



# Ethical Control of Unmanned Systems

Ethical Control of Unmanned Systems using  
Formal Mission Ontologies for Undersea Warfare

Don Brutzman and Curt Blais

Naval Postgraduate School (NPS)

4 May 2020

# Ethical Control



## Project synopsis

- Mission orders must be understandable by both humans and machines
- Logical testing and command trust possible using Semantic Web standards
- Ready for cross-disciplinary thesis work in USW and other majors

## Projections for future work

- Rich set of scenarios are in hand with initial set of logical queries, now ready for more sophisticated ethical queries and mission adaptation
- Promoting ethical thinking: NPS needs a Center for Ethical Warfighting

## Acknowledgements

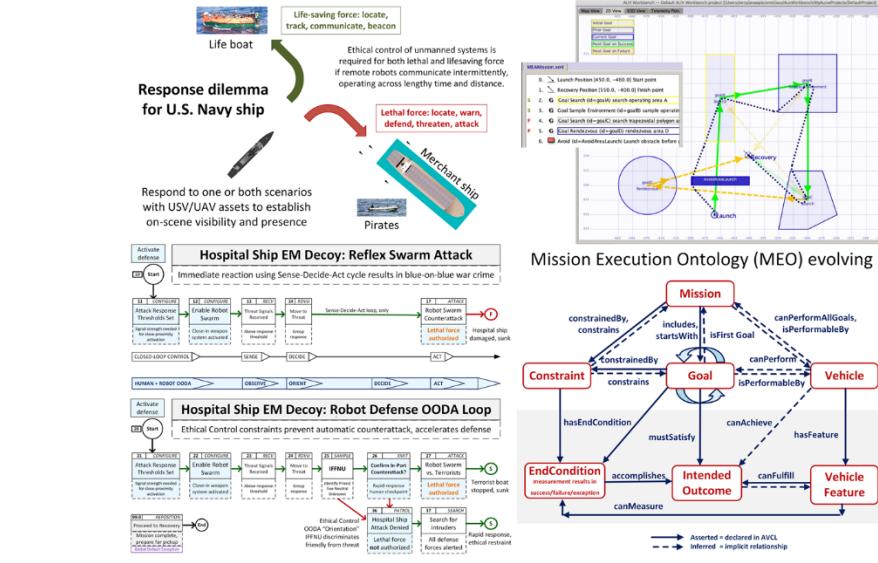
- Thanks to [CRUSER](#) and [Raytheon](#) past support for this many-year project
- NPS-Raytheon CRADA partnership offers new possibilities for influence

# Synopsis: Ethical Control of Unmanned Systems

- **Project Motivation:** ethically constrained control of unmanned systems and robot missions by human supervisors and warfighters.
- **Precept:** 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*.
- **Project Goal:** apply Semantic Web ontology to scenario goals and constraints for logical validation that human-approved mission orders for robots are semantically coherent, precise, unambiguous, and without internal contradictions.
- **Long-term Objective:** demonstrate that no technological limitations exist that prevent applying the same kind of ethical constraints on robots and unmanned vehicles that already apply to human beings.

Paraphrase: can qualified robots correctly follow human orders?

# Ethical Control of Unmanned Systems: Keeping Warfighters in Charge of Autonomy



## Why / Objectives

- Ethical control of unmanned systems can be accomplished through structured mission definitions that are trusted, consistently readable, validatable, repeatable and understandable by humans and robots.
- Orders must be lawful. Unmanned systems must behave ethically and comprehensibly if they are to support manned military units effectively.
- Well-structured mission orders can be tested and trusted 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*.
- Demonstrate that no technological limitations exist that prevent applying the same kind of ethical constraints on robots and unmanned vehicles that already apply to humans, in lethal and life-saving scenarios.

<https://savage.nps.edu/EthicalControl>



## Building on Concept Demonstrations

## Milestones and Transitions

- CRUSER development led to first project selection under CRADA with Raytheon Missile Systems (RMS).
- Successful progress on test missions entering TRL 5 with simulation and Web-shareable 3D visualization.
- Expressing multiple robot mission plans consistently, coherently for diverse UAV, USV, UUV platforms.
- Use Semantic Web Standards to support warfighters.
- Evaluate NAVSEA Unmanned Maritime Autonomy Architecture (UMAA) evolution for robot qualification.

## What / Deliverables

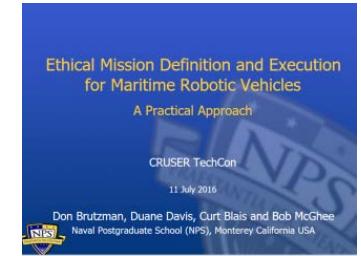
- Update Mission Execution Ontology (MEO) concepts demonstrated in tests and simulation, building to perform field experimentation (FX).
- Supervise thesis work to explore canonical exemplar missions that are expected to utilize unmanned systems, looking across the full range of Naval warfare communities. Example scenarios include UAV for sailor overboard, UAV for refugee/lifeboat escort, and adept scouts. All must observe Law of Armed Conflict (LOAC), Rules of Engagement (ROE), and moral guidance of commanders despite long durations/distances.
- Define, simulate, and test combination of real-world goals and ethical constraints to robot mission tasking across set of canonical scenarios.
- Illustrate how human-robot teams meet moral and legal requirements if deploying unmanned systems with potential for lethal, life-saving force.

Principal Investigator: Don Brutzman  
[brutzman@nps.edu](mailto:brutzman@nps.edu) 1.831.656.2149

Co-Investigator: Curtis Blais  
[cblais@nps.edu](mailto:cblais@nps.edu) 1.831.656.3215



# Key Insights regarding Human Ethical Control



- 1. Humans 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 \(LOAC\)](#).
- 2. Ethical behaviors don't define the mission plan. Instead, ethical constraints inform the mission plan.**
- 3. Naval forces can only command mission orders that are**
  - Understandable by (legally culpable) humans, then
  - Reliably and safely executed by robots.

*Reference:* CRUSER TechCon Overview 2016

<https://gitlab.nps.edu/Savage/EthicalControl/tree/master/documents/presentations>

# Autonomous Vehicle Command Language (AVCL)

- AVCL is a command and control language for humans supervising autonomous unmanned vehicles.
  - Clarity arises from close correspondence to human naval terminology.
- Structured vocabulary defining terms and relationships for mission planning, execution, conduct, recording and replay across diverse robot types.
- Common-ground XML representations for
  - Mission agenda plans, mission scripts, and post-mission recorded telemetry results.
  - ***Future work:*** defining unit tests and expected results for verification and validation.
- Operators have single archivable, validatable format for robot tasking, results
  - directly convertible to and from a wide variety of different robot command languages.

