2016-09

Investigating a linked open data model for unmanned system research and development portfolio management

Blais, Curtis L.

Monterey, California, Naval Postgraduate School

https://hdl.handle.net/10945/50695

NPS-MV-16-003

NAVAL
POSTGRADUATE
SCHOOL

MONTEREY, CALIFORNIA

INVESTIGATING A LINKED OPEN DATA MODEL FOR UNMANNED SYSTEMS RESEARCH AND DEVELOPMENT PORTFOLIO MANAGEMENT

by

Curtis L. Blais

September 2016

Approved for public release; distribution is unlimited

Prepared for: Office of the Secretary of Defense Joint Ground Robotics Enterprise
3090 Defense Pentagon, Room 5C756, Washington, DC 20301

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
30-09-2016

2. REPORT TYPE
Technical Report

3. DATES COVERED *(From-To)*
8 Oct 2015-30 Sept 2016

4. TITLE AND SUBTITLE
Investigating a Linked Open Data Model for Unmanned System Research and Development Portfolio Management

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Curtis L. Blais

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER
RVJYM

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) AND ADDRESS(ES)
Naval Postgraduate School
1 University Circle
Monterey, CA 93943

8. PERFORMING ORGANIZATION REPORT NUMBER
NPS-MV-16-003

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Secretary of Defense Joint Ground Robotics Enterprise
3090 Defense Pentagon, Room 5C756, Washington, DC 20301

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

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

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The Office of the Secretary of Defense Joint Ground Robotics Enterprise tasked the Naval Postgraduate School to perform research exploring (1) enhancement of robotics education; (2) improved representation of robotic systems in combat simulations; and (3) interoperability standards for military robotics systems. As part of this work, NPS investigated the application of state-of-the-art information modeling techniques to create a conceptual and technical knowledge representation of unmanned systems architectures, research activities, development efforts, and operational employment to help better understand unmanned systems and the breadth and coverage of DoD research, development, and education projects. The chosen approach leverages Semantic Web technologies as demonstrated by the Linking Open Data project activities being performed on the World Wide Web. This report provides an introduction to the technical approach and lays out a plan of action for creating the desired knowledge representation and for populating the resulting model with data from Department of Defense unmanned system programs.

15. SUBJECT TERMS
Linked open data; Resource Description Framework; web technologies; semantic web

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Curtis Blais
Unclassified	Unclassified	Unclassified	UU	41	**19b. TELEPHONE NUMBER *(include area code)*** 831-656-3215

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

THIS PAGE INTENTIONALLY LEFT BLANK

NAVAL POSTGRADUATE SCHOOL
Monterey, California 93943-5000

Ronald A. Route President	Steve Lerman Provost

The report entitled “Investigating a Linked Open Data Model for Unmanned Systems Research and Development Portfolio Management” was prepared for and funded by the Office of the Secretary of Defense Joint Ground Robotics Enterprise.

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

This report was prepared by:

Curtis L. Blais
Faculty Associate - Research

Reviewed by:	**Released by:**
Imre Balogh, Director MOVES Institute	Jeffrey D. Paduan Dean of Research

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF FIGURES

I. INTRODUCTION

A. BACKGROUND

Acquisition and deployment of unmanned systems in the Department of Defense (DoD) has grown substantially over the past fifteen years. Indeed, in 2001 the United States Congress mandated significant levels of employment of unmanned systems in DoD operations, as reiterated in the House of Representatives DoD Appropriations Bill for 2017 (House of Representatives Committee on Appropriations 2016, 104):

> "Section 220 of the Floyd D. Spence National Defense Authorization Act for Fiscal Year 2001 (Public Law 106-398) mandated a goal, regarding unmanned advanced capability combat aircraft and ground combat vehicles, that by the year 2010, one-third of the aircraft in the operational deep strike force fleet would be unmanned, and that by year 2005, one-third of the operational ground combat vehicles would be unmanned."

Furthermore, this House of Representatives report calls for an update from the DoD on progress toward these congressionally mandated goals by no later than 15 September 2016, requesting a briefing that "shall include an assessment of progress towards meeting the goals identified for the subset of unmanned air and ground systems established in Section 220 of Public Law 106-398, as well as an assessment of existing, viable unmanned ground vehicle technologies that can be economically used for making significant progress toward the achievement of the 2001 goal within the next 5 years." Further illustrating the commitment to unmanned systems, this budget authorizes funds to address such issues as manned-unmanned system teaming, human-machine autonomous command and control environment, carrier-based operations for unmanned aerial vehicles, immersive operator control stations for unmanned systems, unmanned advanced capability combat aircraft and ground combat vehicles, armed robotic platforms deployable with manned platforms, and many other related areas. The Defense Science Board Summer Study on Autonomy (Defense Science Board 2016) broadened the discussion to include *autonomy at rest* (e.g., systems operating virtually, in software, to include planning systems and expert advisory systems) and *autonomy in motion* (e.g., unmanned vehicles operating in the air, on the ground, on the sea surface, or undersea):

> “Rapid global market expansion for robotics and other intelligent systems to address consumer and industrial applications is stimulating increasing commercial investment and delivering a diverse array of products. At the same time, autonomy is being embedded in a growing array of software systems to enhance speed and consistency of decision-making, among other benefits. Likewise, governmental entities, motivated by economic development opportunities in addition to security missions and other public sector applications, are investing in related basic and applied research. Applications include commercial endeavors, such as IBM’s Watson, the use of robotics in ports and mines worldwide, autonomous vehicles (from autopilot drones to self-driving cars), automated logistics and supply chain management, and many more. Japanese and U.S. companies invested more than $2 billion in autonomous system in 2014, led by Apple, Facebook, Google, Hitachi, IBM, Intel, LinkedIn, NEC, Yahoo, and Twitter.” (Defense Science Board 2016, 6)

The United States Office of the Secretary of Defense (OSD) Joint Ground Robotics Enterprise (JGRE) is the principal organization in DoD for providing oversight, policy, and program direction to establish definitive robotics operational requirements and to pursue critical technologies to satisfy those requirements. The organization focuses on interoperability, modeling and simulation, and test and evaluation. In Fiscal Year 2016 (FY16), JGRE funded the Naval Postgraduate School (NPS) to explore the following topics: (1) enhancement of robotics education; (2) improved representation of robotic systems in combat simulations (Blais and McGregor 2016); and (3) interoperability standards for military robotics systems (Blais 2016). To address the research and education goals of the project, NPS identified the need for a conceptual framework that can facilitate better understanding by educators, researchers, and program managers of unmanned system architectures, research activities, development efforts, and operational employment.

