Seven Degrees of Separation: The Importance of High-Quality Contractor Data in Cost Estimating

Crickett Petty, Missile Defense Agency

Dr. Christian Smart, Missile Defense Agency

James Lawlor, Computer Sciences Corporation

Abstract

The popular notion that any two people are linked to one another on average via a chain with "six degrees of separation" is based on a sample of only 64 people (Travers & Milgram, 1969). In 2007 Microsoft researchers discovered the average degrees of separation is closer to seven, based on an analysis of 30 billion conversations among 180 million people. This more in-depth analysis also yielded insights into the maximum degrees of separation. This highlights the need for complete and thorough data in any analysis. In particular, the quality of data is critical to a cost estimate's credibility. Poor data lead to poor estimates, as encapsulated in the maxim "garbage in, garbage out." The best methodology and analysis are futile efforts without a firm basis in sound, applicable, and well-documented data. However, the collection of accurate, high-quality contractor cost data is challenging. These challenges are discussed in detail. We present the Missile Defense Agency's (MDA's) data collection and validation process, including the MDA's collaboration with the Office of the Secretary of Defense's Cost Assessment and Program Evaluation (OSD CAPE) office.

Introduction

The popular notion that any two people are linked to one another on average via a chain with "six degrees of separation" is based on a sample of only 64 people (Travers & Milgram, 1969).

This notion is entrenched in popular culture; it has inspired a play, television shows, and films. There is even a game involving film actor Kevin Bacon. However, in 2007, researchers at Microsoft discovered that the average degree of separation is closer to seven (Leskovec & Horvitz, 2007). They examined 30 billion conversations among 180 million distinct users on Microsoft Messenger instant-messaging network during a one-month period. This more in-depth analysis also yielded insights into the maximum degrees of separation, which they found to be equal to 29. The researchers also discovered that people have more and longer conversations with those who are similar to themselves in terms of age, language, and location. They also discovered that cross-gender communications are more frequent and longer in duration that same-gender communications.

This highlights the need for complete and thorough data in any analysis. More data allow for deeper insights not possible with smaller quantities. Also, the quality of data is critical to a cost estimate's credibility. Poor data lead to poor estimates, as encapsulated in the maxim "garbage in, garbage out." The best methodology and analysis are futile efforts without a firm basis in sound, applicable, and well-documented data. The MDA has a well-documented process for collecting and validating contractor cost data. However, the collection and validation of accurate, high-quality data are challenging. These challenges are discussed in detail. The MDA's work with the Office of the Secretary of Defense's Cost Assessment and Program Evaluation (OSD CAPE) office is also discussed.

The Importance of Data

At the 2014 ICEAA workshop, one of the authors presented a paper on the application of Bayesian techniques to develop accurate estimates with limited data. (Smart, 2014). The lack of applicable data available to develop cost estimates for government programs motivated the research. It is intended as a measure to use until more data can be collected, or in cases where there is little direct history and thus few truly applicable predecessors. It was not intended as a replacement for cost estimating relationships based on large data sets, since as noted by Google researchers in 2009, "Invariably, simple models and a lot of data trump more elaborate models based on less data" (Halevy, Norvig, & Pereira, 2009).

The quality of data is critical to the estimate's credibility. Lack of data often leads to poor estimates, as encapsulated in the maxim "garbage in, garbage out." The best documentation, methodology, risk analysis, etc., are futile without a basis in sound, applicable, well-documented data. Depending on the data quality, an estimate can range from a mere guess to a highly defensible cost position.

Sound data are critical to a quality estimate because the quality of a cost estimate depends largely on the quality of the data collected. Data collection is the single most important component of cost estimation. The history of cost estimating demonstrates repeatedly that successfully defending a cost estimate requires:

- The use of sound, quantitative cost data–not conjecture,
- An understanding of the data in detail (costs, prices, technical, and programmatic aspects),
- An awareness of problems or weaknesses in the data, and
- A willingness to discuss or share the data with critics.

Data provide: information on cost trends over a variety of related programs; insight into cost structures; and the information used to develop cost estimating relationships (CERs), parametric estimates, and models. The paradigm that "what has happened in the past repeats itself" is the basis for all cost estimating.

Data can be viewed as the lowest level of abstraction from which information and knowledge are derived. The relationship between data and information can be represented as a pyramid with data as the base rising through ground rules, assumptions, and analysis to the cost estimate at the apex. See Figure 1.

