Use of Natural Language Processing (NLP) to Extract Technical Competency Frameworks from Maintenance Task Analyses (MTA)

Wayne Gafford
OPNAV N12
Arlington, Virginia
wayne.gafford@navy.mil

Jeanne Kitchens
Credential Engine
Washington, DC
jkitchens@credentialengine.org

Fritz Ray
Eduworks
Portland, Oregon
fritz.ray@eduworks.com

ABSTRACT

Maintenance, operations, and troubleshooting competencies for engineered systems must be derived from authoritative data sources much like maintenance task analyses (MTA). Data sources are constantly updated through changes to engineering design and support solutions. Competencies link systems and the work to people, producing essential sources for human and system performance analysis. In the Navy, there are gaps between authoritative data sources used in logistics and technical curriculum that if closed, can greatly improve and accelerate readiness by actively deriving competency frameworks from MTA using natural language processing (NLP).

To bridge the gaps, the Navy and Credential Engine signed a Cooperative Research and Development Agreement (CRADA) to map the GEIA 0007 Logistics Product Data and the S3000L Logistic Support Analysis specifications to the Credential Transparency Description Language-Achievement Standards Network (CTDL-ASN) specification. The CRADA team developed software to convert MTA data managed in the specifications into Linked Open Data (LOD). LOD extracts data from the respective tasks, then uses NLP to mold content into learning objective grammar (audience, behavior, condition, degree) structured in CTDL-ASN for use in training and Product Lifecycle Management (PLM) systems. The resulting competency frameworks mirror product structures and the associated tasks in the logistics specification and are linked through system-unique identifiers to form a "digital thread". The software analyzes the competency framework in the CTDL-ASN and outputs a corresponding course structure in S1000D, an international technical data specification.

Manual job duty task analysis (JDTA) is substituted by an iterative cycle of decision support driven by NLP for competency framework development. The process binds system design and support work to learning using data standards enabling real-time identification and turnaround of curriculum impacted by engineering design and support changes. This paper describes the process.

ABOUT THE AUTHORS

Wayne Gafford is a Program Analyst at OPNAV N12, Total Force Manpower, Training, & Education Division. He currently leads innovations, projects, and improvements through Navy initiatives that integrate training needs analysis with product support analysis. It involves interfacing workforce, training, and Navy system data through data standards, semantic technologies, artificial intelligence, and commercial products. Ensuring credentials and occupational standards are connected to training and system requirements are fundamental.

Fritz Ray is the Director of Engineering at Eduworks Corporation, and Lead Contributor to the open source Competency and Skills System (CaSS) Project. He currently specializes in Linked Open Data and HTML5-based solutions for education, training, research, application development, modeling, security, federated systems, and has a deep yearning to see information de-siloed. Mr. Ray is a periodic contributor to several learning specification bodies including IEEE LTSC, Schema.org, xAPI, and CTDL/ASN. He holds a B.S. in Software Engineering Technology from the Oregon Institute of Technology. Portland, Oregon is home.

Jeanne Kitchens is the lead technology integrator at Credential Engine. Over the past 20 years, she and her teams have designed and implemented career and workforce development programs using technologies to address workforce development challenges. Current initiatives and programs include: project management for Credential Engine technologies including Credential Registry Services and the Credential Transparency Description Language; facilitating a data workgroup for the Job Data Exchange (JDX); managing the open data standards project for T3 Innovation Network, and developing the Talent Pipeline Management Web Tools for US Chamber of Commerce, and managing the Illinois workNet Web Portal System and Illinois Open Educational Resources for State of Illinois.

Use of Natural Language Processing (NLP) to Extract Technical Competency Frameworks from Maintenance Task Analyses (MTA)

Wayne Gafford
OPNAV N12
Arlington, Virginia
wayne.gafford@navy.mil

Jeanne Kitchens
Credential Engine
Washington, DC
jkitchens@credentialengine.org

Fritz Ray
Eduworks
Portland, Oregon
fritz.ray@eduworks.com

INTRODUCTION TO THE PROBLEM

Technical training covering maintenance, operation, and troubleshooting of engineered systems represents sixty percent (60%) of Navy curriculum. Curriculum is developed in text, graphics, multimedia, models, simulations, and other formats. Training must translate to system, human, and mission readiness. Readiness is enhanced by directly linking learning data to the related work. This paper asserts that the birthplace for Navy technical training requirements for engineered systems is in the task analysis documentation. However, authoritative product data in industry standards is not documented in the form of technical learning objectives. MTA contains terms, short phrases, and single character identifiers to communicate maintenance and support concepts. NLP mining MTA assist training analysts with decision support to accurately identify and build all relevant learning objectives. This paper describes how NLP can translate MTA data in industry data standards into sentence structures that represent performance and learning objectives. Product lifecycle management practices are essential for NLP to mine MTA data for competency framework generation. A short primer of PLM is needed before discussing how industry data standards and NLP work together.