The NPS Consortium for Robotics and Unmanned System Education and Research (CRUSER) program, funded through the Office of Naval Research, sponsors numerous research and education projects each year across NPS faculty, students, and outside collaborators in industry, government, and academia. In certain respects, the CRUSER program can be seen as a microcosm of efforts that occur across DoD, at least with respect to technical, academic investigations if not actual system acquisitions. Even within CRUSER there is a need for consolidation of information to enable program administrators to gain an understanding of what has been accomplished over the past several years, in what areas of robotic system architecture or employment, and with what level of success in order to make ongoing decisions about what new work should be funded. Such questions as “How does the proposed/performed work relate to prior CRUSER projects?” or “What aspects of unmanned systems are not receiving attention through the CRUSER program?” cannot readily be answered at present, since no one has complete corporate knowledge of work that has been performed. The result is that numerous initiatives continue to be worked and awarded without an understanding of what has gone before (both within and outside the CRUSER program) and how current or new work relates to specific aspects of unmanned system architectures, operations, control, etc.

This problem in understanding the scope of unmanned systems research, development, acquisition, and employment is likely several orders of magnitude larger when considered across the DoD enterprise. A common knowledge structure, including vocabulary and semantics, can provide a foundation for knowledge representation and management at the system and enterprise levels to improve understanding of the breadth and application of DoD-funded programs within the context of the program itself as well as its contributions to the field at large. A common knowledge representation can also support educational objectives by creating a conceptual foundation facilitating deeper understanding of technical components and operational concerns. A common knowledge representation scheme also can be applied to such diverse areas of advanced unmanned system operations as mission representations, vehicle situational awareness, data sharing, rules of ethics and law, automated reasoning, mission rehearsal, mission monitoring, and post-mission analysis. In short, the CRUSER program and the DoD unmanned system program at large need a cohesive conceptual framework to relate the variety of efforts completed, ongoing, and projected in order to provide a foundation for unmanned systems management, acquisition, engineering, and education.

Even considering the size of the problem in DoD, the knowledge management problem pales when compared to understanding information across the World Wide Web (WWW, or just “Web”). One approach being applied in the Web domain leverages Semantic Web (Berners-Lee, Hendler and Lassila 2001) technologies[1] to describe and

---

[1] World Wide Web Consortium: http://www.w3.org

relate information across vastly diverse data sets. The initiative is called the Linking Open Data (LOD) project.[2] In the summer of 2015, the concept was presented to JGRE representatives who immediately recognized the potential for helping program managers to better characterize and understand DoD's unmanned systems research and development portfolio. The FY16 tasking from JGRE to NPS included an investigation of this approach in the statement of work, but at a modest level of effort with expectation of matching funds from the CRUSER program. Unfortunately, the partner proposal to CRUSER was not selected for FY16 funding. In the absence of the matching funds, only a preliminary effort to investigate the application of state-of-the-art information modeling techniques to create a knowledge representation of the unmanned systems research, engineering, acquisition, and employment domain could be performed in the FY16 JGRE project.

This report describes the technical foundation for the proposed knowledge representation work and lays out a plan of action for creating and populating the knowledge model with data from DoD unmanned systems programs if follow-on funding becomes available through one or both organizations.

B. OBJECTIVES

The objective of this initial effort is to investigate state-of-the-art information modeling techniques to create a conceptual and technical knowledge representation of unmanned systems architecture, research, development, and employment to use in better understanding unmanned systems and the extent of DoD research and education projects (of which CRUSER efforts are a subset).

C. ORGANIZATION OF THIS DOCUMENT

Chapter 1 of this report provided background, objectives, and an overview of the document. Chapter 2 introduces the technical approach and how the methodology is being used in industry and academia to address complex issues and knowledge interrelationships. Chapter 3 outlines a conceptual framework for the knowledge representation in the context of unmanned system research and development. Chapter 4 provides a plan of action and recommendations for continuation of the work to fully

---

[2] Linking Open Data project: http://linkeddata.org and http://lod-cloud.net

structure and populate the information model. The appendixes provide supporting material, including a glossary of abbreviations and acronyms and a list of references.

II. TECHNICAL FOUNDATION

A. INTRODUCTION

This chapter provides a brief introduction to the Semantic Web, its associated technologies, and application of the technologies in the Linked Open Data Project. It concludes with a discussion of what this technical approach can offer to the unmanned systems community.

B. SEMANTIC WEB

The Semantic Web was conceived by Tim Berners-Lee as a means of adding meaning to information on the web:

> "The Semantic Web is 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. The first steps in weaving the Semantic Web into the structure of the existing Web are already underway. In the near future, these developments will usher in significant new functionality as machines become much better able to process and 'understand' the data that they merely display at present." (Berners-Lee, Hendler and Lassila 2001)

> "The Semantic Web is about two things. It is about common formats for integration and combination of data drawn from diverse sources, where on the original Web mainly concentrated on the interchange of documents. It is also about language for recording how the data relates to real world objects. That allows a person, or a machine, to start off in one database, and then move through an unending set of databases which are connected not by wires but by being about the same thing." (World Wide Web Consortium 2013)

Whereas for many years information on web pages was laid out for screen presentation through stylized interpretations of Hyper Text Markup Language (HTML) tags, the Semantic Web has led to the encoding of meaning to information stored on the web, facilitating more intelligent access, usage, and reasoning on that information by computer software.

The Semantic Web is now characterized by a set of information representation technologies, exhibited in the Semantic Web stack shown in Figure 1.

