



**REFINING AIR FORCE ASSET MANAGEMENT STRATEGY: CONNECTING
AIR FORCE INFRASTRUCTURE TO CORE MISSIONS**

THESIS

Joseph C. Buyer, Captain, USAF

AFIT-ENV-MS-17-M-177

**DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY**

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

DISTRIBUTION STATEMENT A.
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

The views expressed in this thesis are those of the author and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the United States Government. This material is declared a work of the U.S. Government and is not subject to copyright protection in the United States.

AFIT-ENV-MS-17-M-177

REFINING AIR FORCE ASSET MANAGEMENT STRATEGY: CONNECTING AIR
FORCE INFRASTRUCTURE TO CORE MISSIONS

THESIS

Presented to the Faculty

Department of Systems and Engineering Management

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the
Degree of Master of Science in Engineering Management

Joseph C. Buyer, BS

Captain, USAF

March 2015

DISTRIBUTION STATEMENT A.
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

AFIT-ENV-MS-17-M-177

REFINING AIR FORCE ASSET MANAGEMENT STRATEGY: CONNECTING AIR
FORCE INFRASTRUCTURE TO CORE MISSIONS

Joseph C. Buyer, BS
Captain, USAF

Committee Membership:

Lt Col. Vhance Valencia, PhD, PE
Chair

Maj. Jason Freels, PhD
Member

Dr. Jared Astin, PhD, PE
Member

Abstract

At this time, Air Force organizational objectives are not tied to the Air Force asset management program. Additionally, the Air Force is unable to objectively say with any confidence which infrastructure assets are mission and which ones are support. While the Air Force Strategic Master Plan gives the example of choosing between a dormitory and a space launch facility, the distinction is not always so easily made. While difficult, there is value in adjusting Air Force asset management strategic objectives by connecting Air Force investments in infrastructure to organizational objectives. Such a connection would allow senior leaders to more confidently make difficult decision related to infrastructure in this fiscally uncertain environment.

This research focuses on improving Air Force asset management strategic objectives; specifically, it investigates infrastructure assets and Air Force core missions to determine if linkages can be made between them. Qualitative and quantitative methods are used to identify keywords from Air Force strategic documents and build a model which identifies infrastructure asset groupings related to Air Force core missions. The model breaks infrastructure assets into two groups, (1) core – those assets which are directly related to Air Force core missions – and (2) non-core. This method of identifying infrastructure assets related to core missions can be used by the Air Force to objectively differentiate between mission and support facilities. Additionally it provides senior leaders with additional information, reducing the amount of uncertainty related to asset management funding decisions.

Acknowledgments

I would like to express my sincere appreciation to my faculty advisor, Lt Col Vhance Valencia, for his guidance and support throughout the course of this thesis effort. The insight and experience was certainly appreciated. I would, also, like to thank my sponsor, the Air Force Civil Engineering Center, for both the support and latitude provided to me in this endeavor. Additionally, I could not have done this without the love and support of my wife at home. Her love and support during long days made this a much easier hill to climb. Finally, I would like to thank my classmates of 17M. While I enjoyed my time learning and growing with them, it was their help, insight, mentorship, and guidance that I will keep with me and treasure moving forward.

Joseph C. Buyer

Table of Contents

	Page
Abstract	iv
Table of Contents	vi
List of Figures	ix
List of Tables	x
List of Equations	xi
I. Introduction	1
Background.....	1
<i>The State of Infrastructure</i>	2
<i>The Fiscal Environment</i>	3
Infrastructure Asset Management.....	4
<i>Air Force Asset Management Program</i>	5
<i>Air Force Asset Management Strategy</i>	6
Problem Statement.....	7
Research Objective	7
Methodology.....	8
Assumptions & Limitations.....	9
Implications	10
Summary.....	11
II. Literature Review	12
Chapter Overview.....	12
Decaying Infrastructure	12
<i>National infrastructure</i>	13
<i>Federal Infrastructure</i>	14
<i>Air Force Infrastructure</i>	15
<i>Fiscal Environment</i>	16
Defining Asset Management	17
<i>Strategic Asset Management Plan</i>	20
<i>Asset Management Program</i>	21
<i>Risk Analysis</i>	22
Air Force Asset Management.....	23

<i>Air Force Asset Management Program</i>	24
<i>Shortfalls</i>	27
<i>Refining Air Force Asset Management Strategy</i>	29
Literature Gap	30
Qualitative Analysis	31
<i>Coding</i>	32
<i>Categories</i>	33
<i>Subjectivity of Qualitative Analysis</i>	33
Conclusions	33
III. Methodology	35
Chapter Overview	35
Purpose	35
Content Analysis	36
Data Description	38
<i>Strategic Documents</i>	38
<i>Operational Doctrine</i>	39
<i>AF Real Property Data</i>	40
<i>DoD FAC codes and AF CATCODEs</i>	41
Conventional Content Analysis	44
<i>Developing Codes and Categories</i>	44
<i>How Codes and Categories are used</i>	45
Summative Content Analysis	46
<i>Textual Data Processing</i>	47
<i>Text Mining</i>	47
<i>Selecting Keywords</i>	48
Model Build	49
<i>Identifying CATCODEs that Contain Keywords</i>	49
<i>Defining Infrastructure Groupings</i>	49
Model Verification	50
Summary	51
IV. Analysis and Results	52
Chapter Overview	52
Brief Description of Results	52
Data Collection	52
Method Implementation: Conventional Content Analysis	55
<i>Coding Strategic Documents</i>	55
<i>Categories</i>	57
Method Implementation: Summative Content Analysis	58
<i>Choosing Operational Doctrine</i>	59
<i>Analysis of Operational Doctrine</i>	59
<i>Selecting Keywords</i>	61

Method Implementation: Model Assembly	62
Model Results	63
Model Verification	65
Summary.....	66
V. Conclusions and Recommendations	68
Chapter Overview.....	68
Conclusions of Research	68
Discussion.....	69
<i>Defining Core and Non-core</i>	70
<i>Thoughts on the Results</i>	71
Limitations of the Research.....	72
Significance of the Research	74
Recommendations for Action.....	75
Recommendations for Future Research.....	76
Summary.....	77
Appendix A: Category Groups	79
Appendix B: Basic Categories	80
Appendix C: Complete List of Intentionally Joined Words	83
Appendix D: R Code for Summative Content Analysis	84
Appendix E: Keyword Counts Found in FAC/CATCODE Descriptions.....	92
Appendix F: Basic Categories Not Selected and >50%	93
Appendix G: R Code for Shiny Application Model	94
Bibliography	98
Vita.....	102

List of Figures

	Page
Figure 1: Air Force asset management process	25
Figure 2: Proposed research method diagram.....	36
Figure 3: Example subset of real property data.	41
Figure 4: Illustration of the breakdown of CATCODEs.....	43
Figure 5: Linear research method flowchart.....	53
Figure 6: Statistical analysis comparing mean MDI's of core and non-core	65

List of Tables

	Page
Table 1: Complete list of operational doctrine.	40
Table 2: Count by classifications.	42
Table 3: Complete list of Facility Classes.	43
Table 4: Sample of codes for Global Strike core mission	57
Table 5: Example of codes and categories for Global Strike.....	58
Table 6: Selected operational doctrine.....	59
Table 7: Complete list of intentionally joined words.....	60
Table 8: Complete list of selected keywords.	62
Table 9: Complete list of Basic Categories found to be related to core missions.	64
Table 10: Basic Categories not selected as core with high average MDI values.	66
Table 11: Results shown by Facility Class.	72

List of Equations

	Page
(1) Air Force asset management total project score equation.....	25
(2) Air Force asset management probability of failure equation	26
(3) Air Force asset management consequence of failure equation	27

REFINING AIR FORCE ASSET MANAGEMENT STRATEGY: CONNECTING AIR FORCE INFRASTRUCTURE TO CORE MISSIONS

I. Introduction

The state of the nation's infrastructure has received increasing scrutiny, specifically in the area of maintenance and repair (GAO, 2015; Hastak & Baim, 2015; NRC, 2012). The federal government has a maintenance and repair backlog in the tens of billions of dollars and a budget that is continually shrinking (GAO, 2008a; NRC, 2012). To address these issues, government agencies have increasingly turned to asset management (Executive Order No. 13327, 2004). The goal of asset management is to extend the useful life of the physical and/or built environments (Vanier, 2001). To complicate the matter, the federal fiscal environment has come under increased scrutiny, increasing pressure to justify budgetary needs and decreasing dollars available for maintenance and repair (GAO, 2008b; NRC, 2012). While the Air Force does have an asset management program, it lacks the connection between Air Force strategic objectives and asset management objectives. This research focuses on bolstering the strategic asset management plan by linking Air Force core missions and Air Force infrastructure assets.

Background

The United States has recognized for several years that its infrastructure – facilities and utilities – is in need of serious attention (Executive Order No. 13327, 2004). The American Society of Civil Engineers (ASCE) has given national infrastructure a

grade of D+ in its Report Card for America's Infrastructure indicating that the United States is barely meeting maintenance and repair needs (ASCE, 2013). Additionally, multiple Government Accountability Office (GAO) reports cite the ongoing maintenance problem the United States is facing at the national and federal levels (GAO, 2008a, 2008b, 2015). The demand for and the use of infrastructure has significantly outpaced capacity in areas such as transportation and aviation (GAO, 2008b). Aging infrastructure is a chronic issue with potentially devastating consequences when assets fail and it will require significant attention and money from federal, state, and local government, as well as the private sector. Estimates to repair, replace, or upgrade US infrastructure, as well as to expand capacity to meet the increased demand, are in the hundreds of billions of dollars (GAO, 2008a).