Mission  
Tasking

<https://savage.nps.edu/Savage/AuvWorkbench/AVCL/AVCL.html>

# Example AVCL mission agenda, as pseudo-code XML

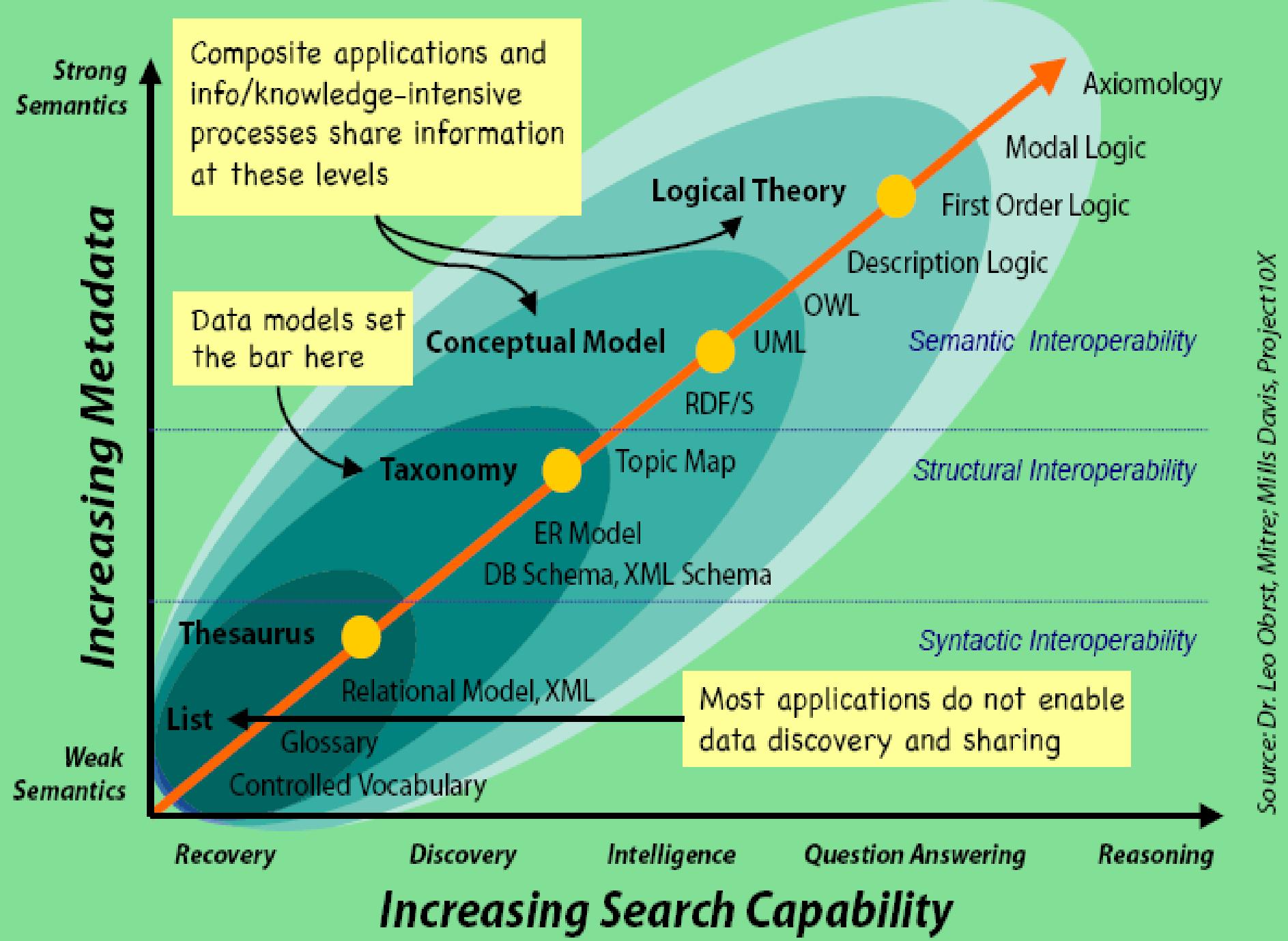
```
<?xml version="1.0" encoding="UTF-8"?>
<UUVMission>
  <GoalSet>
    <Goal area="A" id="goal1">
      <Search nextOnSuccess="goal2" nextOnFailure="goal3"/>
    </Goal>
    <Goal area="A" id="goal2">
      <SampleEnvironment nextOnSuccess="goal3"
        nextOnFailure="recover"/>
    </Goal>
    <Goal area="B" id="goal3">
      <Search nextOnSuccess="goal4" nextOnFailure="goal4"/>
    </Goal>
    <Goal area="C" id="goal4">
      <Rendezvous nextOnSuccess="recover" nextOnFailure="recover"/>
    </Goal>
    <Goal area="recoveryPosition" id="recover">
      <Transit nextOnSuccess="missionComplete"
        nextOnFailure="missionAbort"/>
    </Goal>
  </GoalSet>
</UUVMission>
```

AVCL is readable by human or robot,  
captures logic of mission tasking

XML ensures syntactically correct,  
well-defined, numerically valid

Needed: semantic representation  
to check ethical, logical consistency

# Improving Semantic Representation



# Semantic Web Stack

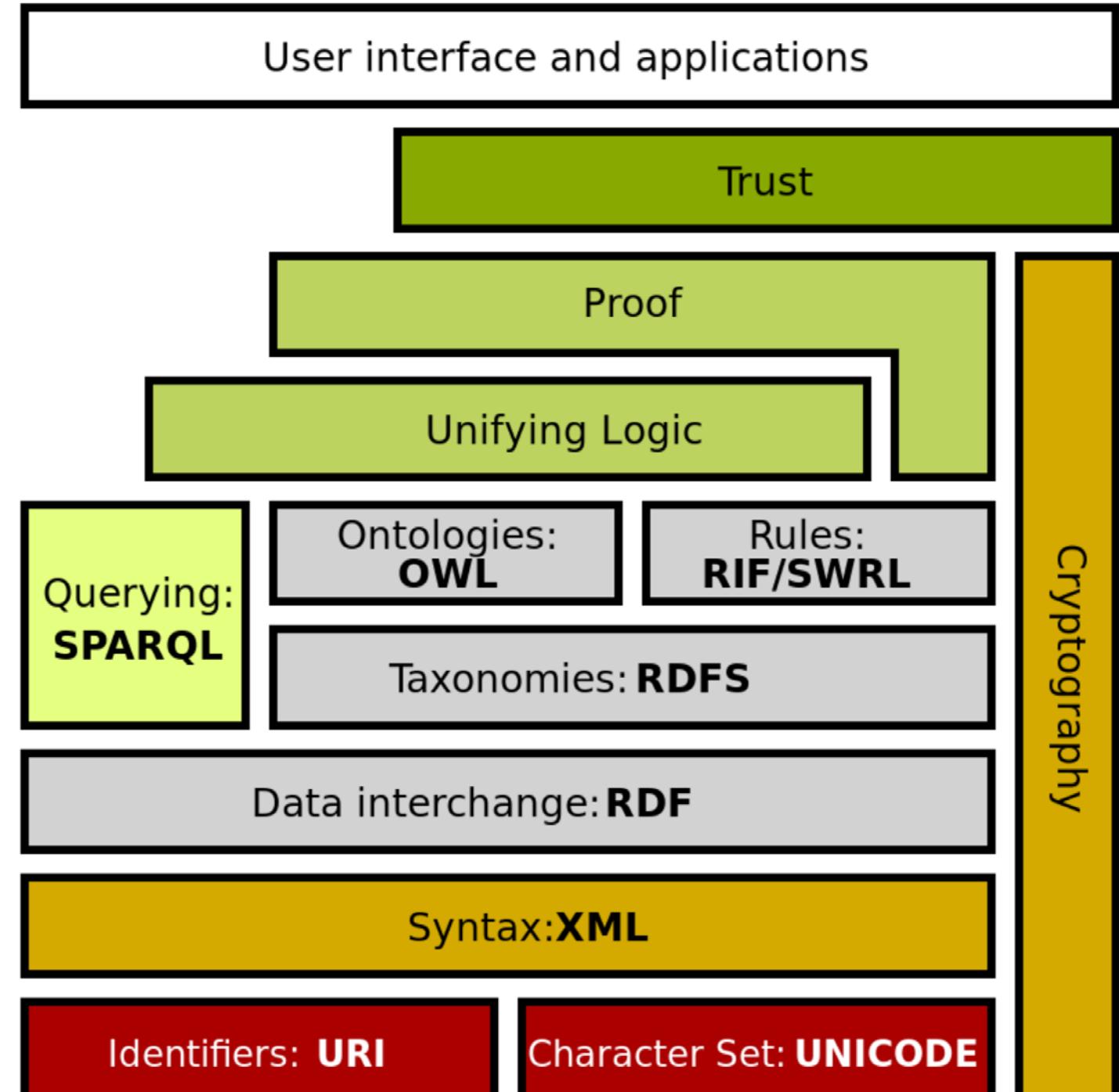
Extends larger Web architecture

- All of these data languages are approved W3C standards
- Proof and unifying logic are mathematically well defined

Trusting derived (composed) statements arises from

- Encryption + digital signature confirms trusted data sources
- Formal logic is basis for deriving new information
- Wikipedia: [Semantic Web Stack](#)

Of note: this project is exercising every layer of Semantic Web stack.



