## Air Force Institute of Technology

## **AFIT Scholar**

**Faculty Publications** 

10-2018

# Text Mining Analysis of Acquisition Reforms and Expert Views

Amanda L. McGowin

Jonathan D. Ritschel Air Force Institute of Technology

R. David Fass Air Force Institute of Technology

Bradley C. Boehmke

Follow this and additional works at: https://scholar.afit.edu/facpub



Part of the Government Contracts Commons, and the Legislation Commons

#### **Recommended Citation**

McGowin, A. L., Ritschel, J. D., Fass, R. D., & Boehmke, B. (2018). Text Mining Analysis of Acquisition Reforms and Expert Views. Defense Acquisition Research Journal, 25(3), 288-323.

This Article is brought to you for free and open access by AFIT Scholar. It has been accepted for inclusion in Faculty Publications by an authorized administrator of AFIT Scholar. For more information, please contact richard.mansfield@afit.edu.



Capt Amanda L. McGowin, USAF, Jonathan D. Ritschel, Robert D. Fass, and Bradley C. Boehmke

Defense acquisition reforms have a long history due to perceived cost, schedule, and performance problems. Dozens of experts and senior leaders within the acquisition community have published their notions on the reasons for



**DOI:** https://doi.org/10.22594/dau.18-802.25.03

Keywords: Text Mining, Grounded Theory, Defense Acquisition, Cost Growth, Acquisition Reform

It is not a revelation to state that the Department of Defense (DoD) acquisition system is portrayed as inefficient, if not "broken," due to perceived cost, schedule, or performance problems (Burden, 2010; Government Accountability Office, 2015). Confronted with these imperfect outcomes, the natural reaction is a call to action to rectify the problems. In defense acquisition, this action is manifested in what is commonly referred to as acquisition reform. Indeed, over the past 60 years, the U.S. Government has enacted multitudes of reforms in attempts to improve the acquisition system, focusing primarily on the military acquisition processes and organizational structure (Eide & Allen, 2012; Jackson, 2011). The omnipresence of acquisition reform as the solution is demonstrated by the fact that nearly every year some form of legislation or policy is enacted. For example, from 1971 to 2010—a span of 39 years—50 acquisition reforms and initiatives were implemented. Among these 50 reforms, five are considered major transformations (Ritschel, 2012):

- 1. The Nunn-McCurdy Provision of the 1982 Defense Authorization Act
- 2. The President's Blue Ribbon Commission on Defense Management of 1986, more informally known as the Packard Commission
- 3. The Defense Acquisition Workforce Improvement Act (DAWIA) of 1990
- 4. The Federal Acquisition Streamlining Act (FASA) of 1994



While some analyses of the five major reforms show that at least minor improvement of program outcomes exist (Rich & Dews, 1987; Smirnoff & Hicks, 2008), Major Defense Acquisition Programs (MDAPs) reflect no significant decrease in aggregate cost and schedule overruns (Eide & Allen, 2012). Despite the ineffectual historical record, Congress intends to introduce more acquisition reform initiatives, with Representative Mac Thornberry, House Armed Services Committee Chairman, 115th Congress, articulating plans to focus on innovation and organizational restructuring (Mitchell, 2017). Conversely, former Under Secretary of Defense for Acquisition, Technology, and Logistics Frank Kendall sent a warning to Congress in his farewell address stating that bureaucracy and regulation are not good tools to achieve the desired results and, in reality, burden the system (Serbu, 2017).

The purpose of this research is to identify and analyze trends within the enacted legislation encompassing five major defense acquisition reform efforts in comparison to a compendium of views from 32 leaders within the defense acquisition community on the efficacy of acquisition reform.

The methodological approach of prior research analyzing the efficacy of acquisition reforms has been predominantly quantitative. However, given that reforms are unstructured text-based documents, qualitative analysis through Text Mining is an appropriate alternative approach (Patten, 2009; Yu, Jannasch-Pennell, & Digangi, 2011). As a result, this research is intended to be an initial, exploratory analysis that will shed light on the utility of the Text Mining technique in the defense acquisition arena. Text Mining is a process that extracts useful information from textual data through the identification of patterns (Feldman & Sanger, 2006; Yu et al., 2011). This relatively new analytic technique emerged in the late 1990s and is becoming increasingly more prevalent (Witten, n.d.). The method has been applied in pharmaceutical drug discovery, survey analysis, capability engineering framework, and within the government for counterterrorism, scientific research, and problem detection in defense acquisition programs (Grimes, 2007; Kirk & Monarch, 2008; Losiewicz, Oard, & Kostoff, 2003; Miller, 2012). Although Text Mining has been employed within government research, it has not yet been applied to defense acquisition reform.

Thus, the purpose of this research is to identify and analyze trends within the enacted legislation encompassing five major defense acquisition reform efforts in comparison to a compendium of views from 32 leaders within the defense acquisition community on the efficacy of acquisition reform. The analysis is designed to provide insight on the commonalities and differences of the five major acquisition reforms. These results are then juxtaposed against the recommendations from the 32 expert essays. Through the Text Mining technique, the desired outcome is to provide a historical understanding of the emphasized areas of reform from the major enacted legislation in relation to the acquisition community's leading experts' views on the root causes of cost and schedule growth. From this, gaps can be identified as recommended areas of emphasis for future actions to mitigate cost and schedule growth in MDAPs.

## **Background and Data**

The history of defense acquisition is extensive, comprising entire books and a mountainous accumulation of research dating back to the seminal studies of Peck and Scherer (1962). Recent works by Fox (2011) have detailed an extensive chronological history of these defense reforms. This article, however, focuses on the five major acquisition reform efforts as identified in the literature and summarized in Table 1.

Congress introduced an amendment in the 1982 Defense Authorization Act to ameliorate the problem of cost growth in defense weapon systems. The amendment, known as the Nunn-McCurdy Act, establishes cost growth thresholds that require Congressional reporting on programs that breach growth over 15%, and program termination for growth over 25% (unless the Office of the Secretary of Defense deems the program essential to national security) (Fox, 2011). The provision became permanent in 1983. The Nunn-McCurdy legislation remained relatively unchanged until the 2006 National Defense Authorization Act (NDAA). The 2006 NDAA provides two alterations to the breach criteria. First, it defines categories of cost growth as either "Significant" or "Critical." Second, it adds another baseline from which the breaches are calculated. Breaches are now determined from both the "current" (at 15% for significant and 25% for critical) and the "original" (at 30% for significant and 50% for critical) baselines (Schwartz, 2010).

In the mid-1980s, President Ronald Reagan revitalized the President's Blue Ribbon Commission on Defense Management, also known as the Packard Commission, by Executive Order. The overarching purpose of the commission was to reduce inefficiencies and streamline the defense acquisition process (Christensen, Searle, & Vickery, 1999). Further, the

commission was responsible for the organizational leadership structure of Program Executive Officers (PEO); Service Acquisition Executives (SAE); the Under Secretary of Defense (Acquisition), or USD(A), which later became the Under Secretary of Defense for Acquisition, Technology, and Logistics, or USD(AT&L), also serving as the Defense Acquisition Executive; and the Joint Requirements Management Board. [Note that in 2018, USD(AT&L) was split into two new undersecretariats: Under Secretary of Defense for Research and Engineering and Under Secretary of Defense for Acquisition and Sustainment.] Lastly, the Packard Commission implemented rigorous testing of prototypes prior to production, more frequent use of off-the-shelf products, and the continuation of civilian management (Fox, 2011).