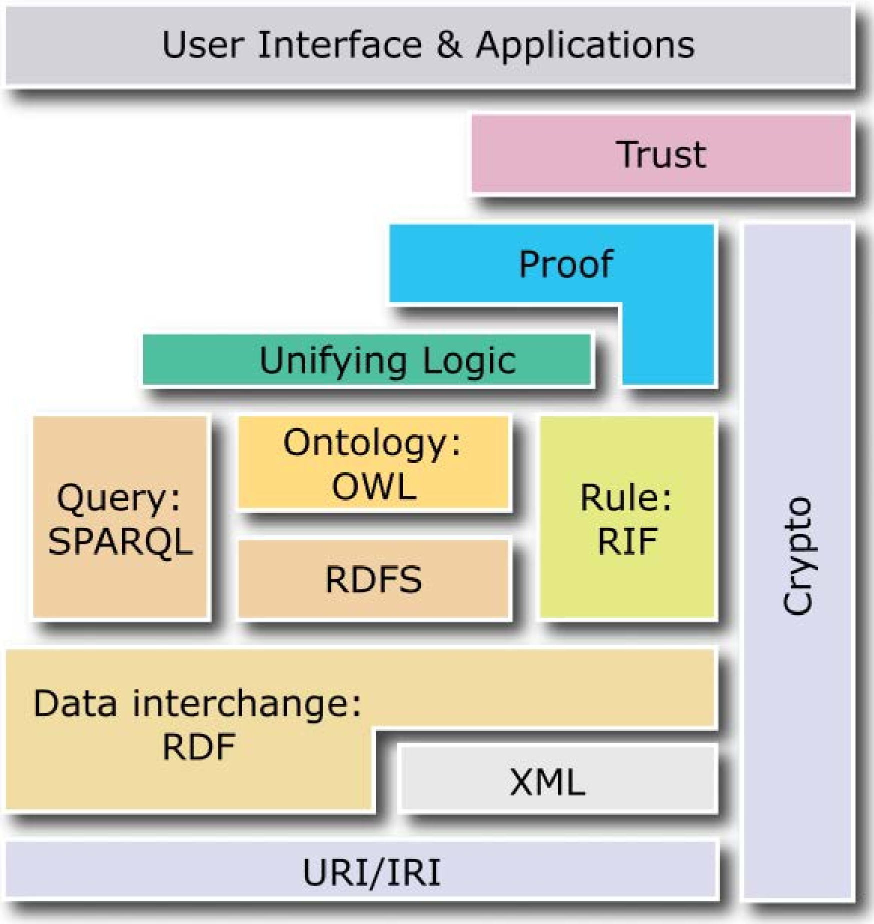

Figure 1. Semantic Web technology stack (from (World Wide Web Consortium 2007))

It is beyond the scope of the present report to describe each of these technologies in detail. A brief description of these components is provided below.

URI/IRI (Uniform Resource Identifier / Internationalized Resource Identifier) (World Wide Web Consortium 2001). The URI and IRI are standardized identification schemes for labeling information resources on the Web, to include documents, images, downloadable files, services, electronic mailboxes, and other objects.[3] The URI has two forms: (1) the Uniform Resource Name (URN) – "intended to serve as persistent location-independent resource identifiers" (Internet Engineering Task Force Network Working Group 1997); (2) the Uniform Resource Locator (URL) is the commonly used labeling format for

---

[3] http://www.w3.org/Addressing/

describing the logical or physical location of a resource on the web used by protocols such as the Hyper Text Transfer Protocol (HTTP) to access the resource. The IRI permits the use of international characters in a URI. The official registry of URI scheme names is maintained by the Internet Assigned Numbers Authority (IANA).[4]

XML (Extensible Markup Language) (World Wide Web Consortium 2016). XML is a text labeling language that is used to describe information in a document. Whereas HTML provided tags to indicate how a web browser should display information to a user, XML permits users to create their own tags to describe the information in a web document, thus creating a separation of data presentation from data content that enables greater automation in software processing of web-based information content.

RDF (Resource Description Framework) (World Wide Web Consortium 2014). RDF is a standardized structure (subject-predicate-object) for making assertions about information on the Web. Each of the three components of an RDF statement are given URI/IRI labels. When components of multiple RDF statements use the same labels, the statements can be linked together, forming graphs that can be traversed by software to respond to queries or discover relationships across diverse sets of information.

RDFS (RDF Schema) (World Wide Web Consortium 2014). RDF Schema (RDFS) is an extension to RDF to enable the specification of new vocabulary to use in describing information and data interrelationships, such as in specifying a class of objects based on a set of properties, and specifying subclasses of those classes. This is useful in organizing information for better understanding and processing.

Query. The Query layer of the Semantic Web stack provides the capability for users to search for information on the web using data content, relationships with other information, or other conditional phrases. The SPARQL Protocol and RDF Query Language (SPARQL)[5] is a W3C Recommendation for "querying and

---

[4] http://www.iana.org/assignments/uri-schemes

[5] The name is a "recursive acronym" (DuCharme 2011, 1).

manipulating RDF graph content on the Web or in an RDF store" (World Wide Web Consortium 2013).

Rule (World Wide Web Consortium 2010). A rule is an expression of a condition and an action to perform when the condition is true (a rule of this form is referred to as a *production rule*). Alternatively, a *declarative rul*e can be a statement about the world, such as "if P is true, then Q is true." Rules have long been a primary mode of knowledge expression in artificial intelligence research and development, particularly in the development of expert systems (Hayes-Roth, Waterman and Lenat 1983). The Rule Interchange Format (RIF) has been specified by the W3C for this part of the stack, with a mapping to RDF (World Wide Web Consortium 2013).

OWL (Web Ontology Language) (World Wide Web Consortium 2012). OWL builds upon RDF and RDFS to formally specify concepts and concept interrelationships (i.e., for specifying an *ontology*), as well as for expressing specific assertions about those concepts and concept interrelationships. As a formal specification following the tenets of an area of formal logic called *description logics*, (Brachman and Levesque 2004) (Davis 2014) software programs called *reasoners* can interpret OWL expressions to infer new knowledge from the assertions provided in a knowledge base.

Unifying Logic. Given the expression of information in a well-understood (standard) syntax, with well-understood (standard) semantics, a variety of software applications can be developed to process and manipulate the information across many information sources to achieve the original vision of the Semantic Web. For example, several reasoners available for OWL are identified in (Davis 2014, 67-68).

Proof. The Proof layer builds upon the application of the lower layers to enable automated and transparent reasoning on the information: "The Proof layer involves the actual deductive process as well as the representation of proofs in Web languages (from lower levels) and proof validation" (Antoniou and van Harmelen 2004, 18).

Crypto (Cryptography) (World Wide Web Consortium 2013). The Crypto layer represents a cross-cutting component that provides the means to associate information security mechanisms (identity, authentication, encryption, etc.) with the information content on the Web.

Trust (Berners-Lee 2006). Trust in web-based information is attained when there is well-understood provenance to the data (digital signature, certification, certification authorities, integrity, etc.). The application of the layers of the Semantic Web stack to address such properties as information markup, logical consistency, formal semantics, and information security contribute to the establishment of trust in the accessed information.

User Interface and Applications. The layer represents the use of the technologies to achieve user functional requirements.

As we will see in subsequent sections of this chapter and in the next chapter of this report, we will focus primarily on lower levels of the stack, specifically URI/IRI, XML, RDF, RDFS, and OWL, for the purposes of this study.

C. LINKING OPEN DATA PROJECT INITIATIVE

As more and more information becomes available on the Web using Semantic Web principles and technologies, researchers have seized the opportunity to exploit the data by examining the complex interrelationships that exist. Linked Data is an initiative inspired by a set of four rules for linked data (Berners-Lee 2006)[6]:

(1) Use URIs as names for things.
(2) Use HTTP URIs so that people can look up those names.
(3) When someone looks up a URI, provide useful information, using the standards RDF and SPARQL.
(4) Include links to other URIs so they can discover more things.

As described by (Bizer, Heath and Berners-Lee 2009, 2):

> "Linked Data refers to data published on the Web in such a way that it is machine-readable, its meaning is explicitly defined, it is linked to other external data sets, and can in turn be linked to from external data sets."

