Explosive Analysis

Using Data to Hold Warfare Centers Accountable

Authors

Ryan Webster
Robel Semunegus
Taylor Fountain
Daniel Larison
Wendy Cassidy

Abstract

The Joint Service Explosive Ordnance Disposal Publications program creates and maintains critical documents for the safe handling of ordnances. This effort is managed by a Naval Warfare Center. Historically, Navy senior leadership has funded these efforts without the ability to evaluate reasonableness of annual funding requests. Augur has recently obtained publications system data, resulting in valuable analysis of historical efforts. This data is being leveraged to develop a planning calculator capable of estimating ranges of labor hours based on ordnance type, country of origin, and other complexity drivers derived through regression analysis and other visualization techniques. This tool and the accompanying insights will enable senior leadership to negotiate with the Warfare Center and more easily measure performance.

International Cost Estimating and Analysis Association (ICEAA)

2024 Professional Development & Training Workshop
Date: 2024-02-20

Table of Contents

List of Figures 3

List of Tables 3

1. Introduction & Overview 4

1.1 An overview of Ordnance "Publications" Process 4

1.2 Executive Summary 5

2.0 Data Phases 6

2.1 Phase I 6

2.1.1 Hours per Publication and Subject 7

2.1.2 Number of Publications and Subjects 10

2.1.3 Manual Time Reporting Activities 11

2.1.4 Labor Cost 11

2.1.5 Limitations 12

2.2 Phase II 12

2.2.1 Regression Analysis 13

2.2.2 Regression Results 17

2.2.3 Methodology: Development of a nuanced calculator 17

2.3 Phase III (Future Objectives) 23

2.3.1 Enhanced Calculator 23

2.3.2 Baseline and Performance Management 23

2.3.3 Claiming Credit & WIP 24

2.3.4 Dashboard 24

3.0 Conclusion 25

List of Figures

Figure 1: Model of the Publications Workflow 4
Figure 2: RDT&E Publication Time vs. Subject Time 8
Figure 3: Coverage Level vs. Average Total Publication/Subject Time 9
Figure 4: RDTEN Bronze Total Publication Time Distribution 9
Figure 5: Quantity Projections and Actuals from FY18 - FY22 10
Figure 6: Example Rubric of Subject ID Definition 12
Figure 7: Subject and Publication Time by Appropriation 14
Figure 8: Cumulative Distribution of Subject Time by Coverage Level 14
Figure 9: Box Plot of Subject Time by Ordnance Type 15
Figure 10: Box Plot of Subject Time by Country 16
Figure 11: Breakout by Nature of Manual 16
Figure 12: Tools Publications 16
Figure 13: Results of Regression Analysis 17
Figure 21: Prototype of Publications Power BI Dashboard 25

List of Tables

Table 1 User Inputs for Calculator 20
Table 2 Labor Rate Inputs 20
Table 3 Projected Total Hours 21
Table 4 Projected Total Costs 21
Table 5 Total Publication Cost by Confidence Level 22
Table 7 Credit Claimed Model 24

1. Introduction & Overview

1.1 An overview of Ordnance "Publications" Process

Explosive Ordnance Disposal (EOD) Publications (referred to as "Pubs") is a Joint Service, Navy-led PMS 408 Program of Record (PoR) without traditional acquisition milestones. The focus of this project is to support the development, hosting, and maintenance of publishing procedures to be used by EOD technicians.

This process is managed by a Naval Warfare Center staffed with highly specialized labor (primarily former EOD technicians). During the publications process, these experts develop and maintain excerpts, known as subjects, on an ordnance or tool which contain specific information like the size, place of origin, etc. Some subjects then go through a packaging process to form a publication. As publications age, they are reviewed to determine if updates are required or if the information is no longer relevant to the Warfighter.

A specialized workflow management tool called the Solution Business Manager (SBM) is used to develop and maintain publications and subjects. To work on a publication or subject, the researcher checks it out of the SBM and then returns it when the tasks are complete. The time between a subject or publication being checked out and checked back in is recorded within SBM and categorized based on the work done during that time. This data can be extracted from the workflow tool and is referred to later in this paper as workflow data.

Historically, this information was disseminated via pocket-sized, printed booklets but has evolved to an online portal analogous to a Wikipedia of EOD manuals with countless links to components, fuses, and other information that may be shared across ordnances and tools. This very complex process (as outlined in Figure 1) accounts for large amounts of instruction, references, sources, and graphics.

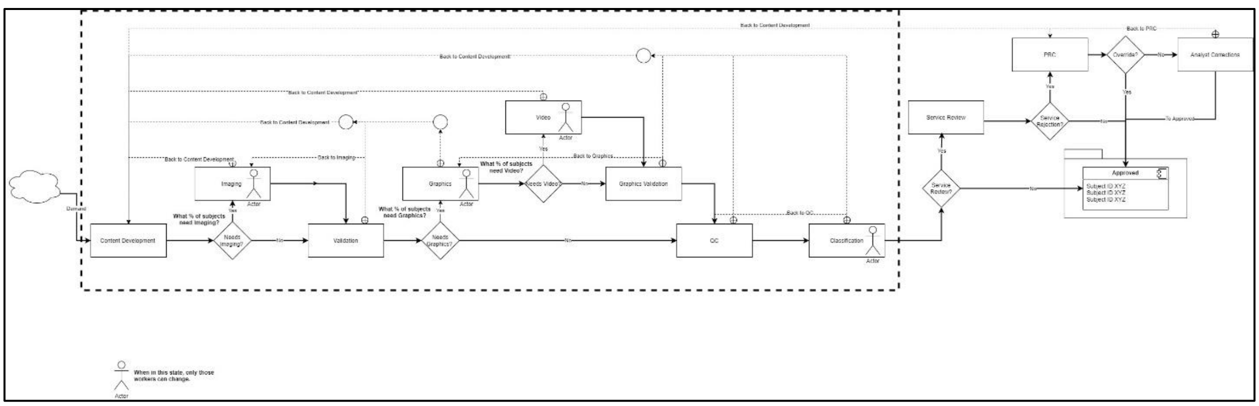

Figure 1: Model of the Publications Workflow

A variety of information is generated in the publications process. For the purposes of this paper, the term "Information Set" refers to any of items that are published throughout the publication process, including:

1) **Subjects** – *Store all information (i.e., nomenclature, measurements, definition, etc.) graphics, and tables on an item (i.e., bomb, missiles, weapons of mass destruction, etc.). Subjects may also cover broader concepts such as general EOD information, tools and equipment, synthesis, weapons systems, or explosives and EOD-related hazardous materials.*

2) **Munitions data reports -** *Digital copies of ordnance details including CT scanned images which enable 3-D rendering of each component. These details are used to develop render safe procedures and exploitation techniques. An example of an exploitation technique would be to reverse engineer a threat.*
3) **Render Safe & Disposal Procedures** – *This product is a result of performing tests and establishing steps to enable an ordnance safe for handling. A simple example would be cutting or removing a fuse followed by disposal of the explosive component. Much of the required effort will not show up in the workflow tracker as it precedes documentation.*
4) **Publications –** *The formal product consisting of packaging, labeling, bundling, and publishing any combination of the prior listed information sets. Not all information sets require this step. A standalone Subject can be published without being bundled as a publication.*

It is important to note that these various information sets are all published to the portal for the user community, regardless of whether or not the information set is bundled and packaged as a formal "publication".

1.2 Executive Summary

The publications process presents challenges in assessing fair and reasonable pricing. The publications teams consist of highly specialized writers, testers, and other areas of expertise, most of which are former ordnance technicians. This unique expertise has prevented competition from materializing around the publications scope. Nonetheless, despite the lack of competition, analytical rigor within the process and accountability of labor must be implemented. Over the last few years, Augur, in coordination with Navy PMS 408, has made major strides to streamline the publications process and add clarity to the effort required for the various types of information sets.

Augur has improved the ability to estimate the cost range for any given ordnance process. No two publications are alike, thus, capturing the various cost drivers that contribute to a single information set requires accounting for uncertainty ranges of various spreads. Additionally, Warfighter requirements change and are re-prioritized frequently so an automated tool to capture the cost of new scope will be valuable in program planning. The publications team has the priority of ensuring uninterrupted progress and is understandably less interested in robust cost estimation efforts every year.

Because of their focus on performing and the inherent complexity of estimation required, the Warfare Center's position has historically been that these efforts are "impossible to estimate". Augur has proven that the data can be used for planning purposes. The models show clear trends showing predictive variables that can be used to plan the scope in question, though uncertainty ranges are often high.

Augur has developed two estimation methods based on the data available at the time. Recently, Augur obtained real time access to the workflow tool used by the content developers, which has resulted in a greater understanding of the data fields and the scope they represent. This enhanced understanding has recently improved estimation techniques and will continue to improve methods implemented in the coming months.

Over the years, Augur has developed various methods of sizing the annual requirement for information sets based on the information available. This paper will cover three stages of analysis:

- **Phase I:** The initial, previous approach based on limited data. Estimates were based on a single primary effort driver (Coverage Level). During this phase, publication details were largely unknown, so this single variable method was used to estimate high quantities of publications that were typically unknown.

