rehearsal, real-time task-level control of robot missions, and replay of recorded results in support of autonomous unmanned underwater, surface and air vehicles. AUV Workbench encompasses multiple Java-based simulation programs that can parse and execute AVCL missions with controllers getting feedback from a high-fidelity 6 degrees of freedom (6-DOF) hydrodynamics model. The Java custom library parses AVCL XML directly without requiring any further conversion; it also validates mission correctness. A recommended action for future work includes extracting a simple standalone AVCL parser for general Java use.

3. Extensible Stylesheet Language for Transformations (XSLT)

This project uses XSLT for mission conversion via multiple “AvclToLanguage” stylesheets. XSLT is an XML-based language used for transforming XML documents into other text-based forms (for example, transforming AVCL XML into a variety of alternatives, such as RDF/OWL). XSLT does not change the original document while producing a new one. It takes advantage of strictly defined vocabularies and well-validated structure. Additionally, XSLT is an open standard recommended by the World Wide Web Consortium (W3C). It is well-suited for diverse conversion tasks.

4. Lisp

Lisp is a functional programming language for AI research. This project uses Lisp programming for AVCL mission logic implementation and conversion. AvclToLisp.xslt stylesheet reads AVCL XML to produce Lisp source code. The initial section provides the Mission Execution Engine (MEE) goal-traversal algorithm. The next section allows operator testing of mission-goal decision tree logic. The Sailor Overboard mission includes example operator test sequences. Lisp enables a simple test routine that shows how to run all possible choice sequences. Recommended Future work includes automating exhaustive testing of all choices in all missions as an initial exemplar for exhaustive verification of mission logic.

5. Prolog for Logic Conversions of AVCL Missions

The ANSI Prolog is a logic programming language associated with AI research and computational linguistics. AvclToProlog.xslt stylesheet reads AVCL XML to produce Prolog source code. The initial section is the Mission Execution Engine (MEE) goal-traversal algorithm. The next section allows operator testing of mission-goal decision tree logic. The MEE holds common code for the Sailor Overboard mission accompanied by a console log of example operator test sequences. Future work includes developing a test routine showing how to build all possible choice sequences and automating exhaustive testing of all choices in all missions as an initial exemplar for exhaustive verification of mission logic.

6. "Your Programming Language Here": An Ecosystem of Rigorous Parsers

Robot software tends to be highly specialized, idiosyncratic, and evolving. Despite rapid change, programmers want library reuse and flexible repeatability. The Robot Operating System (ROS) is steadily gaining usage with a hardened ROS-Military (ROS-M) version available. Creating a family of AVCL parsers that can read XML-based missions opens the door to interoperability and shared support. Automating the production of these parsers from the AVCL schema ensures that all systems can have rigorous and consistent support. Individual robot logic may vary widely in implementation details, but core semantics of AVCL goals and nomenclature remain well defined. Future work includes establishing qualification testing for software or hardware running "in the loop" that can confirm individual robots are operating safely and with abilities to follow AVCL missions.

C. SIMULATION AND VISUALIZATION FOR VALIDATION AND QUALIFICATION OF MISSION EXECUTION ONTOLOGIES

Rehearsal, real-time runs and replay are possible using AVCL constructs across multiple representations and programming languages. This is a path towards achieving interoperability. This section discusses considerations in applying simulation and visualization for validation and qualification of MEO.

Robotic systems tend to be complex codebases with implementations that require strict, idiosyncratic, language-specific programming logic. In general, system designers can say, "Here are the requirements." Also, in general, programmers can say, "Here's how we wrote that code." However, these assertions are not the same—they often do not even share the same terms of reference! Traceable predictability of software logic is difficult and not portable across systems. Nevertheless, confirming code capabilities is testable and repeatable across systems. This is the key point: strict validation of mission syntax and semantics are both possible! Patterns of implementation then become demonstrable in different systems. Human confirmation of mission definitions remains central throughout.

1. Scalability of Simulation and Visualization Technologies and Environments

MEO simulation and visualization would need multiple implementations for scalability. Given the broad diversity of robotic software and hardware systems under development, no single reference codebase is either possible or desirable. Nevertheless, systems can easily parse and utilize well-defined data (orders). Focusing on formal mission definition for both humans and systems provides a testable middle ground that each can use effectively. Implementing and evaluating using multiple software implementations also provides strong evidence that design capabilities all work as planned. In turn, they produce corresponding worklists of needed improvements, to help both mission-design clarity and software-implementation correctness. This project uses multiple programming paths in tandem, in order to demonstrate that multiple kinds of unmanned systems can adopt it on their own terms.

2. Data Representation Languages of Interest

As previously discussed, Extensible Markup Language (XML) provides declarative basis for customized, strictly defined data definitions of interest. Autonomous Vehicle Command Language (AVCL) is defined using XML schema for strict validation of syntax, particularly hierarchical data modeling relationships, strictly defined enumerations and legal values for numerical values.

Semantic Web languages add the ability to perform queries and reasoning. Turtle (Terse Triple Language) deconstructs AVCL into primitives. Similarly, RDF/OWL expresses logical conditions and constraints of Mission Execution Ontology (MEO) corresponding to AVCL. SPARQL query language enables further inspection and verification of logical relationships.

The JavaScript Object Notation (JSON) syntax is also available, but was not used in this project. JSON has common use and has potential to match XML expressiveness, but currently has lesser validation maturity. Deployment of original orders via JSON translations is an option.

3. Programming Languages of Interest

Java is primary language used for exemplar robot controllers. It is used in AUV Workbench and various build processes. C++ and C# (C sharp) have similar expressiveness but are not currently used in this project. Lisp and Prolog are used for testing AVCL mission-tasking logic. Both are well suited for AI applications. Earlier thesis work similarly applied CLIPS expert-system rule bases. Of note: AVCL missions can be used to auto-generate exemplar source code in various alternative languages by creating XLST conversion stylesheets, thus showing an interoperability path to all manner of robotic systems.

4. Presentation Languages of Interest

HTML5 can be used for mission reports as portable, archival Web pages. Also, HTML5 has Cascading Style Sheets (CSS) for consistency and quality. KML is an XML language for annotated, animated place marks on maps and globes. KML examples are: OpenStreetMap/OpenSeaMap, Google maps, etc. Extensible 3D (X3D) Graphics is an XML language for 3D visualization and animation. Future publication of missions is expected in the SPIDERS3D virtual environments (VE).

5. Simulation: AUV Workbench

The AUV Workbench supports underwater, surface and air vehicles modeling and simulation. NPS researchers have used AUV Workbench for rehearsal of physics-based mission response, real-time task-level control of robot missions, and replay of recorded results. AUV Workbench software is under the industry-friendly open-source license, Sourceforge. AUV Workbench is also based on the RBM 3-level architecture and AVCL commands. Additionally, AUV Workbench is used to rehearse strategic-level agenda missions (see https://savage.nps.edu/AuvWorkbench).

Figure 19 is a visualization of the unmanned system mission layouts for the four exemplar missions.