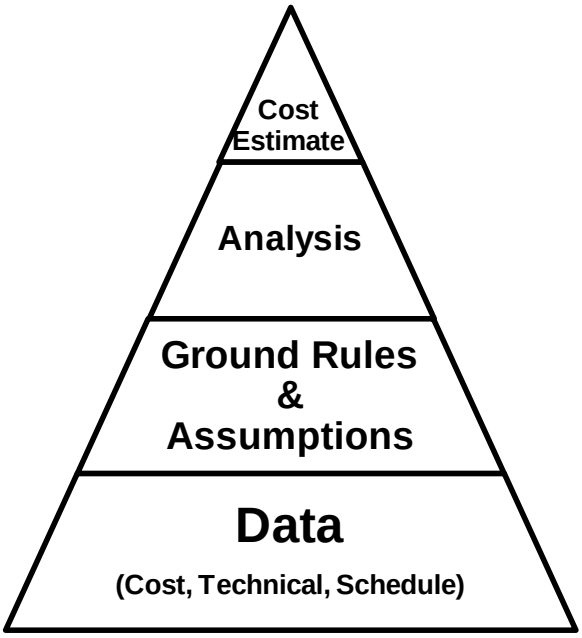

Figure 1 (Missile Defense Agency [MDA], 2012)

Large quantities of data are needed to develop accurate, sophisticated cost models with a large number of variables. When developing regression-based CERs, for example, we need a large number of data points. This need grows as we add more variates to the regression model in an attempt to get a better fit. A general rule of thumb is that we need at least 10 data points in the training set for every variable, and one-third of the data should be held out for a test set. For a five-variate regression model we would need at least 75 data points: 50 for the training set and 25 for out-of-sample testing to make sure we don't overfit. For a 10-variate model, we need at least 150 data points; 100 for the training set and 50 for out-of-sample testing.

The need for quality data is compounded by the fact that the government produces specialized hardware with few directly applicable analogies. For example, NASA has developed only a few human-rated launch systems, and the Missile Defense Agency has developed only a few kill vehicles. In such cases, lower levels of detail, such as at the subsystem and component level, can

enable us to use components that have been used in other systems to be included in data sets for developing more sophisticated cost estimating relationships. For example if we only have cost and technical data at the kill vehicle level, there are only a handful of applicable data points. However, if we have lower levels of details, such as cost and technical data at the attitude control system level or power system level, we can include data from other systems in developing a Kill Vehicle cost model. See Figure 2 for the Kill Vehicle WBS proposed for inclusion in MIL-STD-881.

MIL-STD-881C Missile System with Kill Vehicle

C.3 WORK BREAKDOWN STRUCTURE LEVELS

WBS#	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6
1.1.8			Kill Vehicle			
1.1.8.1				Kill Vehicle Structure and Harnesses		
1.1.8.2				Divert and Attitude Control System (DACS)		
1.1.8.2.1					DACS Integration, Assembly, Test and Checkout	
1.1.8.2.2					Divert Subsystem	
1.1.8.2.3					Attitude Control System	
1.1.8.2.4					Gas Generator/Structure	
1.1.8.2.5					Controller Electronics	
1.1.8.2.6					Ordnance Initiation Set	
1.1.8.2.7					Flight Termination System	
1.1.8.2.8					DACS Software Release 1…n	
1.1.8.3				Power and Distribution		
1.1.8.3.1					Power and Distribution Integration, Assembly, Test and Checkout	
1.1.8.3.2					Primary Power	
1.1.8.3.3					Power Conditioning Electronics	
1.1.8.3.4					Distribution Harness	
1.1.8.3.5					Power and Distribution Software Release 1…n	
1.1.8.4				Guidance and Control Processing		
1.1.8.4.1					Guidance and Control Processing Integration, Assembly, Test and Checkout	
1.1.8.4.2					Seeker Assembly	
1.1.8.4.3.1						Seeker Integration, Assembly, Test and Checkout
1.1.8.4.3.2						Optical Telescope Assembly
1.1.8.4.3.3						Focal Plane Array
1.1.8.4.3.4						Cooling Assembly
1.1.8.4.3.5						Electronics
1.1.8.4.3.6						Gimbal Assembly
1.1.8.4.3.7						Seeker Software Release 1…n
1.1.8.4.3					Guidance Computer	
1.1.8.4.4					Guidance and Control Processing Software Release 1…n	
1.1.8.5				Navigation		
1.1.8.5.1					Navigation Integration, Assembly, Test and Checkout	
1.1.8.5.2					Sensor Assemblies	
1.1.8.5.3					Navigation Software Release 1…n	
1.1.8.6				Communications		
1.1.8.6.1					Communications Integration, Assembly, Test and Checkout	
1.1.8.6.2					Communications Subsystem	
1.1.8.6.3					Antenna Assembly	
1.1.8.6.4					Communications Software Release 1…n	
1.1.8.7				Reserved		
1.1.8.8				Kill Vehicle Integration, Assembly, Test and Checkout		
1.1.8.9				Systems Engineering		
1.1.8.10				Program Management		
1.1.8.11				System Test and Evaluation		
1.1.8.12				Peculiar Support Equipment		
1.1.8.13				Common Support Equipment		