- **Phase II:** The current approach (still in process) is based on more complete data and a better understanding of the requirements process. This stage includes the development of a cost calculator for various effort drivers including coverage level, country of origin, ordnance type, and nature of the manual.
- **Phase III:** This future analysis includes further refinement of the data analyzed in Phase II. Additional analysis is being performed to capture activities that fall outside of the workflow tool (such as asset testing to enable the creation of a procedure). This full picture estimate will be used to baseline and measure performance for FY25. Once baselined, custom fields will be established in the workflow tool to enable an automated and agreed upon method for claiming credit for Work In Process (WIP), as well as other quality metrics.

2.0 Data Phases

2.1 Phase I

For the initial estimate of direct publications support, it was assumed the cost of personnel directly involved in the creation and revision of publications and subjects was captured by the workflow data. Because the data captured past actuals, it served as the basis of estimate of this labor. The understanding of the publications process was that subjects are the "building blocks" of a publication. All data presented in this report has been scrubbed and cleaned of any real programmatic data. Relationships and trends are accurately reflected, but actual values, labels, and any CUI data have been removed. It is important to note that these controls are for demonstration purposes only for the sanitized data in this paper. In the CUI version of the government calculator, the expected coverage levels are highly individualized and not linear as demonstrated here.

The workflow data was divided into two subsets. The first subset of data captured multiple metrics for subjects not tied to a publication that were created and revised from CY16 – CY21. Of the metrics included for subjects, the two most critical to the estimate were the "touch time" and "non-touch time." The other subset captured multiple metrics for publications that were created and revised from CY16 – CY20. Of these metrics included for publications, the two most critical were "publication time" and "subject time". Term definitions are captured below:

Touch Time – Captures time spent working directly on subjects that are not tied to a publication. For example, writing and formatting a subject.

Non-Touch Time – Captures time spent on any ancillary work to support a subject not tied to a publication. May include research, testing, and other work needed to develop a subject.

Publication Time – Captures time spent working directly on publications; does not include any time spent working on subjects tied to the publication. For example, publication time captures the formatting of the publication and process to package subjects together that form the publication. Typically, publication time is shorter than subject time for a given publication.

Subject Time – Captures time spent developing or revising subjects that are tied to a publication. The subject time for a single publication captures the sum of the Touch Time and Non-Touch Time for all subjects tied to that publication. This includes research, testing, and other work necessary to create and revise a subject. Subject Time captures the majority of time needed to create or revise a given publication.

Together Touch Time, Non-Touch Time, Publication Time, and Subject Time capture the total labor needed in the publication process. For the purposes of explaining the methodology of the estimate, the sum of publication time and subject time will be referred to as "Total Publication Time." Likewise, the sum of subject touch time and non-touch time will be referred to as "Total Subject Time."

Direct publications support was estimated by splitting the whole into its smaller components and estimating at the lowest level. The general steps to estimate direct publications support are as follows:

- Estimate the number of hours required to create or revise a single publication or subject
- Estimate the number of publications and subjects that are worked on each year
- Estimate the number of hours required for manually reported activities (i.e. meetings)
- Use the metrics estimated above to derive the total number of hours required for direct publications support per years

The subsequent sections detail each step in this process.

2.1.1 Hours per Publication and Subject

The initial workflow data details the time used to create/revise publications and subjects from CY16 – CY20 and CY16 – CY21 respectively. The spreadsheets received were split into the four following categories: RDT&E publication data, O&M publication data, RDT&E subject data, and O&M subject data. All RDT&E funded efforts capture the creation of a new publication or subject, while all O&M funded efforts capture subsequent revisions to publications and subjects. Since the two sets of data capture different aspects of the publication process, they were analyzed and estimated separately. Grouping publications together with similar characteristics and estimating the average total publication time produced an estimate with greater fidelity.

The first relationship explored for both subsets of data was between subject time and publication time. The analysis included deriving statistics for each metric and plotting the Publication Time vs. Subject Time, as pictured in Figure 2 (where each point represents a single RDT&E publication). Note that high-complexity outliers were excluded from the figure. The publication data showed that publication time and subject time are independent of one another (i.e., variance in one metric does not affect the other). Similar trends were found for O&M publications and subjects not tied to a publication.

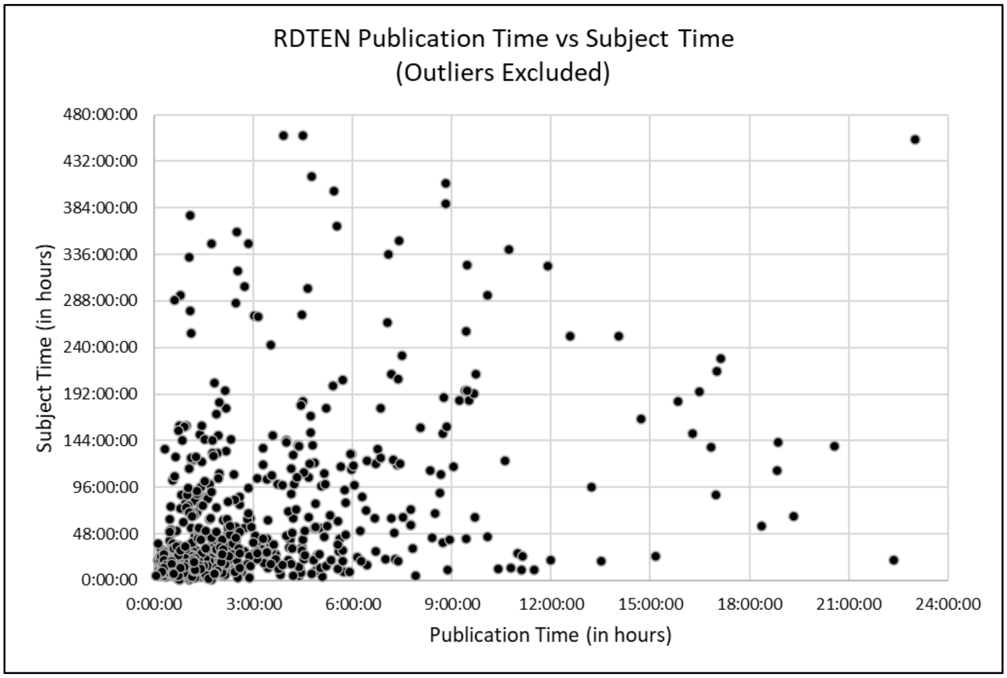

Figure 2: RDT&E Publication Time vs. Subject Time

The initial analysis made it clear that the remaining available metrics needed to be used in the estimate. Among those metrics are the submit and close dates, the coverage level, and the number of subjects tied to a publication. The last two metrics are not applicable to subject data. The definition of each metric provided by the service provider is below.

Submit Date – Date that kicks off the publication workflow.

Close Date – The calendar year in which the publication is finished.

Coverage Level – Describes the amount of EOD information included in the publication. For this paper's purposes, the three coverage levels are Bronze, Silver, and Gold.

Number of Subjects – Quantity of subjects that are tied to a publication.

To determine which metric to use, the following relationships were explored:

- Calendar Year vs. Total Publication Time
- Coverage Level vs. Total Publication Time
- Number of Subjects vs. Total Publication Time

Among the trends observed, the most significant relationship was Coverage Level vs. Total Publication Time (true for both RDT&E and O&M efforts). The analysis showed that as coverage level increases, Total Publication Time increases, as shown in Figure 3.

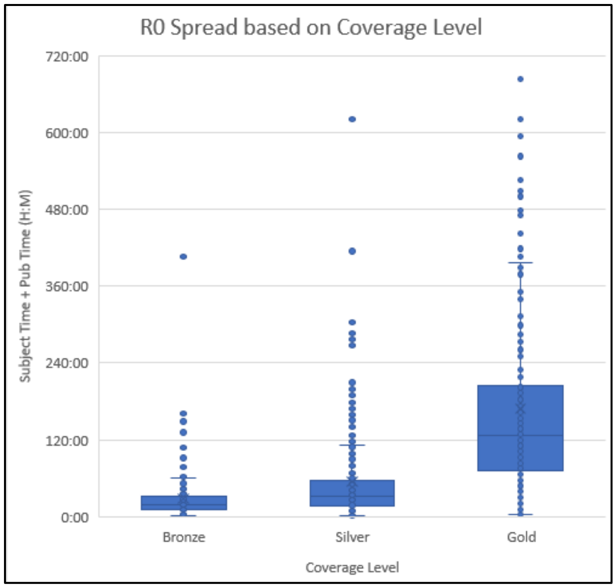

Figure 3: Coverage Level vs. Average Total Publication/Subject Time

This observation led to the division of the RDT&E and O&M publication data by coverage level. To complete the division, subjects not tied to a publication were grouped together as a pseudo-coverage level.

The maximum, minimum, and average total publication/subject times were derived for each of the subsets. The average of each subset was used as the estimate of total publication/subject time needed for its corresponding set of data. In addition, histograms were created for each of the subsets to visualize the distribution of the data. The histogram for RDT&E Bronze Total Publication Time is pictured below in Figure 4.