# Mission Execution Ontology (MEO) for Ethical Control of Unmanned Systems in Surrogate Scenarios

Summary of relationships

- **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.
- **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.

# Mission Execution Ontology (MEO) 3.0



Asserted = explicitly declared in AVCL mission  
Inferred = implicit by semantic relationships



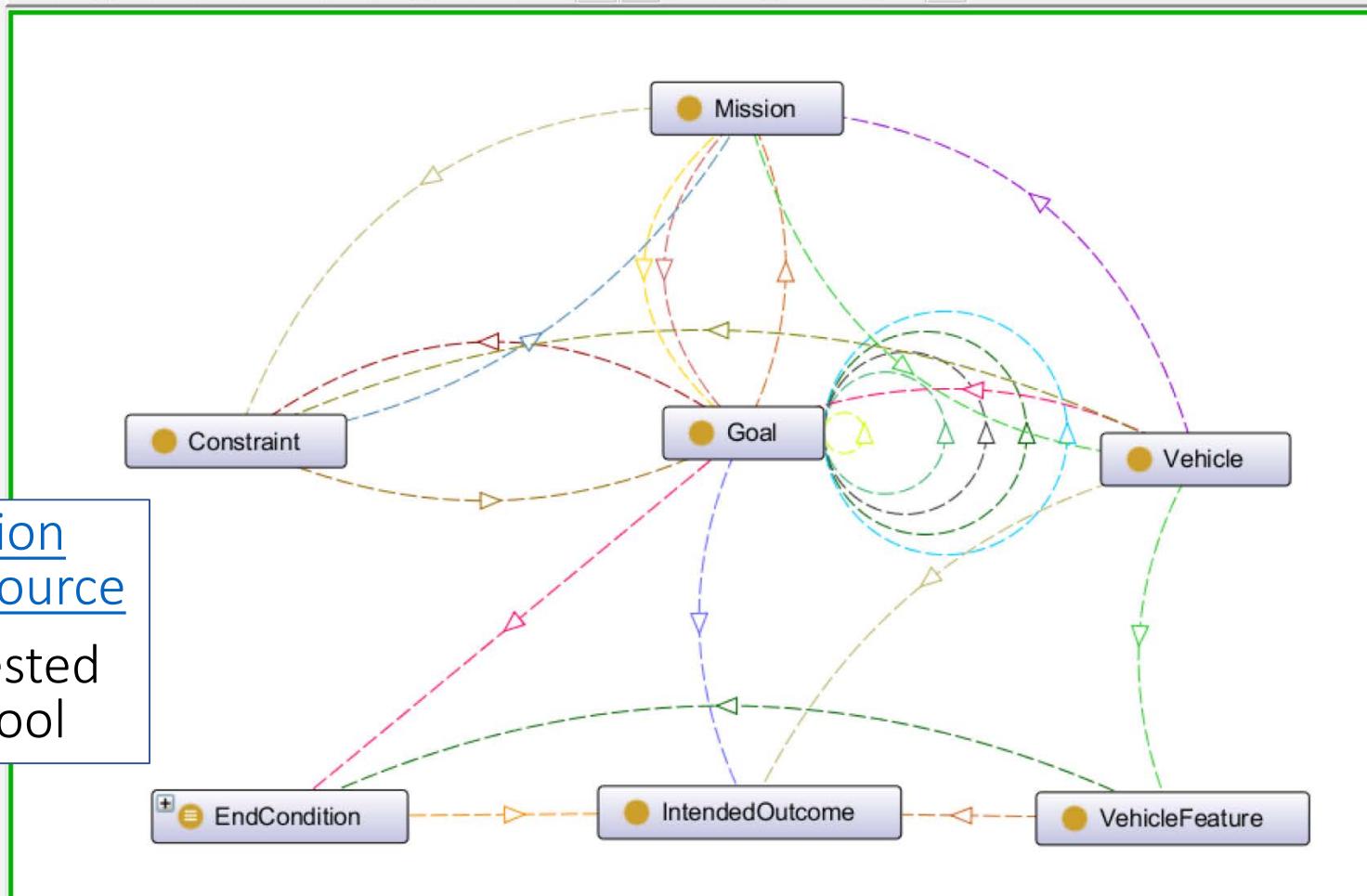
- owl:Thing
  - Constraint
  - EndCondition
  - EXCEPTION
  - FAILURE
  - SUCCESS
  - Goal
  - IntendedOutcome
  - Mission
  - Vehicle
  - VehicleFeature

Search:

contains

Search

Clear



<input checked="" type="checkbox"/> Arc Types
<input type="checkbox"/> <b>accquires</b> (Domain>Range)
<input checked="" type="checkbox"/> <b>canAchieve</b> (Domain>Range)
<input type="checkbox"/> <b>canAchieve(Subclass all)</b>
<input checked="" type="checkbox"/> <b>canFulfill</b> (Domain>Range)
<input type="checkbox"/> <b>canFulfill(Subclass all)</b>
<input checked="" type="checkbox"/> <b>canIdentify</b> (Domain>Range)
<input type="checkbox"/> <b>canIdentify(Subclass all)</b>
<input checked="" type="checkbox"/> <b>canMeasure</b> (Domain>Range)
<input type="checkbox"/> <b>canMeasure(Subclass all)</b>
<input checked="" type="checkbox"/> <b>canMeet</b> (Domain>Range)
<input type="checkbox"/> <b>canMeet(Subclass all)</b>
<input checked="" type="checkbox"/> <b>canPerform</b> (Domain>Range)
<input type="checkbox"/> <b>canPerform(Subclass all)</b>
<input checked="" type="checkbox"/> <b>constraintsGoal</b> (Domain>Range)
<input checked="" type="checkbox"/> <b>constraintsMission</b> (Domain>Range)
<input type="checkbox"/> <b>has individual</b>

Mission Execution  
Ontology (MEO) source

implemented, tested  
using Protégé tool

# Core Considerations for Artificial Intelligence (AI)

- Effective 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, 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*.
- Successful Ethical Control of unmanned systems must be testable.

# Mission clarity for humans – and robots

- Simplicity of success, failure, and (rare) exception outcomes encourages well-defined tasks and unambiguous, measurable criteria for continuation.

Confirmable beforehand: can a tactical officer (or commanding officer) review such a mission and then confidently say

- “yes I understand and approve this human-robot mission” or, equivalently,
- “yes I understand this mission and my team can carry it out themselves.”

Converse:

- if an officer can't fully review/understand/approve such a mission, then likely it is **ill-defined** and needs further clarification anyway.

Added benefit: missions that are clearly readable/runnable by humans and robots can be further composed and checked by C2 planning tools to test for group operational-space management, avoiding mutual interference, etc.

# Wrong question, right question

Wrong question to ask first when planning a tactical operation:

- “What are my *robots* doing out there?”

Right question to ask first when planning a tactical operation:

- “What is my *human-robot team* doing out there?”