---

[6] Wikipedia defines Linked Data as "a method of publishing structured data so that it can be interlinked and become more useful through semantic queries. It builds upon standard Web technologies such as HTTP, RDF and URIs, but rather than using them to serve web pages for human readers, it extends them to share information in a way that can be read automatically by computers." (https://en.wikipedia.org/wiki/Linked_data)

The four rules for linked data "provide a basic recipe for publishing and connecting data using the infrastructure of the Web while adhering to its architecture and standards" (Bizer, Heath and Berners-Lee 2009, 2). Publishing a data set as Linked Data on the Web involves the following steps (Bizer, Heath and Berners-Lee 2009, 6):

(1) Assign URIs to the entities described by the data set and provide for dereferencing these URIs over the HTTP protocol into RDF representations.
(2) Set RDF links to other data sources on the Web, so that clients can navigate the Web of Data as a whole by following RDF links.
(3) Provide metadata about published data, so that clients can assess the quality of published data and choose between different means of access.

Going a step further, Linked *Open* Data (LOD) is defined by Berners-Lee as "Linked Data which is released under an open licence, which does not impede its reuse for free." Applying these rules, the Linking Open Data Project has created a vast collection of inter-related information resources, as pictured in Figure 2. This information is being leveraged by numerous programs (e.g., pharmaceutical studies; see (Samwald, et al. 2011)) to understand important relationships that previously have been buried in the layers of complexity and vast quantities of data stored on the web.

In light of the growing success of this initiative[7], the W3C has established a W3C Recommendation for the Linked Data Platform 1.0 (World Wide Web Consortium 2015). As the amount of linked data has grown, the community has developed and made available several tools for publishing linked data; see (Bizer, Heath and Berners-Lee 2009, 8-9).

Resources identified with a URI can be dereferenced by looking up the address using the HTTP protocol. HTTP provides the linkages for information interconnection on the Web. The content of the resources is expressed as RDF triples—subject-predicate-object assertions, where each term is represented as a URI. The predicate expresses the relationship between the subject and the object. For example, an RDF triple could describe a *has_component* relationship between a robotic system and a sensor component. In turn, there may be a *manufactured_by* relationship between that sensor and a corporation. The chaining or linking of the terms is achieved through use of a

---

[7] A 2010 compilation of statistics on the LOD showed a total of over 19 billion RDF triples in this "web of data". See, for example, https://www.w3.org/wiki/TaskForces/CommunityProjects/LinkingOpenData/DataSets/Statistics

particular URI in multiple expressions (or even through the use of a “same-as” relationship indicating two distinct URIs refer to the same concept or resource). RDFS and OWL provide the means to create vocabularies describing and relating concepts, and the terms in these vocabularies are linked through the use of RDF triples (in fact, all expressions in RDFS and OWL are essentially RDF triples in structure).

At the heart of this process is the selection of vocabularies (i.e., metadata) to describe information. The initiative has led to establishment of Linked Open Vocabularies (http://lov.okfn.org/dataset/lov/), but there is very little available there regarding robotics or unmanned systems (there is a generic reference to robots as “ComputationalAgents”). There are many taxonomies and ontologies in use on the Web that can be useful to the effort to describe DoD unmanned systems. We will discuss this in more detail below and in the next chapter.

D. APPLICATION TO THE UNMANNED SYSTEMS DOMAIN

Given the value of semantic representation of information exhibited in the Linking Open Data initiative, we can consider how this approach could apply to the unmanned systems domain. In this case, we have thousands of activities in research and development of systems, but we lack the means to characterize and examine the work that is being performed. Even reducing the scope to just DoD programs, we still face a daunting challenge of describing the work being funded, systems being procured, and operations being conducted. However, one advantage we have is the possibility of creating policy to require programs to provide information about projects in a specific format that would enable the creation of a “web of data” about unmanned systems in the DoD. What if a funding agency in DoD could examine the areas of funded research against the architectural components of unmanned systems in order to identify where focused attention is needed? What if operational commanders could examine availability of unmanned systems best suited to particular missions or to determine what missions can be best conducted by unmanned systems? These and many other perspectives could be addressed by encoding information about unmanned systems research, engineering, procurement, and employment in a manner that would enable representation of the complex interrelationships of the many efforts that have been completed, are underway, or are envisioned.