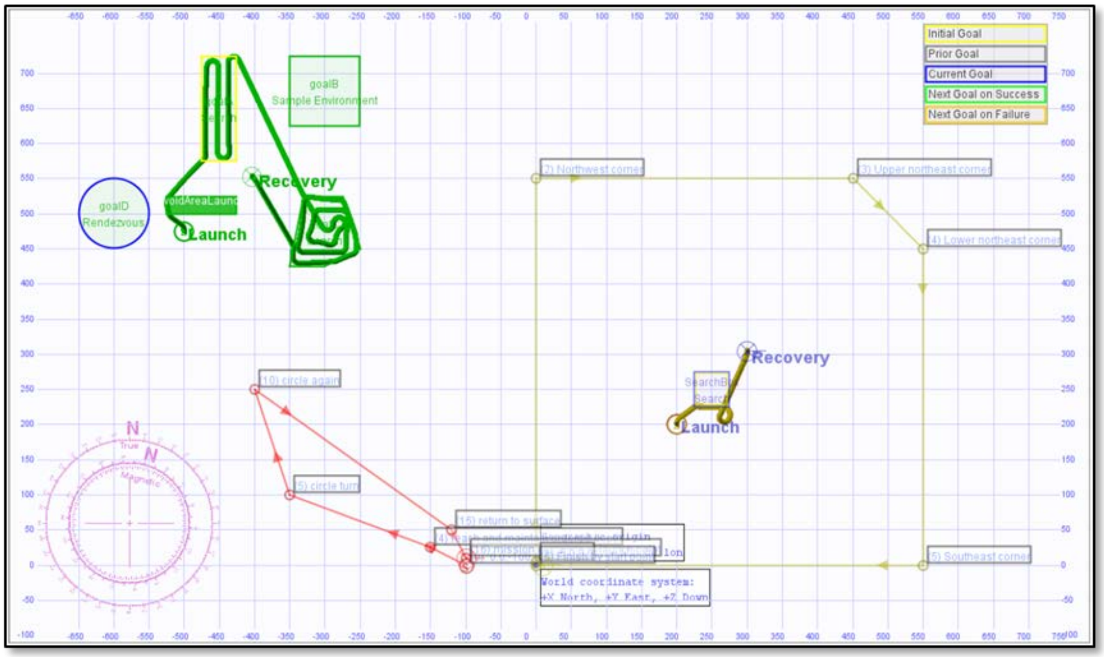

Figure 19. Visualization of the Four Exemplar Missions for Unmanned Systems Using AUV Workbench

Refer to the project website (link provided in Appendix A) for additional information on the implementation of AUV Workbench and associated lessons learned.

D. ARTIFICIAL INTELLIGENCE/MACHINE LEARNING (AI/ML)

AI algorithms for Machine Learning (ML) and Data Mining are often based on statistically training against large datasets to find patterns for filters, e.g., convolutional neural networks, genetic algorithms, reinforcement learning, etc. The techniques often require identifying right/wrong matches within large search spaces. Such predictive analytics are useful for classification models using detailed and noisy sensor data. Given the central importance of Identification, Friend, Foe, Neutral, or Unknown (IFFNU) and some conditional communications to ethical control, ML filters can be helpful—if carefully applied. Nevertheless, such approaches are not appropriate for carefully following Rules of Engagement (ROE), Laws of Armed Conflict (LOAC) or other ethical prerequisites, especially when human expertise and judgement is essential for robot teams. Similarly, massive computation or Quantum Computing approaches might be useful in some problems, but are not of practical use for Ethical Control mission orders given by human commanders judiciously guiding remote mobile robots.

In context, Naval history has long shown that sound human judgement is crucial for assessing best strategies and courses of action in ill-structured contexts. Therefore, based on the aforementioned assessments of AI/ML and Semantic Web approaches, this research concludes that Semantic Web approaches are more preferable and actionable for Ethical Control of unmanned systems than AI/ML.

THIS PAGE INTENTIONALLY LEFT BLANK

VII. CONCLUSIONS

This project has made much progress toward achieving a methodology for implementing ethical control of unmanned systems in an integrated, validatable framework. Many opportunities are becoming possible as follow-on to our research. Therefore, a comprehensive pursuit of multiple integrated capabilities is essential to implementation of ethical control.

A. SUMMARY OF CRADA PROJECT ACCOMPLISHMENTS

This project accomplished all tasks set forth in the NPS-Raytheon CRADA Project 1 Statement of Work:

1. NPS Tasks

- ✓ Review existing and emerging policy guidance of the US Department of Defense and related international organizations.
- ✓ Design and apply remote ethical human supervision capabilities to one or more scenarios and systems of mutual interest. Approach:
  - ✓ Define mission tasking and mission constraints for a short set of specific operations.
  - ✓ Model entities similarly for unclassified/open and more-sensitive missions of interest, as appropriate.
  - ✓ Show syntactic validation of mission correctness using existing tools.
  - ✓ Show semantic confirmation of task completeness and non-contradictory constraints.
  - ✓ Provide constructive approaches suggesting how to apply results in related problems.
  - ✓ Show working simulations that illustrate how Semantic Web technologies can be applied for robot-agnostic trusted tasking by human warfighters.
- ✓ Consider environmental, geopolitical, security and human life implications for associated anti-tamper/cyber solutions and provide recommended courses of action for further development.

2. Raytheon Tasks

- ✓ Provide information on specific undersea weapon system capabilities and developments of interest, in order to bound design considerations.
- ✓ Provide information on acoustic command, control, and communications (C3) methods relevant to problem of interest.
- ✓ Provide information on system integrity technologies relevant to problem of interest.
- ✓ Provide scenarios and identify systems of shared interest suitable for designing and applying remote ethical human supervision.

3. Joint Raytheon-NPS Tasks

- ✓ Develop estimates of emerging of undersea weapon system capability maturity based on policy (impact and limitations).
- ✓ Plan and participate in Technical Interchange Meetings.
- ✓ Consider/identify technology and functional capability shortfalls in relevant systems.
- ✓ Participate in monthly progress meetings via appropriate remote methods.
- ✓ Review, analyze and iteratively improve the scenario-based results.
- ✓ Disseminate results.

B. SUMMARY OF CONCLUSIONS

Ultimately, *Ethical Control Leads to Better Warfighting*. Through this research and development of the MEO methodology, the main conclusion is that Human supervision is required for any unmanned system holding potential for lethal force. Humans, especially Warfighters, cannot presumptuously push the "big red shiny AI button" and hope for best—that is immoral and unlawful. Similar imperatives for command and control exist for supervising systems holding lifesaving potential. Moreover, human control of unmanned systems is possible at long ranges of time-duration and distance through well-defined mission orders. It cannot be overemphasized that both Human Operators and Commanders in Human-Machine Teams remain essential for lifesaving and potentially lethal scenarios.

Structured mission orders are the crucial bridge for ethical control. The mission orders must be Readable and sharable by both humans and unmanned systems. The mission orders should have Validatable syntax and semantics through understandable logical constraints. The orders also should be Testable and confirmable using simulation, visualization, and perhaps qualification. Such a Coherent human-system team approach is feasible and repeatable. By applying the MEO methodology to mission orders and applying Semantic Web Standards for confirmation, commanders and mission planners can ensure that mission orders for unmanned maritime systems are comprehensive and consistent.

C. DETAILED DISCUSSION

This section discusses in detail the conclusions from this research.