**Human-robot team mission has to be understood first!**

- Robots complement humans, who must remain in charge throughout.
- If you don’t have an OODA loop, you don’t have a competent plan.

# Canonical Ethical Mission Development 1

Progressive sophistication to test and evaluate Ethical Control design.

- a. *Sailor Overboard* for life-saving force under close coordination,
- b. *Lifeboat Tracking* for life-saving force under remote conditions,
- c. *Pirate Boats Seizing Merchant* for steady escalation to lethal force,
- d. *Robot Swarm Attacks Hospital Ship* to prevent hostile provocation of sense-decide-act vulnerabilities through Ethical Control safeguards.  
Two variants show fundamental importance of ethical constraints.

Demonstrating a range of functionality tests majority of Ethical Control capabilities that are currently envisioned. These canonical capabilities also “set the stage” for further work on advanced problems of interest.

# OODA Loops for Ethical Control Canonical Missions

Ethical Control OODA Loops	Observe	Orient	Decide	Act
Sailor Overboard	Find Sailor	Report status	Avoid interference	Track sailor until rescued or relieved
Lifeboat Rescue	Find Lifeboat	Report status	Two-way communication	Track life raft until relieved
Pirate Seizure of Merchant Ship	Find merchant ship, pirate small boats	Identity 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 Swarm Attack	EM threat signals detected	(no orientation step in Sense Decide Act)	Reflex-response weapons attack	Mistaken attack on friendly = war crime
Hospital Ship Defense detects spoofing anti-pattern	EM threat signals detected	IFFNU including correlation	Human requirement for lethal force unmet, attack avoided	Report threat alert, commence search for hostile actors

## **Response dilemma for U.S. Navy ship**



Life boat

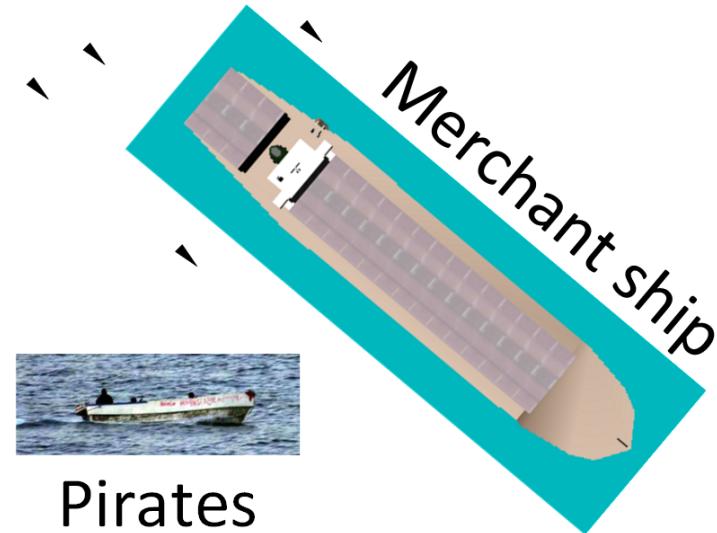


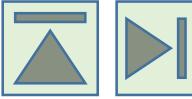
Respond to one or both scenarios  
with USV/UAV assets to establish  
on-scene visibility and presence

**Life-saving force: locate,  
track, communicate, beacon**

Ethical control of unmanned systems is required for both lethal and lifesaving force if remote robots communicate intermittently, operating across lengthy time and distance.

**Lethal force: locate, warn,  
defend, threaten, attack**





# Sailor Overboard Mission: Design

## Motivations

- Explore how *life-saving force* is complementary to *lethal force*, with many similar considerations for remote supervision.
- First demonstrate human-system teaming in close proximity to ship, where direct override of robot control by human operator is possible.
- No temporal delays, all actions and reactions must be immediate.
- Show that mission design can complement shipboard procedures.

## Lessons learned

- Good design patterns to visually represent AVCL mission satisfactorily.
- Temporal flow (left to right, e.g. [Gantt chart](#)) best suited for diagram.

# Sailor Overboard Mission: Description

## Purpose

- Life saving: single unmanned air/surface vehicle actions to complement human responses when performing “*SAILOR OVERBOARD*” operations.
- Carried out in direct concert with formal *shipboard emergency procedures*.
- Multiple UAVs/USVs might be employed in parallel with ships and aircraft, avoid mutual interference by each following deconflicted mission orders.

## Phases

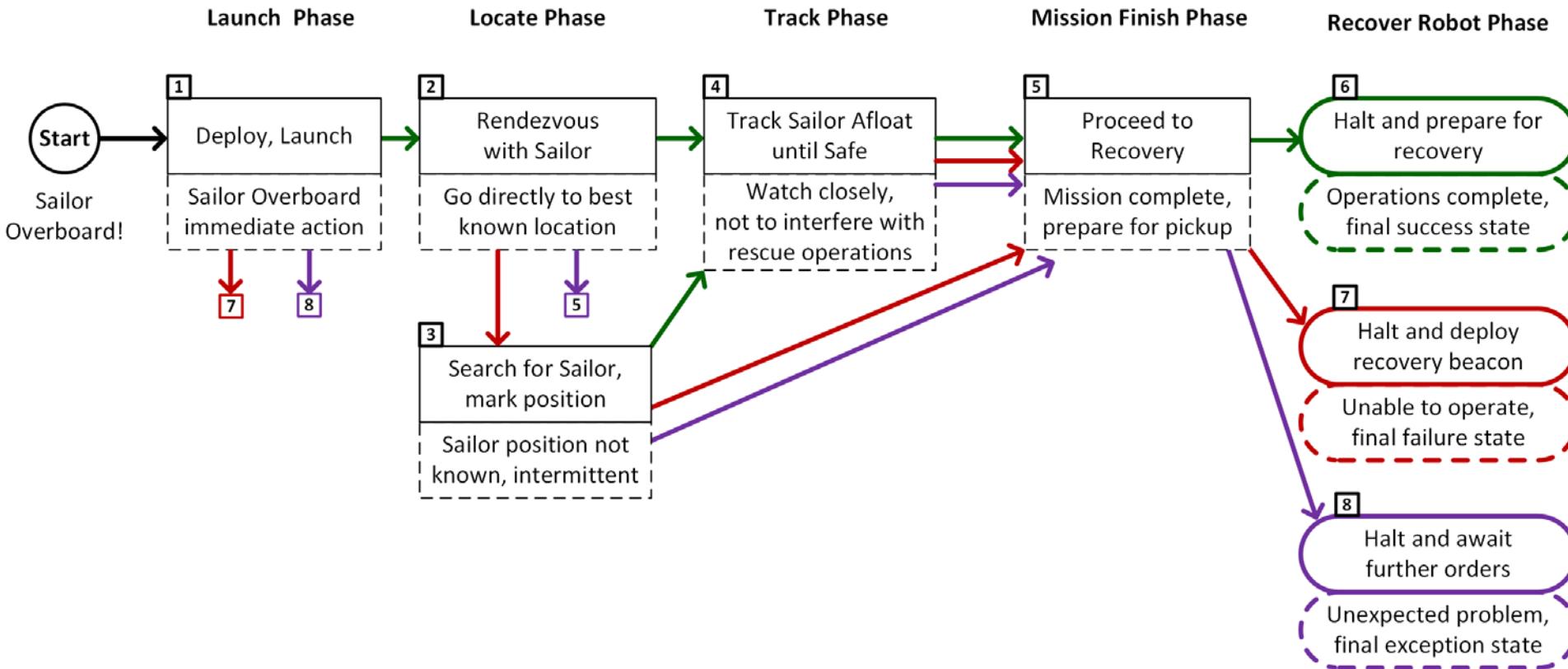
- Deploy/Launch, Rendezvous, Track Sailor until Safe, Return/Recovery.