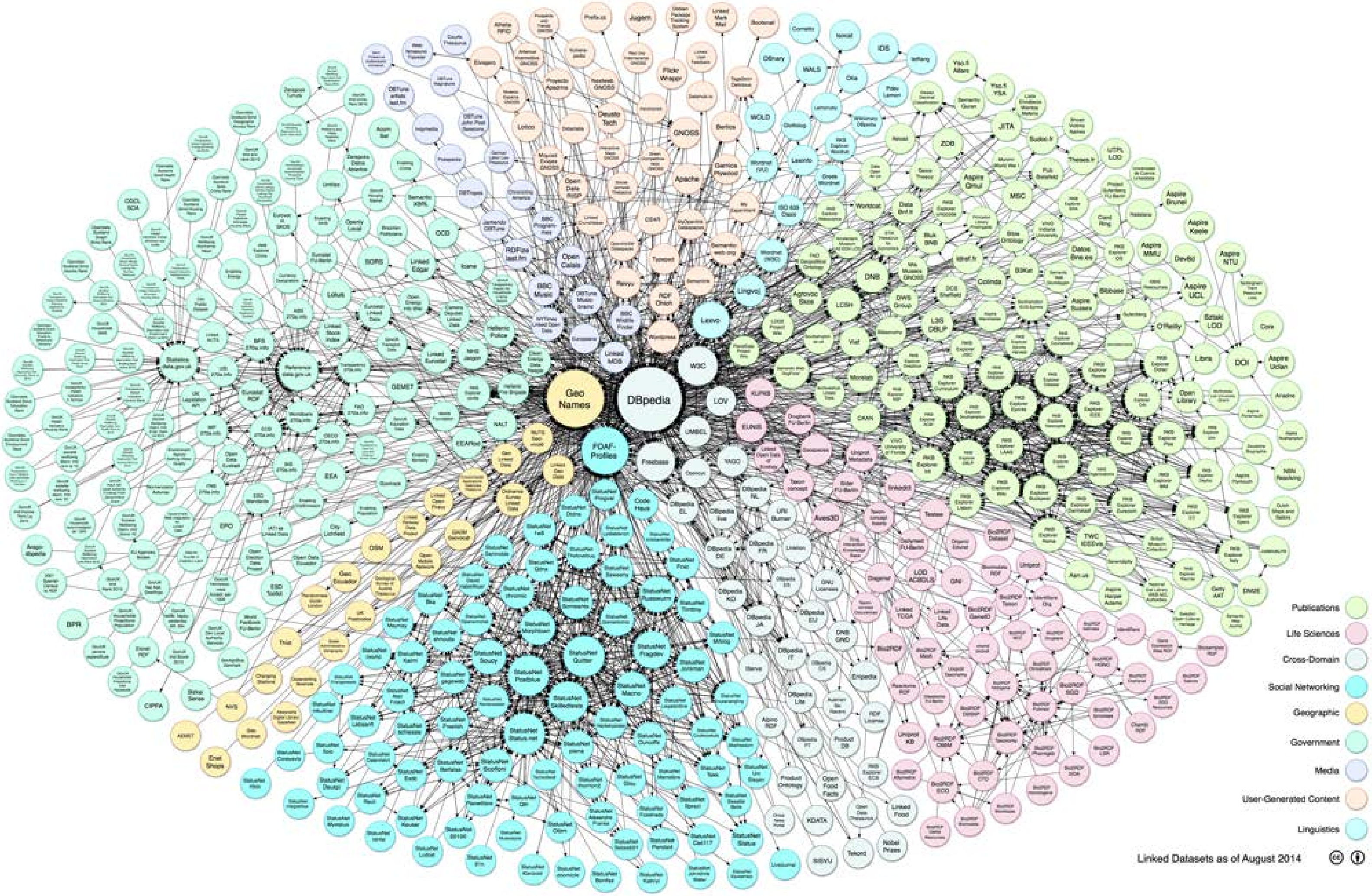

Figure 2. Example of the application of Linking Open Data (from (Schmachtenberg, et al. 2014))

The Linking Open Data technical approach recognizes that it is not possible to create a huge database accumulating all of the information about some topic, but that information sources can describe their work in ways that can identify links to related information objects. The starting point for this effort in the context of DoD unmanned systems is establishment of a vocabulary (or set of vocabularies) for characterizing what we want to know or be able to discover. This activity is discussed in the next chapter.

THIS PAGE INTENTIONALLY LEFT BLANK

III. INFORMATION MODELING

A. INTRODUCTION

There are many ways to approach development of the information model to begin accumulating data on unmanned systems research, engineering, acquisition, and employment. In this chapter, we discuss various possible starting points for the knowledge model, laying a foundation for follow-on work to create and begin populating the model. Guiding the discussion throughout is practicality—we do not want to "boil the ocean" by creating a perfect, all-encompassing knowledge model up front, but seek to lay the groundwork for an initial model upon which we can apply a process of iterative, incremental improvement as experience with the knowledge structure, data extraction to populate the model, and utility in applying the model evolve.

B. CONSTRUCTING AN INITIAL KNOWLEDGE FRAMEWORK

The fundamental question is, "What do we want to know about DoD unmanned systems?" There are several perspectives one can take. On the technical side, we are interested in unmanned system architectures—the various components that make up the system, how they are connected, what standards are used, etc. This can include not only the physical components, but also the logical components, such as computational elements (sensor data fusion, path-finding, etc.) and intercommunications (inter-component, across like systems, and across different systems). In looking at research projects, we can then relate the projects to the architectural elements—what components are addressed in the project, what is the purpose of the project (replacement, improvement, addition). From a funding and procurement perspective, we are interested in who is doing the work, who manufactured a component, who integrated a sub-system or system, what are or were the costs and schedule. From an operational perspective, we are interested in what missions or tasks the unmanned system can perform, what human operator support is needed, what are the training, logistics, and maintenance requirements. All of these perspectives are suggestive of an information model driven by practical considerations—what we want to know about DoD unmanned systems.

So, how to start? There are numerous guidelines for developing an ontology. For example, (Noy and McGuinness 2006) outline the following steps:

- *Determine the domain and scope of the ontology*. For our case, we are concerned about DoD unmanned systems, but from a number of different perspectives, as we discussed above.
- *Consider reusing existing ontologies*. For example, the information provided in the DoD Discovery Metadata Specification (DDMS)[8] and Intelligence Community Information Classification Markings (IC-ISM)[9] are good starting points for metadata about a project.
- *Enumerate important terms in the ontology*. The terms we consider important, several of which were suggested in the discussion above, relate to the information we want to know about the domain of DoD unmanned systems.
- *Define the classes and class hierarchy*. This step imposes a formal structure (e.g., using RDFS and OWL) onto the information to permit queries and reasoning on the information.
- *Define the properties of classes*. This step further defines interrelationships across concepts and provides values to attributes considered important for our domain, such as identifying the employment (e.g., *has_employment_mode*) of an unmanned system as ground, surface, sub-surface, or air.
- *Define the facets of the properties*. This step further refines the definition of properties to indicate their cardinality and other aspects (e.g., symmetric, transitive, etc.).
- *Create instances*. This step creates specific information about specific unmanned systems or unmanned system projects in accordance with the defined information model structure.

In light of our comments in Section A above, those authors make additional points that are worth keeping in mind should JGRE decide to pursue this effort:

> “1) There is no one correct way to model a domain— there are always viable alternatives. The best solution almost always depends on the application that you have in mind and the extensions that you anticipate.
>
> 2) Ontology development is necessarily an iterative process.
>
> 3) Concepts in the ontology should be close to objects (physical or logical) and relationships in your domain of interest. These are most likely to be nouns (objects) or verbs (relationships) in sentences that describe your domain.” (Noy and McGuinness 2006, 4)

---

[8] https://metadata.ces.mil/dse-help/DDMS/index.html

[9] https://www.dni.gov/index.php/about/organization/chief-information-officer/information-security-marking-metadata

In considering the encouragement to find and use existing ontologies, one would think the DoD Data Services Environment (DSE) would be a good resource. For many years, DoD requirements for data sharing have driven the creation of enterprise-level information architectures to ensure data are “visible, accessible, understandable, visible, trusted, and interoperable throughout their lifecycles for all authorized users” (Department of Defense Chief Information Officer 2013, 2). Communities of Interest (COIs) have been established to determine information commonalities and data sharing needs in various domains within the DoD. However, a quick examination of the COIs registered with the DSE[10] reveals no COI focused on the growing area of robotic/unmanned systems. Arguably, of course, robotic systems represent a cross-cutting concern, of interest to communities such as Logistics and Undersea Warfare. On the other hand, there is a COI for Modeling and Simulation that is also cross-cutting, so it is reasonable to take the position that unmanned systems should have a similar focus in the enterprise. There are references to, for example, the term “unmanned” in various taxonomies and schemas relating to air operations, joint air and missile defense, CBRN (chemical, biological, radiation, nuclear), intelligence, and command and control, but they are scattered and, in several cases, found in “retired” or “deprecated” metadata resources. One of the primary sources of relevant metadata found on the DSE is the Autonomous Vehicle Command Language (AVCL) developed at NPS and posted several years ago to the metadata registry. Overall, though, we would conclude at the moment that there appears to be no coherent representation of unmanned systems in general in the metadata registry providing a basis for common understanding and information processing automation across the enterprise. We can certainly leverage what is available for purposes of the linked open data vocabulary, while also considering the benefit of providing our results back to the DSE for others to discover and use.