Product Life Cycle Management (PLM): Where MTA and NLP Work Together

Technical curriculum can be managed like all product data. The term “digital thread” is a popular way to conceptualize PLM. It is a method of linking engineering and readiness models to MTA and other technical data such that changes to one inform the next. Technical competencies and related curriculum should be threaded into the fabric of PLM. **Figure 1** illustrates where NLP fits in the conceptual product lifecycle model. Training needs analysis, curriculum development, and analytics are displayed in red. Each level within the conceptual model is further elaborated in the following subsections and describes the links between MTA and its competencies.

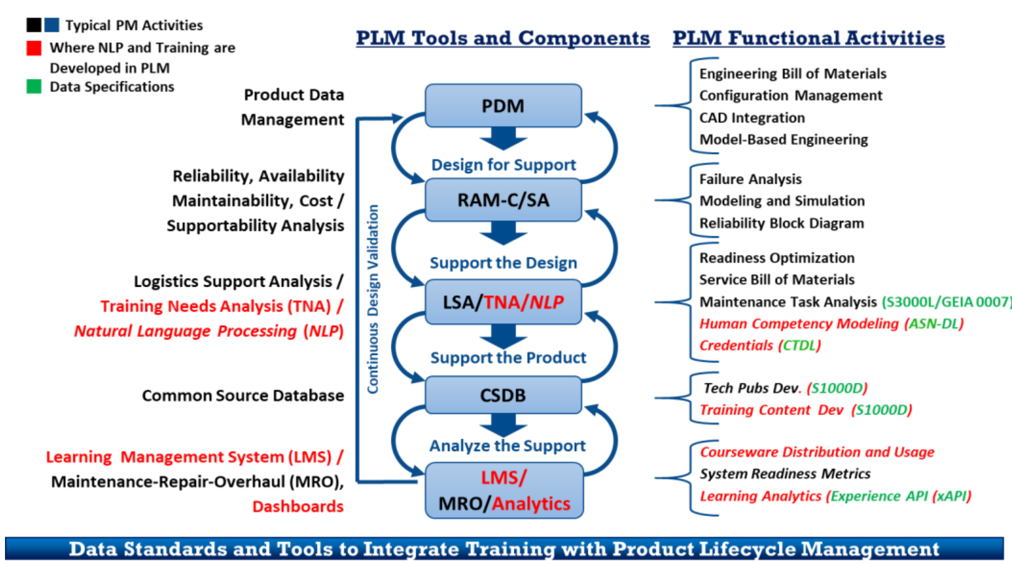

Figure 1. Technical Training Requirements and Development in Product Lifecycle Management

Product Data Management: ***Design for Support***
The top-level PLM component is product data management (PDM). This consists of activities that include: engineering models, bills of material, and configuration development. From a system readiness perspective, PDM is where engineers design for support. Modular components designed for easy access, maintenance, and replacement of parts and assemblies are essential to designing for support. The computer-aided-drawing (CAD) models are at the root of the "digital twin", a digital representation of the system.

Reliability, Availability, Maintainability, and Cost (RAM-C): ***Design for Support, Continued***
The second level PLM component goes deeper into system readiness considerations by analyzing and reaching optimal RAM-C targets. Using the digital twin to perform system fault, failure, and repair analysis, PLM steps closer to training needs analysis (TNA) and NLP by establishing how the system's design is to be maintained for optimal performance. TNA originates in the operation and maintenance of a system's functional purpose. Not until maintenance, operation, and troubleshooting requirements are documented in the next PLM level can NLP be used to extract competency frameworks.

Logistical Support Analysis (LSA): ***Support the Design***
The third level PLM component, LSA, is where the support for the design is defined and where TNA using NLP begins. The essential knowledge, skills, and abilities required for a maintainer is identified in the MTA through references to tools, personnel, skills, tasks, subtasks, and other systems. At the heart of this paper, algorithms and NLP select MTA data properties to derive content that is directly molded into the grammar of performance, terminal, and enabling learning objectives. When fully assembled, the knowledge, skills, and abilities (KSAs) represent a *competency framework*. The MTA properties used in the algorithms and in the NLP to generate competency frameworks.

The Common Source Database: ***Support the Product***
The fourth level PLM component is the common source database (CSDB). The model pivots from supporting the design in the LSA, to supporting the product with technical documentation. Just as the MTA is used as an authoritative source for the NLP-generated competency model and corresponding curriculum, it is also the source for the technical manuals. This paper advocates that technical curriculum and technical manuals be developed and managed together in the CSDB. Currently, there is no real-time notification of system change proposals to the Navy training community. It causes built-in latency and inaccuracies in the curriculum life cycle process.