## Human Supervisory Role

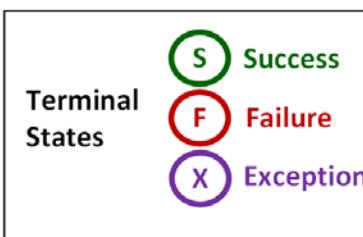
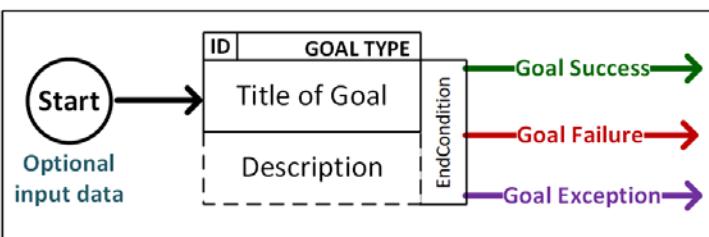
- Order standoff if interfering, manual control is possible due to proximity, can communicate to sailor via loudspeaker or beacon light.

# Sailor Overboard, 8 Phases – Mission Execution Automaton (MEA)

Single unmanned air/surface vehicle actions to complement human response when performing “SAILOR OVERBOARD” operations, carried out in concert with **shipboard emergency procedures**.  
Multiple UAVs/USVs can be employed in parallel with ships/aircraft, each following mission orders.



## Legend



Don Brutzman and Bob McGhee  
Mission upgrade 19 NOV 2019

# SailorOverboard.xml in gitlab.nps.edu version control

GitLab Projects Groups More

Search or jump to... 10 4 ?

E SailorOverboard.xml 5.63 KB

Edit Web IDE Replace Delete

```
<?xml version="1.0" encoding="UTF-8"?>
<AVCL version="3.0" vehicleName="RescueDrone" vehicleType="UAV" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:noNamespaceSchemaLocation="../../../../avcl/AVCL.3.0.xsd">
    <head>
        <meta name="title" content="SailorOverboard.xml"/>
        <meta name="created" content="19 November 2019"/>
        <meta name="modified" content="31 December 2019"/>
        <meta name="creator" content="Don Brutzman"/>
        <meta name="reference" content="https://wiki.nps.edu/display/NOW/Ethical+Control+of+Unmanned+Systems"/>
        <meta name="reference" content="https://gitlab.nps.edu/Savage/EthicalControl/raw/master/missions/SailorOverboard8PhaseMission.png"/>
        <meta name="identifier" content="https://gitlab.nps.edu/Savage/EthicalControl/tree/master/missions/avcl/SailorOverboard.xml"/>
        <meta name="generator" content="Altova XMLSpy, https://www.altova.com"/>
        <meta name="generator" content="Apache NetBeans, https://netbeans.apache.org"/>
        <meta name="generator" content="NPS Autonomous Unmanned Vehicle (AUV) Workbench, https://savage.nps.edu/AuvWorkbench"/>
        <meta name="license" content="../license.html"/>
    </head>
    <body>
        <MissionPreparation>
            <UnitsOfMeasure distance="meters" angle="degrees" mass="kilograms" time="seconds"/>
            <AgendaMission>
                <LaunchPosition id="LaunchPosition" description="Ship position when robot is launched">
                    <LatitudeLongitude latitude="0" longitude="0"/>
                </LaunchPosition>
                <RecoveryPosition id="RecoveryPosition" description="Ship position when ready to recover robot">
                    <LatitudeLongitude latitude="0" longitude="0"/>
                </RecoveryPosition>
                <GoalList>
                    <Goal id="Goal1" title="Deploy, Launch" description="Sailor Overboard immediate Action" nextOnSucceed="Goal2" nextOnFail="Goal2" nextOnException="Goal2">
                        <Rendezvous description="">
                            <TargetVehicleID value="101" description="Sailor in water"/>
                        </Rendezvous>
                        <OperatingArea>
                            <Point>
                                <LatitudeLongitude latitude="36.62" longitude="121.506"/>
                                <!--36°36'11"N 121°53'37"W-->
                            </Point>
                        </OperatingArea>
                        <Duration value="300"/>
                    </Goal>
                </GoalList>
            </AgendaMission>
        </MissionPreparation>
    </body>

```



XML  
Co... edited with XMLSpy v2020 (x64) (http://www.altova.com) by Don Brutzman (Naval Postgraduate School)

## AVCL

```
= version 3.0
= vehicleName RescueDrone
= vehicleType UAV
= xmlns:xsi http://www.w3.org/2001/XMLSchema-instance
= xsi:noNames AVCL.3.0.xsd
```

&lt;&gt;head

body

## MissionPreparation

```
<> UnitsOfMeasure distance=meters angle=degrees mass=kilograms time=seconds
```

## AgendaMission

- <> LaunchPosition id=LaunchPosition description=Ship position when robot is launched
- <> RecoveryPosition id=RecoveryPosition description=Ship position when ready to recover robot

## GoalList

## Goal (6)

	= id	= description	= nextOnSucceed	= nextOnFail	= nextOnException	<> Rendezvous	<> Search	<> MarkTarget	<> MonitorTransmission	<> Reposition	<> OperatingArea	<> Duration
1	Goal1	Proceed to initial known location of sailor overboard	Goal2	Goal2	Goal6	<> Rendezvous description					<> OperatingArea	<> Duration
2	Goal2	Search for sailor in water	Goal3	Goal2	Goal6		<> Search datumType=				<> OperatingArea	<> Duration
3	Goal3	Mark sailor location	Goal4	Goal2	Goal6			<> MarkTarget description			<> OperatingArea	<> Duration
4	Goal4	Track sailor, monitor comms	Goal5	Goal2	Goal6				<> MonitorTransmission		<> OperatingArea	<> Duration
5	Goal5	Mission complete, reposition to ship	RecoveryPosition	Goal2	Goal6					<> Reposition description	<> OperatingArea	<> Duration
6	Goal6	Exception, mission abort, return to ship if possible	LaunchPosition							<> Reposition description	<> OperatingArea	<> Duration

# Sailor overboard mission orders AVCL

Text Grid Schema WSDL XBRL Authentic Browser

SailorOverboard.xml

## Messages



File C:\x-nps-gitlab\NetworkOptionalWarfare\ethicalcontrol\avcl\SailorOverboard.xml is valid.

missions (https://www.nps.edu/ontologies/MissionExecutionOntology/missions) : [C:\x-nps-gitlab\NetworkOptionalWarfare\ethicalcontrol\ontologies\CanonicalMission.ttl]

File Edit View Reasoner Tools Refactor Window Mastro Ontop Help

missions (https://www.nps.edu/ontologies/MissionExecutionOntology/missions) Search... Goal

Active ontology Entities Classes Individuals by class OWLViz DL Query OntoGraf Debugger Snap SPARQL

**Class hierarchy: Goal**

owl:Thing  
 Constraint  
 EndCondition  
**Goal**  
 GoalRequirement  
 Mission  
 Vehicle  
 VehicleFeature

Asserted Annotations Usage

Usage: Goal

Show:  this  disjoint  named sub/superclasses

Found 69 uses of Goal

- appliesTo
  - appliesTo Range Goal or Mission
- canMeet
  - canMeet Range Goal

Constraint

- DisjointClasses: Constraint, EndCondition, Goal, GoalRequirement, Mission, Vehicle, VehicleFeature
- Constraint SubClassOf appliesTo only (Goal or Mission)
- Constraint SubClassOf appliesTo min 1 Goal