The State of Infrastructure

This section provides a snapshot into the need for, and cost of, maintenance and repair in a few areas as evidence of the infrastructure problem facing the nation. As of 2013, America's dams had an average age of 52 years old with a need for \$21 billion for maintenance and repair. There are 240,000 breaks a year with the cost to replace them topping \$1 trillion. One in nine bridges is rated as structurally deficient and would require \$20.5 billion annually to correct them by 2028 (ASCE, 2013). The private sector is responsible for the majority of national assets, but the federal government owns a sizable portion (Hemme, 2015).

Federal facilities are owned and managed by more than 30 different government agencies (NRC, 2012). In 2008, the federal government operates more than 407,000 buildings, totaling more than 3.3 billion square feet (GSA, 2009). These facilities,

buildings and structures combined, were valued at over \$1.5 trillion and cost the government \$47 billion to operate (GAO, 2008a). Of the more than 30 different government agencies responsible for federal infrastructure, the Department of Defense is one of the largest which includes the Air Force (AF).

According to the GSA real property report from 2012, the Air Force has a portfolio containing 54,873 buildings and 73,055 structures. It maintained the 594 million square feet and 8.4 million acres of land at an operating cost of \$4.64 billion (GSA, 2013). This is troubling when one considers the AF and other government agencies continue to struggle with mounting backlogs of maintenance and repair (GAO, 2008a). While the size and condition of the infrastructure assets are problematic enough, there is not enough money to address them all.

The Fiscal Environment

To further compound issues, multiple government reports cite future strain on fiscal funding and a poor budgetary outlook (GAO, 2008b; NRC, 2012). Additionally the AF recognizes the need for a cost-conscious mindset in this fiscally constrained environment (USAF, 2015e). This indicates that funds that used to be available to perform maintenance and repair will begin to evaporate while the portfolio of AF infrastructure assets stays the same or grows. All government agencies, including the AF, will have to optimize the way it chooses infrastructure projects in order to address the most critical projects first.

Infrastructure Asset Management

Infrastructure is a national asset (NRC, 2012; PPD-21, 2013). The maintenance and repair of an infrastructure asset is critical for keeping it in service and for reaching its maximum life. Planning the maintenance and repair of multiple infrastructure assets requires a management program. The NRC defines asset management as a systematic process for maintaining, upgrading, and operating physical assets cost effectively (NRC, 2008).

The government and its many agencies have increasingly relied on asset management tools to help manage, and even predict, maintenance and repair with the US Army Corps of Engineers (USACE) leading the way (Oostermeijer & Boss, 2010). In most cases, almost all asset management tools will seek to clarify what assets an organization owns, what condition each asset is in, and communicate the risk associated with that asset if it fails (Faber & Stewart, 2003; Marlow, Beale, & Mashford, 2012). However some asset management models fall short by failing to connect organizational objective to asset management objectives through the strategic asset management plan (SAMP).

The SAMP is arguably the most important part of an asset management program (IFMA, 2009; Infrastructure Asset Management Exchange, 2015; ISO, 2014; The Facilities Society, 2014). It provides guidance and direction to ensure that the organization is able to achieve its goals. ISO 55000 (2014), an internationally recognized asset management manual, defines a SAMP as the connecting tissue between organizational objectives and asset management implementation.

Moreover, as the International Facility Management Association (IFMA) and The Facilities Society points out, differentiating between assets supporting core business functions and assets supporting non-core business functions must lie at the heart of a SAMP (IFMA, 2009; The Facilities Society, 2014). This allows the organization to distinguish between competing requirements and allocate money to assets that further organizational objectives. Both emphasize that while asset management should focus on core assets, “it must not lose sight of the supporting services – its non-core business” (The Facilities Society, 2014).

Air Force Asset Management Program

The AF began implementing asset management following an executive order from President Bush in 2004 (Executive Order No. 13327, 2004). The executive order mandated that all federal agencies begin to establish policies, goals, and objectives in an effort to responsibly manage United States real property. Since then the AF asset management program has undergone several changes, all aimed at improving the process.

The current AF model consists of three criteria. They are probability of failure (PoF), consequence of failure (CoF), and a savings to investment ratio (SIR) (AFCEC, 2015). Each criterion addresses a different piece of an asset and, together, assists leaders in determining where to allocate resources to avoid critical mission failure. All assets are scored and combined to a single list called the Integrated Priority List (IPL). The IPL is used to select which maintenance and repair, demolition, and consolidation projects are funded in a given year.

The list is intended to give senior leaders the ability to compare projects from across the enterprise. Ideally, the top scoring projects are those that both need to be

addressed (PoF) and are critical to mission assurance (CoF). However the idea of mission assurance is general and does not apply to a particular mission. In short, the current AF asset management program lacks strategic objectives that connect AF core missions to asset management principles.

Air Force Asset Management Strategy

The AF, just like any other organization, has organizational objectives. Additionally, the AF Strategic Master Plan (SMP) stresses the importance of being able to distinguish between mission and support infrastructure (USAF, 2015c). The organizational objectives are found in the form of AF core missions. The core missions include (1) air and space superiority; (2) intelligence, surveillance, and reconnaissance; (3) rapid global mobility; (4) global strike; and (5) command and control. These five core missions define what the AF aims to accomplish. They are not, however, the organizational objectives used in the asset management program. Moreover, the AF does not have a way to identify infrastructure asset as mission or support. In this way, the AF SAMP is lacking and leaves room for improvement.

Currently the AF asset management program uses a numbered system developed by the Navy called the mission dependency index (MDI). The AF borrowed this methodology as a temporary way to indicate which infrastructure assets were most important. It does not relate to AF organizational objectives or strategy in any real, tangible way and does not distinguish between mission and support infrastructure. More on this subject, including its strengths and weaknesses related to asset management, can be found in Literature Review.

Problem Statement

At this time, AF organizational objectives are not tied to the AF asset management program. Additionally the AF is unable to objectively say with any confidence which infrastructure assets are mission and which ones are support. While the SMP gives the example of choosing between a dormitory and a space launch facility, the distinction is not always so easily made. While difficult, there is value in adjusting AF asset management strategic objectives by connecting AF investments in infrastructure to organizational objectives. Such a connection would allow senior leaders to more confidently make difficult decision related to infrastructure in this fiscally uncertain environment.

Research Objective

This research focuses on improving AF asset management strategic objectives; specifically, it investigates infrastructure assets and AF core missions to determine if linkages can be made between them. A mix of qualitative and quantitative skills are used to both identify keywords from AF strategic documents and build a model to identify groupings of infrastructure related to AF core missions. The model breaks infrastructure into two classifications: (1) core – directly related to AF core missions – and (2) non-core. The two classification will allow senior leaders to differentiate between mission and support infrastructure while reducing the amount of uncertainty related to project funding.

Methodology

This research effort investigates the links between AF core missions and AF infrastructure assets by searching infrastructure descriptions for keywords that describe core missions. The keywords are identified and selected using conventional content analysis of strategic documentation and summative content analysis of AF strategic doctrine. Facilities were then broken into two groups: mission (or core) infrastructure and support (non-core) infrastructure.

This research uses a mix of qualitative and quantitative methods to build a model that will connect infrastructure to AF core missions. To begin, the researcher uses conventional content analysis to understand core missions through the creation of codes and categories. That understanding allows the researcher to (1) appropriately select operational doctrine related to the AF core missions and (2) select keywords based on their ability to fit into the codes and categories.

Next, summative content analysis is used to examine the selected operational doctrine to determine which words and phrases were used with the highest frequency. Some of those high frequency words will be intentionally selected as “keywords” for use in the model. A word is selected based on its ability to fit into the codes and categories developed during the conventional content analysis. The model then searches through CATCODE and FAC code titles and descriptions to determine if they contain any keywords and, thus, relate to the core missions. The model calculates the percentage of unique CATCODEs that contain keywords within each unique Basic Category, or

collection of digits. Any Basic Categories having 75% or more unique CATCODEs containing keywords will be considered core, or directly relating to AF core missions.

Assumptions & Limitations

The methodology does make some assumptions, notably in the summative content analysis portion. It assumes that the words used most frequently will reflect the areas of greatest concern or importance. This may not always be the case. Weber (1990) points out that synonyms may be used for stylistic reasons leading researchers to underestimate the importance of a particular term. Additionally, certain words may have more than one meaning (Stemler, 2001). When performing summative content analysis, word frequencies will be used to identify words of potential interest. Codes and categories developed through conventional content analysis are used to ultimately determine if a word is selected as a keyword.

This research also assumes that the AF core missions are an adequate representation of the organizational objectives. Most internal, AF organizations have their own missions and the overall missions of “Fly, Fight, and Win!” is too generic. However upon a review of unclassified material and strategic doctrine the researcher determined that using the core missions was appropriate in both size and breadth while also encapsulating the mission of the AF.

There are several limitations. Some of them are listed here but all of them will be explained further in Conclusions and Recommendations chapter. The first is that of the documents chosen for summative content analysis. They all appear to be written by different sources and vary substantially in length and depth of information. The second

limitation is the descriptions associated with each CATCODE. Only 313 CATCODEs contained an AF description leaving 606 without one. While this seems like a large limiting factor, all CATCODEs did have a FAC description. The third limitation is the strategic value of the documents used in the research. The documents used for this study were meant for public consumption and therefore remained fairly generic.

Implications

The AF SMP calls for the ability to differentiate between mission and support facilities. This research effort serves as the first time the AF has investigated their infrastructure to determine objectively which infrastructure assets are directly relate to the AF mission and which ones are support. Currently, the MDI communicates the relevant importance of a facility or utility. The methodology used to derive the MDI score does not account for the mission of the AF and therefore falls short as a way to distinguish between mission and support. This research is the first step in answering the SMP's call to distinguish between mission and support facilities.