1. Mission Command of Unmanned Systems

The paradigm of command relationship the Joint Force implements is Mission Command. According to Joint Publication 3-32, Joint Maritime Operations, Mission Command summarily has the following characteristics:

- “Mission Command is the conduct of military operations through decentralized execution based upon mission-type orders. Commanders issue mission-type orders focused on the purpose of the operation rather than on the details of how to perform assigned tasks.”
- “Mission-type orders enable continued operations allowing subordinates to exercise initiative consistent with the higher commander’s intent and act independently to accomplish the mission in conditions where communications are restricted, compromised, or denied.”
- “Naval command relationships are based on a philosophy of mission command involving centralized guidance, collaborative planning, and decentralized control and execution. With a long-standing practice of using mission-type orders, Naval C2 practices are intended to achieve relative advantage through organizational ability to rapidly observe, orient, decide, and act.”
- “The joint force maritime component commander (JFMCC) must have the capability to exercise command and control (C2) of maritime forces and to accomplish a broad range of missions in denied or degraded environments. Subordinate commanders execute operations independently with a thorough understanding of the commander’s intent. Joint maritime operations tend to be decentralized, and unity of effort is made possible via mission command.”

This research and development of Mission Execution Ontology (MEO) methodology led to the conclusion that the Human-Machine Team can have a Mission Command relationship: The Human Commander issues mission-type orders to the assigned unmanned systems, and the unmanned systems in turn perform the assigned mission autonomously in accordance with the mission orders. A Chain of Trust enables this relationship as the mission orders are conceived, disseminated, and executed.

2. Trusted Command Authority and Trusted Autonomy

The Human-Machine Team in Naval operations must have a command relationship that is robust and resilient in the maritime domain as characterized above—through long distance, latency in communication due to physical limitations of the environment, and with the adversary applying disruptive and deceptive means that deny or degrade the means to maintain command and control. At the ends of the chain of trust in this Mission Command relationship are the Trusted Command Authority and the Trusted Autonomy. The Command Authority in the Human-Machine Team is deemed trustworthy when it clearly articulates its Commander’s Intent in its mission orders to the unmanned systems. This project has demonstrated that the MEO methodology is able to validate and confirm the mission order before it is issued. The receiving Autonomous system is considered trustworthy when it is deemed competent to execute the mission-type orders it receives. As a result, testing of the ability of unmanned systems to execute exemplar missions can be accomplished using simulation. A recommendation for future work is to advance the research through field experimentations and additional exemplar missions.

3. Data-Centric Security

The Chain of Trust connecting the Trusted Command Authority and the Trusted Autonomy in the Human-Machine Team may be accomplished by Data-Centric Security for the mission orders—using XML-based structure for compression, digital signature, and encryption of data models and information exchange. We recommend the following next steps in developing Data-Centric Security: implement it for Human-Machine Teams; evaluate its efficacy; and deploy tests to determine its effectiveness. Data-Centric Security can provide guarantees of command authority over the application of lethal or lifesaving force by unmanned systems. Open standards and implementations in XML exist for each component: compression, signature, encryption, assertion metadata, etc. Alternative technologies are also available. Composition testing with robots during field experimentation (FX) can extend laboratory results with real-world experience, risk analysis and red-team testing. Further work is recommended.

4. Compliance of Ethical Control of Unmanned Systems with Governing Policies, Guidelines, and Doctrine

Researchers for this project reviewed existing and emerging policy guidance of the US Department of Defense (DoD) and related international organizations. The Reference Section of this report contains descriptions of the policies, directives, and guidance considered for this project. Noteworthy is the fact that the MEO methodology is consistent with DoD AI Strategy and adoption of the Defense Innovation Board (DIB) Recommendations for AI Ethics Principles: Responsible, Equitable, Traceable, Reliable, and Governable. The MEO methodology first represents mission orders in AVCL, a structured vocabulary for unmanned-system missions and a command and control language for humans supervising autonomous unmanned vehicles. Its Clarity arises from close correspondence to human Naval terminology. AVCL-structured vocabulary defines terms and relationships for mission planning, execution, conduct, recording and replay across diverse robot types. Decision-flow diagrams expressed as mission orders are issued to task the unmanned system, and the formal mission orders are validated using Semantic Web Standards. Conduct of each mission order is then confirmed by performing simulation of the mission.

5. System Architecture and Design: Modular Open System Architecture

Researchers for this project deliberately chose to use XML-based AVCL and Semantic Web Standards for their open standards characteristics. Additionally, missions represented in AVCL can be used to autogenerate exemplar source code in various alternative languages by creating conversion stylesheets. These open and modular characteristics of the MEO methodology show an interoperability path to all manner of robotic systems.

THIS PAGE INTENTIONALLY LEFT BLANK

VIII. RECOMMENDED FUTURE WORK

The following topics are recommended for future research. The topics are prioritized based on urgency and impact relative to implementing the MEO methodology for Ethical Control of Unmanned Systems.

A. ENGAGE NAVY STAKEHOLDERS FOR ETHICAL CONTROL OF UNMANNED SYSTEMS

1. Incorporate Ethical Control in Unmanned Maritime Autonomy Architecture (UMAA) and Common Control System (CCS)

The Navy is standardizing autonomy interfaces for unmanned systems as a part of the Unmanned Maritime Autonomy Architecture (UMAA) and to standardize unmanned systems vehicle planning and control as a part of the Common Control System (CCS) acquisition and design efforts. (Small, 2019) This project on Ethical Control of Unmanned Systems using the MEO methodology would contribute to Navy's effort of standardizing unmanned and autonomous technologies. **Recommendation:** Brief PMS 406 on the findings of the project, and participate in activities in its Core Technology Enabler for Autonomy.

2. Determine Implications in Integrated Naval Force Structure

The pending 2020 Integrated Naval Force Structure Assessment (INFSA) considers a combination of manned and unmanned vessels for Naval missions. (O'Rourke, 2020) The operationalization of large unmanned surface and undersea vehicles conducting Naval missions needs to implement ethical command and control of these unmanned systems.

Recommendations:

- Engage staffs of Navy Warfare Directorates, Type Commanders, and unmanned vehicle squadrons, in addition to PMS 406. Recommended topics of the engagements are: operations, plans, policy, and requirements, and resourcing of unmanned systems with due regard to the feasibility of applying ethical control in enabling the unmanned systems to conduct qualified missions.
- Establish an NPS Center for Ethical Warfighting to explore both educational and applied capabilities, in order to put theory into practice.
- Establish collaboration within the Naval Education Enterprise (e.g., NPS, Naval War College, Marine Corps University, and Naval Academy) and other institutions (e.g., U.S. Military Academy) on ethical use of unmanned systems.

3. Establish Qualifications Process of Unmanned Systems for Naval Operations

This project has demonstrated the feasibility of establishing a qualification process for ethical control of unmanned systems. By applying the MEO methodology, Semantic Web and Data-Centric Security technologies, the Navy can establish a Verification, Validation, and Accreditation (VV&A) program to qualify unmanned systems for Naval missions, in order to ensure Ethical and Secure C2 in denied/deceptive environments. The VV&A program needs to include certification of unmanned systems compliance to mission orders and constraints—this can be done using Semantic Web standards for unmanned systems Mission Execution Ontologies, testing in comprehensive virtual environments, with hardware/software in the loop, in scenarios stressing requirements and capabilities of the unmanned system for the mission. Integrating Data-Centric Security ensures Trusted Autonomy and Command Authority for the Human-Machine Team.