EndCondition

- DisjointClasses: Constraint, EndCondition, Goal, GoalRequirement, Mission, Vehicle, VehicleFeature

Goal

- Goal SubClassOf requires only GoalRequirement
- DisjointClasses: Constraint, EndCondition, Goal, GoalRequirement, Mission, Vehicle, VehicleFeature
- Goal SubClassOf hasEndCondition max 1 EndCondition
- Goal rdfs:comment "A Goal is an operational objective the mission is attempting to achieve."
- Goal SubClassOf hasEndCondition only EndCondition

Description: Goal

Equivalent To +

SubClass Of +
 

- ((hasEndCondition only FAIL)  
 and (hasNextOnFail exactly 1 Goal)) or ((hasEndCondition only SUCCEED)  
 and (hasNextOnSucceed exactly 1 Goal)) or ((hasEndCondition only VIOLATE)  
 and (hasNextOnViolate exactly 1 Goal))
- hasEndCondition max 1 EndCondition
- hasEndCondition only EndCondition

# Sailor overboard mission .ttl Turtle

Git: master To use the reasoner click Reasoner > Start reasoner Show Inferences

## SPARQL query response

### [SailorOverboardConverted. MissionQuery\\_01\\_GoalBranches.rq.txt](#)

goal	nextOnSuccess	nextOnFailure	nextOnException	isPartOfPhase	description
:Goal11	:Goal12	:Goal12	:Goal15	"Launch"	"Deploy, Launch: Sailor Overboard Immediate Action"
:Goal12	:Goal14	:Goal13	:Goal15	"Locate"	"Rendezvous with Sailor: Go directly to best known location"
:Goal13	:Goal14	:Goal15	:Goal15	"Locate"	"Search for Sailor: Sailor position not known, intermittent"
:Goal14	:Goal15	:Goal15	:Goal15	"Track"	"Track Sailor afloat until safe: Watch closely, not to interfere with rescue operations"
:Goal15	:Goal16	:Goal12	:Goal16	"Mission Finish"	"Proceed to Recovery: Mission complete, prepare for pickup"
:Goal16				"Recover Robot"	"Halt and prepare for recovery: Operations complete, final success state"
:Goal17				"Recover Robot"	"Halt and deploy recovery beacon: Unable to continue, final failure state"
:Goal18				"Recover Robot"	"Halt and await further orders: Unexpected problem, final exception state"

## SPARQL query response

### [SailorOverboardConverted. MissionQuery\\_02\\_InitialGoal.rq.txt](#)

Mission	InitialGoal	isPartOfPhase	goalDescription
:SailorOverboard	:Goal11	"Launch"	"Deploy, Launch: Sailor Overboard Immediate Action"

# Sailor overboard simulation AUVW

Initial Goal  
Prior Goal  
Current Goal  
Next Goal on Success  
Next Goal on Failure

UPDATE IN PROGRESS

Launch



SearchBox  
Search

RendezvousSailor  
Rendezvous

A blue rectangular search box centered at approximately (300, 150) with a dotted boundary. It contains a circular path with a dot in the center.

MarkLocationBox  
Mark

A blue rectangular search box centered at approximately (450, 100) with a dotted boundary.

TrackSailorBox  
Search

A blue rectangular search box centered at approximately (600, 100) with a dotted boundary.

ReturnToShipBox  
Search

A dark blue rectangular search box centered at approximately (300, -10) with a solid boundary.

StayOutBox

A dark blue rectangular search box centered at approximately (550, 0) with a solid boundary.

RotatedBoxTest

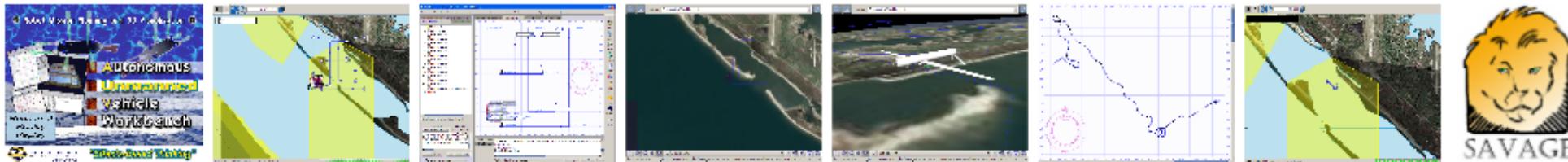
Geographic origin  
36.61 lat, -121.89 lon

World coordinate system:  
+X North, +Y East, +Z Down

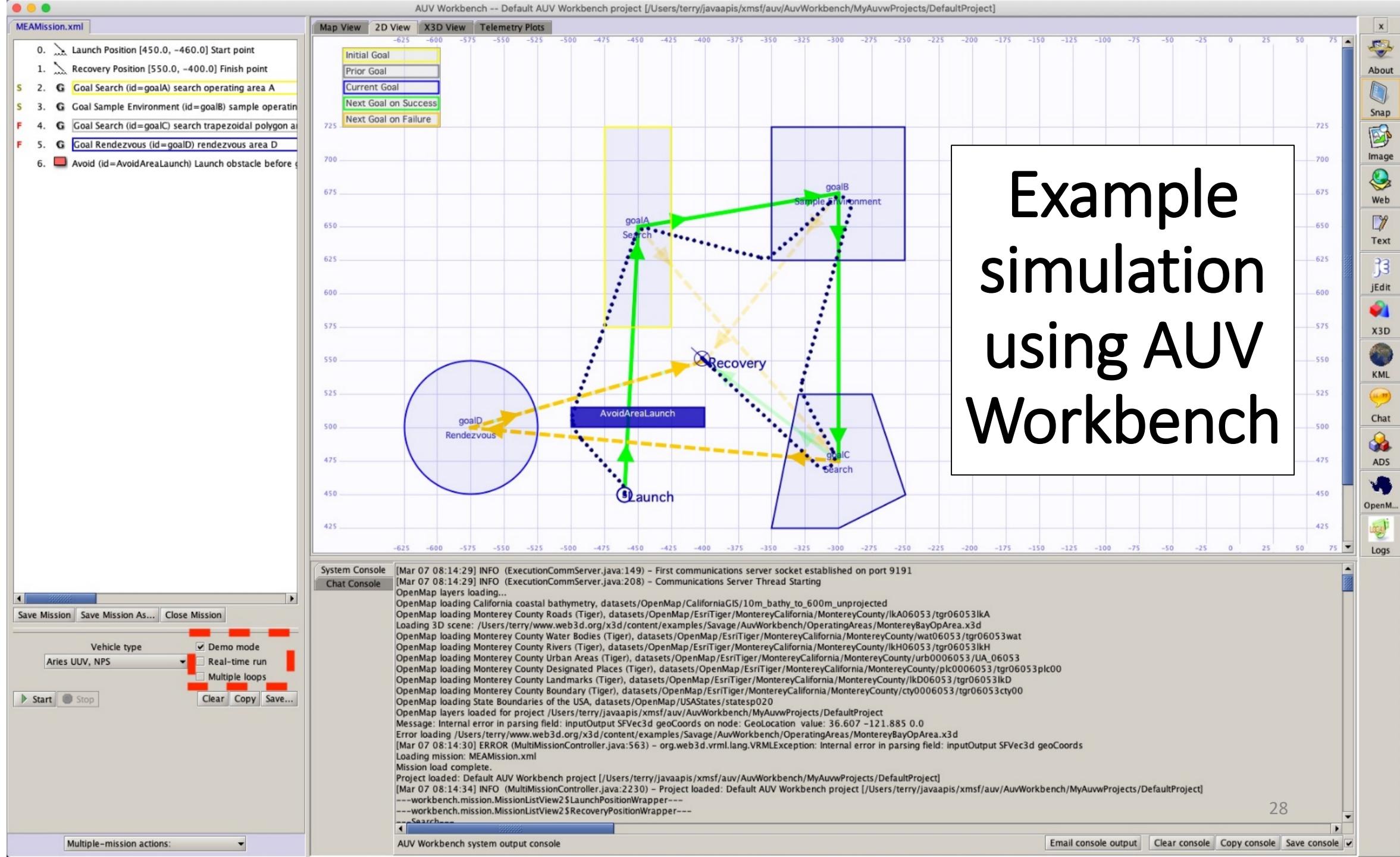
# AUV Workbench

Example  
open-source  
implementation

- Autonomous Unmanned Vehicle (AUV) Workbench supports underwater, surface and air vehicles
  - Rehearsal of physics-based mission response
  - Real-time task-level control of robot missions, and
  - Replay of recorded results
  - Industry-friendly open-source license, Sourceforge
  - Basis: RBM 3-level architecture, AVCL commands
- Used to rehearse strategic-level agenda missions
  - <https://savage.nps.edu/AuvWorkbench>