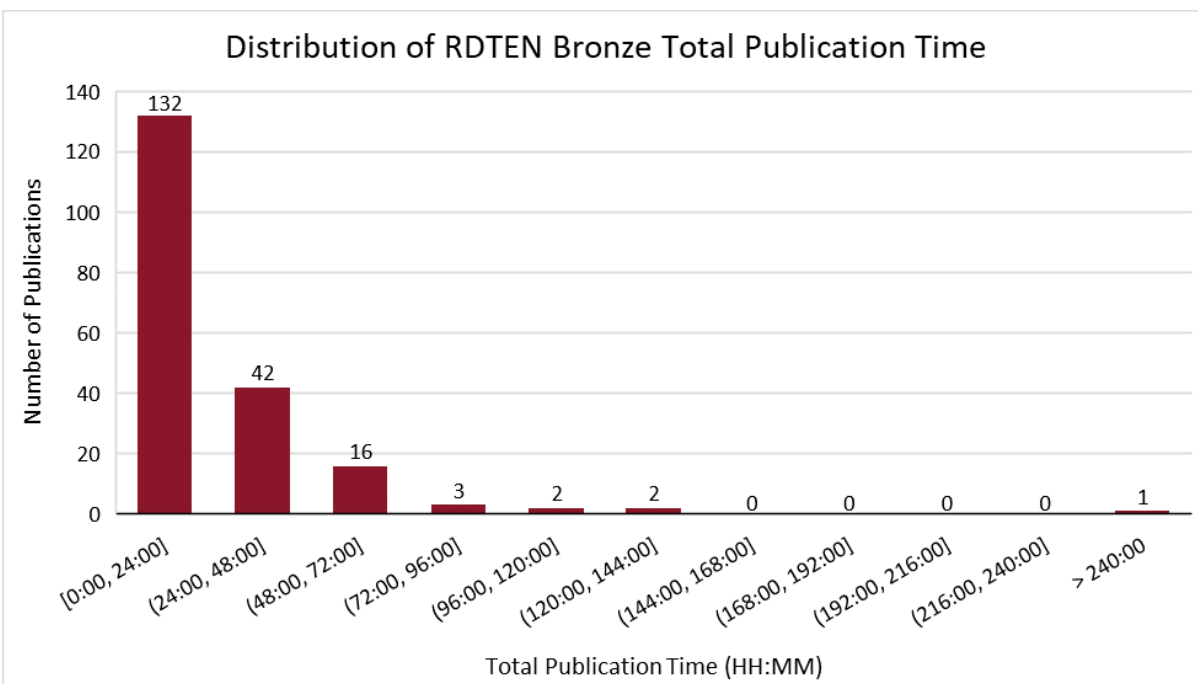

Figure 4: RDTEN Bronze Total Publication Time Distribution

Risk and uncertainty were incorporated to complete the estimates for average Total Publication/Subject Time. Adding risk and uncertainty creates a more accurate estimate and mitigates concerns expressed by the service provider about the variability of the data. In terms of modeling, this means applying a probability density function to each element. To create each function, the statistical computing software *R* was used to perform a simple random sample on each of the subsets of data. Essentially, the program randomly selected a subset of data from the appropriate population, where the number of data points selected was based on the number of publications/subjects worked on each year for each of the eight subsets. The average Total Publication/Subject Time of that subset was then derived and recorded. Repeating this process 10,000 times formed a robust distribution used to capture risk and uncertainty in the cost model. By applying the functions to the Total Publication and Subject Time elements, their "most likely" values were forecasted by the model.

2.1.2 Number of Publications and Subjects

To calculate the labor required for developing publications, the expected number of publications and subjects that are worked on each year was estimated. The estimate was informed by actuals provided by PMS 408 where the data for the number of publications and subjects was from FY18 – FY21 and included projections for FY22 (shown in Figure 5).

Using multiple years of data ensured that no single point was projected onto the remainder of the program lifecycle. This was critical for the estimate since it was clear from the data that the number of publications and subjects worked on each year varies.

The total number of RDT&E and O&M publications and revisions per year was estimated first, then divided into the three coverage levels and subjects, and finally split into the various funding lines using the available data as shown in Figure 5. This allows for the combination of time per publication and number of publications that is discussed in a subsequent section.

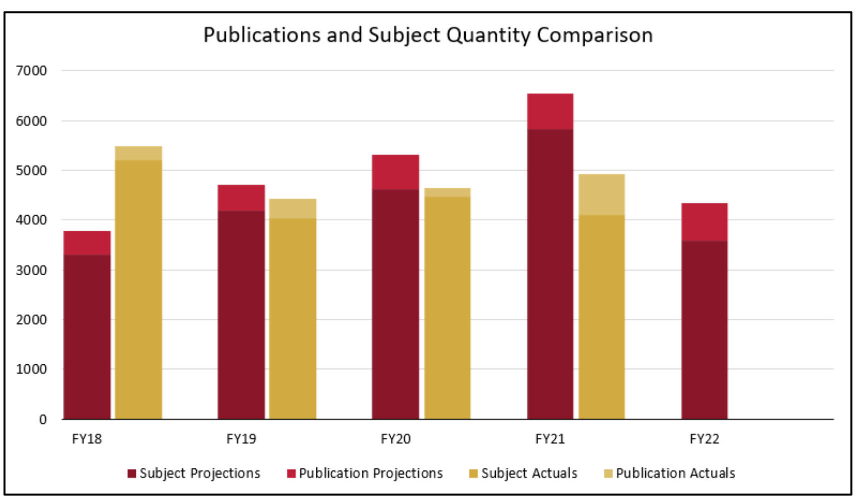

Figure 5: Quantity Projections and Actuals from FY18 - FY22

The average of the past actuals and projections was used to estimate the total number of publications. Similarly, the ratios used to partition the total were estimated using the ratios observed in the past actuals and projections. For example, to estimate the number of RDT&E Bronze Procedure publications worked on per year, the total number of RDT&E publications was estimated using past actuals and then divided by coverage level based on the ratios observed from FY18 – FY20. The result of this process was the number of RDT&E publications that are Bronze Procedures.

2.1.3 Manual Time Reporting Activities

In addition to the direct publications support labor captured in the data, a portion of the labor is captured outside of the workflow using manual time reporting. These activities include training, meetings, and research that is not focused on a specific publication/subject. To determine the appropriate estimate for these activities, various methodologies were used including examining the delta between the FY21 NEPS reports and the workflow data, examining past actuals from analogous programs, and holding discussions with the service provider.

The FY21 NEPS reports include the headcounts requested for Pubs, while the workflow environment solely captures the time spent working directly on publications and subjects. The assumption was that the difference between the headcounts requested in the NEPS reports and the workflow data is the manually reported labor hours. After separating the direct publication support from the rest of the data in the FY21 NEPS reports and comparing it to the workflow data for FY18 – FY20, the difference between the headcounts indicates that 50 – 80% of a person's time would be spent on manually reported activities while the remaining time is spent working on publications and subjects. Some of these headcount differences were likely influenced by events such as the COVID-19 pandemic, however these activities are generally considered administrative/overhead activities, which typically make up 20% of the work week in analogous PMS 408 programs.

The 50 – 80% proved to be unreasonable, as during a previous briefing of the cost estimate where the Warfare Center indicated that the original assumption of 30% was too high. As a result of these conversations with the service provider and analysis of other programs, the estimate assumed that 20% of the labor required for Pubs is needed for manually reported activities. For example, suppose 100 hours are needed per year to develop and revise publications. Then 20 hours would be estimated for manually reported activities, meaning 120 hours are required in total.

2.1.4 Labor Cost

To fully capture the cost of operating the Publications Program, the below equation used to estimate the total FTEs needed per year for publications support. The hours per publication come from the estimate outlined in and the estimate for publications per year is outlined

2.1.2 Number of Publications and Subjects. The percentage for manual time reporting is outlined in

2.1.3 Manual Time Reporting Activities.

$$\frac{Hours}{Publication} \times \frac{Publications}{Year} \times (1 + manual\ time\ reporting\ percentage) \times \frac{Labor\ Cost}{Hour} = \frac{Labor\ Cost}{Year}$$

For example, multiplying the RDT&E Bronze Total Publication Time by the number of RDT&E Bronze Render Safe Procedures made in a year produces the number of hours required to work directly on RDT&E Bronze Render Safe Procedures. The total is increased by 20% to account for manually reported activities. To complete the estimate for direct publications support, the level of effort required was multiplied by the Warfare Center fully burdened hourly labor rate.

2.1.5 Limitations

This methodology enabled high-level sizing on the level of effort required for a given publication. However, while the very high uncertainty ranges were an accurate reflection of the data provided and our current understanding of the problem, it was apparent that both the high (> 1000 hours) and low (< 1 hour) bounds for a given publication could not provide much insight for short-term planning. Additionally, estimating with the mean as the primary method was sensitive to high-complexity outliers. While this was sufficient at the time, and valuable for long range planning, there was a need to refine the analysis for detailed near term planning, which led to the additional methods described in Phase II.