The results produced by this research may allow the AF to make better, more informed decisions related to the funding of infrastructure projects. The research allows decision makes to create two groups: (1) mission related facilities, and (2) support facilities. This distinction would allow senior leaders to compare like facilities and make decisions based on the condition and MAJCOM input. It also adds a level of confidence knowing which dollars are supporting the AF mission. At the very least, such a distinction would reduce the amount of uncertainty that exists when comparing competing requirements.

Summary

National infrastructure is in a state of disrepair and in need of serious maintenance and renewal. To further compound the issue, the federal fiscal environment is restrictive with not enough money to fund all infrastructure requirements. Combating these issues begins with a good asset management program that connects the organizational objectives to the strategic asset management plan. The AF's most recent program helps identify which maintenance and repair requirements should be done first. Currently, however, the AF is unable to clearly differentiate between its mission and support infrastructure. The research described in this thesis creates a model that allows senior leaders to objectively differentiate between mission (or core) and support (non-core) infrastructure. This is accomplished by establishing linkages between infrastructure assets and the AF core missions.

II. Literature Review

Chapter Overview

The purpose of this chapter is to give an understanding of the background and provide some deeper insight into the AF asset management program. First it will provide a full picture of the state of infrastructure in the United States, starting at the national level and working down to the Air Force level. Second, it will define asset management and give a general understanding of what is required. Lastly, this chapter will explore the Air Force's asset management program in order to understand how it works and how it can be improved.

Decaying Infrastructure

Infrastructure is an all-encompassing term. Generally the public envisions roads, bridges, utilities, and facilities when talking about infrastructure. It does, however, cover quite a bit more than just these few areas. The American Society of Civil Engineers (ASCE) includes the following in its report on infrastructure (2013) published every four years:

- | | | |
|-------------------|-------------|----------------|
| • dams | • aviation | • transit |
| • drinking water | • bridges | • public parks |
| • hazardous waste | • waterways | • schools |
| • levees | • ports | • energy |
| • solid waste | • rail | |
| • wastewater | • roads | |

As illustrated, the term infrastructure can refer to, and encompass, a great number of areas. Infrastructure refers to anything that the public uses or relies on, either directly

or indirectly, to carry out its day-to-day operations. Infrastructure requires maintenance and repair to avoid its failure. While some is given more attention than others, the United States as a whole does a poor job in maintaining its infrastructure assets (ASCE, 2013) and this is an expensive problem. The next few sections take a deeper look into the infrastructure problem starting at the national level and ending at the AF enterprise level.

National infrastructure

They are split between the public and private sector with the federal government owning about 15% (Hemme, 2015).

The national inventory of infrastructure includes some 84,000 dams, over 100,000 miles of levees, 4 million miles of roads, 117,000 miles of rail, 610,749 bridges, 11,000 miles of transit lines, 500 train stations, 300 ports, 19,000 airports, 55,000 community drinking water systems, and 30,000 wastewater treatment facilities (ASCE, 2013; GAO, 2008a, 2015). It is expansive and covers many different areas. In its most recent report card, ASCE gave American infrastructure a letter grade of "D+" (ASCE, 2013).

As of 2013, the average age of dams was 52 years old. Many have been reclassified as high-hazard dams, requiring physical upgrades, because of increased urban development downstream. These upgrades are in addition to the countless maintenance and repair items that are also required (ASCE, 2013). Of the over 1 million miles of water mains, it is estimated that there are 240,000 breaks per year. Additionally, much of the system is near the end of its useful life (ASCE, 2013). Furthermore, one in nine bridges is rated as structurally deficient. Collectively, the nation's bridges have an average age of 42 years old (ASCE, 2013).

These are just three examples of the issues plaguing national infrastructure. The ASCE report is littered with additional examples with similar bleak outlooks. While the nation's infrastructure is vast, much of it is critical to the nation's economy and affects the daily lives of the public (GAO, 2008b). Therefore, it is important that these issues be addressed to ensure the continued, uninterrupted growth of the economy, not to mention public safety.

While the sheer size and condition of infrastructure is daunting, the problem is more than just size; it is also cost. The association of State Dam Safety Officials estimated that the required maintenance and repair on the nation's dams would cost \$21 billion. According to the American Water Works Association, the cost to replace every water main in the country would top \$1 trillion. The Federal Highway Administration estimates that to complete the nation's bridge maintenance backlog by 2028, \$20.5 billion would be required annually. Their estimate is almost double the current expenditures of \$12.8 billion.

While this is just a snapshot of the cost of maintenance and renewal in a few areas, it is further evidence of the infrastructure problem facing the nation. The total cost to remedy the nation's maintenance and repair backlog is estimated to be \$3.6 trillion (Hemme, 2015). It is undeniable that the size and costs to remedy nation infrastructure are large and in need of attention. Some of that responsibility falls on the federal government.

Federal Infrastructure

Approximately 85% of the nation's infrastructure is owned and operated by the private sector (Hemme, 2015). However, even at 15%, the federal government is still

responsible for a sizable portfolio. Federal assets are owned and managed by more than 30 different government agencies with the Department of Defense (DoD), the Department of Veterans Affairs, and the General Service Administration (GSA) managing the greatest number of facilities (NRC, 2012).

In 2008, the federal government operated more than 407,000 buildings, totaling more than 3.3 billion square feet. This is in addition to another 488,000 structures (GSA, 2009). These facilities - buildings and structures - were valued at over \$1.5 trillion and cost the government \$47 billion to operate (GAO, 2008a). This does not include the more than 43 million acres the federal government manages. This land includes those owned and leased by the government and amount to an additional \$231 million in annual operating costs (GSA, 2013).

Furthermore, it is estimated that as many as 45,000 facilities are considered to be excess or underutilized (NRC, 2012). This excess creates a drain on the federal budget. The money spent on these excess facilities represents a lost opportunity to decrease a federal, deferred maintenance backlog in the tens of billions of dollars (NRC, 2012). In 2012 it was noted that the inventory of federal facilities has remained relatively stable while the cost to maintain them rose 2% (GSA, 2013). Of the 168,000 buildings operated by the federal government, over 11,000 have a utilization rate below 50%. While not the largest entity in the federal government, the AF is responsible for a sizable amount.

Air Force Infrastructure

According to the GSA real property report from 2012, the Air Force has a portfolio containing 54,873 buildings and 73,055 structures. It maintained the 594 million square feet and 8.4 million acres of land at an operating cost of \$4.64 billion

(GSA, 2013). With a plant replacement value (PRV) of \$890 billion in 2015, the AF is spending less than half a percent of the PRV on maintenance and repair. Industry suggests between two and four percent to address annual maintenance and repair (Civil Engineering Research Foundation, 1996). This is made worse when considering the AF and other government agencies continue to struggle with mounting backlogs of maintenance and repair (GAO, 2008a). While the size and condition of the infrastructure assets are problematic enough, there is not enough money to address them all.

At less than half a percent, the AF is well below the 2-4% recommended for maintenance and repair (Civil Engineering Research Foundation, 1996; NRC, 1996). It should also be noted that 2-4% would not be adequate with such a significant amount of deferred maintenance. In truth, the 2-4% estimate would only be enough to cover the annual maintenance and repair requirements. With a PRV of \$890 billion, the AF would need to spend between \$17.8 and \$35.6 billion annually to keep up with routine maintenance and repair. Unfortunately, congressional committee briefings from 2009 indicate that no agency, including the military, has reported a level of investment as high as 2% and that underinvestment remains a huge issue (GSA, 2009; NRC, 2012).

Fiscal Environment

To further compound the issues, multiple government reports cite strain on fiscal funding and poor budgetary outlook for the nation's infrastructure assets (GAO, 2008b; NRC, 2012). This is significant because, as described above, the nation has a large backlog of maintenance and repair. While the available budget for maintenance and repair will continue to shrink, infrastructure maintenance and repair backlogs will continue to grow. The AF also acknowledged the fiscally constrained environment and

the need for a cost-conscious mindset in the most recent Strategic Master Plan (USAF, 2015e). It further indicates that funds which used to be available for maintenance and repair will begin to evaporate while the portfolio of AF infrastructure assets stays the same or grows. All government agencies, including the AF, will have to optimize the way it chooses infrastructure projects in order to address the most critical projects first.

With an estimated cost of \$3.6 trillion, the \$47 billion spent on the 995,000 facilities the federal government operates is not enough. It is for these reasons that responsible spending and infrastructure management reform have become important focus areas (GAO, 2008b; NRC, 2012). The eminent age, growing costs, mounting backlogs, and decreasing budgets have combined to create a massive infrastructure issue. Fortunately, infrastructure asset management can help.

Defining Asset Management

The infrastructure dilemma is not new to the federal government. In 2004 an Executive Order was signed mandating that all federal agencies begin to establish policies, goals, and objectives in an effort to responsibly manage United States real property (Executive Order No. 13327, 2004). This was the first step in combating the issue. These policies, goals, and directives would pave the way for infrastructure asset management programs within the federal government. This section describes what an asset management program is and what it hopes to achieve.

Vanier (2001) suggests that an asset management program should seek answers to six basic questions:

- What do you own?
- What is it worth?

- What is the deferred maintenance?
- What is its condition?
- What is the remaining service life?
- What do you fix first?

The federal government views asset management as a systematic process for maintaining, upgrading, and operating physical assets cost effectively (NRC, 2008). The US Army Corps of Engineers (USACE) defines asset management as “a disciplined corporate approach for the management of an asset portfolio requiring integration and collaboration with all corporate organizations and programs, and their respective activities and contributions” (USACE, 2011). Publically Available Standard 55 (PAS-55), an internationally recognized guide on implementing asset management, defines it as “systematic and coordinated, activities and practices, through which an organization optimally and sustainably manages its assets and asset systems, their associated performance, risks, and expenditures over their life cycles for the purposes of achieving its organizational strategic plan” (The Institute of Asset Management, n.d.).