Figure 2. Missile Defense Agency Kill Vehicle WBS (Tarin, Tetrault, & Smart, 2014).

Analysts spend much of their time developing and honing tools and techniques and documenting estimates, but should spend the majority of their time and effort collecting a large quantity of high-quality data.

Collecting Contractor Cost Data

Financial data are complex. Depending on the source, reported amounts may represent the government purchase price of the end item or the actual cost and effort expended in the design and production of a product. Cost can be recorded at a high level, such as at the system level, with no level of insight below it (e.g., "kill vehicle"); or it can be recorded at a low level, such as the subsystem or component level (e.g., "guidance control processing," "focal plane array"). Component costs may be broken down further based on functional categories including material and labor or non-recurring and recurring. Data may or may not include general and administrative costs. The source may or may not separate fee, profit or loss. Low levels of detail are not always needed. For example, a quick estimate could use a systems-level analogy; however, when planning budgets, lower levels of detail are required to make optimum use of prior actuals. Having access to lower levels of detail will enable the analyst developing a cost estimate to assess the complexity of the cost components of the previous actuals, develop a more accurate estimate with more insight into the composition of past programs, and ensure comparability to the proposed system being estimated. Therefore, the context of the data must be understood before it is used in cost estimation. Correct understanding of the composition of the costs is crucial to accurate estimates.

For example, Contract Line Item Numbers (CLINs) from contracts and contract modifications are prices, not actual costs. Just because a price is negotiated on a single contract does not mean the contractor expends the CLIN price in the design or production of a product. Negotiated price data are not necessarily useful for inclusion in future cost estimates. Analysts must consider that the negotiated CLIN price is not based solely on the work required to deliver the product; it is confounded by the contractor's and government's projection of cost and risk throughout the contract (and the accuracy thereof), budget stability, profit and performance incentives, and overall program context at the time of negotiations. Furthermore, CLINs are not always product-oriented and may represent varying scopes. To optimize contract management efforts, the MDA enacted policy in 2014 to clarify and narrow CLINs on future contracts to be more singularly related to products.

This negotiation is most easily illustrated in the compromise of Firm-Fixed Price (FFP) contracts and CLINs. FFPs are characteristically a negotiated price the government agrees to pay a contractor for work and/or goods. When examining FFP from contracts, the cost estimator cannot clearly extract the general and administrative expense, the cost of money, or the loss or profit the contractor made from the sale. Without that knowledge, the estimator cannot

accurately predict the cost of comparable future efforts, particularly if an analogous future effort requires a different contract type.

Earned value data are similarly complex. Integrated Program Management Reports (IPMRs), formerly known as Contractor Performance Reports (CPRs), focus on the value and management of high-risk contract items/efforts and are not all inclusive of the work on a contract. Earned value data are often based on a product-oriented Work Breakdown Structure (WBS). However, the entire WBS for a contract is often not represented in IPMRs or CPRs and the WBS often does not have enough depth to identify cost drivers for future cost estimation. General and administrative costs are not segregated from the individual WBS elements, further complicating the use of earned value data by cost estimators.