TABLE 1. SUMMARY OF MAJOR REFORM IMPLEMENTATION AND FOCUS			
Major Reform	Year	Focus	
Nunn-McCurdy	1982 (2006)	<ul> <li>Establishment of cost growth thresholds</li> <li>Requirement for MDAP Congressional reporting and potential program termination for threshold breaches</li> </ul>	
Packard Commission	1986	<ul> <li>Establishment of current acquisition organizational structure: PEOs, SAEs, DAE</li> <li>Establishment of the Joint Requirements Management Board</li> <li>Requirements for prototype testing prior to production, more frequent use of off-the-shelf products, continuation of civilian management, and improved training for acquisition personnel</li> </ul>	
DAWIA	1990	<ul> <li>Improvement of the quality of the acquisition workforce</li> <li>Creation of the Defense Acquisition University</li> <li>Requirements for formal education, training, and certification</li> <li>Establishment of program manager career paths</li> </ul>	
FASA	1994	<ul> <li>Promotion of commercial purchases</li> <li>Simplification of the contract award and management processes while reducing requirements placed on commercial firms</li> <li>Modernization of business practices</li> </ul>	
WSARA	2009	<ul> <li>Improvement of the early phases of weapon system development through use of systems engineering principles</li> <li>Establishment of CAPE and the Office of Developmental Test and Evaluation</li> <li>Implementation of technological maturity reviews</li> <li>Integration of combatant commanders into the requirements generation process</li> </ul>	

**Note.** CAPE = Cost Assessment and Program Evaluation; DAE = Defense Acquisition Executive; MDAP = Major Defense Acquisition Program; PEO = Program Executive Officer; SAE = Service Acquisition Executive.

The Packard Commission also laid the groundwork for improving training for acquisition personnel. This foundation eventually prompted the implementation of the Defense Acquisition Workforce Improvement Act (DAWIA) in 1990, which focuses on improving the quality of the acquisition workforce. DAWIA created the Defense Acquisition University, formalized education and training programs for acquisition personnel, enacted certification requirements, and established career paths for program managers (Layton, 2007; Pope, 1997).

The 1990s brought additional reform as Congress passed the Federal Acquisition Streamlining Act (FASA). FASA promotes commercial purchases and simplifies the contract award and management process. It also reduces requirements placed on commercial firms in the bidding process. Most importantly, FASA showed the acquisition workforce that Congress was committed to modernizing business practices and empowering federal agencies to make decisions about how to manage their programs (Smirnoff & Hicks, 2008).



In 2009, Congress passed the Weapon Systems Acquisition Reform Act (WSARA). WSARA focuses on the early phases of weapon system development, stressing the importance of sound systems engineering principles. It establishes offices for Cost Assessment and Program Evaluation (CAPE), Developmental Test and Evaluation, and Systems Engineering. WSARA also implements technological maturity reviews for major programs and integrates combatant commanders into the requirements generation process (Schwartz, 2013, 2014).

The reforms discussed above comprise one portion of the dataset for analysis. In addition, a compendium of views from 32 experts was compiled for analysis. These experts provide their views on the defense acquisition process and how defense procurement can be improved. A list of the 32 essays is provided in Table 2.

TABLE	2. ACQUISITION EXPERT DATA- EXPERT VIEWS	–A COMPENDIUM OF LEADING
Year	Document	Experts
2008	Testimony of the Honorable James I. Finley, Deputy Under Secretary of Defense (Acquisition and Technology) Before the United States House of Representatives Committee on Oversight and Government Reform and Subcommittee on National Security and Foreign Affairs	The Honorable James I. Finley
2014	Defense Acquisition Reform: Where Do We Go from Here? A Compendium of Views by Leading Experts	Brig Gen Frank J. Anderson, USAF (Ret.) The Honorable Norman R. Augustine Mr. David J. Berteau Mr. Irv Blickstein Gen James Cartwright, USMC (Ret.) The Honorable Thomas P. Christie Mr. Jonathan Etherton The Honorable Christine H. Fox Dr. J. Ronald Fox Mr. Paul Francis The Honorable Dr. Jacques S. Gansler The Honorable Dr. J. Michael Gilmore The Honorable Daniel I. Gordon Mr. William C. Greenwalt Mr. Todd Harrison The Honorable Tina W. Jonas The Honorable Paul G. Kaminski The Honorable Frank Kendall III The Honorable Elizabeth McGrath Dr. David L. McNicol The Honorable Dr. Jamie Morin The Honorable Dr. Jamie Morin The Honorable David Oliver ADM Gary Roughead, USN (Ret.) Ms. Katherine Schinasi Gen Norton A. Schwartz, USAF (Ret.) The Honorable Sean J. Stackley Mr. Michael J. Sullivan VADM David J. Venlet, USN (Ret.) Lt Col Daniel Ward, USAF The Honorable Dr. Dov Zakheim
2017	Getting Defense Acquisition Right	The Honorable Frank Kendall III

A plethora of previous research has examined the effectiveness of defense acquisition reforms on cost overruns and growth using a variety of methods as delineated in Table 3. Most results indicate that the major reforms had little to no significant effect on reducing cost and schedule growth or overruns. Note that these methods are predominantly quantitative and geared toward analyzing the specific effects on cost and schedule.

TABLE 3. PARTIAL LIST OF QUANTITATIVE ANALYSES OF PAST MAJOR ACQUISITION REFORMS			
Author	Year	Торіс	Method
Christensen, Searle, & Vickery	1999	Impact of Packard Commission on Cost Overruns	Means Comparison
Holbrook	2003	Analysis of Acquisition Reforms & Contract Cost Variance	Means Comparison & Time Phasing
Giacomazzi	2007	Impact of Acquisition Reforms & External Factors on Schedule Growth	Regression
Smirnoff & Hicks	2008	Impact of Acquisition Reforms & Economic Factors on Cost Overruns	Regression
Ritschel, Lucas, White, & Mrla	2019*	Impact of WSARA on the Cost of Air Force Weapon Systems	Regression

Note. \*Manuscript submitted for publication

In contrast, the goal of this article is to understand the relationships between the stated aims and goals of the reforms in relation to those areas that are considered most important by leading defense experts. Given that reforms are unstructured, text-based documents, qualitative analysis is an appropriate alternative approach (Patten, 2009). Although Text Mining has been employed within government research as shown in Table 4, it has not yet been applied to defense acquisition reform—a gap that this research intends to fill.