Data-Driven Analytics: ***Analyze the Support***
The fifth level PLM component is the most elusive *and* the most desired for training analytics. To date, evaluation of training effectiveness takes the form of student surveys. Surveys do not collect performance data in the fleet that can be compared to learning data collected in the classroom. The main challenge is to build activity stream architectures that deliver raw data to learning record stores for comparing competency models generated through NLP.

Solution Development in a Formal Mechanism
The previous section articulated the problem space in Navy technical training and where NLP fits into the PLM conceptual model. A formal organization was needed with the right leadership and subject matter experts to come together and work toward an effective solution to link MTA and discrete learning objectives. The next section describes how a Navy product support office in California supporting fleet system life cycles collaborated with a non-profit in Washington, DC supporting credential and competency metadata to explore solutions.

COOPERATIVE RESEARCH AND DEVELOPMENT AGREEMENT (CRADA) TO SUPPORT NLP

To help solve the training data management problem, the Naval Surface Warfare Center Port Hueneme Division (NSWC PHD) entered into a CRADA with Credential Engine in April 2018. A CRADA is an agreement to contribute resources for analysis and software development that will lead to a technology transfer to the Navy and to the public. The CRADA has evolved into a formal Navy initiative to improve the overall lifecycle management of Navy technical curriculum by integrating it into PLM systems. The CRADA signatories are described below.

Credential Engine

Credential Engine pursues its mission by promoting an open applications marketplace through maintenance of the open-licensed Credential Registry and the *Credential Transparency Description Language* (CTDL) (Sutton & Kitchens, 2019). CTDL is a metadata schema based on Resource Description Framework (RDF) principles. The CTDL Profile of Achievement Standards Network (CTDL-ASN) is a companion to the CTDL providing a metadata schema for descriptions of KSAs. CTDL-ASN links credentials to competencies.

Naval Surface Warfare Center Port Hueneme (NSWC PHD)

NSWC PHD supports the U.S. Navy's combat and weapon systems. The Navy's PLM initiatives originated at NSWC PHD and has now grown into an ambitious overhaul of system lifecycle management referred to as Model-Based Product Support (MBPS). Training Support Agents (TSAs) at NSWC PHD are currently reviewing and updating what they need to know to operate within the Navy's new PLM business processes. The CRADA initiative will roll up into MBPS. NSWC PHD actively promotes data standards that improve system life cycle management.

Data Standards Descriptions Used in the CRADA Project

Using industry data standards is foundational for NLP processes to reliably and repeatedly extract competencies for air, land, and sea system maintenance tasks. **Table 1** identifies and describes the public sector data standards for MTA, competencies, curriculum, and credentialing used in the CRADA. Within the CRADA project, these data standards are used integratively along the PLM digital thread to solve data management issues in Navy technical training.

Table 1. Industry Data Standards Used by NLP to Generate and Support Competency Frameworks

Data Standards	*Publisher*	*Description*
S3000L	AeroSpace and Defense Industries Association of Europe (ASD) Aerospace Industries Association (AIA)	The ASD/AIA S3000L specification describes the LSA process, which is one of the most important processes to realize the requirements of Integrated Logistics Support (ILS). S3000L defines requirements for product support analysis and configuration management. It includes an information model and data exchange format governing the performance of LSA throughout the lifecycle of any complex technical system. See http://www.s3000l.org. NLP mines and extracts data from S3000L to form competency statements.
GEIA 0007	SAE International	GEIA 0007 defines logistics product data generated during the requirements definition and design of a system, end item, or product. This standard is for use by industry and government activities. There are 108 entities and 600 data types in the standard. See https://www.sae.org/standards/content/geiahb0007/. NLP mines and extracts data from GEIA 0007 to form competency statements.
CTDL/ CTDL-ASN	Credential Engine	The *Credential Transparency Description Language (CTDL)* is a schema comprised of terms that are useful in making assertions about a Credential and its relationships to other entities such as competencies. The vocabulary is a set of terms, a set in which the members are properties, classes, concept schemes, and/or data types. The CTDL is modeled as a directed graph using the W3C's Resource Description Framework (RDF) for describing data on the Web and for application processing. See https://credreg.net/ctdl/handbook. The NLP-generated competency frameworks are deposited in CTDL-ASN

Data Standards	***Publisher***	***Description***
S1000D	ASD AIA ATA e-Business Program	S1000D is an international specification for the production of technical publications. The S1000D steering committee introduced support for technical curriculum in 2008 and later updated in 2010. Currently, its uses include: Defense systems – including land, sea, and air products; Civil aviation products; Construction industry products, and Ship industry products. See www.s1000d.org. Course structures and content are developed in S1000D.

Figure 2 illustrates the S3000L product support model used by NLP to encode refrigeration unit learning objectives, one of the CRADA sample data sets. Data from the properties in the model is reformulated into a hierarchy of learning objectives that link to the corresponding maintenance task statements.