To resolve the issues of contract and earned value data, requirements for Contractor Cost Data Reports (CCDRs) should be placed on contracts to fully capture the actual costs needed for future estimation. CCDRs, also known as the DD Form 1921 series, are essential. In accordance with DoDI 5000.02 (Operation of the Defense Acquisition System), the MDA policy requires that CCDRs are placed on contracts and subcontracts above $50 million and are optional on any contract above $20 million with high-risk or high-technical interest (Under Secretary of Defense (AT&L), 2015). CCDRs should follow the product-oriented Work Breakdown Structures defined in MIL-STD-881C. The WBS should also be consistent with the WBS presented in earned value reports, though the CCDR may request reporting at a much lower level than the IPMR or CPR. Most importantly, the Data Item Descriptions authorizing the collection of CCDRs require contractors submit “actual cost data,” further defined in the Cost and Software Data Reporting (CSDR) Manual (DoDM 5000.04-M-1) as “the costs sustained in fact, on the basis of costs incurred, as opposed to standard or predetermined costs. Estimated actual costs may be used for actual costs that have not been recorded in the books of record, when based on verifiable records such as invoices and journal vouchers that have not yet been accrued in the books of record, to ensure all valid costs are included. Actual costs to date include cost of direct labor, direct material, and other direct charges specifically identified to appropriate control accounts as incurred, and any costs and general administrative expenses allocated to control accounts” (Department of Defense [DOD], 2011).

Contractor Cost Data Reports provide data not readily available from other reporting sources. The collection of accurate, valuable CCDRs for cost estimating purposes hinges on three phases: contract setup, validation of data, and the accessibility of authenticated data.

Contract Setup

As mentioned previously, the MIL-STD-881C does not account for Kill Vehicles. However, with the OSD CAPE-coordinated Kill Vehicle Work Breakdown Structure, the majority of the MDA Contract Work Breakdown Structures (CWBS) on new efforts should comply with the military standard (Tarin, Tetrault, & Smart, 2014). The most important issue then becomes coordination,

both within the procuring agency and between the customer and contractors. In the MDA, the Cost Estimating and Analysis Directorate (DOC) must preplan requirements and the CWBS with the Earned Value Directorate, the Acquisition Directorate, and the Program Office. Why must all the players coordinate?

1. The Program Office should confirm the proposed CWBS captures the entire effort. Engineering opinion is pertinent in making sure high-risk or highly-technical efforts are fully defined to an adequate precision in the CWBS. Furthermore, engineering opinion is critical to ensure that the CWBS structure is correctly partitioned into a complete set of non-overlapping, product-oriented subsystems.
2. The Earned Value Directorate CWBS should match the CCDR WBS at a high level to ensure efficiency in the contractors' accounting system. If the WBS is different between earned value and cost, the contractor will be compelled to map the accounts between reporting structures, creating room for error and analyst subjectivity and increasing the cost of the data deliverables.
3. The Acquisition Directorate must include the appropriate cost Contractor Data Requirements List (CDRLs, DD Form 1423-1), Cost and Software Data Reporting Plan (CSDR Plan, DD Form 2794), and Resource Distribution Table (RDT) as early in the contracting process as possible. The acquisition strategy should be made clear to all stakeholders in order to create a foundation for accurate, efficient, and traceable cost reporting. As with most items in a contract, early planning ensures better compliance and is much cheaper than last minute requirements.

Once the contractor and subcontractors have been identified, a post-award conference with the reporting contractors can simplify reporting and identify any edits that should be made to the reporting plans. An ideal post-award conference should include: the contracting officer; business and engineering representatives from the prime contractor and all reporting subcontractors; engineering, earned value, and cost representatives from the Program Office; and Defense Cost and Resource Center representatives from the OSD CAPE. In the absence of a post-award conference, the CSDR Plan (DD Form 2794) should be circulated to the above parties as early as possible to resolve any misconceptions or inefficiencies with the cost data reporting requirements before deliverables are due.

Validating Contractor Cost Data

Once the CCDR requirements are in place on a contract, the cost estimator must be vigilant in the collection and validation of the data. At this time, the MDA does not have a single Agency-wide CDRL-management system. This requires the Cost Estimating and Analysis Directorate (DOC) to be flexible with each Program Office, contractor, and subcontractor in order to collect and communicate information in a timely manner.

For example, some Program Offices have a CDRL collection protocol which notifies reviewers. Other Program Offices use Microsoft SharePoint sites with special permissions. Other prime contractors have proprietary systems for delivering data. Some prime and subcontractors deliver using the Cost and Software Data Reporting Submit-Review site hosted by the Cost Assessment and Program Evaluation office (OSD CAPE). Other contractors simply provide data and/or notification through email to the csdr@mda.mil address managed by DOC's Research Team. Still other contractors have insisted on delivering data by CD via the mail!

Once the data are provided to the DOC Research Team, the team uses the OSD CAPE's CSDR Planning and Execution Tool (cPet) to perform simple checks for summation of children to parents, non-recurring and recurring to totals, and consistency across the multiple DD 1921 Forms.