TABLE 4. USE OF TEXT MINING METHODS WITHIN DEFENSE ACQUISITION			
Author	Year	Topic	Method
Miller	2012	Acquisition Program Problem Detection Using Text Mining	Latent Dirichlet Allocation (LDA)
Freeman	2013	Text Classification Approach to Cost Growth in Acquisition Programs	Naïve Bayes Text Classification
Brown	2017	Measuring Relevance of Cost Estimating Through Text Analytics	Word Frequency, Regression, & Correlation

## **Methods**

Text Mining is a process that extracts useful information from unstructured text through the identification of patterns (Feldman & Sanger, 2006; Witten, 2003; Yu et al., 2011). While structured text such as a Hypertext Markup Language (HTML) uses various codes and tags to deliminate portions of the documents (e.g., titles, headers, paragraphs, etc.), unstructured text only has semantic and syntactical structure (e.g., white space, punctuation, special characters, etc.). With the use of Text Mining, a researcher can distill features such as characters, words, terms, concepts, and sentiments existing within a body of text (Feldman & Sanger, 2006; Grimes, 2007). While dozens of Text Mining methods are available, some of the most common are used in this research: word relationships (Losiewicz et al., 2003; Silge & Robinson, 2017), sentiment analysis (Feldman, 2013; Silge & Robinson, 2017), and clustering (Feldman & Sanger, 2006; Losiewicz et al., 2003).

Word relationship analysis examines which words tend to follow others, or that co-occur within documents or across the corpus. The analysis is conducted by providing and analyzing a count of pairs or groups of words. Word relationships use a "token." In general textual analysis, a token is a feature of the document text such as individual words, terms, or sentences. The word relationship analysis here uses a token called the n-gram, which is a sequence of n words that compose these pairs (bigrams) or word groupings. Typically, word frequency and word relationship analyses are used within the data exploration phase of research (Silge & Robinson, 2017).

One of the more sophisticated Text Mining methods is sentiment analysis, which attempts to extract the meaning or emotional intent of a document. In the R programming language, prebuilt sentiment datasets utilize three of the most popular lexicons (a lexicon is a list of words mapped to a sentiment score) for single words (Silge & Robinson, 2017): AFINN, created by Finn Årup Nielsen; BING by Bing Liu and collaborators; and NRC by Saif Mohammad and Peter Turney.

Each of the three available lexicons assigns positive and negative scores to each individual word within the dataset. One challenge of this method is that it is difficult to apply to documents containing multiple paragraphs since the positive or negative sentiment of the document may vary

throughout, effectively averaging to zero. For this reason, sentence- and paragraph-sized analysis is used throughout this research (Feldman, 2013; Silge & Robinson, 2017).

The final Text Mining method this research employs is clustering, which is used to identify and define different categories, or concepts, within text (Feldman & Sanger, 2006; Grün & Hornik, 2011; Losiewicz et al., 2003). Silge and Robinson (2017) call the clustering method "Topic Modeling" and specifically use Latent Dirichlet Allocation (LDA) for fitting a topic to a document. The LDA package utilized is fitted using Gibbs Sampling. LDA is driven by two principles: Every document is a mixture of topics and every topic is a mixture of words.

The first LDA principle essentially states that a document is composed of a certain percentage of Topic A and a certain percentage of Topic B (e.g., 70% Topic A, 20% Topic B, and 10% Topic C). The second principle is used to identify the most commonly used words within one of the document's topics. Said more simply, the LDA model shows "how words are associated with topics and how topics are associated with documents" (Silge & Robinson, 2017). The R Programming package Idatuning provides a function to accomplish this. The function uses four metrics to estimate the number of topics, two of which attempt to optimize by determining the minimum number of topics likely within the data (Arun, Suresh, Veni Madhavan, & Narasimha Murthy, 2010; Cao, Xia, Li, Zhang, & Tang, 2009) while the other two use maximization (Deveaud, Sanjuan, & Bellot, 2014; Griffiths & Steyyers, 2004).



Lastly, as a validation of the Text Mining results, Grounded Theory is performed on a subset of the reform and expert opinion texts. This validation is conducted only on the LDA portion of the analysis due to the time and labor-intensive nature of the process. Grounded Theory is a strategy for systematically analyzing data in an exploratory manner for the development of theory (Glaser & Strauss, 1967). It allows for the identification of a pattern within the data, and from that pattern, the discovery of the core category or foundation of the theory. The guiding principle is to let the data derive the theory, as opposed to fitting data to a predisposed assumption.

Because Grounded Theory is a manual process, it explicitly incorporates the "human brain" in the process. Classic Grounded Theory design—the method being applied herein—utilizes the constant comparative analysis method. This process involves assigning codes or categories to each line of data, and constantly comparing those codes to related codes across the document (Glaser & Strauss, 1967). The process of coding continues until a core category and related concepts emerge, and all possible categories are exhausted (Holton, 2010).



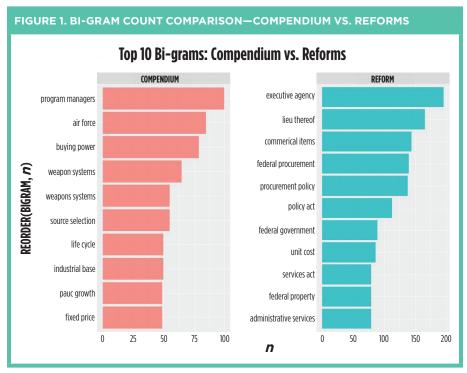
Grounded Theory has been used successfully for decades in areas such as sociology (Glaser & Strauss, 1967), health care (Piko, 2014; Williams, King, & Fox, 2016), information technology (Wiesche, Jurich, Yetton, & Krcmar, 2017), and consumer behavior (Goulding, 2000; Johnson, 2015), among others. It has not, however, been used in analysis of defense acquisition. The results from the Grounded Theory subset are compared to the LDA Text Mining results as a validity check.

## Results

As previously discussed, various Text Mining methods are applied to the dataset in an attempt to identify and analyze trends. The specific Text Mining methods utilized include word relationships, sentiment analysis, and topic modeling. This section discusses the results of these methods.

#### **Word Relationships**

To garner a top-level glimpse into the content and trends within the major reforms and the compendium of expert views, the analysis begins with an examination of word relationships. Word relationship analysis provides more context than an examination of the frequencies of individual words utilized throughout the document. Through the use of bi-grams, word relationships identify pairs of words that are consistently used together. Collectively, and within the compendium, the experts mention weapon systems, program managers, and buying power most frequently, while the major reforms utilize higher level terminology such as executive agency, federal procurement, and procurement policy, as shown in Figure 1.



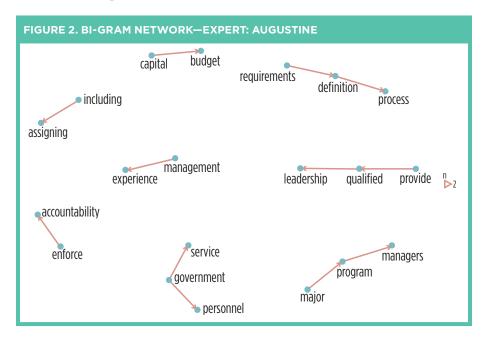
Note. pauc = Program Acquisition Unit Cost.

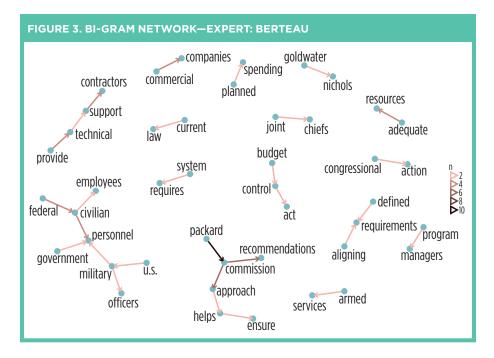
The term relationship results can also be disaggregated to allow insight into the contents of each of the individual major reforms. For example, WSARA discusses *systems engineering*, *developmental tests*, and *cost assessments*, which indicate themes related to ensuring that the weapon systems are both

sound and affordable. Conversely, the Packard Commission uses terminology related to preventing fraud and implementation of punitive actions such as *suspension debarment*, *voluntary disclosure*, and *false claims*.