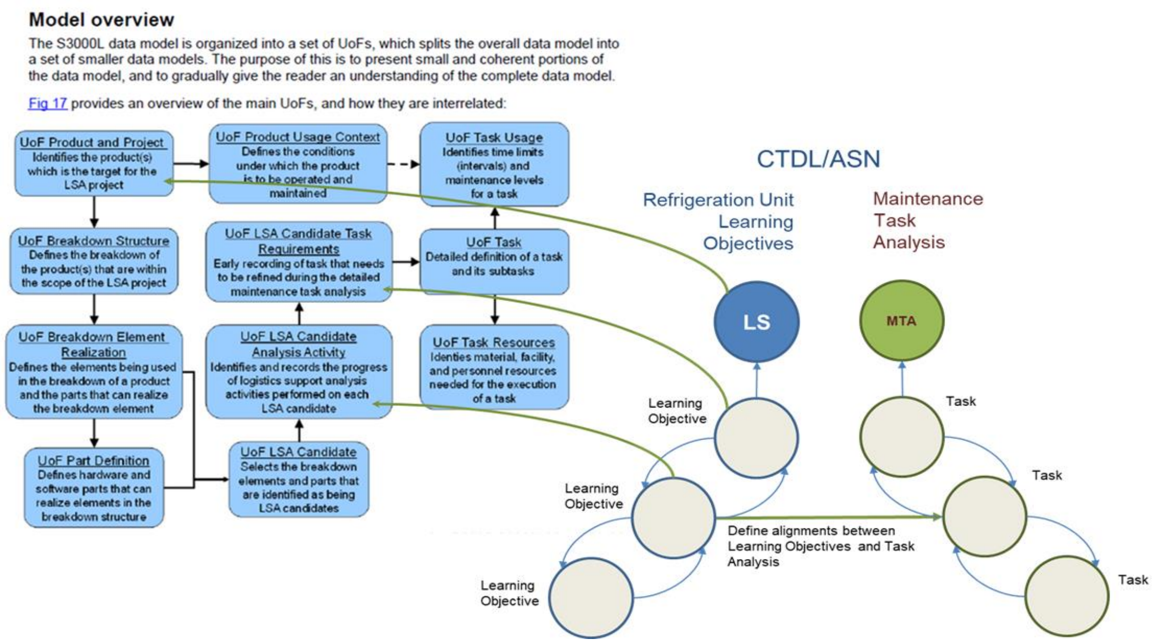

Figure 2. S3000L Data to CTDL-ASN with Alignment to Maintenance Task Analysis

CRADA Objectives, Mappings, and Outputs

Based on the assertion that technical learning objectives *emerge* from MTA, Credential Engine and NSWC PHD created the following CRADA objectives:

(a) Identify GEIA 0007 & S3000L classes and properties that contain information about systems, subsystems, tasks, tools, time, and other information determined that are relevant to generating learning objectives and express those objectives in CTDL-ASN classes and properties.

(i) Use Logistical Control Numbers (LCN) in derived competency statements to maintain the links back to task statements.

(ii) Encode and store generated frameworks as Linked Open Data (LOD), a World Wide Web Consortium (W3C) specification for machine-actionable data on the Semantic Web.

(b) Convert GEIA 0007, S3000L, and S1000D to LOD and provide an interface to explore and download the GEIA 0007, S1000D, and S3000L data standards.

(c) Generate learning objectives expressed as Terminal Learning Objectives from maintenance task analysis in GEIA 0007 and S3000L by using algorithms and Natural Language Processing (NLP).

(d) Generate course frameworks in S1000D for curriculum development from the NLP-derived learning objectives in the CTDL-ASN.

GEIA 0007 and S3000L Classes and Properties Used to Formulate Learning Objective Statements

After careful analysis, the CRADA team determined that the following classes and properties in GEIA 0007 and S3000L hold content that can be regarded as *audience, behavior, condition, and degree* components of a technical learning objective. **Tables 2 and 3** list data elements from the GEIA 0007 and S3000L used by NLP to write learning objectives that are expressed as the CTDL-ASN and encoded as LOD. GEIA 0007 and S3000L are similar but separate product support specifications owned by different organizations. The CRADA used both specifications independently to illustrate the processes.