Asset management can be broken into two distinct parts. The first is maintenance and repair; the second is capital renewal. Maintenance and repair is defined as an intervention to ensure an asset reaches its optimal service life (Hughes, 1991). Maintenance and repair is critical if an asset is to reach its maximum life. Whether the asset is a road or an HVAC unit, regular maintenance and repair is a requirement. This not only maximizes its life, but also is necessary for the optimal and consistent performance of any asset.

Changing the filters on a heating, ventilating, and air conditioning (HVAC) unit is an example of required routine maintenance. If the filters are never changed, debris will slowly clog it which will cause the fan to work harder to draw in air. The primary effects of neglecting to change a filter are dramatic reduction in the effectiveness of the unit. If unchanged for too long, it will also decrease the useful life of that asset and perhaps require that the fan motor or condensing unit be replaced. A secondary result might be employees comfort level which, in turn, could affect their productivity. Both results will cost the owning agency more money than if the filter had been replaced.

Capital renewal is defined as a comprehensive action to completely replace an asset with a new one, potentially with a different function or location (Vanier, 2001). This happens for two reasons. The first is due to the end of an assets life. In the case of the HVAC unit, it could be beyond repair, or cheaper to replace than repair, after many years of continued use. The second case is due to organizational expansion or a change in function. For example, a company may have outgrown a particular facility and plans to build a new one.

Both maintenance and repair, and capital renewal are required parts of an asset management program and are necessary to ensure that organizational objectives continue to be met. However both things require planning. In the example of building a new facility, an organization will have planned and budgeted for capital renewal long before the building begins. They also would have verified that such an investment was in the best interest of the company and further organizational objectives. The maintenance and repair of assets should be prioritized in a similar fashion. If the budget does not allow for

the maintenance and repair of all assets, the organization must decide which ones to fund. This is where a strategic asset management plan (SAMP) becomes a necessity.

Strategic Asset Management Plan

Strategy is arguably the most important part of an asset management program (IFMA, 2009; Infrastructure Asset Management Exchange, 2015; ISO, 2014; The Facilities Society, 2014). It provides direction. ISO 55000 defines a strategic asset management plan (SAMP) as the "documented information that specifies how organizational objectives are to be converted into asset management objectives, the approach for developing asset management plans, and the role of the asset management system in supporting achievement of the asset management objectives" (ISO, 2014). In short, a SAMP is the key linking factor between organizational objectives and asset management implementation (Oostermeijer & Boss, 2010). It ensures that organizational objects continue to be met and go uninterrupted through the implementation of asset management. In the case of the AF this means that the mission is neither interrupted nor degraded due to a lack of maintenance and repair or capital renewal.

In a recent survey of asset management practitioners and solution providers, both ranked building a coordinated asset management strategy as a top priority (Infrastructure Asset Management Exchange, 2015). In order to do so, an organization must be able to identify its organizational objectives and find a way to connect them to the way it manages its assets. It is important for an organization to have the ability to differentiate between its core business facilities, and its non-core facilities (IFMA, 2009; The Facilities Society, 2014). Without such a connection, the organization would be unable to prioritize its maintenance and repair to ensure that organizational objective continue to

be met. An additional, equally as significant, benefit is the ability to compare like-needs (core to core) instead of competing needs (core to non-core).

This distinction is a very important part of working organizational objectives into the asset management program. The SAMP should connect core, organizational strategies and their core facilities to ensure that organizational objectives are met. While it is important for an organization to consider all their needs, it must be able to distinguish between competing requirements. This guarantees that the organization will be able to continue functioning effectively, even in difficult economic environments.

Asset Management Program

The SAMP is just one part of an asset management program. This section investigates what else a successful asset management program requires. Almost universally, asset management programs will seek to clarify what assets an organization owns, what condition each asset is in, and the risk to the organization if that asset fails (Faber & Stewart, 2003; Marlow et al., 2012). Such information forms the data base of an asset management program. While the information is important, knowing what to do with it and how to use it to make informed decisions is also important. The key to efficient and timely infrastructure asset management is good decision making tools (Hastak & Baim, 2015; Infrastructure Asset Management Exchange, 2015; NRC, 2008; Vanier, 2001). The tools allow for cost-effective decision making over the whole life-cycle of an asset. Together with the right policy, asset management tools can help managers direct money and energy appropriately while also helping to plan for the future.

Currently, there is no universally accepted tool or model. While there are universally accepted principles spelled out in both ISO 55000 and the International

Infrastructure Management Manual (IIMM), each organization is free to adopt whichever tools or models they deem necessary to achieve their asset management objectives. The goal of any asset management model should be to optimize the maintenance, repair, and renewal investments while minimizing the risks to organizational objectives.

While the SAMP, and subsequent policy, dictate how an organization will achieve asset management objectives, the asset management model is one of the most important tactical pieces for decision makers. The model uses the asset management data base and connects them in various ways to show which assets need attention first. The effectiveness of any model hinges on the quality of the data set (Provost & Fawcett, 2013). Two significant pieces of data are the condition of an asset, and the risk associated with that asset. When combined with a decision analysis tool they produce the basis of an asset management model (Elsawah, Guerrero, & Moselhi, 2014). Since the purpose of this research is not condition assessments, it will not be further investigated. The next section will cover risk and how it relates to the asset management program and the SAMP.

Risk Analysis

Communicating risk is an important part of any asset management program and it has been extensively researched (Faber & Stewart, 2003; Hastak & Baim, 2015; Marlow et al., 2012). It helps organizations compare the cost of performing maintenance and repair against the cost or impact of deferring it in favor of another (Water Research Foundation, 2012). Risk "R" is defined as the expected consequences "C" associated with the probability "P" of a given activity (Faber & Stewart, 2003).

$$R = C * P$$

The goal of risk analysis within the framework of infrastructure asset management is to reduce uncertainty related to the cost of performing maintenance and repair against the cost, or impact, of not performing it (Marlow et al., 2012). Risk can be hard to capture because it is difficult to quantify. The aforementioned definition is an attempt to quantify risk and could be in terms of dollars, human fatalities, or hours of interrupted service. In most cases there will be an acceptable and unacceptable level of risk defined by the organization. By including risk in an asset management model, program managers have a more complete picture, enabling them to more effectively prioritize maintenance, repair and capital renewal.

One way to connect risk to organizational objectives is through the consequence “C” of an assets failure. However, defining the consequences associated with risk as a metric is difficult. This specific area lacks the research found in many other asset management criterion (Salman & Salem, 2012). In rudimentary asset management models, the consequence are simply the immediate and local effects of an assets failure. Through the development of a quality SAMP, however, the consequence shifts to an inability to achieve organizational objectives. In this way, risk becomes a vital metric in the way an organization performs asset management.

Air Force Asset Management

This section is meant to give an understanding of the current AF asset management program. It first defines how the AF views asset management and the different criterion used to make infrastructure investment decisions. Second, it lists the shortfalls of the current program, specifically the mission dependency metric. Finally, the

section will end with a discussion on the lack of a strategic asset management plan and how developing one would benefit the AF.

Air Force Asset Management Program

Each organization defines asset management a little differently. The AF defines asset management as “the expertise and processes required to facilitate a centralized approach to built and natural asset resource allocation decisions, taking into account such things as mission impact, facility conditions, strategic policy directives, and safety deficiencies” (USAF, 2012b). The goal of AF asset management is to “develop investment plans that support mission-critical needs while addressing the USAF's most urgent requirement first, minimizing the risk to mission and Airman” (USAF, 2012a). In short, the AF employs asset management to assure they have an enterprise-wide picture when making investment decisions so as to minimize the risk of mission degradation or failure.

The AF Programming Plan (P-Plan) for implementing Program Action Directive (PAD 12-03) lays out the AF's asset management process. Before its implementation, the AF had asset management projects scattered throughout several different lists. Each list was specific to a program or asset. For example, there was a list for capital renewal, and a separate list for maintenance and repair. There were even separate lists for project types such as energy efficiency and environmental. In its most recent form the AF consolidates all infrastructure requirements into a single list, the Integrated Priority List (IPL). The AF moved to a centralized process to ensure senior leaders had complete visibility of the enterprise while also giving them the ability to compare requirements across all programs to ensure the best investments were made (AFCEC, 2015). To do

this, the AF had to develop a scoring system that would compare projects of different types and rank them from most urgent to least.

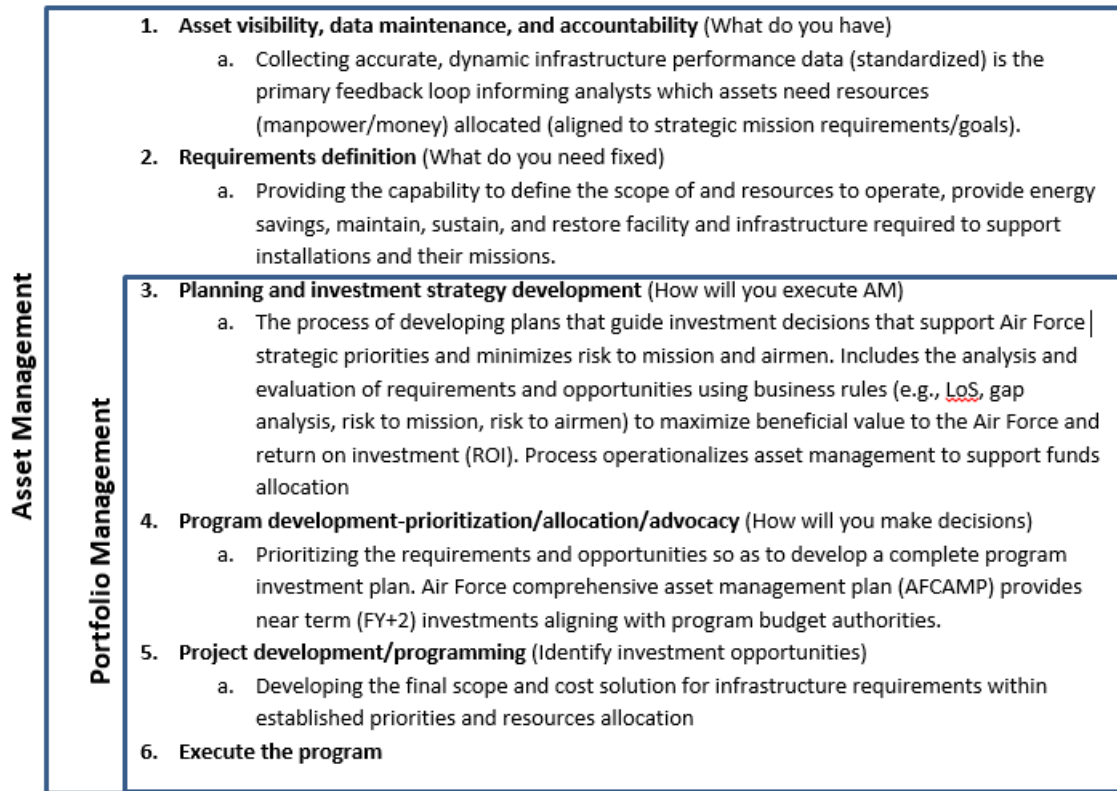


Figure 1: Air Force Asset Management Process.