While insight is gained through this method for the major reforms, document size is a challenge when examining the expert's opinions. For some of the experts, such as James Finley and Frank Anderson, their essays are just large enough to gain a glimpse into their views, but they have relatively low frequencies for their most frequently used terms (for example, Finley's highest term frequency was only two). Other experts, such as Norman Augustine and David Berteau did not have more than a few term frequencies above one, making their results difficult to interpret. To solve this issue, the most frequently used terms for each expert (and each of the major reforms) is plotted in word maps that show a directional connection of the words that comprise each term. Furthermore, the bi-gram maps are converted into network maps identifying some of the most used phrases.

Examining Norman Augustine's [see Figure 2] and David Berteau's [see Figure 3] network maps, the identification of potential themes within their views is more easily garnered. For example, Augustine's essay contains phrases such as *provide quality leadership* and *requirements definition process*, while Berteau's essay contains *budget control act* and *Packard Commission requirements*.





#### **Sentiment Analysis**

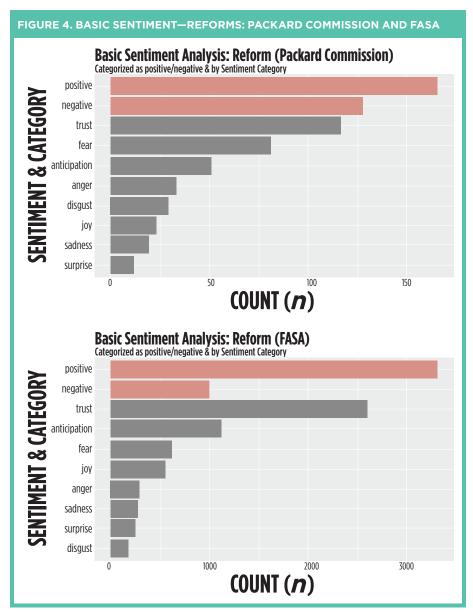
While the use of bi-grams in the word relationship analysis provides insight into the context of word usage, sentiment analysis portrays the feeling or emotion contained within each document. The emotion can simply be either positive or negative and will fall into one of eight categories: anger, anticipation, disgust, fear, joy, sadness, surprise, or trust (Feldman, 2013). An important component to account for when conducting sentiment analysis is the use of negation words (e.g., no, not, without, and never). When negation words precede a positive word (i.e., not greater), the term should be counted as negative, but has actually been counted as negative-positive, equating to a neutral net sentiment score. For this reason, the potential exists for the results of the basic sentiment analysis to be stated as more positive than it is in actuality.

Examination of all reforms and expert opinions found only one instance (FASA) of significant positive overstatement. In addition, the presence of negation words preceding positive words has an effect on the sentiment categories by either the total count contributing to that category or by the shifting of that category's position based on frequency. However, within the NRC lexicon utilized for this portion of the sentiment analysis, neither *not* nor *greater* were associated with a type of sentiment (i.e., fear, anger, trust, etc.). Therefore, the presence of negation words preceding positive words

did not have an effect on the order of the sentiment type classifications. Sentiment analysis of each of the major reforms reveals similar sentiment categorizations, with the exception of the Packard Commission. The majority of the reforms use largely positive vocabulary with very little negativity, thus falling into the trust and anticipation categories.



Conversely, the Packard Commission, with its positive vocabulary still outweighing the negative, uses much more negatively associated words than the other reforms. In addition, its top sentiment categories are trust and fear. At the third level, all of the reforms have the categories of trust, anticipation, and fear as their top three emotions. Figure 4 displays the sentiment of the Packard Commission and FASA, which is similar to the remaining reforms.



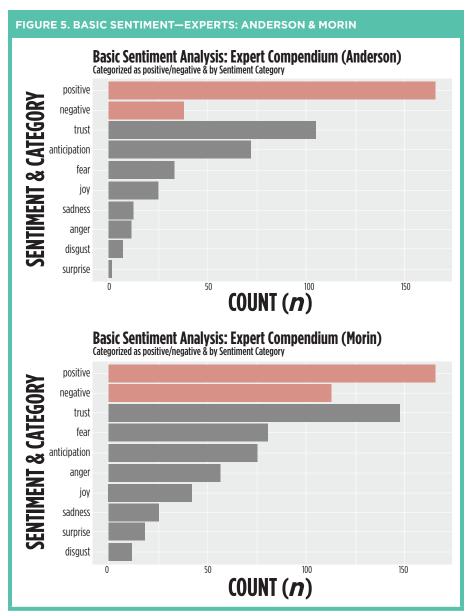
Note. Common acquisition words removed

The sentiment of a majority of the experts closely resembles the emotion found in the majority of the reforms. Most of the experts use approximately twice as much (or more) positive language than negative, with only four exceptions: Gansler, Harrison, Lehman, and Morin, whose opinions are slightly more than half negative. Analysis of their biographies reveals some

minor overlaps in experience (Harrison and Morin both were part of the Center for Strategic and Budgetary Assessment), but otherwise no significant commonalities are noted between the four.

When examining the emotion categories within each of the expert views, results are similar to the reforms. For the majority of the experts (21 of 32), trust and anticipation are the top two categories, while the remaining experts top two emotions are trust and fear. For all of the experts, trust and anticipation are within the top three. Figure 5 displays the sentiment of Frank Anderson (representative of the majority of experts who fall into the largely positive, trust/anticipation category) and Jamie Morin (representative of those experts using more than average amounts of negative vocabulary and falling into the trust/fear category).





Note. Common acquisition words removed

A further technique to analyze the sentiment of each document is to determine how the emotion changes throughout while applying a sentiment score (utilizing the AFINN lexicon) to identify positive or negative segments. Figure 6 displays the progression through the reforms: red indicates a net negativity for that section in the document, blue indicates positivity, while

the absence of color is an indication of neutrality. Additionally, the saturation of each bar represents how positive or negative that section of the document is. Recalling the sentiment analysis from Figure 4, the Packard Commission displayed more negative sentiment than the other reforms. Figure 6 shows a considerable amount of red as the Packard Commission progresses. However, the red has a fairly light saturation, indicating that it may not be as negative as initially portrayed. Similarly, comparing FASA from Figure 6 to Figure 4, it demonstrates an abundant amount of dark blue, which is an indication that FASA is actually a very positive document.

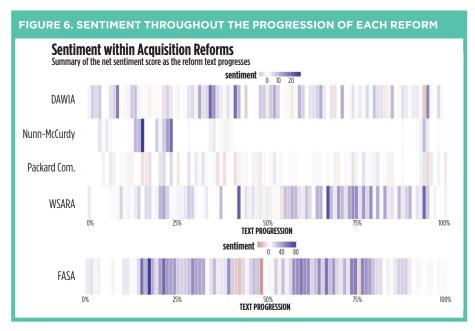


Figure 7 displays the sentiment progression through each of the expert's views. While Frank Kendall's essay looks extremely positive, the remainder of the experts all appear similar, including the four experts (Gansler, Harrison, Lehman, and Morin) who had higher negative sentiment counts than the rest. Interestingly, each of the experts tends to end their essays on a positive-negative-positive note—something that was lacking in the reform legislation, which likely utilizes more formal language than the experts.