Table 2. GEIA 0007 Task Classes and Properties Used by NLP to Formulate Learning Objectives

XA **End Item Acronym Code**	**XB** **Logistics Support Analysis Control Number Indentured Item**	**CA** **Task Requirement**	**CB** **Subtask Requirement Narrative**	**CL** **Task Subtask Associated Narrative**	**HA** **Item Identification**
• End Item Acronym Code	• Logistics Support Analysis Control Number • Alternate Logistics Support Analysis Control Number Code • Logistics Support Analysis Control Number Type • Logistics Support Analysis Control Number Nomenclature	• Task Code (7 encoded meanings) • Task Identification • Hardness Criticality Code • Hazardous Maintenance Procedure Code • Mean Man Hours • Mean Elapsed Time • Task Criticality Code • Hazardous Maintenance procedure Code • Task Frequency • Number of Subtask Personnel	• Subtask Number • Subtask Identification • Subtask Description	• Subtask Warnings • Subtask Cautions	• Reference Number/ Commercial and Government Entity Code • Item Name

Table 3. S3000L Task Classes and Properties Used by NLP to Formulate Learning Objectives

Product / Product Variant	**Breakdown / Breakdown Element**	**Part As Designed**	**Task**	**Subtask**	**Task/Subtask Personnel Resource**	**Task/Subtask Resource**
• Product Identifier • Product Variant Identifier	• Breakdown Type • Breakdown Revision • Breakdown Element Identifier • Breakdown Element Revision • Breakdown Element Essentiality • Breakdown Element Name • Breakdown Element Realization • Breakdown Element In Zone	• Part Identifier • Part Name • Part As Designed Parts List • Substitute Part	• Task Identifier • Task Revision • Task Name • Information Code • Allocated Maintenance Level/Location • Task Personnel Safety Criticality • Task Product Integrity Criticality • Task Operability Impact • Warnings and Cautions • Task Frequency • Applicability	• Subtask Identifier • Subtask Timeline • Subtask Maintenance Location • Subtask In Zone • Subtask Acceptance Parameter • Applicability	• Task Total Labor Time • Task Number Of Personnel Resource	• Material Resource (Support Equipment) • Facility Resource • Applicability

Learning objective generation requires the GEIA 0007 and S3000L data standards and parsing their data exchange files produced from maintenance task analysis software. The sample files meant for testing purposes used system and task information for a refrigeration unit and an engine. **Figure 3,** Specification Mapping Between Maintenance Task Analysis, Competency Framework Structure, and Course and Content Structure, is a high-level product data mapping model. The learning objective framework mirrors the product and task structure. Any training asset in any format identified through S1000D can be referenced by the S1000D course framework. The course structure mirrors the learning objective framework. Each data item is linked on the digital thread through the logistical control number (LCN) in the MTA.

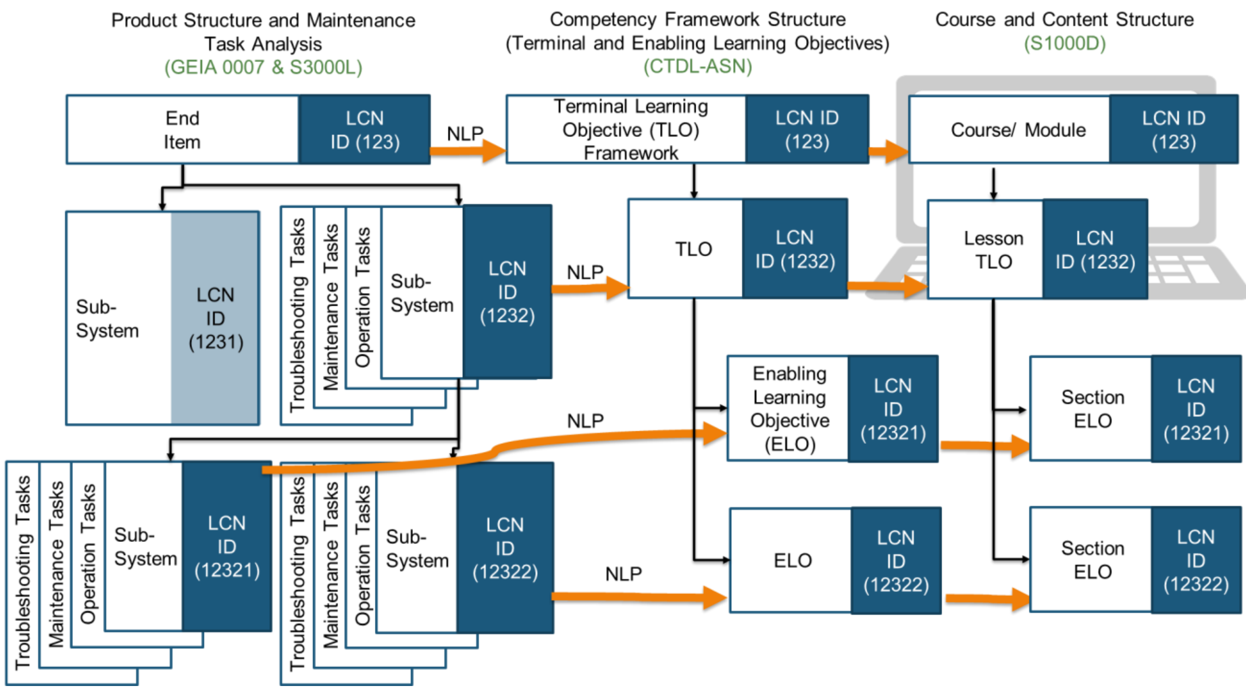

Figure 3. Specification Mapping Between Maintenance Task Analysis, Competency Framework Structure, and Course Content Structure Linked by Identifiers