Regarding the architecture of unmanned systems, (Paull, et al. 2012) identify generic components of autonomous systems, as shown in Figure 3, arguing that any ontology of autonomous systems must represent these components. Similarly, the Advanced Explosive Ordnance Disposal Robotics System (AEODRS) (Del Signore 2015) identifies generic components (modules) and their interconnections (see Figure 4),

---

[10] As of September 2016.

although more specifically for EOD unmanned ground vehicles (UGVs). Either of these approaches (or a combination) can be used as a starting point that can be adapted to a for purposes of describing DoD unmanned systems.

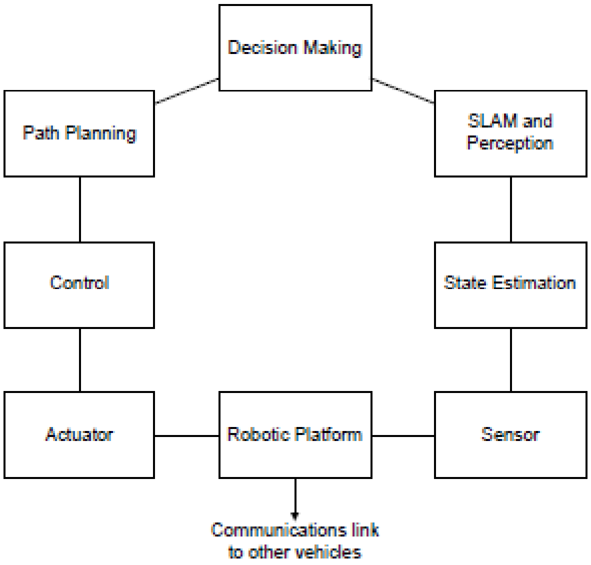

Figure 3. Components of an autonomous vehicle architecture (from (Paull, et al. 2012, 1360))

The Navy Objective Architecture provides a generic structure for Naval systems that can be a starting point for a general architectural perspective on unmanned systems. In (Johnson and Blais 2008), the authors use OWL to create a taxonomy describing components of the Navy Surface Combat System Top-Level Objective Architecture, as well as an ontology describing computational components (Common Systems Function List). A top-level view of the objective architecture and its representation in an ontology is shown in Figure 5. While not specifically addressing the unmanned systems domain, the approach taken can certainly be adapted to the DoD unmanned systems domain.

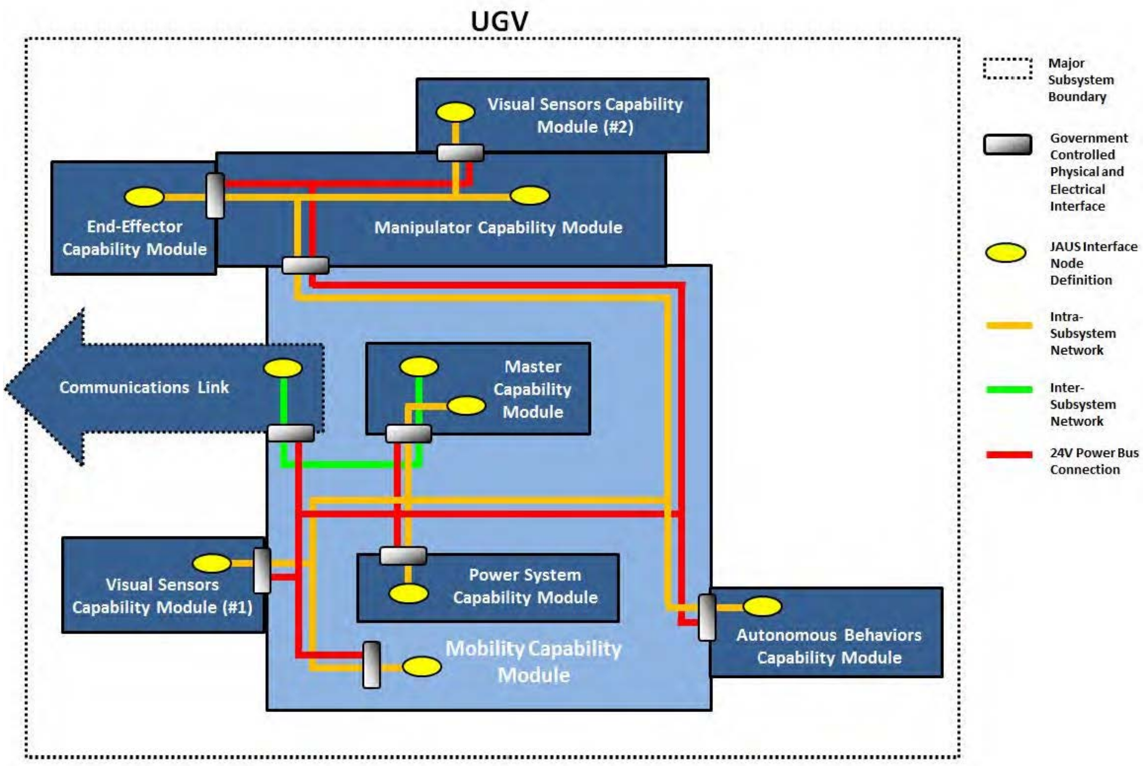

Figure 4. AEODRS common architecture (from (Del Signore 2015, 10))

Creation of the ontology can include description of the level of autonomy of the system, as well as characterizations of the role of the system in working with human operators, following the definitions provided in (National Institute of Standards and Technology October, 2008). As mentioned previously, the DDMS and IC-ISM standards can be applied to capture much of the metadata regarding agents and organizations performing the work and classification restrictions on the activities, products, or applications (employment). With respect to employment of the systems, established approaches such as the Universal Joint Task List (UJTL)[11] may provide a starting point for identifying the kinds of activities performed by the unmanned systems in military operations. We can also expect to create cross-relationships, such as which components of a system are used in particular missions. The creation and use of ontologies are very flexible—if the concepts of interest are defined in multiple ontologies, concepts from one ontology can be referenced in another.