Another interesting occurrence is the presence of either very dark blue (positivity) or very dark red (negativity) segment saturation within several of the experts' essays (circled in Figure 7). By extracting the text associated with each of the segments, it is possible to identify what each expert is saying in each of those instances. The "Ultra Negative" or dark red segments

displayed in Table 5 come from four experts. A close inspection of these text passages reveals that the experts have a truly negative tone at that time. The lone exception is Gilmore, who does use negative language, but is talking about the prevention of loss of life through proper system performance, which may be a positive message.

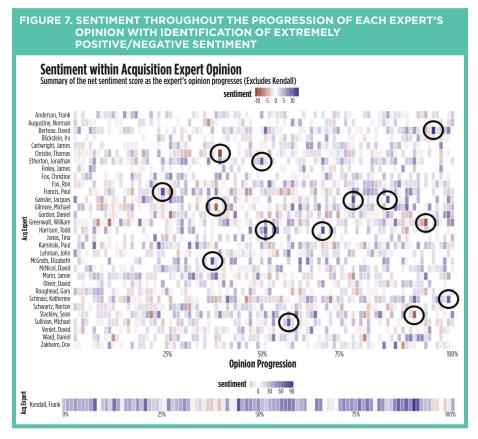




TABLE 5. ULTRA NEGATIVE—DARK RED SEGMENTS FROM SENTIMENT PROGRESSION			
Expert	Sentiment	Sentence	Text
Christie, Thomas	-11	38	By the time the technical and cost issues finally become known in the current system, few, if any, of those involved initially are still around, and those who are, refuse to admit they had been wrong, to cut their losses before the problems worsen, or to discipline the system by making an example of program officials and their contractors who have sold the department and the taxpayers a bill of goods.
Gilmore, Michael	-9	37	The substantive purpose of a test and evaluation program is to characterize system capabilities across the intended operational conditions so that problems with system performance are not discovered at the worst possible time—in combat when lives will be lost if operational performance is not fully understood.
Greenwalt, William	-10	126-127	Past reactions to failure and fraud have made success even unlikelier as risk-averse behavior and mind-numbing bureaucratic processes have increased waste and destroyed creativity and innovation.  Sometimes the best course of oversight action in reaction to the scandal of the day is to not legislate, but to ensure that criminals are going to jail and that there is enough flexibility in the system to buy what the warfighter needs.
Stackley, Sean	-9	81	The penalty for too much oversight is ever-increasing costs and impediments to execution that have no ceiling; the penalty for too little oversight is the costs and risks of rework for unforced errors.

The "Ultra Positive" or dark blue segments are listed in Table 6. After examining the text, seven of the 10 segments are clearly positive messages. The first segment in question is from McGrath. The text is a restatement of the original question she is asked in the survey dispersed to each of the experts. The segment itself used positive language, but did not actually include her opinion. The next two segments in question are from Harrison. In sentences 35 and 45, he discusses how split awards can decrease competition later in the life cycle if there is a considerable amount of learning that occurs, which is a mostly negative message. This segment was likely categorized as positive due to his use of words like award, winner, advantage, and greater.

TABLE 6. ULTRA POSITIVE—DARK BLUE SEGMENTS FROM SENTIMENT PROGRESSION			
Expert	Sentiment	Sentence	Text
Berteau, David	13	104-105	It is important to point out problems and to highlight possible corrective actions, but is equally important to highlight successes and progress. Congress can do better in this regard, selecting successful programs and managers for constructive oversight attention in hearings, speeches, commentary, and reports.
Etherton, Jonathan	10	51-52	The department should improve requirements development by sustaining centers of expertise in requirements analysis and development, and agencies should ensure that all acquisitions of complex services (e.g., information technology or management) occur only with express advance approval of requirements by the program manager, user, and the contracting officer, regardless of the type of acquisition vehicle used. While some acquisition workforce and cultural reforms may not have enjoyed hoped-for success in this, others were quite successful.
Francis, Paul	11	35-36	The answers to these questions will not necessarily be found in acquisition policy nor encourage good acquisition practices. While individual participants see their needs as rational and aligned with the national interest, collectively, these needs create incentives for pushing programs and encouraging undue optimism, parochialism, and other compromises of good judgment.
Gansler, Jacques	11	67	There are two required (industrial base) changes: the removal of the barriers to the DoD buying from commercial or foreign firms (when they offer the best value), the removal of the barriers to firms integrating their commercial and defense operations in the same facilities (in order to gain the cost and performance benefits from the economics of scale of the higher volume; and, to gain the performance and cost benefits from the technology transfer between the sectors).
Gansler, Jacques	10	75	When the DoD decided to harden their soldiers-carrying vehicles against roadside bombs (the largest killer and maimer of fighting men and women in Iraq and Afghanistan), they found that the best armor came from Israel; the best shock absorbers came from Germany; the best tires came from France; and the best design for the undercarriage (against mines) came from South Africa.

TABLE 6. ULTRA POSITIVE—DARK BLUE SEGMENTS FROM SENTIMENT PROGRESSION (CONTINUED)			
Expert	Sentiment	Sentence	Text
Harrison, Todd	10	35	If the split in award is large enough (i.e., the winner gets a much larger share) and the learning curve steep enough (i.e., unit costs decline rapidly as more units are built), the company that loses in the first round may never be able to overcome the cost advantage of its competitor in subsequent rounds.
Harrison, Todd	11	45	A lower learning percent means learning happens faster, giving a greater advantage for the company that wins the first round of competition and potentially making competition less effective.
McGrath, Elizabeth	11	22	What steps would you recommend to help ensure that top performers within the acquisition workforce are rewarded for their performance and empowered to manage programs with success?
Schinasi, Katherine	10	178- 179	supported by a robust technology process and talented individuals who are rewarded for success. Micromanagement has not brought success and will not as long as advocacy is combined with the responsibility for execution.
Sullivan, Michael	10	41	improve program management by attracting, training, and retaining professionals and providing them more rewarding career tracks. There have been many acquisition reform studies aimed at the need for improving the program management workforce to achieve improved acquisition outcomes.

Utilizing topic modeling as a Text Mining tool first requires the researcher to know the number of topics that are contained within the data. A predetermined knowledge of the topics or number of topics within this dataset did not exist. For that reason, before the data could be fit to an LDA topic model, the number of topics is estimated utilizing the R programming package ldatuning (Nikita, 2016). The ldatuning results identify the range of expected number of topics for each data subset as displayed in Table 7.