2.2 Phase II

The methodology developed in Phase I remains the standard methodology for long range planning, including the Program Objective Memorandum (POM) Process. For near term planning, there is a need to detail plan and baseline the assigned scope. In 2023 there were rapid changes in publications priorities as well as a demand for a greater traceability to detailed tasks in the publications process. This led Augur to Phase II, which required re-engaging with the Warfare Center to discuss the data fields tracked within the workflow data. This quickly led to the discovery that there were meaningful data fields previously ignored by regression techniques. Initially, the publication ID was interpreted as a serial number, however, once the procedure manual was provided, Augur realized that this serial number actually had embedded descriptive details for of the information set in question. Once this rubric was provided, Augur was able to move forward by analyzing each of the data fields uniquely to dissect the data further. An example is shown below in Figure 6.

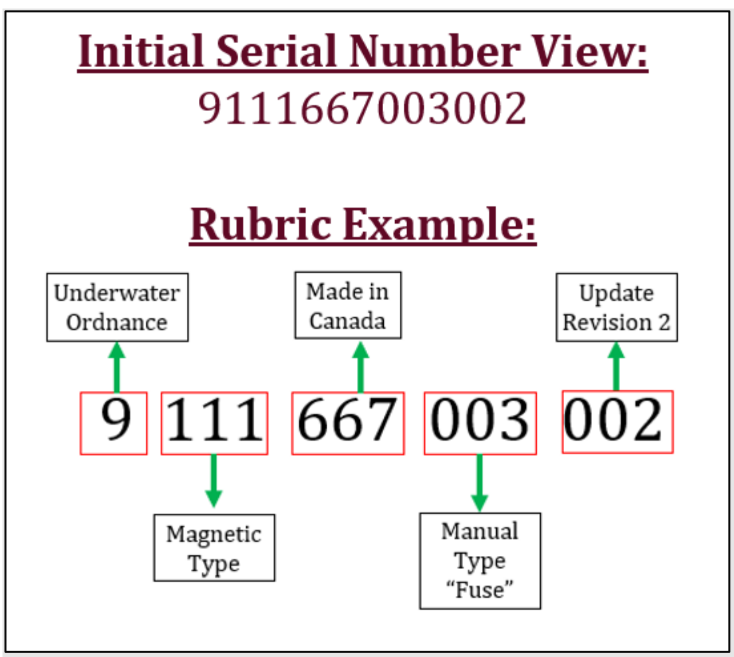

Figure 6: Example Rubric of Subject ID Definition

2.2.1 Regression Analysis

Later in the year, an updated dataset from the Workflow was received, capturing subject time and publication time for over 2,000 initial publications and revisions dating from FY16 to FY23. With renewed interest in a data-driven estimate for direct publications support and input from the Warfare Center on how to interpret certain fields that had previously been discarded, the Augur team went about developing a more granular estimate from Phase I that could support near-term planning and hold the Warfare Center accountable year over year.

The original methodology for developing an estimate for direct publications support from this dataset was to conduct regression analysis. As the only continuous variables in the dataset were the ones predicted, the regression analysis was categorical in nature, requiring different encoding techniques such as one-hot, target, and ordinal encoding. Exploratory data analysis was conducted in Power BI as well as in Python using the open-source pandas and matplotlib packages, while the Python packages pandas and sklearn were used for regression analysis.

Three possible independent variables were determined for the regression analysis: the subject time, publication time, and total time for a given publication. Initial exploratory analysis in Python found that subject time and total time were highly correlated, with a Pearson correlation coefficient of .99. However, while subject time and publication time were correlated, neither the Pearson nor the Spearman correlation coefficients indicated an impactful relationship between the two. Further discussions with the Warfare Center revealed that the two efforts required different skillsets with separate labor rates. Therefore, total time was not evaluated, and publication time and subject time were evaluated separately.

Appropriation, coverage level, ordnance type, country of origin, and nature of the manual were identified as possible features from evaluating the original workflow data. The updated data pull contained three additional fields, each with a cardinality of two: 'AIP', 'Platform', and 'Tools'. Discussions with the Warfare Center identified AIPs as a condensed workflow to rapidly develop publications during the COVID-19 pandemic, and that it would not be used going forward. Thus, publications marked as AIPs were removed from the dataset. While the Platform field – indicating the platform the publication was published under - was originally converted to a dummy variable, instances of the second platform made up less than 80 records and did not appear significantly different from the first. Thus, the feature was dropped but the records for both platforms were kept. Finally, Tools, a Boolean field narrowing down the content of the publication, was used as a dummy variable.

Each feature was then examined using visualization tools in Power BI and Python. First, looking at appropriations, the median level of support for a publication was not significantly different between the initial development (RDTEN) and follow-on revisions (OMN) of a publication for both publication and subject time. However, the average level of support per publication, while significantly higher than the median in all cases, was higher for revisions than for initial development, indicating a greater skew of the sample for OMN-funded efforts (see Figure 7). Due to its binary nature, the appropriation feature was used as a dummy variable for the initial regression analysis.

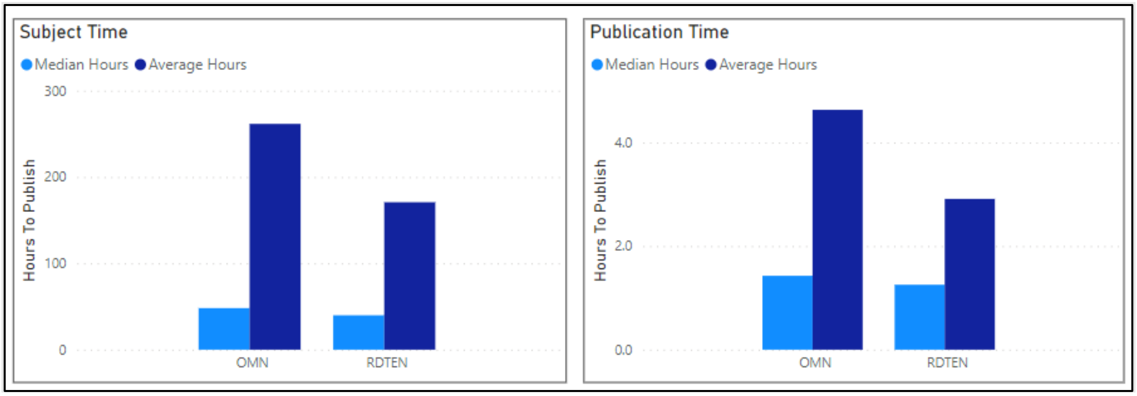

Figure 7: Subject and Publication Time by Appropriation

Next, the coverage level was examined. In line with both the Phase I analysis and domain knowledge, the empirical CDF of each subset indicates predictive power in this feature (as shown in Figure 8). By plotting these on a log-scaled x-axis, it's clear that both the subject and publication time are log-normally distributed for each coverage level. The spread for the Gold coverage level is much wider than for Bronze or Silver, indicating greater variability within that subset. This indicates that there are likely more impactful drivers among Gold publications.

Because this feature is not binary in nature, dummy encoding was not possible. The clear ranking of Bronze, Silver, and Gold made it a good candidate for ordinal encoding; however, ordinal encoding assumes that the relationship between the level of effort required for Bronze and Silver is the same as the relationship between Silver and Gold. Therefore, due to the low cardinality of this feature, one-hot encoding was also considered.

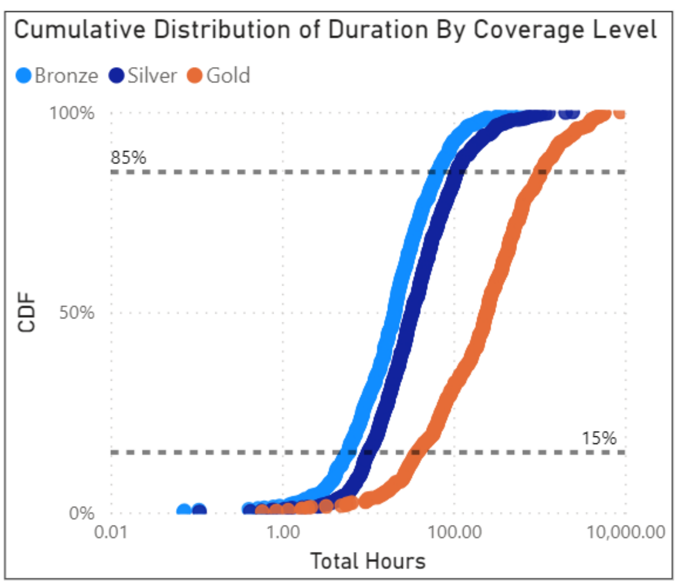

Figure 8: Cumulative Distribution of Subject Time by Coverage Level

The next feature examined was the ordnance type of the publication. The box plot of the natural log of the subject time by publication demonstrates very high cardinality. It also indicates significant variation in subject time, both within a category and between different categories (Displayed in Figure 9).