After the preliminary checks are performed using cPet, the Research Team then coordinates with Program Office cost estimators to ensure accurate content and detailed documentation of the reported costs. The content checks include accurate unit reporting, an understanding of non-recurring versus recurring costs, full and complete definitions are contained within CWBS dictionaries, appropriate functional cost break-out (engineering labor, manufacturing labor, and materials) for each WBS element, and appropriate subcontractor reporting that is traceable and consistent with the Prime Contractor Cost Data Reports.

The MDA has found that a close working relationship between the contractor CCDR preparers and the Research Team reviewers can significantly aid the validation process, particularly during the detailed review stage. Close examination of the cost details often reveals apparent errors or questions that may be resolved with contractor subject matter expertise. For example, an apparent discrepancy between cost or unit quantities between prime contracts and subcontracts may be due to excess hardware transferred as government-furnished equipment from a prior contract. A contracts management WBS reporting manufacturing labor cost, an apparent error, may be due to production engineering consultations during proposal preparation. Often, these questions may be answered by communication with the reporting contractor during the validation process, and the results compiled and documented along with the CCDR. In many cases, the MDA may then request that the contractor include that supporting documentation with subsequent annual CCDRs and avoid the need to require revisions.

All errors, inadequacies, and questions are then compiled in a validation report which is provided to the Contracting Officer for communication back to the Prime Contractor. Subcontractor validation reports for direct-reporting subcontractors are communicated directly to the subcontractor, with a sanitized notification to the Prime Contractor. When corrections are provided by the contractor and/or subcontractor, the entire validation process is performed again, until the CCDR passes validation. The bottom-line question that the Research Team uses to guide validation is: "Will this report alone make sense to a new analyst five years from now?"

Accessibility of Collected Data

Once a CCDR is authenticated, the contractor must upload the data into the OSD CAPE's CSDR-Submit Review system. The data are then approved with a Memo Date and pushed into the OSD CAPE Defense Automated Cost Information System (DACIMS) repository providing access to all government DoD cost community members. In addition, the final CCDR and validation notes are posted to an internal MDA electronic library for access to DOC members.

Successes

DOC initiated its CCDR review process in January 2012. In the ensuing three years, DOC has made significant accomplishments in improving the timeliness of data submissions, the quality of the reported data, the contractual requirements governing the data submissions, and working relationships with the data preparers.

Figure 3 demonstrates the progress made by the MDA in improving the CCDR reporting coverage and quality of said CCDRs on one of its largest and most complex programs. From 2006 through 2014, Program A had a total of ten reporting contracts (or separately-reporting CLINS); prior to 2012, none of these data were formally validated, and a significant number were not available in the MDA electronic library. Beginning with the 2011 and 2012 data, the DOC validated the majority of the contractor cost data. As DOC continued to heighten internal and external focus on contractor cost data, the MDA was able to obtain full CCDR coverage of the major contracts and subcontracts for Program A.

As a result of a strong working relationship between the contractor CCDR preparers, the MDA Program Offices, and the DOC CCDR reviewers, the MDA has achieved particular success in increasing the contractors' documentation within the DD Form 1921 "Remarks" field. Prior to DOC's work with contractors, CCDR remarks would include only a few sentences and documentation on the individual WBS costs were not consistently included. Currently-reported CCDRs, however, commonly include several paragraphs' worth of documentation, which greatly increases the reports' utility to MDA analysts.