USE OF LOD AND NLP TO GENERATE MACHINE-READABLE COMPETENCIES

Linked Open Data (LOD)

LOD is core to the Semantic Web and the CRADA. All MTA statements and competency outputs were encoded as LOD. This is critical, because according to the W3C, "The Semantic Web provides a common framework that allows data to be shared and reused across application, enterprise, and community boundaries" (W3C, 2015). The data entities are encoded to be self-describing (meaning) and self-identifying (location). Links can then be made between separate encoded data entities, such as individual maintenance tasks and corresponding learning objectives. These core principles are used in the Navy-Credential Engine CRADA project. The Navy does not have a competency management system. A system based on W3C practices, standards that use LOD, and software structure data as LOD are highly recommended to the Navy because it provides the necessary interfaces between training and logistics. LOD enables the goal of using machine-readable data for maintaining direct connections from the product data to training and performance over time as product configurations change.

Natural Language Processing (NLP)

NLP is a sub-discipline of artificial intelligence, information science, linguistics, and computer science. It handles the interpretation and encoding of information and knowledge to and from text. NLP scientists create and use linguistic processing tools to identify parts of speech, disambiguate terms, translate languages, enable document search, generate headlines and sports game summaries, and create automated help desk chat systems. NLP can take the form of supervised and unsupervised machine learning methods, ontological based knowledge systems, complex mathematical models, and simple word counts with an understanding of the human languages in the tasks. The challenge for NLP scientists in the CRADA is to reshape data embedded in MTA properties found in **Table 2** and **Table 3** into a competency framework using a sub-discipline of NLP known as Natural Language Generation. The steps are described below.

The Process Steps for Deriving Learning Objectives from Maintenance Task Statements

1. **Convert Specifications to LOD.** To derive learning objectives from task statements, the first step of the process converted the GEIA 0007, S3000L, and S1000D specifications and their schemas to LOD. The output resulted in the ability for computer software to index and traverse the specification using search queries. Each class and property in the specifications are now universally identifiable.

2. **Convert MTA to LOD.** The second step in the process converted the sample refrigeration unit and T850 engine MTA data in GEIA 0007 and S3000L from XML to JSON-LD. The sample MTA data is augmented with specification data converted to LOD in step 1. Alpha-numeric codes used to symbolize property values in GEIA 0007 and S3000L were exchanged with their human-readable words and phrases useful in NLP. This happens in two steps: (1) the data standard converted to LOD is machine-read, and (2) the encoded parameters of all the tasks (e.g., Hazard Code, Facility Requirement Code) are read and the encodings replaced with their respective human-readable descriptions.

3. **Generate Learning Objectives from MTA Using NLP.** The third step of the process extracted data from the MTA, ordered the data into a learning objective sentence structure, and massaged the grammar into readable forms. The CRADA used an established template to form learning objective structures. The structures, specifically, "Actor, Behavior, Condition, Degree" (ABCD) and "Condition, Task, Standard" (CTS) grammatically shape the sentences. The templates are populated with phrases stored in MTA properties by either copying them directly or using a direct translation from source text to desired text.

4. Input data from the MTA properties in **Table 2** and **Table 3** have fixed meaning and are translated via lookup tables. In multi-word phrases and sentences, or where grammar is relevant, parts-of-speech are tagged and transformed to fit the template. Work continues to perfect output grammar from the source as new MTA data is tested. Natural Language Generation techniques, a process for generating natural language descriptors or statements from structured data, govern the construction of the learning objective statement. Content Determination, an NLP subtask, decides what fields from **Table 2** and **Table 3** need to be inserted in the generated text. The ABCD and CTS templates govern the document structure and data aggregation. Fixed lookup tables translate any encoded terms and results are concatenated into a single sentence. The NLP strategy used Modified Constituency Parsing (E & Johnson, 2017), a method to annotate syntactic and semantic sentence structures, to identify candidate properties in the task statement (Baldridge, 2017). Tagged constituents are mapped to the learning objective template. Lastly, the output is passed through a sequential filter to improve syntactic and semantic coherence. As an example, **Table 4** describes the progression in the learning objective generation process with resultant output sentence is ready to be concatenated and delivered.