The AF asset management scoring system consists of three criteria. They are probability of failure (PoF), consequence of failure (CoF), and a savings to investment ratio (SIR) (AFCEC, 2015). Each criterion addresses a different piece of an asset and are added together to determine a projects total score. Projects with the greatest scores indicate a greater need, or a greater risk to the organization if not completed.

$$Total\ Score = POF + SIR + COF \quad (1)$$

The PoF criteria is found uses BUILDER, a program developed by USACE. As assets and its condition are added to the data base, BUILDER calculates the condition for each asset. This score is called a condition index (CI). The CI has the ability to change over time but is fixed for a particular project in a particular year. PoF scores ranges from 0 to 100 points with higher point value indicating that the facility or internal system needs more attention than others with lower point values.

$$PoF = (100 - CI) * 2 \quad (2)$$

The CoF is a combination of a facility's mission dependency index (MDI) and additional points given out by the different AF Major Commands (MAJCOMs). The MDI is a value index developed by the Navy in 2000 as a way of determining which infrastructure assets were most important. The AF adopted the same metric in 2008 (Nichols, 2015). The MDI ranges from 0 to 100 with higher scores indicative of more valuable assets. The MDI was developed to create a link between infrastructure and mission (Antleman, Dempsey, & Brodt, 2008). MDI scores for infrastructure generally remain static from year to year. MAJCOM points are given to particular projects based on a MAJCOM's desire to see certain projects funded. The points range from 0 to 40 with higher point values indicating a higher priority. Each MAJCOM is assigned a fixed number of priorities to which they can assign points based on the plant replacement value (PRV) of all the assets in their particular MAJCOM. Equation 3 shows how MDI and MAJCOM points are combined to produce a CoF score.

$$CoF = MDI * 0.6 + MAJCOM \text{ points} \quad (3)$$

The savings to investment ratio (SIR) is the third criteria used in the scoring model. The SIR's intent is to capture the value associated with an opportunity to reduce costs and improve services. It is determined using data from the Interim Work Information Management System (IWIMS). IWIMS tracks work performed on any asset, including the time and money spent on a particular work order. This data is used to calculate the SIR for a particular project using an excel worksheet provided by the Air Force. The worksheet uses the IWIMS data and the project's estimated cost to calculate an SIR score.

The AF uses the three criteria to rank order all infrastructure projects in a given year to decide which projects to fund. The scoring method allows the AF to see and compare all asset management projects on a single list. The AF asset management model ensures that the AF funds projects that address assets in the worst condition with the greatest impact to mission.

Shortfalls

As mentioned above, the goal of AF asset management is to address mission-critical needs first. It currently does that by using the MDI, which was designed specifically to create a link between infrastructure and mission (Antleman et al., 2008; USACE et al., 2010). However, the MDI has several fallacies and they have been well documented (Kujawski & Miller, 2009; Nichols, 2015).

The MDI was developed by investigating several questions designed to determine the interruptability and relocateability/relocateability of a facility. The interruptability of a facility asks how long the occupants could be interrupted without affecting their mission. The relocateability/relocateability of a facility asks if the occupants could perform their mission in another facility or by setting up a temporary facility. Lastly, the MDI asks how many other facilities on an installation depend on the facility of interest. While these questions might seem valid at face value, there are several things wrong with the assumption that they are able to create a link between infrastructure and mission.

First, the questions do not define what a mission is. Each entity within the AF has a mission, however not all entities contribute directly to the core missions. By assuming that all missions from every entity are important and worth measuring, nothing becomes important. Further, it does not fit with the idea that an organization has core business functions and non-core business functions; nor does it fit with the idea laid out in the SMP that the AF should be able to distinguish between mission and support infrastructure (USAF, 2015c). This statement, by its very nature, implies that there are some entities within the AF that are purely support. The MDI does not differentiate between mission and support. Second, because the MDI is performed and calculated on an installation by installation basis (Antleman et al., 2008; USACE et al., 2010), it is lacking an enterprise-wide look at infrastructure and mission. While high scores are said to be indicative of a critical facility, it can be argued that this is only when looking at a particular installation. Thus MDI scores only convey the relative importance of a facility and not the absolute importance. Finally, the MDI is a metric based almost entirely on subjective measures and responses that are ambiguous (Kujawski & Miller, 2009).

It is also worth noting that the AF has performed almost no MDI studies of their own to determine the true values at its installations (Nichols, 2015). The AF instead used the values that the Navy received from their studies and only slightly tweaked them to fit the AF perspective. In closing of this section, the MDI is not an appropriate measure of consequence “C” in the risk analysis model. It does not adequately represent the AF’s ability to achieve organizational objectives. The AF is also unable to use the MDI to distinguish between mission and support and, therefore, has no way of differentiating between core and non-core infrastructure.

Refining Air Force Asset Management Strategy

Currently the AF does not have a documented SAMP stating how they are ensuring AF objectives through asset management implementation. Creating an effective SAMP would require the AF to connect AF objectives to its assets. While AF asset management policies do aim to sustain the mission, it does not define what that mission is. The AF does, however, have a tangible mission set. They are found in the five core missions. These core missions are central to what makes the AF a unique organization and defines why it exists. They are (1) air and space superiority; (2) intelligence, surveillance, and reconnaissance; (3) rapid global mobility; (4) global strike; (5) and command and control (USAF, 2015a, 2015e; Welsh, 2014). These five AF core missions are the organizational objectives of the AF. Therefore, they must be worked into the SAMP to ensure that asset management is implemented in a way that safeguards organizational success.

Additionally, as mentioned earlier, the AF has recently called for the ability to distinguish between mission and support infrastructure (USAF, 2015e). The SMP gives

the example of choosing between a dormitory and a space launch facility; however the distinction is not always so easily made. This distinction is another important piece that should be included in the SAMP. For the AF, this means separating facilities that contribute directly to the AF core missions from those that merely support them. This would allow the AF to fundamentally differentiate projects from one another. Using the aforementioned example, a space launch facility project should not be competing with a dormitory project. The two are vastly different and it would be unfair to try and equate them, if not impossible. A space launch facility contributes directly to several of the AF's core missions. While an important aspect of taking care of Airmen, housing personnel is not a core mission of the AF and, therefore, should not compete with projects that are.

Literature Gap

While there are many publications that point out the need to separate core business facilities from non-core (IFMA, 2009; Payne, 2000; The Facilities Society, 2014; USAF, 2015c), there are no publications explaining how an organization might do so. The MDI was developed using operational risk management (ORM) methods. However, the ORM method used to develop the MDI is a significant shift from classical approaches and fails to deliver an accurate measure of risk (Kujawski & Miller, 2009). As mentioned earlier, risk is a product of consequences and probabilities and is defined on a continuous spectrum. However, dividing infrastructure into two groups is categorical and binary. In order to use risk to categorize facilities as core or non-core, the organization would have to come up with an acceptable and unacceptable level of risk.

As Hubbard (2014) points out, measuring such a cut-off point can prove difficult and one should avoid picking an arbitrary point estimate.

The literature review has shown that the AF lacks a good measure of risk associated with infrastructure and core missions and, therefore, some other form of data must be used to categorically separate infrastructure assets into mission and support. The AF does have a substantial data base that contains infrastructure titles and descriptions. This data is paired to unique category codes that identify all the various types of infrastructure the AF maintains in its inventory. Qualitative techniques can be used to identify keywords related to core missions. The titles and descriptions of all the infrastructure assets in the AF inventory can then be searched to determine how many contain keywords. The following section is a literature review on qualitative analysis.

Qualitative Analysis

In many fields, such as psychology, sociology, and health, qualitative analysis provides researchers with a way to investigate complicated questions and develop new theories (Savin-Baden & Major, 2013). While many scientific and physical research efforts can use statistical analysis to draw conclusions, qualitative research requires a different set of skills. Qualitative research is any type of research that produces results not arrived to by traditional statistical methods (Strauss, Corbin, & others, 1990). Savin-Baden (2013) describes it as "an ongoing process that involves breaking data into meaningful parts for the purpose of examining them". While there are many different ways to go about qualitative research, most involve a process known as coding.

Coding

The goal of any qualitative research effort is to 'make sense out of the data' (Merriam, 2009). This is done first through a process known as coding. Coding is defined as an analytic process through which data is broken down, conceptualized, and put back together to form a theory (Strauss et al., 1990). A code is simply a system of symbols - letters or short words - that are used to represent a theme. They should be meaningful and provide insight into an idea within a data segment.