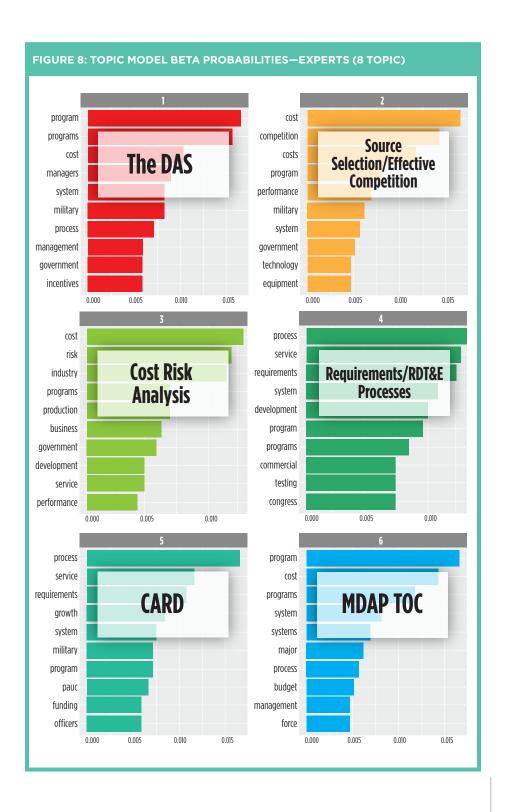
The R Package *topicmodels* (Grün & Hornik, 2017) contains a function to fit data to an LDA model (Grün & Hornik, 2011). As discussed above, that function requires prior knowledge about the number of topics that are contained within the data parameter (k). The range of expected number of topics for each subset, as displayed in Table 7, is used as the input for parameter (k) when fitting the data to an LDA model. Beta probability is the likelihood

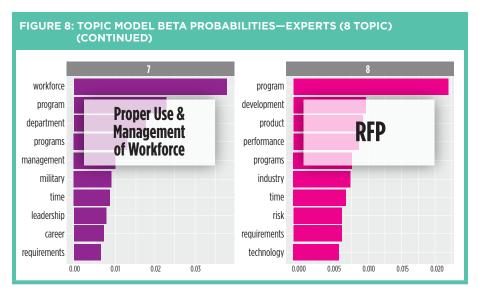
of a word being generated from that topic (Silge & Robinson, 2017) and is utilized to determine the number of expected topics. For the compendium and reforms, the model displaying the highest beta probability is the top of the range of expected topics: eight for the compendium, nine for the reforms.

TABLE 7. EXPECTED NUMBER OF TOPICS WITHIN DATA SUBSETS AS ESTIMATED BY LDATUNING		
Subset Expected Number of Topics		
All Data	6-10	
Compendium	5-8	
Reforms	7-9	

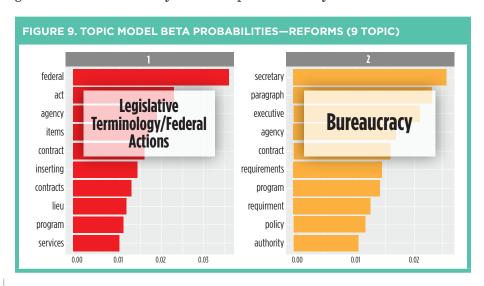
Words are assigned to the topic buckets based on the probability (beta probability) of that word being contained within that topic. One challenge of this method is that the topic buckets do not contain an automatic classification or categorization. As a result, topic names are subjectively applied retroactively based on the top 10 words within each of the buckets. The selection of topic names is applied based on the collective knowledge of individuals currently or previously working in the defense cost analysis field, or currently within the education and training arena focusing on defense cost analysis. It is important to understand, therefore, that the topic names are derived through the lens of defense cost analysts and educators. In qualitative research, this introduces the potential for the researchers' preconceptions from their experiences to enter into the analysis (Tufford & Newman, 2012). If the topic names had been determined by a different group (e.g., acquisition program managers), possibly a slightly different set of topics may have emerged. Thus, it is important to keep the cost analyst lens in mind when interpreting these results.

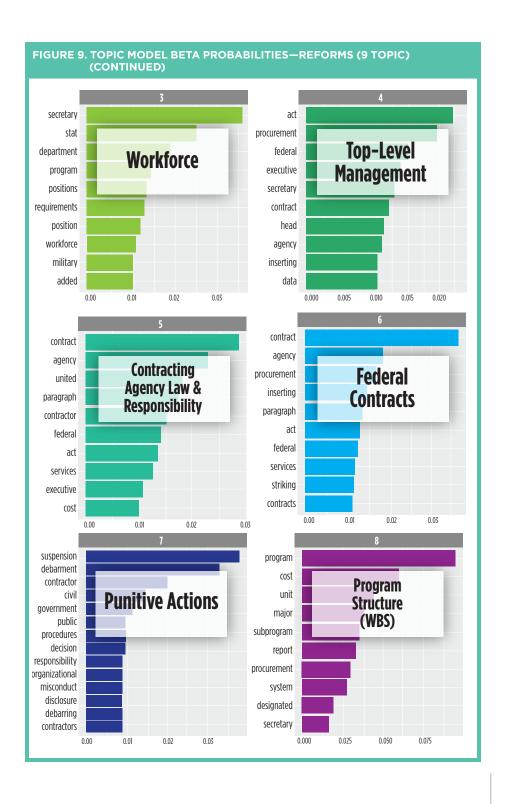
From the compendium subset, an eight-topic model is generated as shown in Figure 8. Based on the top 10 words within each topic, the following categorizations are assigned: the Defense Acquisition System (DAS); Source Selection as a means of Effective Competition; Cost Risk Analysis; the Requirements and Research, Development, Test and Evaluation (RDT&E) processes; various items that are found on the Cost Analysis Requirements Description (CARD); MDAP Total Ownership Cost (TOC); Proper Use and Management of the Workforce; and Request for Proposal (RFP).

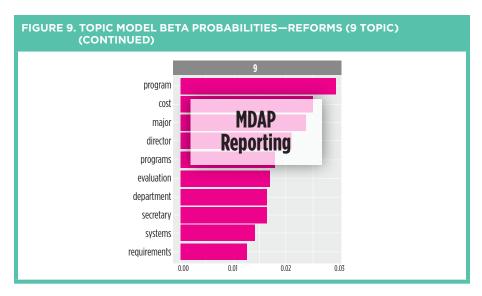




A nine-topic model is generated from the major reforms subset as shown in Figure 9. The topic names assigned are as follows: Federal Actions and Legislative Terminology, Bureaucracy, the Workforce, Top-Level Management, Contracting Agency Law and Responsibilities, Federal Contracts, Punitive Actions, Program Structure or Work Breakdown Structure (WBS) elements, and MDAP Reporting. Based on these results, one tendency noticed is that some of the topics largely encompassed a single reform. For example, Topic 7 is categorized as Punitive Actions, which is largely the focus of the Packard Commission. Additionally, Topic 3, categorized as Workforce may have been predominantly modeled after DAWIA.







While some minor commonalities come to light between the reforms and the compendium, such as discussion of the workforce and management, the two subsets appear to be addressing significantly different issues. The experts predominantly discuss strategies to improve defense acquisition, such as source selection and effective competition, and provide areas to focus improvement, such as the requirements and RDT&E processes. Conversely, the reforms address top-level oversight and impose bureaucracies. From this analysis, the reforms apparently do not address the concerns of the experts.