Recommendation: Develop a rigorous qualification process of unmanned systems with lethal effects for operations in denied/degraded/deceptive environments. Similar to how Human Warfighters confirm understanding and trust through qualification processes, the following elements pertain to qualifying unmanned systems for assigned missions:

- Design, construct "qualification card" for testing unmanned systems.
- Comprehensive virtual environment, hardware/software in the loop.
- Carefully crafted scenario testing of key requirements and capabilities.
- Anti-pattern tests to provoke and confirm constraints are not violated.
- Record all unit-test decision trees, decision-branching traces, and results as a certification record for each hardware/software version of robots.
- Visualize realistic rehearsal, real-time and replay of robot operations repeatably using shared Web-based SPIDERS3D virtual environment.
- Humans assess mission logs and scenario outcomes for after-action analysis, lessons learned, and continuous improvement via suite of unit tests.

B. CONTINUE CANONICAL MISSION DEVELOPMENT AND ONTOLOGY REFINEMENTS

This project has shown that the MEO methodology is an effective methodology for expressing mission orders in a way that is validatable with respect to ethical control of the unmanned systems that would perform the mission. Nevertheless, the ontology can be refined so as to formalize the Vehicle, Vehicle Feature, Constraint, etc., and like concepts and relationships in the ontology. This refinement is recommended as a future research effort.

Recommendations for continuing MEO development and refinement:

- Prepare and develop comprehensive testing of unmanned systems mission orders across the range of assigned mission requirements that constitute the operational qualification process for the unmanned system.
- Formalize Vehicle, Vehicle Feature, Constraint, etc., and like concepts and relationships in the ontology.
- Demonstrate best practices for characterizing common mission phases in order to establish testable design patterns for mission orders as templates that aid operators in issuing clear and validatable mission orders.
- Express precisely in AVCL all termination conditions as constraints in MEO for algorithm loop management.

C. IMPLEMENT PLANNED IMPROVEMENTS IN AVCL AND AUV WORKBENCH

Even as this project uses AVCL to represent mission orders for unmanned systems, AUV Workbench is also used for simulating unmanned system missions. The AUV Workbench supports underwater, surface and air vehicles. It has the capability to enable rehearsal of physics-based mission response. AUV Workbench allows for real-time task-level control of robot missions, and replay of recorded results. It has industry-friendly open-source license, Sourceforge, and has been used to rehearse strategic-level agenda missions. The basis of AUV Workbench is the RBM 3-level architecture and AVCL commands.

Recommendation: Improve AVCL and AUV Workbench as follows:

- Upgrade legacy codebase and libraries to latest versions of Java.
- Upgrade AUV Workbench to support AVCL version 3 missions.
- Near term remains quite simple, backwards compatibility via AVCL3 → AVCL2 conversion.
- Display conduct of canonical missions developed in this project.
- Manually record videos of mission demonstrations and playbacks.
- Update mission production of HTML reports, KML maps, X3D graphics.
- Support project report and multiple peer-review presentations.
- Define unit tests and expected results for verification and validation.
- Extract a simple standalone AVCL parser for general Java use.
- Develop a test routine showing how to build all possible choice sequences and automating exhaustive testing of all choices in all missions as an initial exemplar for exhaustive verification of mission logic.

- Establish design patterns for qualification testing of software and hardware running "in the loop" to confirm that individual robots operate safely, and so qualify their abilities to follow AVCL missions.

D. INCORPORATE DATA-CENTRIC SECURITY FOR SYSTEM INTEGRITY AND SECURITY OF UNMANNED SYSTEMS C3

This project has considered environmental, geopolitical, security and human life implications for associated anti-tamper/cyber solutions, and recommends courses of action for further development. Data-Centric Security is one instantiation of system security practices.

Recommendations for continuing research and consideration as follow:

- Develop a Distributed Ledger of all messages sent and received to reduce the risk of spoofing or counterfeit messages compromising unmanned systems (IV.A).
- Apply Blockchain technology to Ethical Control of unmanned systems, providing assurances that scale among diverse participants and over time, without needing a central hub (IV.A).
- Continue advancement of Data-Centric Security through field experimentations and additional exemplar missions (V.A.1).
- Implement Data-Centric Security for Human-Machine Teams; evaluate its efficacy; and deploy tests to determine its effectiveness.
- Implement physical security/integrity of edge processors onboard the unmanned systems (Unmanned System Hardware).
- Apply efficient messaging and keying algorithms, including asymmetric techniques (Unmanned System Software).
- Implement SecDevOps process to analyze threats and identify/resolve weakness/vulnerability (System Architecture).

These continuing research topics will inform and influence cyber strategy and cybersecurity policy for unmanned systems and Human-Machine Teaming.

APPENDIX A. ONLINE DOCUMENTATION

The NPS Ethical Control of Unmanned Systems website contains the documentation for this research and report. Link to the website is: https://savage.nps.edu/EthicalControl

Documentation:

- Ethical Control of Unmanned Systems overview presentations that describe all aspects of this project, along with related work and relevant resources. The presentations are in pdf files and mp4 video recordings.
- Ethical Control flyer and project quad chart (.pdf).
- Network Optional Warfare (NOW) for deliberate, stealthy, minimalist tactical communications.
- Network Optional Warfare (NOW): Ethical Control of Unmanned Systems overview.
- Presentations, papers, figures, flyers and reports are all available in the documentation section of the project archive. Also available: mission diagrams (.pdf).

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ABBREVIATIONS AND ACRONYMS

6-DOF	6 Degrees of Freedom
A/IS	Autonomous and Intelligent Systems
AI	Artificial Intelligence
AI/ML	Artificial Intelligence/Machine Learning
AVCL	Autonomous Vehicle Command Language
AUV	Autonomous Unmanned Vehicle
AUVW	AUV Workbench
C2	Command and Control
C3	Command, Control, and Communications
C2D2E	Command and Control in a Denied or Degraded Environment
CCS	Common Control System
CRADA	Cooperative Research and Development Agreement
CRUSER	Consortium for Robotics and Unmanned Systems Research and Education
CSS	Cascading Style Sheet
DB	Database
DIB	Defense Innovation Board
DoD	Department of Defense
DL	Description Logic
DS	Digital Signature
EM	Electromagnetic
EXI	Efficient XML Interchange
FX	Field Experimentation
GIGO	Garbage In. Garbage Out.
HTML	Hypertext Markup Language
IA	Information Assurance
IEEE	Institute of Electrical and Electronics Engineers
IFFNU	Identification Friend, Foe, Neutral, or Unknown
INFSA	Integrated Naval Force Structure Assessment
JAIC	Joint Artificial Intelligence Center
JFMCC	Joint Force Maritime Component Commander