There are two types of coding schemes. The first is descriptive coding. It refers to a process of summarizing or describing the text, and tends to involve deriving codes directly from the text. These codes are referred to as inductive codes. The second is analytical coding. It refers to a process of deriving codes based on what the researcher perceives the text is saying. Aside from inductive codes, there are also priori codes, which refers to codes based on prior theory or literature review, and co-occurring codes, which refers to a data segment receiving more than one code. An example of co-occurring might come from a statement of feeling where both anger and sadness are displayed and deserve both codes.

In some cases the researcher may rely only on these open codes. In others, however, the researcher may perform something called axial coding. Axial coding is done by grouping codes under a single similar code. Axial coding can help categories to begin to emerge and is useful in some circumstances. The researcher should define axial codes as they are developed so as to maintain consistency throughout the axial coding process. Axial coding is one step closer to a category.

Categories

Coding data is the first step in qualitative research. The next step is to create categories to group codes into like-kinds. "A category is a group of words with similar meaning or connotations" (Weber, 1990) and is simply a way to organize the codes. The categories must be mutually exclusive and exhaustive while also being on the same conceptual level (GAO, 1996; Savin-Baden & Major, 2013). The goal is to roll all the codes up into groups that will help the researcher in identifying a theme. For example, if the codes were "sadness" and "anger" the category might be "emotions".

Subjectivity of Qualitative Analysis

It is important to note that qualitative research seeks to be as objective as possible. However because the data can be subjective depending on who is viewing it, the codes and categories can be different from researcher to researcher. Savin-Baden (2013) points out that because the social world is more subjective than the natural world, it cannot be understood in an objective reality. The goal of any qualitative analysis is to arrive at a place of intersubjectivity, or mutual agreement. This means that if the research were performed by a peer, they would arrive at similar conclusions. A high quality analysis is one with high intersubjectivity; therefore it should be the aim for any qualitative research effort.

Conclusions

In light of the decay of national infrastructure and decreasing federal budgets, the AF has turned to asset management as a way of ensuring future success. Presently, the AF is unable to tie organizational objective to asset management implementation or

distinguish between its core and non-core infrastructure. Its current model is also flawed in the way it calculates risk. While difficult, there is value in adjusting the SAMP by including AF core missions and connecting them to AF investments in infrastructure. Such a distinction would be the first time the AF has identified which infrastructure assets directly relate to AF core missions and which ones are support. The connection would also allow senior leaders to more confidently make difficult decision related to infrastructure in this fiscally uncertain environment. Crafting a SAMP would indirectly ensure future success of the AF mission by sustaining its infrastructure.

III. Methodology

Chapter Overview

This methodology chapter describes the steps taken to identify infrastructure groupings directly related to Air Force (AF) core missions. This method utilizes qualitative techniques, both conventional and summative content analysis, to identify keywords and phrases. The keywords and phrases are then used in a model containing infrastructure asset data. The model performs a search of real property descriptions to determine which assets contain those keywords or phrases. Any asset that contains a keyword or phrase is flagged. Infrastructure assets are then grouped into preexisting categories and those containing a majority of flagged assets (75% or more) are considered core, or directly related to AF core missions.

Purpose

The purpose of this research is to identify infrastructure groupings that directly affect the Air Force's ability to carry out its core missions of air and space superiority; intelligence, surveillance, and reconnaissance; rapid global mobility; global strike; and command and control. In doing so, the research will meet the intent laid out in the Air Force Strategic Master Plan (SMP) of distinguishing between mission and support infrastructure. This gives senior leaders the ability to make better-informed decisions related to infrastructure maintenance, repair, consolidation, demolition, and new construction.

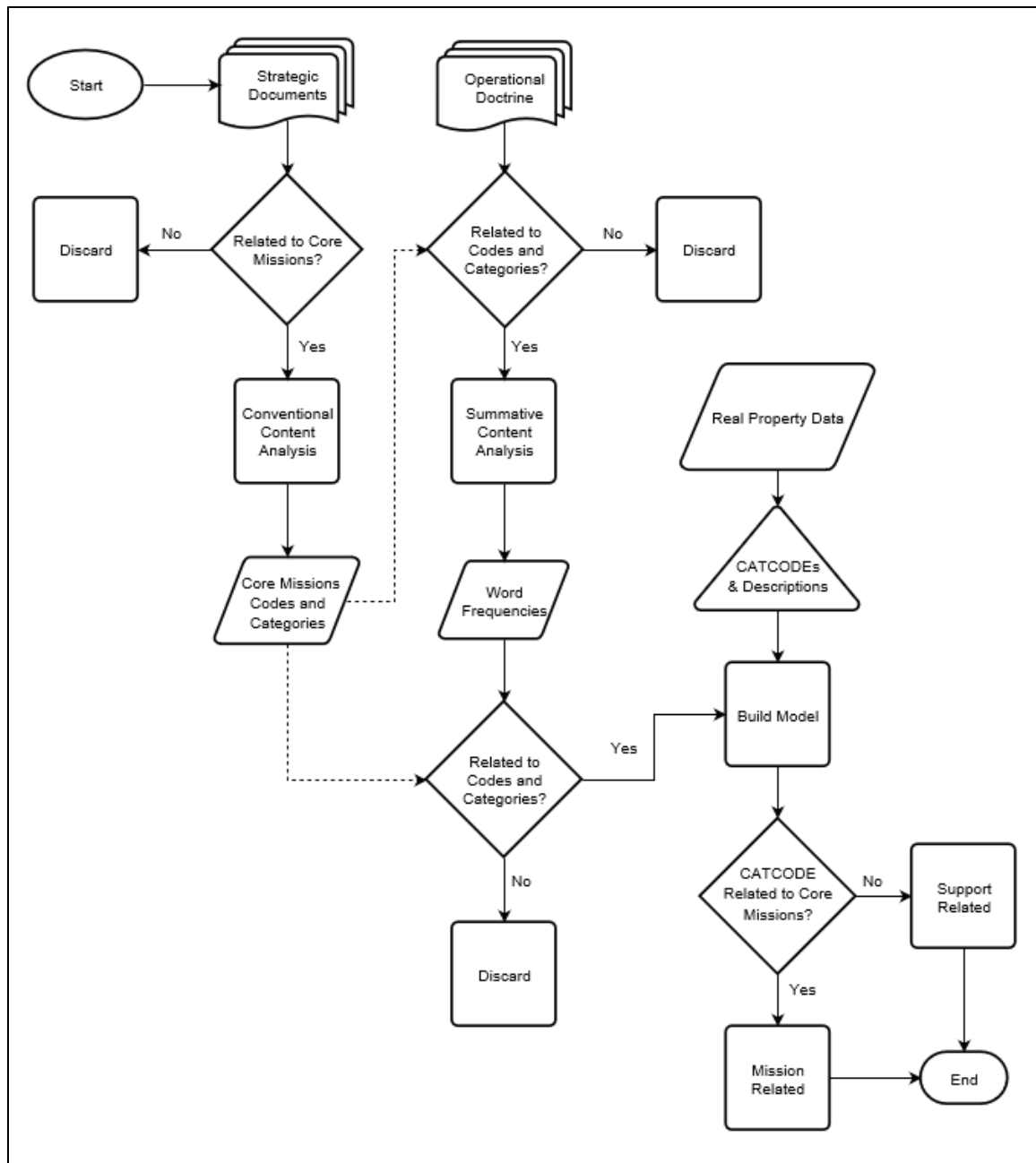


Figure 2: Proposed Research Method Diagram.

Content Analysis

Content analysis is a form of qualitative analysis. The objective of this research in using content analysis is to identify words that appropriately represent the AF core

missions. While other methods, such as the Delphi method, could also work in identifying keywords, this research method chose to utilize content analysis as a means of achieving that goal.

Content analysis is defined as a systematic, replicable technique for compressing many words of text into fewer content categories based on explicit rules of coding (Berelson, 1952; Krippendorff, 1980; Weber, 1990). Content analysis more generally is "any technique for making inferences by objectively and systematically identifying specified characteristics of messages" (Holsti, 1969). Researchers regard content analysis as a flexible method for analyzing textual data (Cavanagh, 1997). Therefore, the data used to support this method is comprised primarily of text. A researcher then uses the data to derive meaning, supporting or rejecting a theory, through a variety of techniques.

While there are several ways to conduct content analysis, two that will be used for this study are conventional content analysis and summative content analysis. In conventional content analysis, the researcher immerses himself in the data and begins to sort through it to find meaning. The researcher uses words and phrases in context to develop a theme (Hsieh & Shannon, 2005). In summative content analysis, the researcher examines word counts as a way of deriving meaning from the text (Savin-Baden & Major, 2013).

Both methods do have limitations. In conventional content analysis the researcher should take care to ignore personal biases. Coding, described in the Literature Review chapter, must be done as objectively as possible. The codes should come directly from the text with the researcher taking special care to ensure that context is preserved.

In summative content analysis the assumption is that words used most frequently will reflect the areas of greatest concern. Weber (1990) points out that synonyms may be used for stylistic reasons leading researchers to underestimate the importance of a particular term. Additionally, certain words may have more than one meaning (Stemler, 2001). For example, the word "space" as a noun could mean an unoccupied area or refer to outer space, while as a verb could mean to position at a distance from one another. The researcher should be aware of such pitfalls and plan around them when possible. When performing summative content analysis, word frequencies should be used to identify words of potential interest. Context will then allow the researcher to select certain terms and disregard others (Stemler, 2001).

Data Description

This portion of the chapter transitions to explaining the various types of data. The two forms of content analysis will utilize two forms of data respectively: (1) AF strategic documents and (2) AF operational doctrine. Although not used in content analysis, Department of Defense (DoD) real property data is also used in this method, and therefore discussed below. The discussion explains how the DoD and the AF categorize their infrastructure assets and how it relates to the research. All different forms of data discussed in this section are necessary in the creation of a model which identifies core infrastructure groupings.