Because these categories were qualitative in nature, this feature was not determined to be a good candidate for ordinal encoding. While it was possible to sort and rank the ordnance types by a central tendency of the target value, it would not necessarily be a defendable assumption. For example, while we can confidently assume that a Type 1 would always be less complex than a Type 16, the same could not be said for Type 5 vs Type 6, as there is significant overlap between their ranges. The high cardinality of this feature indicates that one-hot encoding could result in over-fitting, particularly for categories with very few publications. For example, if there is only one Type 8 publication in the dataset, one-hot encoding would result in the predicted level of effort for a Type 8 publication being exactly the duration of the prior. While target encoding this variable could negate this, the high skew and presence of high-complexity outliers within each category indicates that the mean may not be representative, even amongst publications of the same type.

Plotting the log of the publication time by ordnance category yielded similar results. While dimension reduction techniques could be applied to this field by combining similar ordnance types, defendable groupings of the categories could not be determined. Ultimately, the initial regression analysis examined both one-hot and target encoding for this feature, despite the limitations of each feature encoding method.

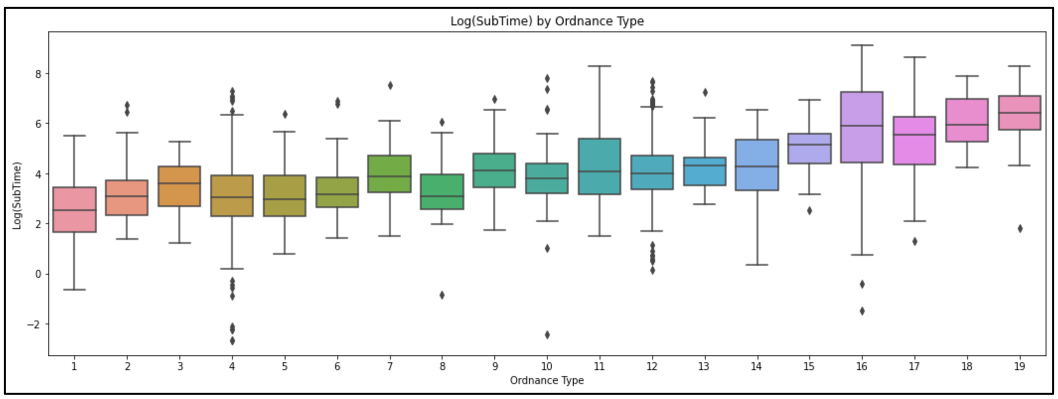

Figure 9: Box Plot of Subject Time by Ordnance Type

The next feature examined was the country of origin (Figure 10). Moreso than the ordnance category, this feature had extremely high cardinality, ruling out one-hot encoding. No defendable ranking could be identified, ruling out ordinal encoding. Ultimately, this feature was determined to be a good candidate for target encoding, though it’s assumed that the predictive power is limited.

Note that there were several countries that were not present in the Workflow data, limiting the applicability of any regression results to publications developed for those countries in the future.

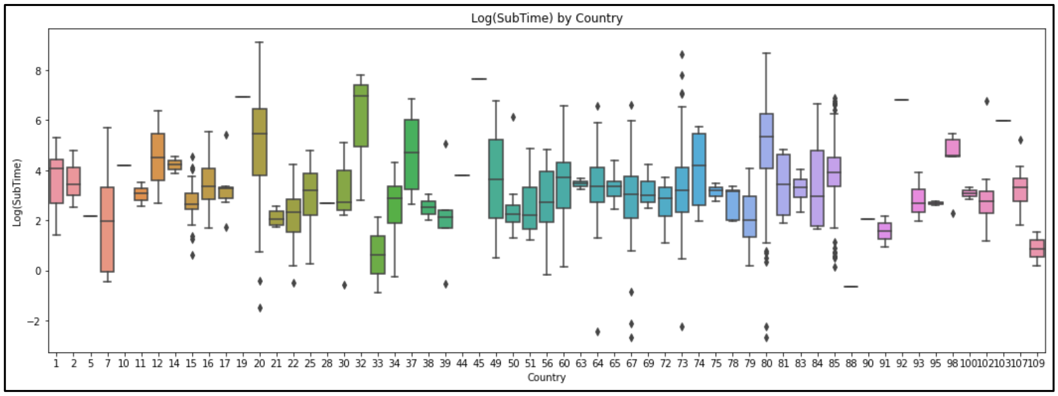

Figure 10: Box Plot of Subject Time by Country

The “nature of the manual” was examined next (Figure 11). This feature has low cardinality, with Type A and Type B manuals exhibiting similar behavior for both publication and subject time, as well as making up most of the publications observed in the workflow dataset. Conversations with SMEs indicated that Type E manuals were no longer being produced; therefore, they were removed from the dataset. Due to their similar behavior, Type A and Type B publications were grouped together. This feature was then one-hot encoded, with dummy variables for Type C and Type D publications.

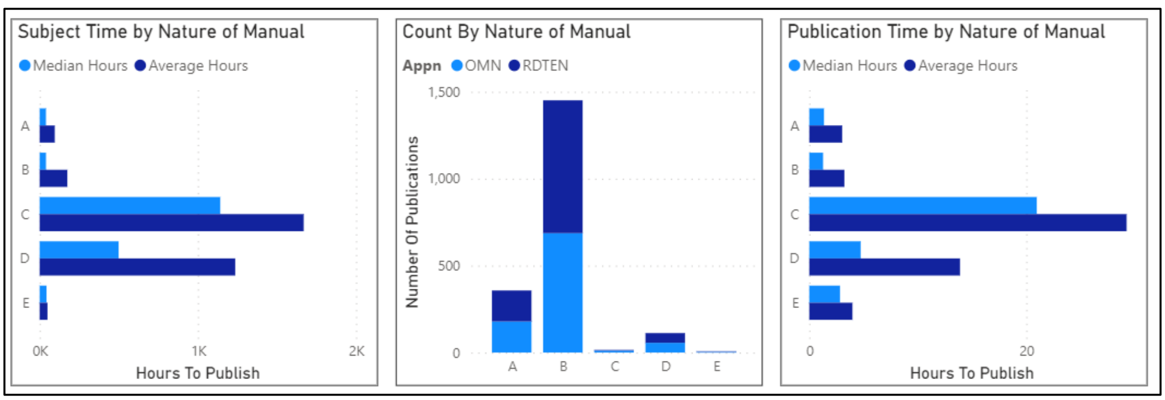

Figure 11: Breakout by Nature of Manual

Finally, it is clear from observation that Tools publications have a higher complexity than non-Tools publications (Figure 12), indicating predictive power. While the number of Tools publications was relatively small compared to the total publications, the dummy variable was included in the analysis.

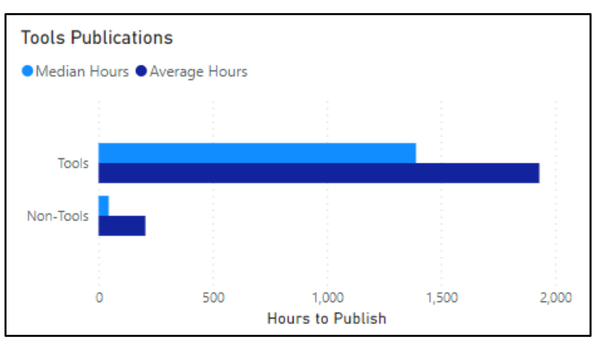

Figure 12: Tools Publications

Due to the apparent log-normal distribution of the responses, both publication and subject time were log-transformed. Observations where the independent variable was zero, or greater than three standard deviations away from the mean in the log-space, were removed. For each independent variable, six cases of linear regression were run to account for the different encoding techniques for coverage level and ordnance type. This analysis was executed using the sklearn package in Python.

2.2.2 Regression Results

A quick look at the results of the regression analysis will show that they did not yield promising results (Figure 13) – the maximum R^2 in the log space was just .5 and transforming back to unit space yielded a lower R^2 value. Additionally, even with the most over-fitted method of over-fitted combination of encoding methods – with both Coverage Level and Category being one-hot encoded, there is limited predictive power.

	CL Encoded	Category Encoded	IV	Samples	Features	R^2
0	Target	Target	Log(PubTime)	1907	8	0.276643
1	Ordinal	Target	Log(PubTime)	1907	8	0.279451
2	One-Hot	Target	Log(PubTime)	1907	9	0.279467
3	Target	One-Hot	Log(PubTime)	1907	25	0.302834
4	Ordinal	One-Hot	Log(PubTime)	1907	25	0.303395
5	One-Hot	One-Hot	Log(PubTime)	1907	26	0.303925
6	Target	Target	Log(SubTime)	1944	8	0.475851
7	Ordinal	Target	Log(SubTime)	1944	8	0.480895
8	One-Hot	Target	Log(SubTime)	1944	9	0.481032
9	Target	One-Hot	Log(SubTime)	1944	25	0.504423
10	Ordinal	One-Hot	Log(SubTime)	1944	25	0.508435
11	One-Hot	One-Hot	Log(SubTime)	1944	26	0.508910

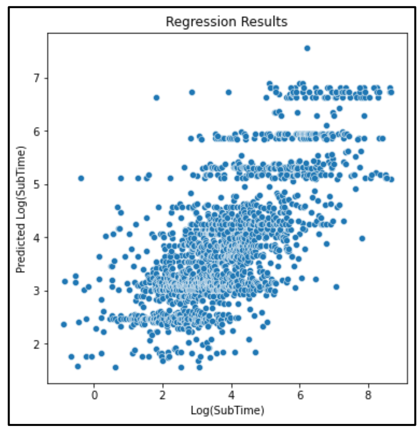

Figure 13: Results of Regression Analysis

However, there are still useful insights that can be extracted from the results. First, we can see that there are other variables involved in how long a publication takes to develop and revise – this makes sense since it's a task executed by people, and human behavior is more difficult to predict than technologies. Further, by looking at the R^2 of the results, we can see that the results are less accurate when the ordnance category is target encoded. This indicates that the mean publication or subject time of an ordnance type does not accurately represent that ordnance type's publication or subject time.