AUV Workbench Demo Mode 2



# Example simulation using AUV Workbench

# Qualification of unmanned systems: extend Verification Validation Accreditation (V V&A)

Proposed  
Future Work

[SISO V V&A](#)  
[Study Group](#)

How to test and certify robots will follow both orders and constraints?  
Humans confirm understanding and trust through qualification processes.

- 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 repeatedly 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.



# Data-Centric Security and Trust

Compression, authentication, encryption, composability, blockchain ledger, and asymmetric advantages enable group communication of secure mission orders and responses.  
This is a Chain of Trust for distributed Command Authority.

# Trust

## Trusted Mission Orders

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

## Trusted Mission Execution

- 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



# Related Ethical Activities

Ethics of lethality and unmanned systems is an active area of work. The following synopses are distilled from each respective resource.

# Unmanned Maritime Autonomy Architecture (UMAA)

Richard R. Burgess, “[Navy Requests Information for Unmanned Maritime Autonomy Architecture](#),” SEA POWER, 20 FEB 19

- “The intent of UMAA is to provide overarching standards that various UUVs and USVs can be built to in order to avoid creating multiple conflicting systems in the future”
- “The UMAA is being established to enable autonomy commonality and reduce acquisition costs across both surface and undersea unmanned vehicles.”
- Topics of interest include Situational Awareness, Sensor and Effector Management, Processing Management, Communications Management, Vehicle Maneuver Management, Vehicle Engineering Management, Vehicle Computing Management, Support Operations

Multiple public NAVSEA documents refer to autonomy efforts and UMAA.

- NAVSEA PMS 406 is Program Office for Unmanned Maritime Systems
- NAVSEA Fact Sheet: Unmanned Maritime Systems Program Office (PMS 406)  
<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, [https://www.navysbir.com/n20\\_1/N201-059.htm](https://www.navysbir.com/n20_1/N201-059.htm)

Proposed  
Critical Path  
Forward



# DoD Directive: Autonomy in Weapon Systems

- [DoD Directive 3000.09](#), 21 NOV 2012 with change 1, 8 May 2017
- Original and update signed by DEPSECDEF Ashton Carter

**CONTROLLING  
REFERENCE  
OVERALL**

## 1. PURPOSE

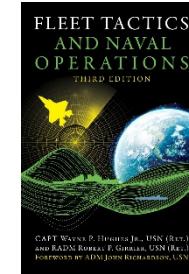
- a. Establishes DoD policy and assigns responsibilities for the development and use of autonomous and semi-autonomous functions in weapon systems, including manned and unmanned platforms.
- b. Establishes guidelines designed to minimize the probability and consequences of failures in autonomous and semi-autonomous weapon systems that could lead to unintended engagements.

## 4. POLICY (excerpted)

- a. 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.
- b. 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). [...]

## 6. RELEASABILITY. Cleared for public release. [...]

# Fleet Tactics and Naval Operations



Wayne P. Hughes, Jr. and Robert P. Girrier, *Fleet Tactics and Naval Operations*, Third Edition, Naval Institute Press, Annapolis Maryland, June 2018.

- <https://www.usni.org/press/books/fleet-tactics-and-naval-operations-third-edition>

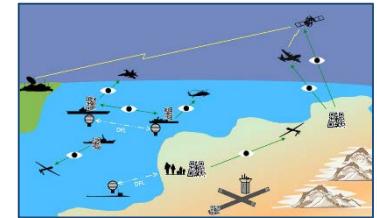
From newly added Chapter 12, A Twenty-First-Century Revolution:

- “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.”

Wayne Hughes coined the term “Network Optional Warfare” after many discussion sessions, directly contrasting it to Network Centric Warfare. Thank you sir.

# Network Optional Warfare (NOW)

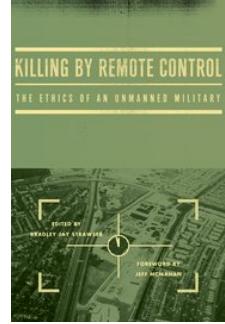
Overall  
Concept



**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-optimal warfare (NOW) precepts include Efficient Messaging, Optical Signaling, Semantic Coherence and Ethical Human Supervision of Autonomy for deliberate, stealthy, minimalist tactical communications.
- <https://wiki.nps.edu/display/NOW/Network+Optional+Warfare>

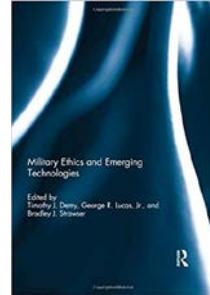
# Killing by Remote Control: the Ethics of an Unmanned Military



[Bradley Jay Strawser](#) editor, [Killing by Remote Control: The Ethics of an Unmanned Military](#), Oxford University Press, 2012

- Addresses topic of contemporary debate: whether use of an unmanned military is ethical.
- The increased military employment of remotely operated aerial vehicles, also known as drones, has raised a wide variety of important ethical questions, concerns, and challenges. Many of these have not yet received the serious scholarly examination such worries rightly demand. This volume attempts to fill that gap through sustained analysis of a wide range of specific moral issues that arise from this new form of killing by remote control. Many, for example, are troubled by the impact that killing through the mediated mechanisms of a drone half a world away has on the pilots who fly them. What happens to concepts such as bravery and courage when a war-fighter controlling a drone is never exposed to any physical danger? This dramatic shift in risk also creates conditions of extreme asymmetry between those who wage war and those they fight. What are the moral implications of such asymmetry on the military that employs such drones and the broader questions for war and a hope for peace in the world going forward? How does this technology impact the likely successes of counter-insurgency operations or humanitarian interventions? Does not such weaponry run the risk of making war too easy to wage and tempt policy makers into killing when other more difficult means should be undertaken?

# Military Ethics and Emerging Technologies



- [Military Ethics and Emerging Technologies](#), edited by Timothy J. Demy, George R. Lucas Jr., and Bradley J. Strawser, Routledge, 2014.
- This volume looks at current and emerging technologies of war and some of the ethical issues surrounding their use. Although the nature and politics of war never change, the weapons and technologies used in war do change and are always undergoing development. Because of that, the arsenal of weapons for twenty-first century conflict is different from previous centuries. Weapons in today's world include an array of instruments of war that include, robotics, cyber war capabilities, human performance enhancement for warriors, and the proliferation of an entire spectrum of unmanned weapons systems and platforms. Tactical weapons now have the potential of strategic results and have changed the understanding of the battle space creating ethical, legal, and political issues unknown in the pre-9/11 world. What do these technologies mean for things such as contemporary international relations, the just-war tradition, and civil-military relations?
- This book was originally published in various issues and volumes of the [Journal of Military Ethics](#).