Strategic Documents

With the goal in mind of selecting words which adequately represent the AF core missions, the first step in the method is to understand the AF core missions. To do so,

codes and categories are developed from strategic documents using conventional content analysis. Strategic documents are a collection of seven publications intended to give Airmen strategic insight. The strategic documents which are used in this research include Global Vigilance, Global Reach, Global Power for America (USAF, 2015a), Senior Leaders Perspective from Gen Mark Welsh III (Welsh, 2014), and the Strategic Master Plan (USAF, 2015e). All contained information on the AF core missions and, therefore, are chosen from among the seven others for conventional content analysis.

Operational Doctrine

There are many levels of AF strategic doctrine and they all have various levels of classification. AF Strategy is developed at the very top in conjunction with the other services and the Secretary of Defense. Some documents are meant for public consumption while others are classified to protect the capabilities and specific tactics of the AF. This research does not examine AF tactics, techniques, and procedures (TTPs) but instead uses operational doctrine. Specific operational doctrine are used in the summative content analysis to identify keywords for use in the model.

Operational doctrine describes the organization and employment of AF forces in the context of distinct objectives, force capabilities, and functional environments (USAF, 2015b). Operational doctrine is found in the Curtis E. Lemay Center for Doctrine Development and Education. The library contains three volumes of doctrine and 29 annexes. A complete list of doctrinal annexes can be seen in Table 1. Only specific annexes are chosen based on their relevance to core missions. The Annexes are used in summative content analysis to identify high frequency words and phrases. Keywords are

then chosen from among the lists of high frequency words and phrases for use within the model.

Table 1: Complete list of Operational Doctrine.

Doctrine Annexes		
Force Development	Cyberspace Ops	Electronic Warfare Ops
Legal Support to Ops	Information Ops	Airspace Control
Global Integrated ISR Ops	Space Ops	Weather Ops
Ops and Planning	Air Mobility Ops	Targeting
Irregular Warfare	Foreign Internal Defense	Public Affairs
Counterair Ops	Homeland Ops	Strategic Attack
Counterland Ops	Command and Control	Nuclear Ops
Countersea Ops	Engineer Ops	Combat Support
Special Ops	Counter-CBRN Ops	Medical Ops
Force Protection	Personnel Recovery	

AF Real Property Data

Infrastructure asset data is an important part of this research and, fortunately, the AF keeps extensive real property data. Each infrastructure asset is categorized in order to differentiate their various functions. The AF refers to this categorization as a category code (CATCODE). Each CATCODE is a collection of six digits with each one having a specific meaning. Figure 4 illustrates how a CATCODE is broken down. As a part of real property data, some CATCODEs have descriptions which accompany them. These descriptions are published in the Air Force Manual (AFMAN) 32-1084. However, because not all CATCODEs have a description, the researcher looked to DoD-assigned facility analysis category (FAC) codes. While not as numerous, every FAC code is accompanied by a description. The final, extracted, real property data used in this research method contains (1) all the unique CATCODEs used in the AF, (2) its respective

CATCODE title and description (if one existed), and (3) its corresponding FAC code title and description. An example of the data can be seen in Figure 3. The section that follows describes the two different categorization levels in more detail. Specifically, how they are related and how the data from two different classifications can be married.

CATCODE	CATCODE Description	FAC Description
136667	TAXIWAY LIGHTING: Taxiway lighting, including edge, centerline, hold and runway exit lighting, and guidance signs, may be provided for all regularly used taxiways.	AIRFIELD PAVEMENT LIGHTING: Lighting on the sides and approaches of airfield pavements. Does not include other airfield lighting, such as perimeter lighting, security lighting, and street and parking lighting.
149621	WIND MEASURING SET: Consists of a transmitter/sensor, indicator, and recorder that provide wind direction and speed for air traffic control operations, weather observations, and forecasting.	AIRCRAFT NAVIGATION FACILITY: A facility, other than a building, that supports aircraft navigation and traffic control operations.
149965	RADAR TOWER: N/A	Miscellaneous Operations Support Facility: A facility, other than a building, that directly supports operational activities and is not included in another Facility Analysis Category.

Figure 3: Example of the Subset of Real Property Data.

DoD FAC codes and AF CATCODEs

The DoD has over 308,000 structures in its inventory (GSA, 2013). To classify all the various types, the DoD uses FAC codes to group them into meaningful sets. A FAC code is a four digit identifier and each FAC code is assigned its own title and description. There are 479 unique, DoD-defined FAC codes; each year FAC codes and their descriptions are reviewed, edited, and republished. FAC codes can be added or removed, and descriptions changed or updated, as the inventory and facility types change. FAC codes are used by the DoD to categorize all facility types across all branches of the military.

FAC codes can be broken down by digits with each additional digit providing more detail about the infrastructure asset. The first digit identifies the Facility Class; the first two, together, identify the Category Group; the first three identify the Basic Category; finally all four make up the Facility Analysis Category, or FAC. A list showing the different classifications and their respective counts can be seen in Table 2.

Table 2: Count by Classifications.

Number of Digits	Facility Grouping	Count of Unique Groups
1	Facility Class	9
2	Category Group	42
3	Basic Category	162
4	Facility Analysis Category	480

As an example, a FAC code with a "1" as the first digit identifies a piece of infrastructure belonging to the Facility Class of *Operations and Training*. A FAC code with the first two digits "11" belongs to the Category Group of *Airfield Pavements*. A FAC code with the first three digits "111" belong to the Basic Category of *Airfield Runways* while a FAC code with the first three digits "112" belongs to the Basic Category of *Airfield Taxiways*. In this way, with each additional digit provides more detail about the facility or utility. A complete list of Facility Classes can be seen in Table 3 and a complete list of Category Groups and Basic Categories can be found in the Appendix A and B respectively.

Table 3: Complete List of Facility Classes.

Code	Title
1	Operation & Training
2	Maintenance & Production
3	Research, Development, Tests & Evaluation
4	Supply
5	Hospital & Medical
6	Administrative
7	Housing & Community
8	Utility & Ground Improvements
9	Land

Just as the DoD uses FAC codes, the AF uses CATCODEs to classify its infrastructure assets. There are currently 919 unique CATCODEs in use. The AF assigns each one a title; however, only 313 CATCODEs contain a description, leaving 606 without one. Each CATCODE is a unique six digit identifier that builds upon DoD FAC codes. Identical to the way FAC codes are built, each collection of digits lends specific meaning; with each additional digit the categories become more exclusive. This is illustrated in Figure 4.

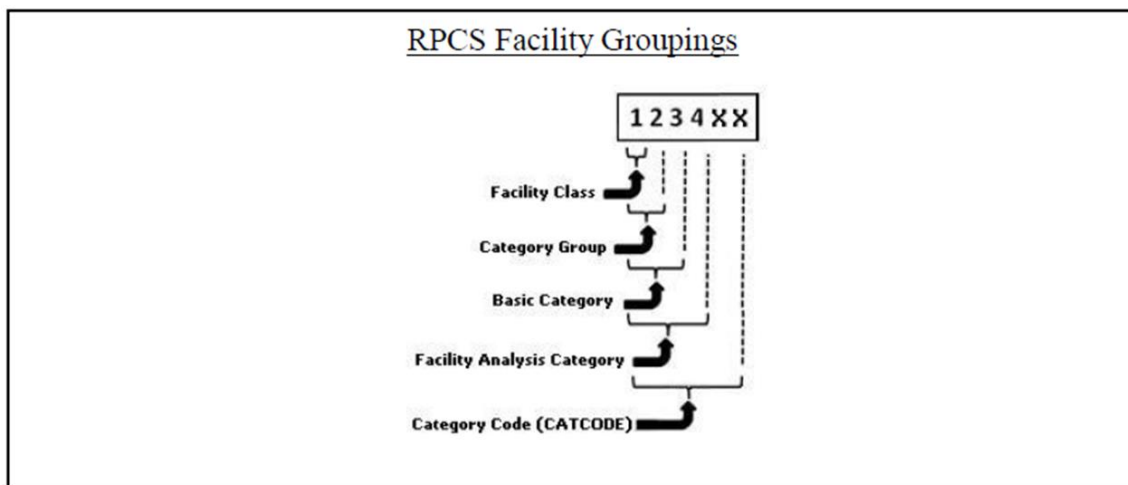


Figure 4: Illustration of the Breakdown of CATCODEs.

In principle, the FAC code can always be found in the first four digits of a CATCODE. In practice, however, it is not always the case. There are some circumstances where the AF CATCODE and DoD-assigned FAC code does not agree. This was discovered by examining the DoD real property publication of FAC codes and descriptions where, in it, the DoD also assigns every CATCODE to a FAC code. In some instances, the first four digits of the CATCODE do not match the DoD-assigned FAC code. While these cases are rare, it is worth noting that the correct, DoD-assigned FAC code cannot always be found in the first four digits of the CATCODE. This research did not investigate why these differences sometimes exist. For this research, all AF CATCODEs are paired with their DoD-assigned, FAC code description.

Conventional Content Analysis

Now that the various types of data have been explained, the next section of the chapter outlines how strategic documents are analyzed using conventional content analysis. It first discusses the process of developing codes and categories before transitioning into how the codes and categories are used. As mentioned in the Literature Review chapter, codes and categories are used to investigate and understand an unknown theme. For this analysis, the theme is known (AF core missions) and codes and categories are developed to provide more insight and describe them in detail.

Developing Codes and Categories

In order to develop the codes, and subsequent categories, the strategic documents are read thoroughly to identify which sections pertain to AF core missions. Once those sections are identified, the researcher examined them closely and began to code them.

Codes are developed for each AF core mission and are inductive (coming straight from the text). Every word and phrase is considered carefully to determine if it is related to core missions. When a word or phrase is found, it is recorded and context is preserved. Each core mission results in a list of words and phrases. This initial list of inductive codes are then axial coded to more broadly define. Definitions are given to each axial code created in order to ensure that the use of a particular axial code is consistent throughout the process.