---

[11] http://www.dtic.mil/doctrine/training/ujtl_tasks.htm

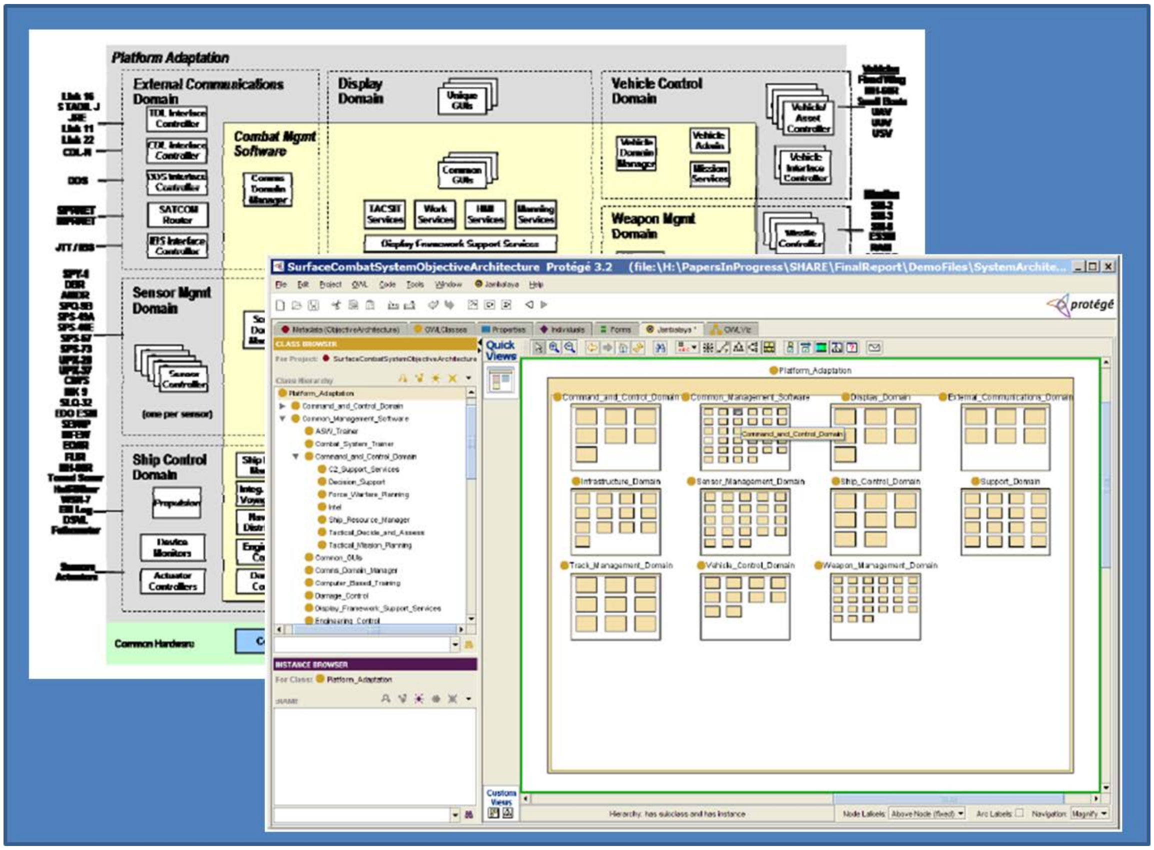

Figure 5. Surface Combat System Top-Level Objective Architecture (background) and an ontology representation in OWL rendered in the Protégé ontology editing tool (foreground) (from (Johnson and Blais 2008, 76-77))

Once the knowledge structures are created in a representation, such as OWL, specific instances describing particular research or development activities or systems and system components can be created and expressed as RDF triples to start populating the web of data. Available tools can be used to publish the data to web sites (whether in the open or in controlled network environments, depending on the nature of the information encoded, such as classification level) and we can begin experimenting with queries and other analytical actions to evaluate the utility of the information and information structure. As these foundations are laid, the process will iterate to build up content and utility over time. The ultimate goal will be to have a web of data describing DoD unmanned systems that can be used by program managers and decision makers to better comprehend the breadth of activities and products in the domain, and to better manage ongoing and future research, engineering, development, acquisition, and employment.

IV. PLAN OF ACTION FOR FOLLOW-ON WORK

A. TECHNICAL APPROACH

This research effort is intended to immediately stimulate the community toward common vocabulary and formal knowledge representation to create a basis for greater understanding across technical and educational endeavors, in order to see individual program efforts as a part of an overall knowledge enterprise rather than as separate, independent efforts. The resulting view of past, current, and future contributions to the unmanned systems field will help identify new directions for education, research, and procurement in DoD.

The technical approach for creating a web of data for DoD unmanned systems involves the following activities:

- Synthesis. Create a taxonomy/ontology (metadata description) of the architecture of unmanned systems, research and development efforts, and operational employment. Creation and iteration on the knowledge structure should consider existing taxonomies (e.g., Dublin Core, DoD Discovery Metadata Specification, Navy Open Architecture) and ontologies relevant to the unmanned system domain (e.g., see (Paull, et al. 2012).
- Implementation. Populate the model to describe specific work performed, possibly beginning with the limited scope of the NPS CRUSER program but building out to the DoD unmanned systems research and development portfolio.
- Demonstration and Evaluation. At determined intervals, evaluate effectiveness of the emerging model and encoded information at conveying information that is of value to program managers and decision makers (e.g., by identifying and applying a set of study questions).
- Iterative Improvement. Iterate on the process, incrementally improving the knowledge structure and continuing to build up the information content.
- Sustainment. Create mechanisms, potentially leading to formal policy, enabling generation of linked data as research, engineering, acquisition, and employment activities occur.

B. RECOMMENDATIONS

This report has laid a conceptual foundation for describing unmanned system architectural components and associated research and development projects to enable program managers and high level decision-makers to obtain a better understanding of the coverage of their funds and planned efforts. This was only a preliminary investigation to possibly generate interest in a more extensive effort to establish the knowledge representation and populate the model with data. We recommend that the JGRE take this to the next level to create the information model and begin populating it with data, perhaps starting with the limited scope of the projects that have been funded through the NPS CRUSER program, in order to reveal vocabulary, links, and usage that demonstrate greatest potential benefit to DoD program managers and decision-makers.

APPENDIX A. GLOSSARY OF ACRONYMS AND ABBREVIATIONS

AEODRS	Advanced Explosive Ordnance Disposal Robotics System
CBRN	Chemical, Biological, Radiation, Nuclear
COI	Community of Interest
CRUSER	Consortium for Robotics and Unmanned System Education and Research
DDMS	DoD Discovery Metadata Specification
DoD	Department of Defense
DSE	Data Services Environment
EOD	Explosive Ordnance Disposal
FY	Fiscal Year
HTML	Hyper Text Markup Language
HTTP	Hyper Text Transfer Protocol
IANA	Internet Assigned Numbers Authority
IC-ISM	Intelligence Community Information Security Marking
IETF	Internet Engineering Task Force
IRI	Internationalized Resource Identifier
JGRE	Joint Ground Robotics Enterprise
LOD	Linking Open Data; Linked Open Data
LOV	Linked Open Vocabularies
NPS	Naval Postgraduate School
OSD	Office of the Secretary of Defense
OWL	Web Ontology Language
RDF	Resource Description Framework
RDF-S	Resource Description Framework Schema
RIF	Rule Interchange Format
SPARQL	SPARQL Protocol and RDF Query Language
UGV	Unmanned Ground Vehicle
UJTL	Universal Joint Task List
URI	Uniform Resource Identifier
URL	Uniform Resource Locator
URN	Uniform Resource Name
W3C	World Wide Web Consortium
WWW	World Wide Web
XML	Extensible Markup Language

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

Antoniou, Grigoris, and Frank van Harmelen. *A Semantic Web Primer.* Cambridge: The MIT Press, 2004.