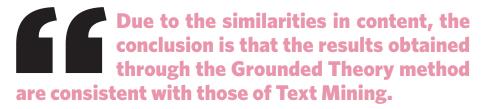
#### **Validation Check with Grounded Theory**

Topic modeling through the Grounded Theory method is conducted on one of the reforms (Title II of WSARA) and one of the expert essays (Jacques Gansler). Grounded Theory requires the researcher to examine the document word by word to derive the theory. Therefore, in order to ensure bias is not introduced, the Grounded Theory method was performed *prior* to any of the Text Mining analysis. Results from the Grounded Theory method are compared to the Text Mining results as a validation check on the Text Mining technique. While the total number of topic categories varies between the two techniques, each of the topic model categorizations from the Text Mining results are found within at least one of the Grounded Theory categories. For example, the Grounded Theory technique, as applied to the Gansler essay, resulted in nine core categories, while the topic modeling technique generated five. In this essay, the Text Mining topic Source Selection Strategy is mapped to the Grounded Theory category of Strategy.

Similarly, the Text Mining topic Effective Competition is mapped to the Grounded Theory category of Competition. Each of the remaining Text Mining topics can be mapped to the Grounded Theory results. Due to the similarities in content, the conclusion is that the results obtained through the Grounded Theory method are consistent with those of Text Mining.

## **Discussion**

The Text Mining technique reveals several interesting findings. First, although the purpose of each major reform is different, commonalities exist across the reforms as they primarily address issues at the contract level or by management (or agency) responsibility. The reforms also share similar sentiment; the verbiage utilized is largely positive or neutral with very little negativity. In addition, each of the reforms is categorized with the same top three sentiment categories: trust, anticipation, and fear.



When comparing the reform legislation to the recommendations of the compendium of experts, commonalities are found in the sentiment and emotions. Aside from this, however, the differences between the reforms and experts are vast. In contrast to the reforms, the experts address issues at the program level and/or with the DAS interactions with the industrial base. The experts predominantly discuss strategies to improve defense acquisition, such as source selection and effective competition, and provide areas to focus improvement, to include the requirements and RDT&E processes. Conversely, the reforms primarily address top-level oversight and impose bureaucracies. Why do cost, schedule, and performance problems persist in the DAS despite decades of reform efforts? Essentially, the reforms are not fully aligned with the issues identified by the experts. This finding is consistent with previous literature (Eide & Allen, 2012; Fox, 2011; Jackson, 2011; O'Neil, 2011; Rich & Dews, 1987; Schwartz, 2013) and is now backed by textual analysis.

As with any research, limitations are associated with the findings. One of the limitations is with the dataset selected for analysis. The selection of only five reforms and 32 essays chosen limits trend analysis and potentially shapes the conclusions. Acquisition history is replete with many more reforms, such as Better Buying Power, that could have been examined. However, given that this effort is a first voyage into the utility of Text Mining for acquisition reforms, limiting subjectivity on the researchers' part by selecting the five "major" reforms, as previously documented in the literature, seemed prudent. Similarly, the selection of essays focused on finding the most recent, complete, and credible sources available. This resulted in the selection of a compendium of views, primarily from the most recent Senate report the researchers discovered.

Limitations are also associated with the techniques employed during the research. As this is an initial, exploratory analysis of the efficacy of Text Mining in defense acquisition, we limited the selection of sentiment databases to preexisting, popular lexicons. Perhaps unique, domain-specific jargon currently embedded in defense acquisition would be better suited to a uniquely derived acquisition lexicon that future researchers can develop. In addition, stemming and lemmatization of the text did not occur. While we do not believe any of the main findings of this analysis would have changed, the opportunity for employing these and other more sophisticated Text Mining techniques is available.

Utilization of the Text Mining technique for defense acquisition analysis is promising, as demonstrated in this exploratory analysis. While the first steps were taken in this article, opportunity abounds for future research. First, additional reforms and policies such as Better Buying Power can be examined as a point of comparison to this effort. These smaller, more recent reforms may provide interesting new insights. Second, bucketing of experts by various types and analysis of speeches, reports, and testimonies could reveal new insights. Third, an investigation into the behavioral economics aspects of bureaucrats' behaviors and attitudes when confronted by the large volume of acquisition legislation and policy is warranted. This endeavor would attempt to ascertain whether bureaucrats focus on remaining within the letter of the law or, instead concentrate on determining the optimal path forward. The possibilities are vast and the technology is readily available to further explore the potential gains to

be garnered through textual analysis.

#### References

- Arun, R., Suresh, V., Veni Madhavan, C. E., & Narasimha Murthy, M. N. (2010, June). On finding the natural number of topics with Latent Dirichlet Allocation: Some observations. In *Proceedings of the 14th Pacific-Asia Conference on Advances in Knowledge Discovery and Data Mining (PAKDD)*. Hyderabad, India (pp. 391–402). Heidelberg, Germany: Springer-Verlag.
- Brown, G. E. (2017, Summer). Measuring the increasing relevance of cost estimating through text analytics. *ICEAA World*, *2*, 32–33.
- Burden, P. W. (2010). Acquisition reform what's really broken in defense acquisition? Carlisle Barracks, PA: U.S. Army War College.
- Cao, J., Xia, T., Li, J., Zhang, Y., & Tang, S. (2009). A density-based method for adaptive LDA model selection. *Neurocomputing*, 72(7-9), 1775-1781.
- Christensen, D. S., Searle, D. A., & Vickery, C. (1999, Summer). The impact of the Packard Commission's recommendations on reducing cost overruns on defense acquisition contracts. *Acquisition Review Quarterly*, 6(3), 251-262.
- Deveaud, R., Sanjuan, É., & Bellot, P. (2014). Accurate and effective latent concept modeling for ad hoc information retrieval. *Lavoisier*, *17*(1), 61–84.
- Eide, P. K., & Allen, C. D. (2012). The more things change, acquisition reform remains the same. *Defense Acquisition Research Journal*, 19(1), 99–120.
- Feldman, R. (2013). Techniques and applications for sentiment analysis. *Communications of the ACM*, *56*(4), 82.
- Feldman, R., & Sanger, J. (2006). *The text mining handbook.* New York, NY: Cambridge University Press.
- Fox, J. R. (2011). *Defense acquisition reform 1960–2009: An elusive goal.* Retrieved from http://history.defense.gov/Portals/70/Documents/acquisition\_pub/CMH\_Pub\_51-3-1.pdf
- Freeman, C. E. (2013). Multivariate and Naïve Bayes text classification approach to cost growth risk in Department of Defense acquisition programs (Master's thesis). Air Force Institute of Technology, Wright-Patterson AFB, OH.
- Giacomazzi, M. P. (2007). An analysis of the impact of defense acquisition reforms and external factors on schedule growth of defense weapon systems (Master's thesis). Air Force Institute of Technology, Wright-Patterson AFB, OH.
- Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research.* Hawthorne, NY: Aldine Publishing Company.
- Goulding, C. (2000). Grounded theory methodology and consumer behavior, procedures, practice and pitfalls. *Advances in Consumer Research, 27*(1), 261-266.
- Government Accountability Office. (2015). Acquisition reform: DoD should streamline its decision-making process for weapon systems to reduce inefficiencies (Report No. GAO-15-192). Washington, DC: U.S. Government Printing Office.
- Griffiths, T. L., & Steyvers, M. (2004). Finding scientific topics. In *Proceedings of the National Academy of Sciences of the United States of America* (Suppl. 1, pp. 5228–35). Washington, DC: National Academy of Sciences.
- Grimes, S. (2007). A brief history of text analytics. *BeyeNetwork*. Retrieved from http://www.b-eye-network.com/view/6311
- Grün, B., & Hornik, K. (2011). topicmodels: An R package for fitting topic models. *Journal of Statistical Software, 40*(13), 1–30.