In this analysis, more regression excursions were run, accounting for different encoding methods, transformation, and implementing various machine learning techniques. While regression analysis was a helpful tool in explaining relationships between variables in the data and the relative effect size of their impact on the publication and subject time, we determined it was not the most effective tool for forecasting the level of effort required for a publication. Therefore, different approaches were considered.

2.2.3 Methodology: Development of a nuanced calculator

To support the publication effort, the Augur team determined a way to forecast the number of hours and associated cost with developing a new publication or revising an existing publication. Historical data from previous publications can provide insight into the amount of "Sub Time" (subject time) and "Pub Time" (publication time) associated with a particular type of publication. To facilitate this forecasting effort, the Augur team developed a calculator designed to output a projection of the range of hours and costs based on a set of user-defined inputs.

The calculator was built based on the following ground rules & assumptions:

- Projected hours within calculator capture direct support for publications only; additional hours calculations are required to capture indirect support and other manually reported activities
- An “RDTEN” appropriation designates a publication as a newly created publication; “OMN” designates it as a revision to an existing publication
- Users will always know the appropriation of the publication being forecasted
- Sub Time and Pub Time labor are assigned separate hourly rates
- The maximum/desired coverage type is based on the Ordnance Type:
    - Any ordnance category from 1 to 8 is expected to have a maximum coverage level of “Gold”
    - Any ordnance category from 7 to 9 is expected to have a maximum coverage level of “Silver”
    - Any ordnance category 10 or greater is expected to have a maximum coverage level of “Bronze”
    - Any publication that exceeds the ordnance’s maximum coverage level is identified by red text in the “Coverage Level” column
- Publications with less than 10 relevant historical data points are shaded red to indicate that there is an insufficient amount of historical data for robust projections. This signals to the user to consider removing a constraint (i.e. primary effort driver) to estimate at a higher level
- “Total Time” columns are based on the summation of the confidence level hours and cost results within “Sub Time” and “Pub Time” columns

With the assumptions above, the calculator seeks to provide users with reliable projections for annual planning and funding purposes. The calculator solicits a set of features that the user must define in order to output the publication’s projected hours and cost (identified by the green shaded section of the calculator). The user must provide inputs to the following features to obtain an output: Publication Name, Appropriation, Ordnance Type, Country of Origin, Nature of Manual, and Coverage Level. To prevent erroneous inputs, data validation was applied to all columns except Publication Name – users can select the necessary inputs from a dropdown. Further, it's important to note that users may not have complete insight into the publication that they are forecasting (i.e. they may not be able to fill out every column in the calculator). The calculator takes this into consideration when outputting projections by allowing users to mark columns (apart from Publication Name and Appropriation columns) as “Unk” (Unknown) to signal the calculator not to consider those features within its projections. The calculator will shade any unknown feature in gray to differentiate it from the known inputs.

Additionally, as stated in the assumptions, each publication has a maximum/desired coverage level based on the ordnance type. In the “Coverage Level” column, the calculator calls out any input that contradicts the ordnance’s maximum coverage level by changing the input's text color to red. Users can also provide the specified quantity to calculate the total funding cost for each publication. Lastly, it’s important to note that the calculator will come pre-defined with the appropriate labor rates for Sub Time and Pub Time to convert the projected hours to costs.

Once the user has provided inputs on all required columns, the calculator will produce several outputs. First, the calculator provides a count of the number of relevant data points from the historical data that were used in projecting the number of hours and costs for each publication. For publications with less than 10 relevant historical data points, the calculator will shade that cell red to indicate that an inadequate number of data points was leveraged in the projection. Additionally, the calculator provides the projected hours and costs in two ways – per publication (quantity 1) and in total (based on the value in the quantity column). Within the calculator, projected hours sections are marked off in blue; projected costs in yellow. Based on the user's inputs, the calculator derives a range of hours for Sub Time, Pub Time, and Total Time (a summation of Sub and Pub Time projections). The calculator leverages historical publication data to provide a projection for the number of hours required to work on each respective publication. The calculator then applies the pre-defined labor rates for Sub and Pub Time onto the projected hours to derive a cost.

The projected hours and costs in the calculator are presented as range of values at the 15/50/85% confidence levels (CLs). Confidence levels provide insight into the level of likelihood of the associated time/cost. For example, the projected value (for either hours or cost) at the 50% CL translates to a 50/50 chance of the actual output being at or below the projected value. For the 85% CL projected value, the interpretation is that there is an 85% chance of the output being at or below the projected value and 15% chance of it being above the projected value (the 15% CL demonstrates the opposite of this). The 50% CL is generally described as the "most-likely" value while the 85% CL is deemed a more conservative approach. Initially, the calculator returned a projection that showcased the min/mean/max for each publication, but the confidence levels were utilized instead to provide more reasonable bounds. While this method is more nuanced and lower spread than simply estimating based on coverage level, the uncertainty spread remains high, once again highlighting the fact that no two ordnances are the same.

Finally, the calculator provides a high-level overview of the projected hours and costs in the "Summary" tab. This tab highlights the crucial elements surrounding each publication including the publication name, quantity required, and number of relevant historical data points. Additionally, the "Summary" tab provides a "traffic light" chart that shows the range of total costs for each publication by the confidence levels specified previously. Finally, a smaller second table presents the total funding costs at each of the specified confidence levels to provide the user with their total funding cost.

Visuals: Screenshots of tool outputs and uncertainty ranges

Step 1: In the "Calculator" tab, user defines the necessary input columns for each publication

- Any "Unk" inputs are marked in gray to differentiate it from the known inputs (see Figure 14)
- Any contradictions to the maximum/desired coverage level are labeled in red font in the "Coverage Level" column (see Table 1)

Table 1 User Inputs for Calculator

	User Inputs						
#	Publication Name	Appropriation	Ordnance Type	Country of Origin	Nature of Manual	Coverage Level	Quantity
1	A1	RDTEN	Unk	Unk	Unk	Unk	1
2	A2	RDTEN	7	Unk	Unk	Unk	5
3	A3	RDTEN	Unk	72 - Cheyenne	Unk	Unk	10
4	A4	OMN	Unk	Unk	B	Unk	7
5	A5	RDTEN	Unk	Unk	Unk	Gold - Procedure	20
6	A6	RDTEN	7	80 - Harrisburg	B	Gold - Procedure	5
7	A7	OMN	17	32 - New York	Unk	Unk	10
8	A8	RDTEN	Unk	20 - Maryland	D	Unk	11
9	A9	OMN	Unk	Unk	A	Bronze - Technical Information	6
10	A10	RDTEN	2	Unk	Unk	Silver - Functioning and/or Appearance	4
11	A11	RDTEN	9	Unk	B	Unk	15
12	A12	OMN	Unk	65 - Baton Rouge	Unk	Gold - Procedure	3
13	A13	OMN	11	80 - Harrisburg	C	Unk	2
14	A14	RDTEN	Unk	10 - Georgia	B	Gold - Procedure	1
15	A15	OMN	8	64 - Austin	Unk	Silver - Functioning and/or Appearance	4
16	A16	OMN	17	Unk	B	Gold - Procedure	1

<u>Step 2: User inserts hourly labor rates for "Sub Time" and "Pub Time" to determine projected costs</u>

- Pub Time labor is assumed to be higher than Sub Time (see Table 2)

Table 2 Labor Rate Inputs

Sub Hourly Labor Rate	**$ 180**	*Pub Hourly Labor Rate*	**$ 200**

<u>Step 3: User reviews the projected hours and costs for each publication based on their inputs</u>

- Calculator calls out any publication that has less than 10 relevant historical data points to develop its projected hours and costs
- Calculator outputs projected hours and costs for a single publication as well as total quantity (as defined by the user). Table 3 and Table 4 depict outputs for total hours and costs, respectively
- Projected hours and costs are shown as a range of values at the 15/50/85% CLs

Projected Total Hours:

Table 3 Projected Total Hours

Projected Hours (Total Hours)								
Sub Time			Pub Time			Total Time		
15% CL	50% CL	85% CL	15% CL	50% CL	85% CL	15% CL	50% CL	85% CL
4.96	29.97	238.73	0.12	1.08	4.58	5.08	31.05	243.31
16.80	153.05	571.17	1.12	5.75	14.91	17.91	158.80	586.08
239.50	239.50	239.50	3.10	3.10	3.10	242.60	242.60	242.60
10.32	124.11	897.50	0.49	5.11	28.27	10.81	129.22	925.76
904.89	5156.20	18321.32	14.80	59.40	154.05	919.69	5215.60	18475.37
410.51	571.65	646.87	3.60	5.80	6.40	414.11	577.45	653.26
7465.34	16400.45	20677.12	51.51	142.80	298.61	7516.85	16543.25	20975.73
318.23	2511.41	10225.93	5.67	30.58	79.70	323.90	2541.99	10305.63
5.20	34.08	218.25	0.36	2.61	10.66	5.56	36.69	228.91
17.58	66.24	89.92	1.12	3.78	8.44	18.70	70.02	98.36
6.30	109.65	681.74	0.00	1.80	58.37	6.30	111.45	740.10
239.37	239.37	239.37	2.52	2.52	2.52	241.89	241.89	241.89
1251.10	3092.32	7580.31	14.37	46.52	103.58	1265.47	3138.84	7683.89
-	-	-	-	-	-	-	-	-
70.04	105.80	173.00	2.93	6.04	6.15	72.98	111.84	179.15
49.91	414.42	1629.68	1.21	4.92	22.62	51.13	419.34	1652.30

Projected Total Costs:

Table 4 Projected Total Costs

Projected Costs (Total Cost)								
Sub Time			Pub Time			Total Time		
15% CL	50% CL	85% CL	15% CL	50% CL	85% CL	15% CL	50% CL	85% CL
$ 893	$ 5,394	$ 42,971	$ 22	$ 194	$ 824	$ 915	$ 5,588	$ 43,796
$ 3,023	$ 27,549	$ 102,811	$ 201	$ 1,035	$ 2,684	$ 3,224	$ 28,584	$ 105,494
$ 43,110	$ 43,110	$ 43,110	$ 558	$ 558	$ 558	$ 43,668	$ 43,668	$ 43,668
$ 1,857	$ 22,340	$ 161,550	$ 88	$ 920	$ 5,088	$ 1,945	$ 23,260	$ 166,638
$ 162,880	$ 928,116	$ 3,297,838	$ 2,664	$ 10,692	$ 27,729	$ 165,544	$ 938,808	$ 3,325,567
$ 73,892	$ 102,897	$ 116,436	$ 647	$ 1,044	$ 1,151	$ 74,539	$ 103,941	$ 117,587
$ 1,343,761	$ 2,952,081	$ 3,721,881	$ 9,271	$ 25,704	$ 53,750	$ 1,353,032	$ 2,977,785	$ 3,775,631
$ 57,281	$ 452,054	$ 1,840,667	$ 1,020	$ 5,504	$ 14,345	$ 58,301	$ 457,558	$ 1,855,013
$ 936	$ 6,134	$ 39,285	$ 65	$ 470	$ 1,918	$ 1,001	$ 6,604	$ 41,203
$ 3,164	$ 11,923	$ 16,186	$ 202	$ 680	$ 1,520	$ 3,366	$ 12,604	$ 17,706
$ 1,134	$ 19,737	$ 122,712	$ -	$ 324	$ 10,506	$ 1,134	$ 20,061	$ 133,218
$ 43,087	$ 43,087	$ 43,087	$ 454	$ 454	$ 454	$ 43,540	$ 43,540	$ 43,540
$ 225,198	$ 556,618	$ 1,364,456	$ 2,587	$ 8,374	$ 18,644	$ 227,785	$ 564,991	$ 1,383,100
-	-	-	-	-	-	-	-	-
$ 12,608	$ 19,044	$ 31,140	$ 528	$ 1,087	$ 1,107	$ 13,136	$ 20,131	$ 32,247
$ 8,985	$ 74,595	$ 293,343	$ 218	$ 886	$ 4,072	$ 9,203	$ 75,480	$ 297,415

Step 4: User reviews "Summary" tab to determine total funding cost

- "Summary" tab outputs the total funding cost at the 15/50/85% CLs (see Tables 5)
- Tab outlines the range of total costs for each publication inputted into the calculator with corresponding quantity and count of relevant historical data points

Table 5 Total Publication Cost by Confidence Level

Total Publications Cost (by Confidence Level)		
15% CL	50% CL	85% CL
$ 2,000,333	$ 5,322,604	$ 11,381,821

Publication Total Funding Summary Table						
Publication Info				Total Funding Range ($)		
#	Publication Name	Quantity	# of Data Points	15% CL	50% CL	85% CL
1	A1	1	1152	$ 915	$ 5,588	$ 43,796
2	A2	5	55	$ 3,224	$ 28,584	$ 105,494
3	A3	10	1	$ 43,668	$ 43,668	$ 43,668
4	A4	7	1049	$ 1,945	$ 23,260	$ 166,638
5	A5	20	254	$ 165,544	$ 938,808	$ 3,325,567
6	A6	5	3	$ 74,539	$ 103,941	$ 117,587
7	A7	10	4	$ 1,353,032	$ 2,977,785	$ 3,775,631
8	A8	11	31	$ 58,301	$ 457,558	$ 1,855,013
9	A9	6	60	$ 1,001	$ 6,604	$ 41,203
10	A10	4	10	$ 3,366	$ 12,604	$ 17,706
11	A11	15	7	$ 1,134	$ 20,061	$ 133,218
12	A12	3	1	$ 43,540	$ 43,540	$ 43,540
13	A13	2	11	$ 227,785	$ 564,991	$ 1,383,100
14	A14	1	0	-	-	-
15	A15	4	3	$ 13,136	$ 20,131	$ 32,247
16	A16	1	48	$ 9,203	$ 75,480	$ 297,415

The example above (Tables 5) clearly shows that there are many instances in which the user should not simply take the output of $5.3M and validate a funding request. The calculator is designed to callout instances where users should use caution when using outputs based on minimal historical data. Using row three as an example, there is only a single datapoint matching the user's parameters (ordnances originating from the city of Cheyenne). The advisable path may be to ask the publications team to crosscheck what data is available for the requirement and see how analogous the new ordnance is expected to be. Using row 14 as an example, clearly zero dollars would be an insufficient and incorrect answer. This example simply demonstrates that no historical analogies are available, so users should develop another estimation method and revalidate that no other effort drivers can be identified. It may be more helpful to generically estimate an expected coverage level "Silver" and move forward with a high level, high uncertainty estimate for this information set.

While Augur is refining the estimation methods, finalizing the calculator will still take months of revisions and close review of the source data before it can be used to baseline the FY25 spend plan. Occasional errors have also been identified in the data including an error that is the result of an automated "check in" time implemented by the workflow tool. A final review of a publication may take only 3 additional hours, but if a user forgets to check in the document, the system will automatically capture the task at 300 hours (for example). These errors are not always obvious when reviewing thousands of datapoints at the total level. After digging into individual reports, these errors have been identified and are being manually adjusted to improve the validity of the data before the calculator is baselined.

2.3 Phase III (Future Objectives)

On the tail end of Phase II, the Augur team participated in a Lean Six Sigma review with PMS 408 leadership and the Warfare Center Pubs team. This event led to even further insight into the publications process. In addition to identifying processes that can be improved and eliminated, these discussions led to an agreement that there are additional details that should be planned, estimated, baselined, and tracked (Publications, Subjects, Munition Data Reports, Render Safe Procedures, etc.). Augur also proposed a baseline procedure in line with a vendor IBR as well as a path to establish meaningful metrics surrounding performance and quality.

2.3.1 Enhanced Calculator

Following the Lean event, Augur was given direct access to the workflow tool used by the writers. With direct access to the system, all data fields are available for historical analysis. This data will enable another revision of the cost calculator based on the latest raw data, pulled and validated by Augur. This will validate current approaches and possibly identify other categories that can be further dissected and analyzed for statistical significance. Examples may include unique flow times by stage (from procedure to graphics to final review) and flow time by driving category (New York ordnances require more time in graphics than Austin ordnances). At this stage, Augur will also develop a nuanced estimation method to discretely account for efforts captured outside of the tool (obtaining assets, scanning assets, testing and exploitation of assets to enable procedures development and documentation). Once finalized, this calculator will be used as negotiation tool with the Warfare Center and will enable resource planning at the PMS 408 level. This tool will be utilized similarly to an Independent Government Cost Estimate (IGCE) used to evaluate fair and reasonable pricing from a vendor prior to contract award. Additionally, this tool will help facilitate discussions and ensure a shared understanding of the scope in question.

Another output of the enhanced calculator is total flow time for each requested information set. This information will be taken to a secondary planning step which looks at resource constraints to enable realistic annual planning. Writers are often highly specialized and therefore similar information sets may be required to be developed serially due to limited expertise available for a given ordnance type. Many tasks cannot be performed concurrently so the overall publications process cannot be perfectly captured with our parametric model without accounting for these limitations. This secondary step will help the PMO balance annual requests for the variety of resources available or justify the need for hiring specialized resources if the annual requirement is not flexible (i.e. hire 2 more Underwater Ordnance SMEs).