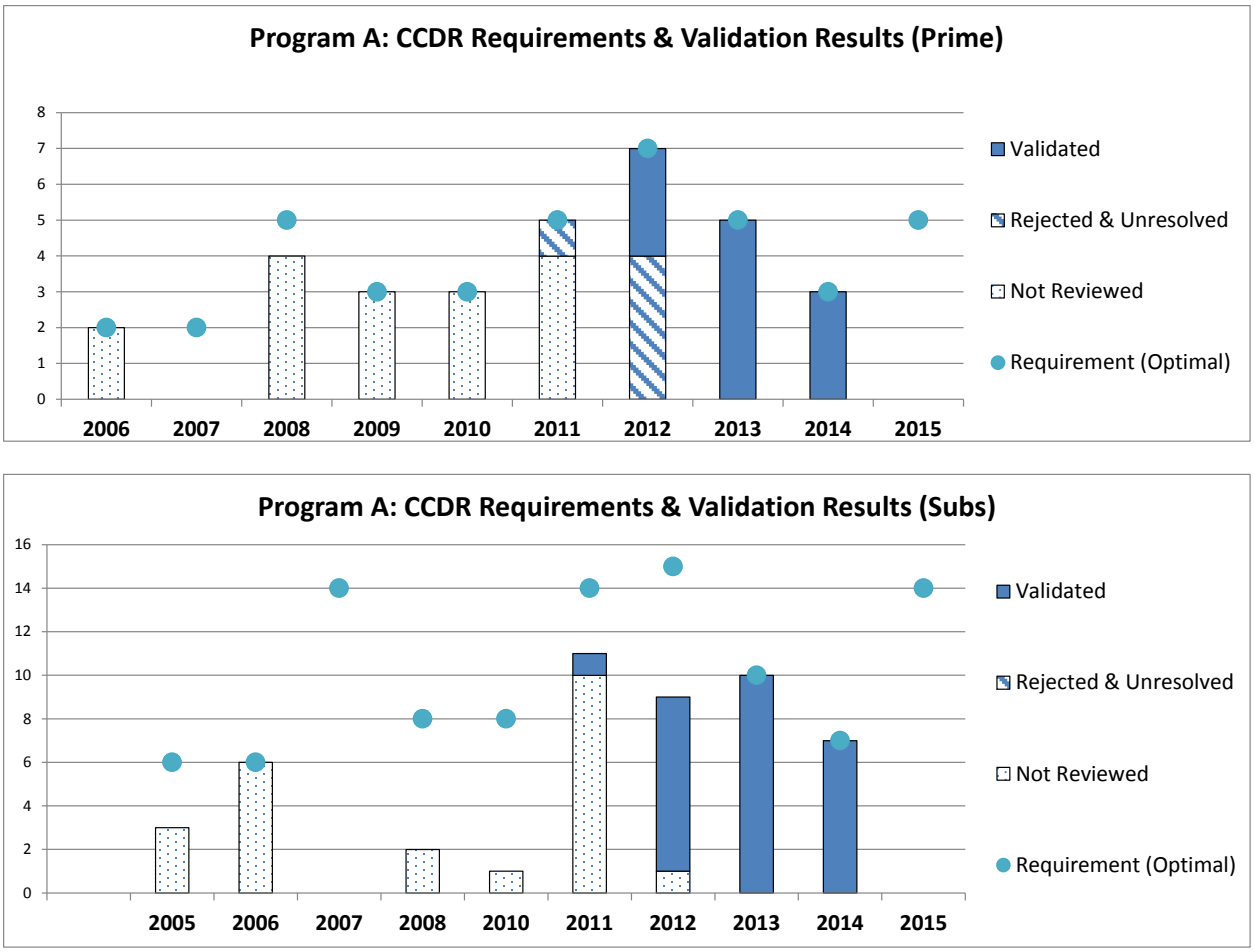

Figure 3:Program A CCDR Requirements & Validation Results 2005 through 2015

When the CCDR preparers ensure data are submitted without fundamental errors, by using the OSD CAPE's cPet software, for example, DOC is able to focus on building a comprehensive understanding of the contractor's data. This evolution can be illustrated by examining the annual validations for one of the MDA's critically important development contracts:

- **2012:** The 2012 data were submitted with a significant number of fundamental errors, making a detailed review of the data impossible. The first revision of the data still contained errors rendering it unusable, including missing elements of the WBS and incorrect functional designation of costs. The second revision was accepted.
- **2013:** The 2013 data were fundamentally correct, save for a single WBS element that did not show the proper cost break-out. The MDA required revision of the report in order to address a significant cost decrease versus the 2012 submission and a small-value CLIN that was excluded from the report. The first revision was accepted.

- **2014:** The 2014 data was free of fundamental errors and well-documented, due to the contractor's improved CCDR business processes and investment into a new capability to automatically generate CCDR data into the OSD CAPE "Flat File" format. As a result, the DOC validation process focused on a thorough understanding of the contractor's delineation between recurring and non-recurring cost and ensuring optimal DD 1921-2 reporting in the future. The report was accepted without revisions.

The data accepted through the MDA validation process has already supported several key analyses within the MDA and the OSD CAPE. In the future, analysts will need less time for extensive research of other records and interviews with experts who participated in the programs at the time of execution, because the CCDR data provides stand-alone documentation explaining the content and scope of historical costs. Due to the importance of sound, historical data to support decision-making, the MDA will continue to prioritize the collection, validation, and use of CCDR.