- Grün, B., & Hornik, K. (2017). *Package 'topicmodels.'* Retrieved from https://cran.r-project.org/web/packages/topicmodels/topicmodels.pdf
- Holbrook, M. A. (2003). An analysis of the implementation of acquisition reform initiatives and contract cost variance (Master's thesis). Air Force Institute of Technology, Wright-Patterson AFB, OH.
- Holton, J. A. (2010). The coding process and its challenges. *Grounded Theory Review*, 9(1), 21-40.
- Jackson, R. A. (2011, November-December). Acquisition reform as performance art: Re-forming acquisition reform. *Defense AT&L, 40*(6), 25–27. Retrieved from http://www.dtic.mil/docs/citations/AD1016048
- Johnson, J. S. (2015). Qualitative sales research: An exposition of grounded theory. Journal of Personal Selling & Sales Management, 35(3), 262-273.
- Kirk, J., & Monarch, I. (2008, October). Mapping acquisition requirements from capabilities in a net-centric enterprise creating a capabilities engineering framework [PowerPoint slides]. Presentation at the National Defense Industrial Association's 11th Annual Systems Engineering Conference, San Diego, CA. Retrieved from http://www.dtic.mil/ndia/2008/systems/7444vankirk.pdf
- Layton, E. (2007). *Training professionals for the acquisition workforce, 1992–2003*. Retrieved from http://www.dtic.mil/docs/citations/ADA483328
- Losiewicz, P., Oard, D. W., & Kostoff, D. N. (2003). Science and technology text mining basic concepts (Technical Report No. ADA415886). Retrieved from http://www.dtic.mil/docs/citations/ADA415886
- Miller, T. P. (2012). Acquisition program problem detection using text mining methods (Master's thesis). Air Force Institute of Technology, Wright-Patterson AFB, OH.
- Mitchell, E. (2017, May 2). GOP chairman to unveil defense acquisition reforms in mid-May. *The Hill.* Retrieved from http://thehill.com/policy/defense/331590-gop-chairman-to-unveil-defense-acquisition-reforms-in-mid-may
- Nikita, M. (2016). Select number of topics for LDA model. Retrieved from https://cran.r-project.org/web/packages/ldatuning/vignettes/topics.html
- O'Neil, W. D. (2011). Cost growth in major defense acquisition: Is there a problem? Is there a solution? *Defense Acquisition Research Journal, 18*(3), 277-294.
- Patten, M. L. (2009). *Understanding research methods* (7th ed.). Glendale, CA: Pyrczak Publishing.
- Peck, M. J., & Scherer, F. M. (1962). *The weapons acquisition process: An economic analysis*. Boston, MA: Division of Research, Graduate School of Business Administration, Harvard University.
- Piko, L. M. (2014). Discovering Glaser: My experience of doing grounded theory. *Grounded Theory Review, 13*(2), 47–55.
- Pope, J. K. (1997). Measuring the effect of the Defense Acquisition Workforce Improvement Act (Master's thesis). Retrieved from http://www.dtic.mil/docs/citations/ADA331749
- Rich, M. D., & Dews, E. (1987). *Thoughts on reforming the military acquisition process* (Report No. P-7352). Santa Monica, CA: RAND.
- Ritschel, J. D. (2012). Efficacy of U.S. legislation in military acquisition programs: Nunn-McCurdy Act unveiled. *Economic Papers: A Journal of Applied Economics and Policy, 31,* 491–500.
- Ritschel, J. D., Lucas, B. M., White, E. D., & Mrla, D. (2019). *The impact of WSARA on the cost of Air Force weapon systems*. Manuscript submitted for publication.

- Schwartz, M. (2010). The Nunn-McCurdy Act: Background, analysis, and issues for Congress. Retrieved from https://fas.org/sgp/crs/misc/R41293.pdf
- Schwartz, M. (2013). Twenty-five years of acquisition reform: Where do we go from here? Retrieved from http://www.crs.gov
- Schwartz, M. (2014). Reform of the defense acquisition system. Retrieved from http://www.crs.gov
- Serbu, J. (2017, January 18). *In final speech, DoD acquisition chief knocks congressional reforms as unhelpful* [Radio newscast]. Retrieved from https://federalnewsradio.com/defense/2017/01/final-speech-dod-acquisition-chief-knocks-congressional-reforms-unhelpful/.
- Silge, J., & Robinson, D. (2017). *Text Mining with R.* Retrieved from http://tidytextmining.com/index.html.
- Smirnoff, J. P., & Hicks, M. J. (2008). The impact of economic factors and acquisition reforms on the cost of defense weapon systems. *Review of Financial Economics*, 17(1), 3-13.
- Tufford, L., & Newman, P. (2012). Bracketing in qualitative research. *Qualitative Social Work, 11*(1), 80–96.
- Wiesche, M., Jurisch, M. C., Yetton, P. W., & Krcmar, H. (2017, September). Grounded theory methodology in information systems research. *MIS Quarterly, 41*(3), 685–701.
- Williams, K., King, J., & Fox, J. R. (2016). Sense of self and anorexia nervosa: A grounded theory. *Psychology and Psychotherapy: Theory, Research and Practice,* 89(2), 211-228.
- Witten, I. (n.d.). *Text mining*. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.74.3588&rep=rep1&type=pdf
- Yu, C. H., Jannasch-Pennell, A., & Digangi, S. (2011). Compatibility between text mining and qualitative research in the perspectives of grounded theory, content analysis, and reliability recommended. *The Qualitative Report, 16*(3), 730-744.

## **Author Biographies**

## Capt Amanda L. McGowin, USAF,

is an operations research analyst for Simulators Division, Air Force Life Cycle Management Center, Wright-Patterson Air Force Base, Ohio. She holds a BS in Computer Science from Florida State University and an MS in Cost Analysis from the Air Force Institute of Technology (AFIT).

 $(E-mail\ address: Amanda.McGowin.1@us.af.mil)$ 



#### Dr. Jonathan D. Ritschel,

is an assistant professor of cost analysis in the Department of Systems Engineering and Management at AFIT. He received his BBA in Accountancy from the University of Notre Dame, his MS in Cost Analysis from AFIT, and his PhD in Economics from George Mason University. Dr. Ritschel's research interests include public choice, the effects of acquisition reforms on cost growth in DoD weapon systems, and economic institutional analysis.

(E-mail address: jonathan.ritschel@afit.edu)



#### Dr. Robert D. Fass

is an assistant professor in the Department of Systems Engineering and Management at AFIT. He holds a BA in Economics, an MBA from the University of New Mexico, and aPhD in Business Administration and Management from New Mexico State University. Dr. Fass's research interests include cost analysis, decision analysis, risk analysis, behavioral economics, organizational behavior, organizational change, and government acquisition policy.

(E-mail address: Robert.Fass@afit.edu)



## Dr. Bradley C. Boehmke,

is a director of Data Science at 84.51° and an adjunct professor at the University of Cincinnati, Wake Forest University, and AFIT. He holds an MS in Cost Analysis and PhD in Logistics from AFIT. Dr. Boehmke's research interests include machine learning, model interpretation, natural language processing, and open source programming.

 $(E{\text{-}mail\,address:}\,BradleyBoehmke@gmail.com)$ 