Table 4. Generation of a Learning Objective Statement from Structured Data in GEIA 0007 Using NLP

GEIA 0007 Property Source	Selected GEIA 0007 Property Source Attribute	NLP Output Phrase / Code	Position (ABCD)
		***Given**	C
CA Task Requirement training_equipment_requirement_code	Y = Required	***the required training equipment,***	C
		***the learner will be able to perform**	A, B
CA Task Requirement task_code, index 1 Task Interval Code	G = Unscheduled	***an unscheduled***	B
CA Task Requirement task_identification	"replace refrigeration unit" <Parse object noun phrase>	***refrigeration unit***	B
CA Task Requirement task_code, index 0	H = Remove and Replace	***replacement,***	B
CA Task Requirement task_code, index 4 Operability Code	A = System Inoperable During Equipment Maintenance	***while inoperable,***	B
		***taking into account**	B
CA Task Requirement hazardous_maintenance_procedure_code	B = Potential severe injury resulting from the incorrect or improper performance of maintenance.	***the prevention of severe injury that can occur from incorrect or improper performance of maintenance,***	B
CA Task Requirement predicted_mean_elapsed_time	0.46	***in less than 0.46 hours***	D

Sentence structure from Content Determination, Modified Constituency Parsing, and Sequential Filtering

Given **the required training equipment**, the learner will be able to perform an **unscheduled refrigeration unit replacement** while **inoperable**, taking into account **the prevention of severe injury that can occur from incorrect or improper performance of maintenance,** in **less than 0.46 hours**.

**Product of sequential filtering*

Utilizing Competency and Skills System (CASS) infrastructure, the learning objective frameworks are now structured in the CTDL-ASN. The competency frameworks mirror the system, subsystem, tasks, and subtasks hierarchy in the CaSS software hosted by Credential Engine. A course outline is exported from the competency framework to the S1000D learning aggregation models depicted in **Figure 3.**

IMPLICATIONS FOR USING NLP TO GENERATE LEARNING OBJECTIVES FROM MTA

There are challenges to ensure predictability and accuracy for maintaining links from product data to training data while providing the right amount of decision support to training analysts. The training analyst and the maintenance planner must work together to populate the required MTA properties for the NLP to process the content. Machine and human iteration are a requirement to ensure learning objective accuracy. The process should be regarded as a practice to enhance job task analysis. Data-driven competency modeling based on authoritative sources is a better alternative to practices conducted today.

Training Needs Analysis Cost Savings Using Industry Data Standards and NLP

If implemented, practiced, and improved regularly, NLP can reduce the time to perform training needs analysis by, arguably, 50 percent (50%). This metric is predicated on the idea that MTA is committed S3000L or GEIA 0007. The NLP software and output are not theoretical. **Table 4** shows the progression from MTA properties to a learning objective. Rigorous measurement practices will determine the cost savings. Not until the practice is perfected and the Navy is committed to share cost data for legacy job task analysis practices will anyone know the exact cost savings. Regardless, using NLP to derive technical competency frameworks from MTA standards is the only data-driven practice that ensures accuracy of all prioritized tasks intended to be taught and learned in the Navy.

These initial recommendations introduce necessary data to support data-driven performance and provide suggestions where training considerations benefit from additional data:

(a) Document business rules to ensure the software systems used to extract product data from GEIA 0007 and S3000L have all the information needed to generate viable learning objectives and course frameworks.
(b) Develop acquisition policies that encourage the development of MTA data in GEIA 0007 and S3000L.
(c) Develop acquisition policies that encourage the management of learning objectives in the CTDL-ASN.
(d) Develop acquisition policies that encourage the management of course outlines and curriculum in S1000D.
(e) Publish the algorithm and NLP specification used to generate learning objectives and course frameworks.
(f) Use technologies that enable LOD to maintain connections between MTA data, learning objectives, course frameworks, related training, and OTJ performance.
(g) Advocate for technical curriculum and technical manuals be developed and managed together in the CSDB.
(h) Adopt W3C practices and standards for LOD.

NEXT STEPS FOR NLP AND THE BROADER NAVY INITIATIVE

The technical success of the CRADA is a long-term promise that LOD and NLP coupled with technical data specifications can improve the art and science of creating learning objective frameworks. Technical learning objectives derived from and linked to authoritative sources is at the heart of integrating technical curriculum with logistics support and engineered systems. The CTDL-ASN, being a flexible schema for describing Linked Data, promises to describe and expose learning objectives to any application that enhances the connections between workforce and systems data. The combination of the data specifications used in the CRADA, use of Linked Data to turn MTA into machine-readable formats, and the ability for NLP to extract learning objectives promises to overcome the problems latent curriculum impose on readiness.