Challenges

Legacy contracts necessitate more analysis and record-keeping to provide useful data to analysts. It is not feasible to change the reporting structures of established contracts. Using legacy reports requires more tacit knowledge of the products and historical performance to ensure estimators use the data accurately.

Consistent contract requirements placed at the appropriate time in the acquisition process requires vigilance and constant communication between DOC, Program Offices, and the Acquisition Directorate. Including prime contractors and subcontractors via coordination of post-award meetings can expedite reporting and clarify any questions.

Program directors, cost estimators, budget financial managers, contractor officers, and contractors must be educated on the importance of collecting Contractor Cost Data Reports throughout the contract. Collection of data after a contract has been let or completed is time-consuming and expensive.

Reporting accuracy, particularly with regard to the fundamental data requirements (such as summation of children to parents, non-recurring and recurring to totals, correspondence between forms, and adherence to the prescribed WBS), remains one of the primary challenges to MDA data collection. A survey of 105 CCDRs accepted or currently in review by the MDA during calendar years 2013 through 2015 reveals that more than half of CCDRs require revisions. Over one-fifth require multiple revisions, which can result in significant delays to obtaining useable data, given the standard review cycle that allows thirty days each for review and for submission of revisions. See Table 1. This trend has remained relatively stable over time, highlighting the need for greater attention to fundamental data quality by the entire cost analysis community. Further analysis of the data demonstrates a significant variance in the need for revisions between the major Program Offices making up the MDA (Table 2). The number of CCDRs requiring

revisions varies from twenty to eighty percent, depending on the Program Office. The variance is driven by factors including program size and complexity, contractor CCDR experience, and contractors' data collection methods and business processes for CCDR preparation.

MDA CCDR Revisions Required	Total CCDR	%
0	**47**	**45%**
1	**34**	**32%**
2	**17**	**16%**
3	**5**	**5%**
4	**1**	**1%**
5	**1**	**1%**
Total	**105**	**100%**
Among MDA CCDR accepted January 2013 through March 2015.		

Table 1. Number of CCDR Revisions Required by the MDA 2013 - 2015

MDA CCDR Revisions Required	Program A CCDR	%	Program B CCDR	%	Program C CCDR	%	Program D CCDR	%	Program E CCDR	%	Program F CCDR	%
0	5	21%	3	60%	7	47%	13	68%	5	83%	13	39%
1	10	42%	2	40%	6	40%	2	11%	0	0%	13	39%
2	7	29%	0	0%	1	7%	1	5%	1	17%	6	18%
3	2	8%	0	0%	1	7%	1	5%	0	0%	1	3%
4	0	0%	0	0%	0	0%	1	5%	0	0%	0	0%
5	0	0%	0	0%	0	0%	1	5%	0	0%	0	0%
Grand Total	**24**	**100%**	**5**	**100%**	**15**	**100%**	**19**	**100%**	**6**	**100%**	**33**	**100%**

Table 2. Number of CCDR Revisions Required by the MDA by Program.

Certain errors are common across programs and contracts. The most common are simple arithmetic errors that result in reports that cannot be used by the MDA (on several occasions,

such errors have been traced to copy-and-paste errors or keystroke errors). DOC has demonstrated that these errors can be greatly reduced by providing CCDR training to contractors and encouraging efficient CCDR business processes. Successful contractor business practices include: using cPet to perform quality control checks before submission to the government; setting up control account structure to match the CCDR and earned value reporting WBS; automatically generating reports directly from accounting systems; participating in post-award conferences and WBS reviews; and reaching out to DOC with questions prior to submission. Frequently, the MDA encounters CCDRs that do not capture all contract costs, commonly due to the misconception that FFP or small-dollar CLINs are excluded from reporting (as is the case with earned value reporting). Similarly, it is common for the MDA to receive CCDRs that do not break out the required summary elements (such as general and administrative cost, fee, and cost of money). Both issues may be solved easily with education and clear communication.

Both the OSD CAPE DACIMS and the internal MDA electronic library use a file-tree structure organized by Program and Contract. After years of experience using existing libraries, both the MDA and the OSD CAPE have found this organizational paradigm cumbersome, making aggregating data across programs labor-intensive. In order to address the data organization challenge and optimize resources, the OSD CAPE is working an initiative called Cost Assessment Data Enterprise (CADE) to aid in the accessibility of all DoD cost data. The MDA is working a complementary initiative called Site Of Useful Records for Cost Estimating (SOURCE) to marry internal MDA CCDR data, earned value Data, and technical parameters to provide an easy to query repository for better analogous and parametric analysis.