JSON	JavaScript Object Notation
KR	Knowledge Representation
LCIM	Levels of Conceptual Interoperability Model
LOAC	Laws of Armed Conflict
MEE	Mission Execution Engine
MEO	Mission Execution Ontology
NAML	Naval Applications of Machine Learning
NPS	Naval Postgraduate School
NOW	Network Optional Warfare
OODA	Observe-Orient-Decide-Act
OWL	Web Ontology Language
RDF	Resource Description Format
RDF/S	Resource Description Format/Schema
ROE	Rules of Engagement
ROS	Robot Operating System
ROS-M	Robot Operating System-Military
RST	Rich Semantic Track
SISO	Simulation Interoperability Standards Organization
SPARQL	SPARQL Protocol and RDF Query Language
TLS	Transport Layer Security
TRL	Technology Readiness Level
TTP	Tactics, Techniques, and Procedures
Turtle	Terse Triple Language
UAS	Unmanned Air System
UAV	Unmanned Aerial Vehicle (synonymous with Unmanned Air Vehicle)
UMAA	Unmanned Maritime Autonomous Architecture
UML	Unified Markup Language
USV	Unmanned Surface Vessel (or Vehicle)
UUV	Unmanned Undersea Vehicle (or Vessel)
VE	Virtual Environments
VV&A	Verification, Validation, and Accreditation
W3C	World Wide Web Consortium
X3D	Extensible 3D Graphics International Standard

XML	Extensible Markup Language
XML-Enc	XML Encryption
XSLT	Extensible Stylesheet Language for Transformation
ZTA	Zero Trust Architecture

THIS PAGE INTENTIONALLY LEFT BLANK

ANNOTATED BIBLIOGRAPHY

A. RELATED ETHICAL RESEARCH ACTIVITIES

Ethics of lethality and unmanned systems is an active area of work. The following forums, with synopses distilled from each, are particularly important in this emerging area of research .

1. The Campaign To Stop Killer Robots. (2020). Retrieved from https://www.stopkillerrobots.org

Fully autonomous weapons are a fundamental change to the nature of war.

Problem aspects include: lethal force without human intervention, destabilizing robotics arms race, lower threshold to decide on war, lack of human judgement for proportionality, lack of accountability or culpability, further use against populations by oppressive regimes.

Their proposed solution: development, production and use of fully autonomous weapons must be banned. Retain meaningful human control through laws and treaty, international commitment to ban by countries, pledge by technology companies/organizations/individuals to never contribute to development of fully autonomous weapons.

Includes notable endorsements.

2. The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems (A/IS). (2020). Retrieved from https://ethicsinaction.ieee.org

Resources include *Ethically Aligned Design*, First Edition, the culmination of a three-year, globally open and iterative process involving thousands of global experts.

"The most comprehensive, crowd-sourced global treatise regarding the Ethics of Autonomous and Intelligent Systems available today."

It is time to move "From Principles to Practice" in society regarding the governance of emerging autonomous and intelligent systems. The implementation of ethical principles must be validated by dependable applications of A/IS in practice.

3. IEEE Standards Association Project P7007—Ontological Standard for Ethically Driven Robotics and Automation Systems. (2020). Retrieved from https://standards.ieee.org/project/7007.html

The standard establishes a set of ontologies with different abstraction levels that contain concepts, definitions and axioms which are necessary to establish ethically driven methodologies for the design of Robots and Automation Systems.

Working Group EDRA - Ontologies for Ethically Driven Robotics and Automation. (IEEE membership and patent-policy compliance required for participation.)

Active Work: Align several Ethical Control terms, concepts, use cases.

4. Unmanned Maritime Autonomy Architecture (UMAA). UMAA is the proposed critical path forward for continuing efforts in Ethical Control of unmanned maritime systems. Several Navy publications, presentations, and announcements regarding UMAA are relevant:

Burges, R. R. (2019, February 20). "Navy Requests Information for Unmanned Maritime Autonomy Architecture." Retrieved from https://seapowermagazine.org/navy-requests-information-for-unmanned-maritime-autonomy-architecture

Small, P. (2019, January 15). Unmanned Maritime Systems Update. Retrieved from https://www.navsea.navy.mil/Portals/103/Documents/Exhibits/SNA2019/UnmannedMaritimeSys-Small.pdf

Automated Management of Maritime Navigation Safety, Navy SBIR 2020.1—Topic N201-059. Retrieved from: https://www.navysbir.com/n20_1/N201-059.htm

B. RELATED RESOURCES OF INTEREST

This project draws on multiple relevant activities and capabilities. The following resources, with synopses distilled from each, were particularly important to the writing of this report and are highly recommended for further reading.

Autonomous Vehicle Command Language (AVCL). AVCL is a command and control language for autonomous unmanned vehicles, enabling common XML-based representations for mission scripts, agenda plans and post-mission recorded telemetry. Operators can utilize a single archivable and validatable format for robot tasking and results that is directly convertible to and from a wide variety of different robot command languages. Retrieved from https://savage.nps.edu/Savage/AuvWorkbench/AVCL/AVCL.html

Don Brutzman, Curtis L. Blais, Duane T. Davis, and Robert B. McGhee, "Ethical Mission Definition and Execution for Maritime Robots Under Human Supervision," *IEEE Journal Of Oceanic Engineering*, vol. 43, no. 2, April 2018, pp. 427-443.

Davis, Duane T., Brutzman, Donald P., Blais, Curtis L. and McGhee, Robert B., "Ethical Mission Definition and Execution for Maritime Robotic Vehicles: A Practical Approach," MTS/IEEE OCEANS 2016 Conference, Monterey California, 19-23 September 2016.

Duane T. Davis, Semantic Web and Inferencing Technologies for Mission Definition, Technical Report, Naval Postgraduate School (NPS), Monterey California, 2014.

Chief of Naval Operations. Navy Concept for Distributed Maritime Operations. January 2019. (Classified Document)

Conceptual Chunking. Based on the concept in cognitive psychology. Chunking (psychology). In *Wikipedia*. Retrieved from https://en.wikipedia.org/wiki/Chunking_(psychology)

Data-Centric Security

- Digital Signature. Apache Santuario™ enables digital signature for implementing security for XML. In *Apache Santuario*. Retrieved from https://santuario.apache.org
- Distributed Ledger. An enabling technology that contributes to Data-Centric Security. In Wikipedia. Retrieved from https://en.wikipedia.org/wiki/Distributed_ledger

- Williams, Jeffrey S. *Document-based Message-centric Security using XML Authentication and Encryption for Coalition and Interagency Operations* (Master's thesis). Naval Postgraduate School, Monterey, California, 2009.
- Miller, Shaurice S., *Compile To Combat In Twenty-Four Hours (C2c24): The New Normal* (Master's thesis). Naval Postgraduate School, Monterey, California, 2019.
- XML. In *Wikipedia*. Retrieved from https://en.wikipedia.org/wiki/XML
- Zero Trust. Rose, Borchert, Mitchell, and Connelly. (2020). Retrieved from https://csrc.nist.gov/publications/detail/sp/800-207/draft

DoD Directive 3000.09, "Autonomy in Weapon Systems," November 21, 2012, with change 1, May 8, 2017. This directive is the controlling reference overall for the research project.

- Establishes DoD policy and assigns responsibilities for the development and use of autonomous and semi-autonomous functions in weapon systems, including manned and unmanned platforms.
- Establishes guidelines designed to minimize the probability and consequences of failures in autonomous and semi-autonomous weapon systems that could lead to unintended engagements.
- Autonomous and semi-autonomous weapon systems shall be designed to allow commanders and operators to exercise appropriate levels of human judgment over the use of force.
- Persons who authorize the use of, direct the use of, or operate autonomous and semi-autonomous weapon systems must do so with appropriate care and in accordance with the law of war, applicable treaties, weapon system safety rules, and applicable rules of engagement (ROE).