The next steps in the CRADA and in the broader Navy initiative go beyond creating a single learning objective for every documented task. Not all tasks are taught. The tasks that are prioritized are designed for different levels of learning. The systems go through lifecycle changes producing variants and revisions. A unique, configured competency model for each revision and variant will be required. Here are the planned next steps for the CRADA to improve the extraction of technical learning objectives from LSA specifications currently funded through March 31, 2020:

1. **Task Prioritization Processing -** Task prioritization is a prerequisite to finalizing learning objectives. CRADA mapped the Ohio State Systematic Curriculum and Instructional Development (SCID) Task Selection Model to corresponding GEIA 0007 and S3000L properties (The Ohio State University, 2019). Difficulty, importance, and frequency levels in the SCID will be determined by the mapped attributes marked

up per task in the authoritative source. The result will be an authoritative subset of prioritized tasks to be trained. If a task is not prioritized as desired, the authoritative source must be modified to cue the algorithm to re-prioritize the task.

2. **Multi-tiered Competency Requirements -** The CRADA has demonstrated the output of a single learning objective per task. The desired result is a multi-tiered framework. The next effort will analyze the authoritative source for clues in the product structure properties that can signal the output of an overall performance framework that includes multiple terminal learning objectives containing multiple enabling learning objectives.

3. **Formatted Course Outline -** The CRADA produced a course outline in the S1000D learning aggregation models. The outline is linked to the competency framework in CTDL-ASN. The outline needs to be formatted for proper visualization and usage in a learning content development tool. The outline represents the course structure as a mirror to the product structure. Mechanisms will be put in place to rearrange the outline while maintaining authoritative links to the product data.

4. **Competency Framework Revision and Variant Control -** A single competency framework representing a new system will be required. Frameworks must support variants. New revisions of the system will require new revisions of the competency frameworks. Each framework revision and variant must be available for different purposes in real-time. The revision and variant control mechanisms in the technical data specifications must be reflected in the CTDL-ASN.

5. **Implementation -** The NLP application must be available for use in a product lifecycle management environment. As of this writing in June 2019, no analysis has occurred to determine if Software as a Service, direct implementation into product lifecycle management environments, or any other strategy is the right way to implement the NLP methodology to extract competency models from maintenance task analysis.

CONCLUSION

The technology described in this paper represents the future of integrating Navy technical training with product support. It will take changes in policy, organizations, cultures, and business rules to ensure its success. The beauty of the CRADA and its core code is it can be implemented in any community of practice outside the Navy so long as the public sector specifications are used as authoritative sources. Different use cases will require re-analysis of the algorithms and the NLP. The flexibility of the standards and technology at play shows that the Semantic Web as a tool can help address one of the Navy's biggest challenges: *Readiness.*

ACKNOWLEDGEMENTS

The authors acknowledge four people who brought technical depth to the CRADA project. Parker Owen, Technical Operations VP at Integrated Support Systems, brought in-depth knowledge of the GEIA 0007 specification. Leif Gyllström, Senior Integrated Logistics Support Advisor at Saab, brought in-depth knowledge of the S3000L specification. Nate Argo, Software Engineer at Credential Engine, converted the three data specifications to LOD. Dr. Stuart Sutton, Associate Professor Emeritus at the Information School, University of Washington, brought in-depth knowledge of information modeling and Semantic Web Knowledge Structures. Their participation enabled the CRADA to be a success.

REFERENCES

115th Congress. (2019, January 14). *Congressional Bills 115th Congress*. Retrieved June 14, 2019, from Congress.Gov: https://www.congress.gov/bill/115th-congress/house-bill/4174/all-actions?overview=closed#tabs

Baldridge, J. (2017). *The Open NLP Project*. Retrieved June 14, 2019, from Open NLP: http://opennlp.apache.org/index.html

E, C., & Johnson, M. (2017, September 30). *BLLIP Reranking Parser*. Retrieved from GitHub: https://github.com/BLLIP/bllip-parser

RDF Working Group. (2014, February 25). *RDF*. Retrieved June 14, 2019, from W3C Semantic Web: https://www.w3.org/RDF/

Sutton, S. A. (2019, March 29). *CTDL Profile of ASN-DL (CTDL0ASN)*. Retrieved June 14, 2019, from Credential Engine Technical Site: https://credreg.net/ctdlasn/handbook

Sutton, S. A., & Kitchens, J. (2019, April 26). *Credential Transparency Description Language Handbook*. Retrieved June 14, 2019, from Credential Engine Technical Site: https://credreg.net/ctdl/handbook

The Ohio State University. (2019, June 2019). *Systematic Curriculum and Instructional Development (SCID)*. Retrieved 14, from The Ohio State University Center on Education and Training for Employment: https://cete.osu.edu/projects/scid

Travers, N. L., Jankowski, N., Bushway, D. J., & Garrison Duncan, A. (2019, May 22). *Learning Frameworks: Tools for Building a Better Educational Experience.* Retrieved June 14, 2019, from Lumina Foundation: https://luminawp.wpengine.com/resource/learning-frameworks/

W3C. (2015). *Linked Data*. Retrieved June 14, 2019, from W3C Semantic Web: https://www.w3.org/standards/semanticweb/data