Path Forward

In order to collect the most complete and accurate contractor costs, DOC must continue to educate internal stakeholders and communicate with prime contractors and subcontractors. DOC has membership on the Integrated Business Processes Team establishing new guidance for policy overlapping acquisition and operations. DOC also serves as a functional data manager, providing guidance for the Acquisitions Policy directorate.

In order to lay a foundation for continued improvement in CCDR quality, DOC must continue to lead industry partners in improving CCDR business processes. DOC works directly with preparers to prevent errors before they occur and ensure that cost data include comprehensive documentation. Furthermore, DOC encourages the MDA to incentivize contractors who provide timely, high-quality cost data. To be truly efficient and provide the most worth to the cost community, DOC must continue to communicate with and improve processes involving the OSD CAPE. DOC has provided Career Development Program participants the opportunity to sit at the DCARC offices in order to manage, troubleshoot, and improve the collection of MDA data in the OSD CAPE's community repository, DACIMS.

DOC also provides a liaison for industry questions in provision of data through the CSDR Submit-Review system. As part of the validation process, DOC is ensuring newly delivered reports are uploaded to the DCARC repository, while also uploading historical records to provide the OSD CAPE and other government analysts the most complete representation of program costs.

Conclusion

Of all major sources of contractor financial data, Contractor Cost Data Reports provide the most utility in estimating future costs. However, to maximize the value of the data, requirements must be understood by the Program Office, the contractors, and Contracting Officers. Requirements must be defined as early in the contracting process as possible and applied consistently across programs. The CCDRs must be validated through a rigorous process to ensure accuracy and precision of the data. Once authenticated, the data, with all supporting explanations and information, should be made easily available to analysts to provide a solid basis for future estimates.

“Cost estimating requires a continual influx of current and relevant cost data to remain credible. The cost data should be managed by estimating professionals who understand what the historical data are based on, can determine whether the data have value in future projections, and can make the data part of the corporate history” (U.S. Government Accountability Office [GAO], 2009).

References

Bishop, C. (2009). *Pattern Recognition and Machine Learning*. New York, NY: Springer Science+Business Media, LLC.

Department of Defense. (2011, November 4). *Cost and Software Data Reporting (CSDR) Manual* (DOD Manual 5000.04-M-1). Washington, DC: Author.
Retrieved from: http://www.dtic.mil/whs/directives/corres/pdf/500004m1.pdf

Halevy, A., Norvig, P., & Pereira, F. (2009, March/April). The Unreasonable Effectiveness of Data. *IEEE Intelligent Systems*, 24(2), 8-12. Retrieved from: http://static.googleusercontent.com/media/research.google.com/en/us/pubs/archive/35179.pdf

Leskovec, J., & Horvitz, E. (2007, June). *Planetary-Scale Views on an Instant-Messaging Network* (MSR-TR-2006-186). Redmond, WA: Microsoft Research. Retrieved from: http://arxiv.org/pdf/0803.0939v1.pdf

Missile Defense Agency (MDA). Director of Operations. (2012, July). *Cost Estimating and Analysis Handbook.* Fort Belvoir, VA: Author.

Office of the Secretary of Defense Cost Assessment and Program Evaluation (OSD CAPE). (2011, May) *Data Item Description: Cost Data Summary Report (DD Form 1921)* (DI-FNCL-81565C). Washington, DC: Author.
Retrieved from: http://dcarc.cape.osd.mil/Files/Policy/2011-1921.pdf

Smart, C. Bayesian Parametrics: How to Develop a CER with Limited Data and Even Without Data. (2014, June). Presented at the 2014 International Cost Estimating and Analysis Association's Professional Development and Training Workshop, Denver, CO.

Tarin, J., Tetrault, P., & Smart, C. (2014, June). Kill Vehicle Work Breakdown Structure. Presented at the 2014 International Cost Estimating and Analysis Association's Professional Development and Training Workshop, Denver, CO.

Travers, J. & Milgram, S. (1969, December) An Experimental Study of the Small World Problem. *Sociometry*, 32(4), 1969.

Under Secretary of Defense (AT&L). (2015, January 7). *Operation of the Defense Acquisition System* (DOD Instruction 5000.02). Washington, DC: Author.

U.S. Government Accountability Office (GAO). (2009, March). *GAO Cost Estimating And Assessment Guide* (GOA-09-3SP). Washington, DC: U.S. Government Accountability Office.
Retrieved from: http://www.gao.gov/products/GAO-09-3SP