Berners-Lee, Tim. *Linked Data.* July 27, 2006. http://www.w3.org/DesignIssues/LinkedData.html (accessed September 10, 2016).

—. *Trust.* July 16, 2006. http://www.w3.org/2000/10/swap/doc/Trust (accessed September 26, 2016).

Berners-Lee, Tim, James Hendler, and Ora Lassila. "The Semantic Web." *Scientific American*, May 2001.

Bizer, Christian, Tom Heath, and Tim Berners-Lee. "Linked data - the story so far." *International Journal on Semantic Web and Information Systems (IJSWIS), Special Issue on Linked Data* 5, no. 3 (2009): 1-22.

Blais, Curtis L. *Unmanned Systems Interoperability Standards.* Technical Report NPS-MV-16-002, Monterey: Naval Postgraduate School, 2016.

Blais, Curtis L., and Don McGregor. *Modeling Manned and Unmanned Systems.* Technical Report NPS-MV-16-001, Monterey: Naval Postgraduate School, 2016.

Brachman, Ronald J., and Hector J. Levesque. *Knowledge Representation and Reasoning.* San Francisco: Elsevier, Inc., 2004.

Davis, Duane. *Semantic Web and Inferencing Technologies for Department of Defense Systems.* Technical Report NPS-CMIS-14-001, Monterey: Naval Postgraduate School, 2014.

Defense Science Board. *Summer Study on Autonomy.* Department of Defense Office of the Under Secretary of Defense for Acquisition, Technology and Logistics, 2016.

Del Signore, Michael. "An open system architecture approach for EOD unmanned vehicles." *Interoperable Open Architecture.* London, 2015.

Department of Defense Chief Information Officer. "Sharing Data, Information, and Information Technology (IT) Services in the Department of Defense." DoD Instruction 8320.02, 2013.

DuCharme, Bob. *Learning SPARQL.* Sebastopol: O'Reilly Media, Inc., 2011.

Hayes-Roth, Frederick, Donald A. Waterman, and Douglas B. Lenat. "An Overview of Expert Systems." In *Building Expert Systems*, edited by Frederick Hayes-Roth, Donald A. Waterman and Douglas B. Lenat, 3-29. Reading: Addison-Wesley Publishing Company, Inc., 1983.

House of Representatives Committee on Appropriations. *Department of Defense Appropriations Bill for 2017.* House Report 114-577, Washington, D.C.: United States Government, 2016.

Internet Engineering Task Force Network Working Group. *URN Syntax.* May 1997. http://www.ietf.org/rfc/rfc2141.txt (accessed September 26, 2016).

Johnson, Jean, and Curtis Blais. *Software Hardware Software Asset Resource Enterprise (SHARE) Repository Framework Final Report: Component Specification and Ontology*. NPS-AM-08-134, Acquisition Research Sponsored Report Series, Monterey: Naval Postgraduate School, 2008.

National Institute of Standards and Technology. "Autonomy Levels for Unmanned Systems (ALFUS) Framework, Volume 1: Terminology, Version 2.0." Edited by Hui-Min Huang. October, 2008.

Noy, Natalya F., and Deborah L. McGuinness. "Ontology Development 101: A Guide to Creating Your First Ontology." *Protege: A free, open-source ontology editor and framework for building intelligent systems.* 2006. http://protege.stanford.edu/publications/ontology_development/ontology101.pdf (accessed September 10, 2016).

Paull, Liam, et al. "Towards an ontology for autonomous robots." *Proceedings of the 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS).* 2012. 1359-1364.

Samwald, Matthias, et al. "Linked open drug data for pharmaceutical research and development." *Journal of Cheminformatics* 3, no. 10 (2011).

Schmachtenberg, Max, Christian Bizer, Anja Jentzsch, and Richard Cyganiak. "Linking Open Data cloud diagram 2014." *The Linking Open Data cloud diagram.* August 30, 2014. http://lod-cloud.net (accessed September 10, 2016).

Schumann, Benjamin, Jim Scanlan, and Hans Fangohr. "A generic unifying ontology for civil unmanned aerial vehicle missions." *Proceedings of the 125h AIAA Aviation Technology, Integration, and Operations (ATIO) Conference and 14th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference.* American Institute of Aeronautics and Astronautics, 2012.

World Wide Web Consortium. *Extensible Markup Language (XML).* October 11, 2016. http://www.w3.org/XML/ (accessed October 15, 2016).

—. *Latest layercake diagram.* 2007. http://www.w3.org/2007/03/layerCake.png (accessed September 20, 2016).

—. *Linked Data Platform 1.0 W3C Recommendation.* February 26, 2015. http://www.w3.org/TR/2015/REC-ldp-20150226/ (accessed September 10, 2016).

—. *OWL 2 Web Ontology Language Primer (Second Edition) W3C Recommendation.* December 11, 2012. http://www.w3.org/TR/2012/REC-owl2-primer-20121211/ (accessed Septemberd 26, 2016).

—. *RDF 1.1 Concepts and Abstract Syntax W3C Recommendation.* February 25, 2014. http://www.w3.org/TR/2014/REC-rdf11-concepts-20140225/ (accessed September 10, 2016).

—. *RDF Schema 1.1 W3C Recommendation.* February 25, 2014. http://www.w3.org/TR/2014/REC-rdf-schema-20140225/ (accessed September 10, 2016).

—. *RIF in RDF (Second Edition) W3C Working Group Note.* February 5, 2013. http://www.w3.org/TR/rif-in-rdf/ (accessed September 26, 2016).

—. *RIF Primer (Second Edition).* 2010. http://www.w3.org/TR/2013/REC-sparql11-overview-20130321/ (accessed September 26, 2016).

—. *SPARQL 1.1 Overview W3C Recommendation.* March 23, 2013. http://www.w3.org/TR/2013/REC-sparql11-overview-20130321/ (accessed September 10, 2016).

—. *URIs, URLs, and URNs: Clarifications and Recommendations 1.0.* September 21, 2001. http://www.w3.org/TR/uri-clarification/ (accessed September 26, 2016).

—. *W3C Semantic Web Activity*. 2013. http://www.w3.org/2001/sw/ (accessed September 20, 2016).

—. *XML Security Working Group*. 2013. http://www.w3.org/2008/xmlsec/ (accessed September 26, 2016).

THIS PAGE INTENTIONALLY LEFT BLANK

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. Shad Reese
   Office of the Secretary of Defense Joint Ground Robotics Enterprise
   Washington, DC

5. Jaret Riddick
   Office of the Secretary of Defense Joint Ground Robotics Enterprise
   Washington, DC