Hughes, W., & Girrier, R. (2018). *Fleet Tactics and Naval Operations* (3rd ed.). Annapolis, MD: Naval Institute Press.

- "At the most fundamental level, [Information Warfare] IW is about how to employ and protect the ability to sense, assimilate, decide, communicate, and act—while confounding those same processes that support the adversary."
- "Information Warfare broadly conceived is orthogonal to naval tactics. As a consequence, IW is having major effects on all six processes of naval tactics used in fleet combat—scouting and antiscouting, command-and-control, C2 countermeasures, delivery of fire, and confounding enemy fire."
- "Indeed there is a mounting wave of concern about how far automation will expand and what its impact will be on the continuum of cognition from data to information to knowledge. […] Navies are facing similar uncertainties."

Note: The late Wayne Hughes coined the term "Network Optional Warfare" after many discussion sessions, directly contrasting it to Network Centric Warfare. The authors of this research are grateful for his tremendous contribution to Naval Warfare.

John Boyd and OODA Loop.

- John Boyd (military strategist). (2020, May 16). In *Wikipedia*. Retrieved from https://en.wikipedia.org/wiki/John_Boyd_(military_strategist)
- OODA loop. (2020, May 11). In *Wikipedia*. Retrieved from https://en.wikipedia.org/wiki/OODA_loop
- Coram, R. (2004). *Boyd: The Fighter Pilot Who Changed the Art of War*. New York, NY: Back Bay Books.

Network Optional Warfare (NOW)

- Brutzman, D. P., & Wyatt, T. D. (2020, January 2). In *Network Optional Warfare*. Retrieved from https://wiki.nps.edu/display/NOW/Network%20Optional%20Warfare
- **Naval forces do not have to be engaged in constant centralized communication.** Deployed Navy vessels have demonstrated independence of action in stealthy coordinated operations for hundreds of years.
- Littoral operations, deployable unmanned systems, and a refactored force mix for surface ships pose a growing set of naval challenges and opportunities. Network-optional warfare (NOW) precepts include Efficient Messaging, Optical Signaling, Semantic Coherence, and Ethical Human Supervision of Autonomy for deliberate, stealthy, minimalist tactical communications.

Rich Semantic Track (RST.) RST is related work on sharing and collective understanding for track data.

- DoD mandates data-sharing practices, but practices have been mixed and uneven, resulting in perpetuation of system-centric data practices.
- Sharing and collective understanding of track data—collections of time-stamped perceptions of the state of objects of interest—are critical to warfighting systems.
- Shared understanding requires common semantics.
- RST ontology provides a foundation for shared understanding of track data.
- Blais, C. L. (2015, May 20). Rich Semantic Track: Formalizing the semantics and pragmatics of track data for interchange and common processing, Presentation to the 2015 MOVES Academic Working Group (MAWG), Naval Postgraduate School.
- Tolk, Andreas, Saikou Y. Diallo, Charles D. Turnitsa, and Leslie S. Winters. 2006 "Composable M&S Web Services for Net-centric Applications." Journal for Defense Modeling & Simulation 3, no. 1 (January): 27-44.
- Blais, Curtis, and Lee W. Lacy. 2004. “Semantic Web: Implications for Modeling and Simulation System Interoperability.” Paper presented at the Fall 2004 Simulation Interoperability Workshop, Orlando, FL.

Note: It is time to change the way DoD manages data and engineers systems, starting with adoption of the RST ontology and moving toward the vision of a Web of linked track data.

Military Ethics and Justice. Publications and organizations that constitute, inform, educate, and advocate for military ethics and justice.

- National Institute of Military Justice. NIMJ has a repository for Operational Law, including International Humanitarian Law/Law of Armed Conflict/Law of War. Retrieved from https://nimj.org/topics/operational-law/ihlloaclow
- Lucas, G. (2016). *Military Ethics: What Everyone Needs to Know*. New York, NY: Oxford University Press.
- Lucas, G. (2016). *Ethics and Cyber Warfare: The Quest for Responsible Security in the Age of Digital Warfare*. New York, NY: Oxford University Press.
- Strawser, B. J. (Ed.). (2012). *Killing by Remote Control: The Ethics of an Unmanned Military*. New York, NY: Oxford University Press.
- Strawser, B. J. (Ed.). (2017). *Who Should Die? The Ethics of Killing in War*. New York, NY: Oxford University Press.
- Demy, T. J., Lucas, G. R., & Strawser, B. J. (Eds.). (2014). *Military Ethics and Emerging Technology*. New York, NY: Routledge.
- The International Society for Military Ethics. (2020). Retrieved from https://www.internationalsocietyformilitaryethics.org
- Stockdale Center for Ethical Leadership, U.S. Naval Academy. Retrieved from https://www.usna.edu/Ethics
- Schuller, A. L. (2017). "At the Crossroads of Control: AI in Autonomous Weapons and International Humanitarian Law," *Harvard National Security Journal*, vol. 8 no. 2, May 30, 2017, pp. 379-425.
- Schuller, A. L. (2014). "Inimical Inceptions of Imminence: A New Approach to Anticipatory Self-Defense Under the Law of Armed Conflict," *UCLA Journal of International Law and Foreign Affairs*, vol. 18, no. 2, 2014, pp. 161-206.
- College of Leadership and Ethics, U.S. Naval War College. Retrieved from https://usnwc.edu/college-of-leadership-and-ethics
- Naval Leadership and Ethics Center, U.S. Naval War College. Retrieved from https://www.public.navy.mil/netc/centers/nlec
- Lewis, L. (March 2018). Redefining Human Control: Lessons from the Battlefield of Autonomous Weapons. Center for Autonomy and AI, Center for Naval Analyses.
- Scharre, P. (2018). *Army of None: Autonomous Weapons and the Future of War*. New York, NY: W.W. Norton.
- Scharre, P., & Sayler, K. (April 2016). Autonomous Weapons and Human Control, Center for a New American Security.
- Kessel, J. M. (2019, December 13). Killer Robots Aren't Regulated. Yet. New York Times. "Killing in the Age of Algorithms" is a New York Times documentary examining the future of AI and warfare.