# Stockdale Center for Ethical Leadership



## Stockdale Center for Ethical Leadership, U.S. Naval Academy (USNA)

- Authorized by the Secretary of the Navy in 1998, the Center for the Study of Professional Military Ethics (CSPME) undertook an ambitious mission – to promote and enhance the ethical development of current and future military leaders. In February 2006, the Superintendent of the Naval Academy directed the expansion of the Center, and the Center was renamed the Vice Admiral James B. Stockdale Center for Ethical Leadership.
- The Center could not have a finer model as its namesake than this most distinguished graduate of the United States Naval Academy. Admiral Stockdale was a man of unsurpassed courage and integrity who clearly understood the gravity of a leader's moment of ethical decision.

# Marine Corps University (MCU) School of Advanced Warfighting



*Mission Statement.* The [School of Advanced Warfighting \(SAW\)](#) develops lead planners and future commanders with the will and intellect to solve complex problems, employ operational art, and design/execute campaigns in order to enhance the Marine Corps' ability to prepare for and fight wars.

- Formulate solutions to complex problems and apply operational art in an uncertain geostrategic security environment.
- Employ knowledge of the operational level of war, the art of command, and ethical behavior in warfighting.
- Quickly and critically assess a situation, determine the essence of a problem, fashion a suitable response, and concisely communicate the conclusion in oral and written forms.
- Demonstrate the competence, confidence, character, and creativity required to plan, lead, and command at high level service, joint, and combined headquarters.
- Design and plan adaptive concepts to meet current and future challenges.

# College of Leadership and Ethics



## College of Leadership and Ethics, U.S. Naval War College, Newport RI

- The College of Leadership and Ethics (CLE) works through three lines of effort: education, research and outreach. The leadership and ethics area of study in the electives program and core curriculum is the key focus of CLE education. CLE's leadership research, assessment, and analysis helps develop various leader development course curricula. CLE outreach includes support for leader development in various Navy communities.
- The College of Leadership & Ethics (CLE) is composed of a group of teaching scholars and practitioners who support Navy leader development initiatives and conduct research on leadership effectiveness. CLE provides an opportunity for students to engage in studies that broaden their educational experience by focusing on leadership and ethics. We develop stronger leaders by preparing them to fulfill their expanding and ever-widening roles as global leaders.

# Joint Artificial Intelligence Center (JAIC) <https://www.ai.mil>

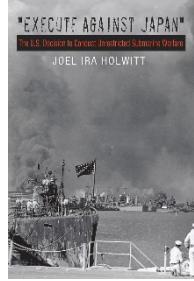


DoD CIO [Joint Artificial Intelligence Center \(JAIC\)](#) tasked to execute DoD AI Strategy.

JAIC's Guiding Tenets:

- AI is critical to the future of United States' national security, and the JAIC is the focal point for the execution of the DoD AI Strategy.
- We exist to create and enable impact for the Armed Services and DoD Components across the full range of their missions – from the back office to the front lines of the battlefield.
- Leadership in ethics and values is core to everything we do at the JAIC.
- The JAIC attracts the best and brightest people from across DoD, commercial industry, and academia. When they get here, they are empowered to bring their expertise to the Department's transformation.

# Execute Against Japan – The US Decision To Conduct Unrestricted Submarine Warfare



- [Joel Ira Holwitt](#), Texas A+M University Press
- “ . . . until now how the Navy managed to instantaneously move from the overt legal restrictions of the naval arms treaties that bound submarines to the cruiser rules of the eighteenth century to a declaration of unrestricted submarine warfare against Japan immediately after the attack on Pearl Harbor has never been explained. Lieutenant Holwitt has dissected this process and has created a compelling story of who did what, when, and to whom.”—The Submarine Review
- “Execute against Japan should be required reading for naval officers (especially in submarine wardrooms), as well as for anyone interested in history, policy, or international law.”  
— Adm. James P. Wisecup, President, US Naval War College (for Naval War College Review)
- “Although the policy of unrestricted air and submarine warfare proved critical to the Pacific war’s course, this splendid work is the first comprehensive account of its origins—illustrating that historians have by no means exhausted questions about this conflict.”—World War II Magazine
- “US Navy submarine officer Joel Ira Holwitt has performed an impressive feat with this book. . . Holwitt is to be commended for not shying away from moral judgments . . . This is a superb book that fully explains how the United States came to adopt a strategy regarded by many as illegal and tantamount to ‘terror’.”—Military Review

# Assessment of Current Thinking

- Human supervision of potentially lethal autonomous systems is a matter of serious global importance.
- Wide consensus is emerging on principles, aspects of the problem, elements of solutions, and need to achieve better capabilities.
- Much philosophical concern but few concrete activities are evident.

Ethical Control of Unmanned Systems project appears to provide a needed path towards practice, with the historic role of warfighting professionals more central than ever as weapons autonomy grows.

# Conclusions

- Human supervision is required for any unmanned systems holding potential for lethal force.
  - Cannot push “big red shiny AI button” and hope for best – immoral, unlawful.
  - Similar imperatives exist for supervising systems holding life-saving potential.
- Human control of unmanned systems is possible at long ranges of time-duration and distance through well-defined mission orders.
  - Readable and sharable by both humans and unmanned systems.
  - Validatable syntax and semantics through understandable logical constraints.
  - Testable and confirmable using simulation, visualization, perhaps qualification.
- Coherent human-system team approach is feasible and repeatable.
  - Semantic Web confirmation can ensure orders are comprehensive, consistent.
  - Human role remains essential for life-saving and potentially lethal scenarios.

# Recommendations for Future Work

## *Continued development*

- Diverse mission exemplars
- Software implementations
- 2D, 3D visualization of results

## *Future capabilities*

- Automatable testing
- Field experimentation (FX)
- “Qualification” of unmanned systems in virtual environments

## *Outreach*

- Presentations, publication review
- Engagement in key ethical forums
- NPS wargame and course support
- NPS thesis and dissertation work

## *Adoption*

- Support for developmental system
- Influence campaign for both C4I and robotics communities of interest

# Recommendations for Future Work

Your NPS efforts matter.

Talk to us about a potential thesis!



# Contact



## Don Brutzman

[brutzman@nps.edu](mailto:brutzman@nps.edu)

<http://faculty.nps.edu/brutzman>

Code USW/Br, Naval Postgraduate School  
Monterey California 93943-5000 USA

1.831.656.2149 work

1.831.402.4809 cell



# Contact



## Curt Blais

[cblais@nps.edu](mailto:cblais@nps.edu)

[home page](#)

Code MV/Bl, Naval Postgraduate School

Monterey California 93943-5000 USA

1.831.656.3215 work

## ETHICAL MISSION DEFINITION AND EXECUTION FOR MARITIME ROBOTS UNDER HUMAN SUPERVISION

- Lethality requires ethical and legal basis, supervised by military teams.
- Executable robot tasking can resemble tactical tasking of humans afloat.
- Careful application of goal constraints makes ethical control feasible.
- Robot missions then complement and extend naval operation orders.
- Semantic Web logic can confirm ethical correctness and completeness.
- Next steps: continue 2 decades of work with realistic scenario testing.

“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.”

Robert McGhee, April 2016

- [IEEE Journal of Oceanic Engineering \(JOE\) paper](#) along with [online references](#).
- Authors Don Brutzman, Curtis Blais, Duane Davis and Robert McGhee, NPS.
- Feedback and recommendations always welcome. Contact: [brutzman@nps.edu](mailto:brutzman@nps.edu)