During this final effort, some visual changes will be made to the calculator to polish the user interface experience. A VBA-enabled control panel is in development which will prompt the user through a “wizard-like” workflow, prompting for the required inputs while hiding the mechanics of the calculator.

2.3.2 Baseline and Performance Management

This tool and the data behind it will be used to perform a baseline event, akin to an IBR. This step is critical, as no parametric model can perfectly capture such a complicated and highly variable process. This discussion will look at resource constraints, priorities, and backlog, and ensure a shared understanding of near-term scope. Methods for claiming performance will be agreed upon up front and custom data fields will be created in the workflow tool, enabling Augur to pull data with ease and generate updated dashboards to show information set progress. This baseline event will also establish a process for scope changes as prioritization changes constantly. This process, performed periodically, will help manage client expectations by showing how many tasks are in process. Typically, when priorities are adjusted, the desire is to finish the tasks that are in process and backlog those which have not begun, so using data to show a realistic timeline to pivot toward new priorities will help once again in facilitating a shared understanding of the complexity of the scope in hand. These events will also show the efficiency lost when priorities change frequently and staffing changes, cross training, and difficulty in surge tasking for such specialized labor.

2.3.3 Claiming Credit & WIP

Previous methods of reporting progress to the PMO and resource sponsors were limited to the number of information sets delivered compared to the quantity of requested information sets. This method did not capture effort applied to WIP nor did it account for varying complexity of the different information sets (i.e. land mine vs nuclear missile). This approach also failed to capture and document the frequency of re-prioritization, as there was no baseline event or clean process to re-baseline or re-measure the scope changes. The future method will normalize to hours required per information set, capturing the nuance of varying complexity. This universal measuring stick will also allow for rapid sizing of scope changes which happen throughout the year.

Preliminary examples of claiming WIP are shown below in Table 6.

Table 6 Credit Claimed Model

Notional Work In Process Credit		
Status	**Examples**	**Credit**
Created	Title and minimum one other field generated (Description or Image)	0%
Started	Created and now prioritized, work started	5%
Documentation 25%	25% of raw documentation generated and sourced	25%
Documentation 50%	50% of raw documentation generated and sourced	50%
Review	Review documentation, graphics included	75%
Approval	Final review of text, sources, and graphics	90%
Published	Achieved Max Coverage Level. Published to portal	100%

Using the rubric above, the percentages (tailored by ordnance and process type) will be applied to the initial expected value of a unique ordnance information set. A method for sizing scope will likely be normalized to expected hours, and an example would be a land mine being "worth" 400 hours and a nuclear missile being worth 20,000 hours.

2.3.4 Dashboard

Augur is in the process of developing a Power BI Dashboard that will compare a baseline plan against monthly actuals. EVM is not applicable, but similar metrics can be used to measure the progress of each deliverable (information set). Augur and PMS 408 are in the process of agreeing upon some creative metrics that are expected to be more meaningful than current methods of simply counting items published. Figure 14 below showcases a Power BI dashboard of the data at hand, which will be leveraged in the future to baseline expected values per task and progress against those tasks.

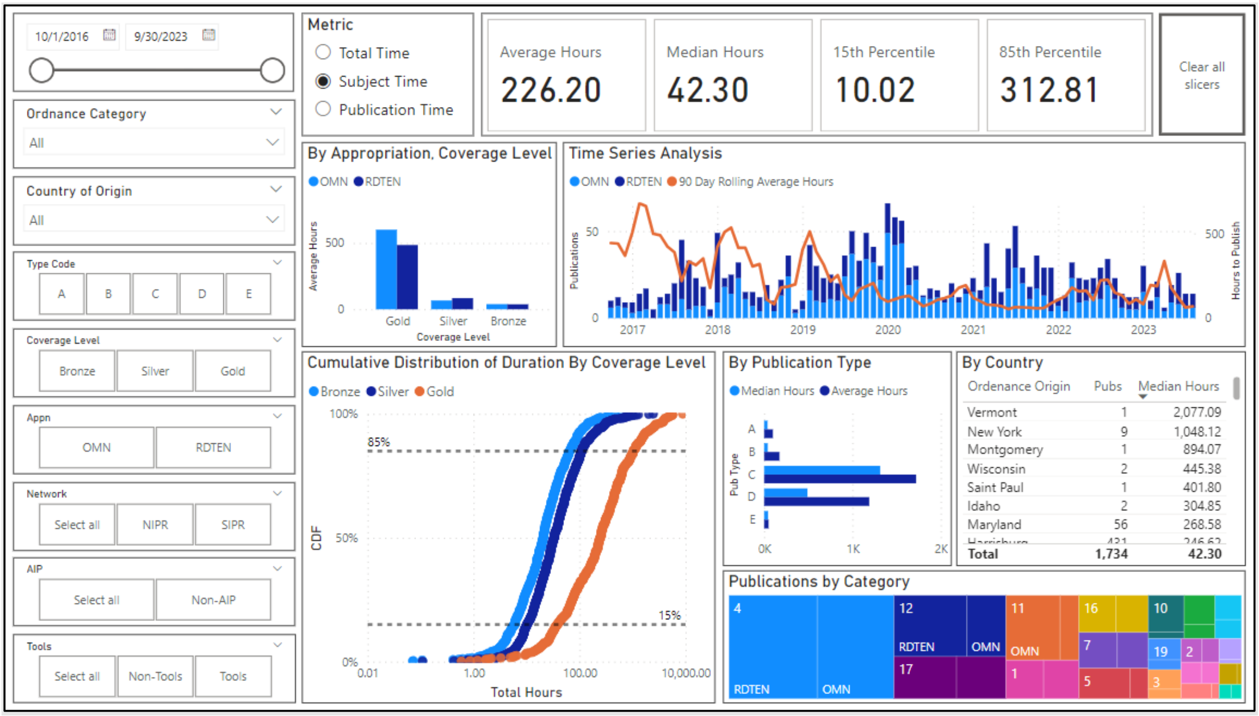

Figure 14: Prototype of Publications Power BI Dashboard

Below are some examples of some of the objectives of the future dashboard:

1. Value of work accomplished (Cumulative Value of completed tasks + value of WIP vs baseline)
2. Hours per Information Set (Cost of labor hours, planned vs actual)
3. Flow Time of each Information Set (Schedule durations, planned vs actual)
4. Backlog burn down
5. Quality: Internal rework instances
6. Quality: External requests for revisions due to lack of completeness or errors

3.0 Conclusion

Data analysis is often an iterative process which has led to continuous refinement to forecast and estimate the publications process as new information is provided. Phase I produced the bedrock of this analysis and created a superior method for long-term planning. The Augur team used Phase I to develop an understanding of the scope through solution provider discussions as well as interpreting the associated data through visualizations. Phase I analysis was constrained due to limited data as well as focusing on a single primary effort driver in coverage levels.

In Phase II, the analysis expanded to include a more complete dataset as well as consideration of several different effort drivers. Phase II analysis focused on short-term planning through an improved analytical process to extract deeper insights into annual planning and the associated funding requests. This effort included tackling the challenge of assessing fair and reasonable pricing by leveraging historical data to forecast the number of hours and costs per publication. To solve this problem, the Augur team developed a dynamic calculator that provides these projected hours/costs based on the primary effort drivers (defined by user input). Furthermore, a recent agreement between the stakeholders established a plan on how to leverage data to size scope, baseline work planned, and visualize performance and other real-time information. Once fully implemented, the Publications program will have a data-informed approach to integrate program management. Phase III expands the robustness of Augur's analysis to bring cost realism, program planning, and accountability to solution providers.

Phase III strives to capture a holistic picture of the publications effort by capturing indirect support efforts. In the end, the methods outlined intend to streamline the ability to forecast the publications effort accurately and refine the process as information becomes available. All of these insights will come together in a dashboard which baselines annual requirements and measures performance in an automated fashion.

Augur will continue to refine the analysis and provide analytical decision support to the Navy client. Never accept an answer of "It is impossible to estimate the work that we do." Always push for access to data. In this case, the analysis performed has helped facilitate a shared understanding of scope and has increased accountability for the solution providers.

Appendix A : Table of Acronyms

Acronym/Abbreviation	Meaning
AIP	Assignment Incentive Pay
CDF	Cumulative Distribution Function
CL	Confidence Level
CUI	Controlled Unclassified Information
CY	Calendar Year
EOD	Explosive Ordnance Disposal
EVM	Earned Value Management
FTE	Full Time Equivalent
FY	Fiscal Year
IBR	Integrated Baseline Review
IDAS	Information Development and Approval Standard
IED	Improvised Explosive Device
IGCE	Independent Government Cost Estimate
NEPS	NAVSEA Enterprise Planning System
O&M	Operations & Maintenance
PoR	Program of Record
PMO	Program Management Office
PMS 408	Program Management Support Expeditionary Missions
RDT&E	Research, Development, Testing & Evaluation
SBM	Solution Business Manager
SME	Subject Matter Expert
VBA	Visual Basic for Applications
WIP	Work in Process
WMD	Weapons of Mass Destruction