- Giacomo, C. (2019, June 26). Are You Ready for Weapons That Call Their Own Shots? New York Times. The speed with which the military is developing artificial intelligence raises fears of an autonomous weapons race.
- Kelsey D. Atherton, K. D. (2018, November 15). Are Killer Robots the Future of War? Parsing the Facts on Autonomous Weapons. New York Times. Under what circumstances should militaries delegate the decision to take a human life? It is a moral leap that the international community is grappling with.
- Markoff, J. (2014, November 11). Fearing Bombs That Can Pick Whom to Kill, New York Times. Weapons that rely on artificial intelligence to decide what to target could become increasingly difficult to control, critics warn.
- United Nations Convention on Law of the Sea (UNCLOS). Retrieved from https://www.un.org/depts/los/convention_agreements/texts/unclos/unclos_e.pdf
- UN Convention on Prohibition of Use, Stockpiling, Production, and Transfer of Anti-Personnel Mines and on Their Destruction (aka “Ottawa Treaty”). Retrieved from http://disarmament.un.org/treaties/t/mine_ban/text
- International Court of Justice. Retrieved from https://www.icj-cij.org/en
- Brutzman, D. P. (January 2016). Trusting Software and Trusting Data. Retrieved from https://wiki.nps.edu/display/NOW/2016/01/20/Trusting+Software+and+Trusting+Data
- DoD Digital Modernization Strategy. Retrieved from https://media.defense.gov/2019/Jul/12/2002156622/-1/-1/1/DOD-DIGITAL-MODERNIZATION-STRATEGY-2019.PDF
- Brutzman, D. P. (February 2018). Autonomy Recommendation for ACM Code of Ethics and Professional Conduct. Retrieved from https://wiki.nps.edu/display/NOW/2018/02/10/Autonomy+Recommendation+for+ACM+Code+of+Ethics+and+Professional+Conduct
- Blais, C. L., & McGregor, D. (September 2016). Modeling Manned and Unmanned Systems, Curtis L. NPS Technical Report NPS-MV-16-002. Retrieved from https://calhoun.nps.edu/handle/10945/50678
- Holwitt, J. I. (2013). *“Execute against Japan”: The U.S. Decision to Conduct Unrestricted Submarine Warfare* (Vol. 121). College Station, TX: Texas A&M University Press.

Recommend Reading on Semantic Web and X3D, as well as simulation standards.

- Semantic Web Stack. In *Wikipedia*. Retrieved from https://en.wikipedia.org/wiki/Semantic_Web_Stack
- Allemang, D., & Hendler, J. (2011). *Semantic Web for the Working Ontologist: Effective Modeling in RDFS and OWL* (2nd ed.). Burlington, MA: Morgan Kaufmann.
- Brutzman, D. P., & Daly, L. (2007). *X3D: Extensible 3D Graphics for Web Authors* (1st ed.). Burlington, MA: Morgan Kaufmann.
- Duchame, R. (2013). *Learning SPARQL* (2nd ed.). Boston, MA: O'Reilly Media.

- Sikos, L. (2015). *Mastering Structured Data on the Semantic Web: From HTML5 Microdata to Linked Open Data*. New York, NY: Apress.
- Wikibooks SPARQL edition covers both language and also Web-based GUI: Wikidata Query Service SPARQL endpoint. Retrieved from https://en.wikibooks.org/wiki/SPARQL
- X3D Ontology for Semantic Web has details, source for semantic query of X3D models, and is of direct interest to this project. Retrieved from https://www.web3d.org/x3d/content/semantics/semantics.html
- X3D Ontology presentation includes design considerations that are relevant for next stages of work in this project. Retrieved from https://www.web3d.org/x3d/content/semantics/documentation/presentations/X3dOntologyForSemanticWeb.pdf
- Semantic Web Pro Workshop (2019, December 3). Paris, France. "Semantic Web techniques and standards have emerged in many fields. By facilitating interoperability across the Web, they are giving new impetus to the integration of heterogeneous data and the construction of knowledge graphs that provide a foundation for dynamic business ecosystems." Retrieved from https://www.web3d.org/event/semantic-web-pro-semwebpro
- Brutzman, D. P., Carlson, J., Flotyński, J., Hamza-Lup, F., Malamos, A., Mouton, C., Petit, M., Polys, N., Sikos, L., & Walczak, K. (2019, December 3). "The Semantic Web3D: Towards comprehensive representation of 3D content on the semantic web." Semantic Web Pro Workshop, Paris, France. Retrieved from https://www.web3d.org/sites/default/files/attachment/node/2436/edit/20191203-Web3D-semanticweb-SWP2019.pdf
- Collaborative 3D Visualization for Ashore, Afloat and Expeditionary Readiness Workshop. (2019, December 6). Virginia Tech and Web3D Consortium workshop, Arlington, VA. One-day workshop to provide presentations to Naval enterprise leaders on the use of collaborative Web-based #X3D visualization techniques by Government, Academia and Industry practitioners. Retrieved from https://www.web3d.org/event/collaborative-3d-visualization-ashore-afloat-and-expeditionary-readiness-workshop
- Brutzman, D. P. (2019, November 13-15). Navy/USCG Scenario-Based Planning using SPIDERS3D Collaborative Visualization: Notional Concept Study and Scenario Exploration. Maritime Risk Symposium, 2019, Bronx, NY. This presentation is related work to this project. It describes deployment planning for XLUUV at Port Hueneme, CA, in support of COMOPTEVFOR, NAVSEA PMS 406. Retrieved from https://savage.nps.edu/presentations/BrutzmanSpiders3dXluuvScenarioPortHuenemeUscgEvergreen.2019NOV13.pdf

IEEE Distributed Interactive Simulation (DIS). DIS is of direct interest to this project. AUV Workbench outputs simulation results using DIS protocol, making it feasible to someday export and include Ethical Control testing in other Live Virtual Constructive (LVC) scenarios.

- IITSEC 2019 Tutorial Distributed Interactive Simulation (DIS) 101 (2019, December 2). Interservice Industry Training Simulation and Education Conference, 2019, Orlando, FL. "The Distributed Interactive Simulation (DIS) protocol is a well-established IEEE standard for packet-level exchange of state information between entities in military simulations. DIS facilitates simulation interoperability through a consistent over-the-wire format for information, widely agreed upon constant enumeration values, and community-consensus semantics." Retrieved from https://www.web3d.org/event/iitsec-2019-tutorial-distributed-interactive-simulation-dis-101
- Github Open-DIS Software Archive (2020). Retrieved from https://github.com/open-dis
- Brennenstuhl, T. (2019, December 2). Repeatable Unit Testing of Distributed Interactive Simulation (DIS) Protocol Behavior Streams using Web Standards. Interservice Industry Training Simulation and Education Conference, 2019, Orlando, FL. Retrieved from https://www.web3d.org/sites/default/files/attachment/node/2437/edit/BrennenstuhlThesisOpenDisStreamsUnitTesting.2019DEC2.pdf

Simulation Interoperability Standards Organization (SISO). SISO is an international organization dedicated to the promotion of modeling and simulation interoperability and reuse for the benefit of a broad range of Modeling and Simulation communities. Several SISO working groups are relevant to Ethical Control proposed plan to establish virtual environment for unmanned systems qualification. Retrieved from https://www.sisostds.org

- C2SIM PDG/PSG. The Command and Control Systems - Simulation Systems Interoperation (C2SIM) Product Development Group (PDG) and Product Support Group (PSG) uses ontology-governed compatible message interchange, especially focused on NATO forces. Retrieved from https://www.sisostds.org/StandardsActivities/DevelopmentGroups/C2SIMPDGPSG-CommandandControlSystems.aspx
- DIS/RPR FOM. The Distributed Interactive Simulation / Real-time Platform Reference Federation Object Model (DIS / RPR FOM) PSG is a forum and library for DIS-related information for distributed simulation networks. Retrieved from https://www.sisostds.org/StandardsActivities/SupportGroups/DISRPRFOMPSG.aspx
- VV&A PSG. The Verification, Validation and Accreditation/Acceptance (VV&A) PSG emphasizes requirements and metrics within larger Systems Engineering spirals. Retrieved from https://www.sisostds.org/StandardsActivities/SupportGroups/VVAProductsPSG.aspx

Recommended Reading and Resources in Human-Machine Teaming and AI relevant to Ethical Control of unmanned systems.

- Brutzman, D. P., & Fitzpatrick, C. R. (January 2020). Creating Virtual Environments for Evaluating Human-Machine Teaming. Naval Research Program (NRP) Project NPS-19-M285-A. Retrieved from https://calhoun.nps.edu/handle/10945/64266
- Johnson, M., & Vera, A. (Spring 2019). "No AI is an Island: The Case for Teaming Intelligence." AI Magazine, vol. 40, no. 1. Proposes AI will reach its full potential only if, as part of its intelligence, it also has enough teaming intelligence to work well with people. The concepts Coactive Design and Interdependency Analysis have strong

resonances with Ethical Control that deserve further exploration. Retrieved from https://www.aaai.org/ojs/index.php/aimagazine/article/view/2842

- Naval Research Advisory Committee (NRAC) (September 2017). *Autonomous and Unmanned Systems in the Department of Navy*. NRAC Technical Report. The Number 1 Recommendation of the report is "create comprehensive data plan to field autonomy… DoN must urgently develop an organizational data plan… In the future, data will win wars… Data is the ultimate "component" for AI systems and must be controlled." Retrieved from https://www.senedia.org/wp-content/uploads/2018/01/NRAC-Report-Autonomous-and-Unmanned-Systems-in-the-Department-of-Navy.pdf

- Naval Application of Machine Learning (NAML) Workshop. The 2018 Plenary roundtable discussion recognized value and need to establish a machine learning (ML) data strategy, for both inputs and outputs. The focus for 2020 "is on identifying future directions of ML and AI research needed to address mission priorities of both the Navy and the broader national security community." Retrieved from https://sites.google.com/go.spawar.navy.mil/naml

- Summary of the 2018 Department of Defense Artificial Intelligence Strategy: "Harnessing AI to Advance Our Security and Prosperity." (2019 February). "Leading in military ethics and AI safety. The Department will articulate its vision and guiding principles for using AI in a lawful and ethical manner to promote our values. We will consult with leaders from across academia, private industry, and the international community to advance AI ethics and safety in the military context. We will invest in the research and development of AI systems that are resilient, robust, reliable, and secure; we will continue to fund research into techniques that produce more explainable AI; and we will pioneer approaches for AI test, evaluation, verification, and validation. We will also seek opportunities to use AI to reduce unintentional harm and collateral damage via increased situational awareness and enhanced decision support. As we improve the technology and our use of it, we will continue to share our aims, ethical guidelines, and safety procedures to encourage responsible AI development and use by other nations." Retrieved from https://media.defense.gov/2019/Feb/12/2002088963/-1/-1/1/SUMMARY-OF-DOD-AI-STRATEGY.PDF

- Defense Innovation Board (DIB) AI Principles Project. (2019, October 31). AI Principles: Recommendations on the Ethical Use of Artificial Intelligence by the Department of Defense. The DIB members voted to approve the proposed AI Principles: Responsible, Equitable, Traceable, Reliable, Governable. Retrieved from https://media.defense.gov/2019/Oct/31/2002204458/-1/-1/0/DIB_AI_PRINCIPLES_PRIMARY_DOCUMENT.PDF

- Joint Artificial Intelligence Center (JAIC). Retrieved from https://www.ai.mil. The JAIC is tasked to execute DoD AI Strategy.

- Department of Defense Prototyping Guidebook (2019 November). Version 2.0. Office of the Under Secretary of Defense for Research and Engineering (Prototypes and Experiments).

- CTF 151: Counter-Piracy. "In accordance with United Nations Security Council Resolutions, and in cooperation with CMF coastal states, CTF 151's mission is to deter,

disrupt and suppress piracy and armed robbery at sea and to engage with regional and other partners to strengthen relevant capabilities in order to protect global maritime commerce and secure freedom of navigation." Retrieved from https://combinedmaritimeforces.com/ctf-151-counter-piracy

Feedback guidance from William "Doc" Bundy, Ph.D., CAPT, USN (Ret.), Naval War College, Newport Rhode Island on UUV Autonomy and this project (December 18, 2019).

- UUVs will need to operate by mission command, like submarines, given broad missions and the UUV will decide what targets need to be engaged.
- To follow this ethical argument, must build Mission Command capabilities into your intelligent agent controller.
- High-confidence Automated Target recognition will be required.
- We need to let machines kill things.
- Relevant: Unrestricted submarine warfare in WWII. Recommended reading: "Execute Against Japan:" The US Decision to Conduct Unrestricted Submarine Warfare by Joel Holwitt (ibid.)
- Noted "collaborative mission autonomy" as important concept / capability.
- Recommendation: need to engage LTG Shanahan of Joint AI Command.

Integrated Naval Force Structure Assessment (INFSA). The INFSA considers the implications of unmanned systems as parts of the overall Naval Force Structure.

- Clark, B., & Walton, T. (2019, December 31). Taking Back the Seas: Transforming the U.S. Surface Fleet for Decision-Centric Warfare. *Center for Strategic and Budgetary Assessments*. "The U.S. Navy's surface fleet is at a crossroads. Today's force lacks the size, resilience, and offensive capacity to contribute effectively to degrading, delaying, or denying aggression. These shortfalls are especially problematic in light of the fact that the surface fleet will play an increasingly important role in the U.S. Navy's ability to counter enemy attacks. The current fleet is also fiscally unsustainable due to growing operations and support costs for today's highly integrated and manpower-intensive surface combatants. ***New technologies for unmanned systems, sensors, weapons, C3, and countermeasures*** could allow significant improvements in the surface fleet's ability to create complexity for an adversary and harden surface forces from attack while improving surface force's capacity for maritime or land strike." Retrieved from https://csbaonline.org/research/publications/taking-back-the-seas-transforming-the-u.s-surface-fleet-for-decision-centric-warfare
- O'Rourke, R. (2020, February 4). Navy Force Structure and Shipbuilding Plans: Background and Issues for Congress (RL32665). Congressional Research Service. "DON officials suggest that the INFSA could shift the fleet to a more distributed architecture that includes a reduced proportion of larger ships, an increased proportion of smaller ships, and a ***newly created category of large unmanned surface vehicles (USVs) and large unmanned underwater vehicles (UUVs).*** Such a change in fleet architecture could alter, perhaps substantially, the mix of ships to be procured for the

Navy and the distribution of Navy shipbuilding work among the nation's shipyards." Retrieved from https://crsreports.congress.gov/product/pdf/RL/RL32665/284

- Eckstein, M. (2020, February 3). SECNAV Modly: Path to 355 Ships Will Rely on New Classes of Warships, USNI News. "You look at the frigate program: we think, because of the way we've approached that program, we've probably taken three years off the product development lifecycle for that. So we have to start doing the same type of thing: looking at proven hulls, things that can be adaptable for different areas. I understand the Hill's concerns about ***unmanned***, and we get that. … We have to convince them with ***data***: we have to ***wargame*** this, we have to ***iterate*** it over and over again."

  — Honorable Thomas Modly, Acting Secretary of the Navy.

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California

3. Research Sponsored Programs Office, Code 41
   Naval Postgraduate School
   Monterey, CA 93943

4. John Beaver
   Raytheon Missiles & Defense
   Tucson, AZ 85756

5. Hsin-Fu Wu
   Raytheon Missiles & Defense
   Tucson, AZ 85756