Once the inductive and axial codes are created, categories are generated. This is a move from seeking particular terms to general terms. Categories are sensitive to the data, exhaustive (enough categories to encompass all relevant data), mutually exclusive (a relevant unit of data can be placed in only one category), and conceptually similar (all at the same conceptual level). Categories are made for each AF core mission and will be specific only to its respective core mission.

How Codes and Categories are used

While the goal of creating codes and categories is to gain more understanding of the AF core missions, they also serve a greater purpose. The insight gained from codes and categories allow for the selection of relevant operational doctrine. Each annex is reviewed carefully to determine if its theme matches any of the codes and categories created for each core mission. All annexes are considered and only those found to be relevant to core missions are used in the summative content analysis.

Summative Content Analysis

Once the appropriate operational doctrine is selected, summative content analysis is used to find which words and phrases are used most frequently. As mentioned earlier in the chapter, summative content analysis assumes that the words used with the most frequency will reflect the areas of greatest concern. In this way, operational doctrine will provide the most insight into the operational aspects of the core missions. Summative content analysis is used in lieu of convention content analysis because of the number of documents and their substantial length.

Three different variations of summative content analysis are performed: (1) single word frequencies, (2) two-word phrase frequencies, and (3) three-word frequencies. In content analysis, phrases are referred to as *n*-grams and defined as a contiguous sequence of *n* items within a text. The bigrams (two-word phrases) and trigrams (three-word phrases) allow the researcher to investigate the context of single words used most frequently. For example, if the word "air" is found to be used most often, the researcher will be able to determine the context from bigrams and trigrams like "air force" or "air traffic control". In this example, the word "air" has very different meanings when put into context.

For all variations of summative content analysis, a statistical software called R Studio (R Core Team, 2016) is used. R Studio allows for very fast processing of large amounts of data. It is an open source coding language which uses packages – a collection of functions – to perform various types of analysis. For this research method, a package called "tm" (Feinerer & Hornik, 2015) will be used to determine single word and *n*-gram frequencies.

Textual Data Processing

Before word counts can be created, the data (AF annexes) are converted to a *.txt* file. Once they are saved as a *.txt* file, they are pulled into R Studio and saved as a corpus. A corpus is a collection of textual data. The corpus is then processed, or cleaned, before the analysis can begin. To clean the corpus, capital letters are made lowercase; punctuations, numbers, and whitespace are removed; common words such as "and", or "the" are removed; and some words are joined together intentionally. For example, the word "air" should always be joined together with "force" when they appear next to each other. A complete list of joined words can be found in Appendix C.

For the single word counts, the corpus will also be stemmed. Stemming a corpus is the act of removing endings such as "es", "tion", and "ly" so that R Studio returns an accurate count of words. Without stemming the corpus, R Studio would return individual counts for the words "operation" and "operations" when they should be counted as a single word. After stemming the corpus, R Studio will return a single count under the word "oper". For bigrams and trigrams the corpus will not be stemmed to preserve context.

Text Mining

Once the textual data has been processed, R Studio then creates a Term Document Matrix (TDM). A TDM is created for each variation of summative content analysis. The first column of a TDM contains every single word or *n*-gram. The subsequent columns belong to each annex and contains the frequencies of the words or *n*-grams found in column one. While a TDM is useful for deeper investigations using the "tm" package, this research transforms each TDM into a data frame for easier manipulation. In total, the

summative content analysis produces three different TDMs that are transformed into data frames: one for single words, one for bigrams, and one for trigrams.

Each data frame is then split into its individual documents and sorted in decreasing order, from most frequent to least frequent, in order to show the words and phrases with the highest frequency. In keeping with the assumption that the words with the highest frequency will reflect the areas of greatest concern, the lists are trimmed to display the top 100 words, top 100 bigrams, and top 100 trigrams for each annex. Finally, each list of words, bigrams, and trigrams for each annex are exported to Excel. The analysis produces a number of lists 3 times the number of annexes that are chosen for analysis.

Selecting Keywords

The three lists for each annex are joined into a single Excel document for quick reference. This produces in a single Excel document for each annex containing single words, bigrams, and trigrams. Words and phrases are then selected from the comprehensive lists based on their relevance to the AF core missions. Codes and categories created in the conventional content analysis assists in determining which words to keep and which to discard. Single words are examined first using the *n*-grams for context. Individual words are chosen as much as possible; however, in any instant where a single word does not provide enough context, a bigram is used. A bigram is also used when a single word has more than one meaning, such as the word "space". The selected words – designated “keywords” – are input into the model for identifying core infrastructure groupings.

Model Build

The basis of the identification model consists of real property data containing (1) all the unique CATCODEs used by the AF, (2) its respective CATCODE title and description (if one exists), and (3) its corresponding FAC code title and description. The keywords will then be added to the model. While each annex may have some keywords that they share with another, a keyword is only entered once. Additionally, the stemmed word should be used whenever possible so as to capture all uses of the keyword within the descriptions of real property data. The following sections describe how the model identifies CATCODEs that are related to core missions and how infrastructure groupings are calculated and considered.

Identifying CATCODEs that Contain Keywords

After each word or phrase is entered into the model, a search of the real property data is performed. The model searches each FAC and CATCODE description for keywords and return a binary "TRUE" or "FALSE" for each one. In the case of a bigram, the model searches for the words in succession. For example, in the case of the phrase "air defense", the model only returns a "TRUE" if the word "air" precedes the word "defense". Once the analysis is complete, the model produces a list of all the CATCODEs that contain a keyword in either the FAC or CATCODE description.

Defining Infrastructure Groupings

This model could be used to identify individual CATCODEs related to core missions. However, this method assumes that some CATCODEs will contain a keyword out of context. Therefore, to eliminate such false positives, the method calculates the percentages within a smaller classification. The assumption in doing so is that the

majority of CATCODEs containing a keyword will, in fact, be related to the core missions. By looking for a majority at a higher categorical level, such as Facility Classes, Category Groups, or Basic Categories, the false positives are nullified. Because of the granularity and count of the different categorical levels, Basic Categories (collection of three digits) is chosen as the infrastructure groupings which will be designated core or non-core. If a Basic Category is reported as having a majority (75% or more), all the CATCODEs within that Basic Category are then assumed to be related to core missions.

As a simple example, the word “nuclear” might be used as a keyword in the model. While this might seem like a word that could only be related to core missions, there is a description within real property data that uses the word out of context. The CATCODE description refers to a nuclear fallout shelter. This is an excellent example because while the word “nuclear” is undoubtedly related to the core missions, a nuclear fallout shelter does not directly contribute to a core mission.

Model Verification

To verify that the model is flagging infrastructure assets that are related to core missions, the missions dependency index (MDI) will be used and compare the resulting list of core to non-core. As mentioned in the Literature Review chapter, the MDI is not without its faults. However, many of the infrastructure assets in the AF that are obviously related to AF missions do have a high MDI score. As an example, a runway has an MDI of 99 while a reviewing stand has an MDI of 39. The opposite is also true in that some facilities that are obviously support have a high MDI. One example might be unaccompanied enlisted dormitories, which has an MDI of 90. The researcher

hypothesizes that the majority of the infrastructure selected by the model as core will contain a high MDI and, therefore, a subsequent higher average when compared to those left as non-core.

It is for this reason that MDI will be used as a way to verify that the model works by comparing the mean MDI between what the model deems as core infrastructure and what is left as non-core infrastructure. If the means of the Basic Categories selected as core are higher than the Basic Categories left as non-core, the model is assumed to be working as the researcher intended. Anomalies in the data will be investigated in an attempt to explain them.

Summary

The methodology presented in this chapter describes how a model, which identifies infrastructure related to core missions, is developed. The different forms of data are discussed as well as the analysis techniques. Conventional content analysis and summative content analysis are used to turn textual data into usable data pertaining to the AF core missions. The model, which uses real property data, identifies CATCODEs that should be designated as core. The model produces the results by Basic Category in an effort to eliminate false positives. The final results are a list of Basic Categories which are considered core and a list considered non-core. This method uses the MDI to verify that the model is working as the research intends. Such a model allows the AF, for the first time, to objectively differentiate between mission and support infrastructure at the Basic Category level.

IV. Analysis and Results

Chapter Overview

This chapter is meant to describe the analysis process and the results. The first section gives a brief description of the results before moving on to specific implementation. The second section describes the data collection effort in terms of a linear process. It then transitions to the method implementation phase which includes: conventional content analysis of strategic documents, summative content analysis of operational doctrine, and the building the model. The fourth section goes deeper to analyze the results while the fifth explains the model verification process.

Brief Description of Results

The model built for this research is meant to identify infrastructure groupings at the Basic Category level that are directly related to Air Force (AF) core missions. The model identifies 34 of the 161 Basic Categories as being related to core missions. The 34 Basic Categories cover 243 of the 919 CATCODEs used by the Air Force. The model is further verified by comparing the mean MDI's between the Basic Categories selected as core to those left as non-core. The means are found to be statistically different, with those selected by the model having a higher mean MDI than those left as non-core. The verification process shows that the model works as the research method intends.

Data Collection

Building a model to identify core infrastructure assets that are related to AF core missions requires three parts. The first part are the strategic documents required to understand the AF core missions. The second are operational doctrine which provide

keywords that are used in the model. The last piece is a combination of AF real property data and DoD real property data. The relevant AF real property data contains CATCODEs, titles, and descriptions and the DoD real property data contains FAC code titles and descriptions. The two data sets are easily merged because CATCODEs and FAC codes are related and the relationship is found in DoD real property data. The model combines this data and provides an output of infrastructure groupings that directly support AF core missions. This linear process can be seen in Figure 5.

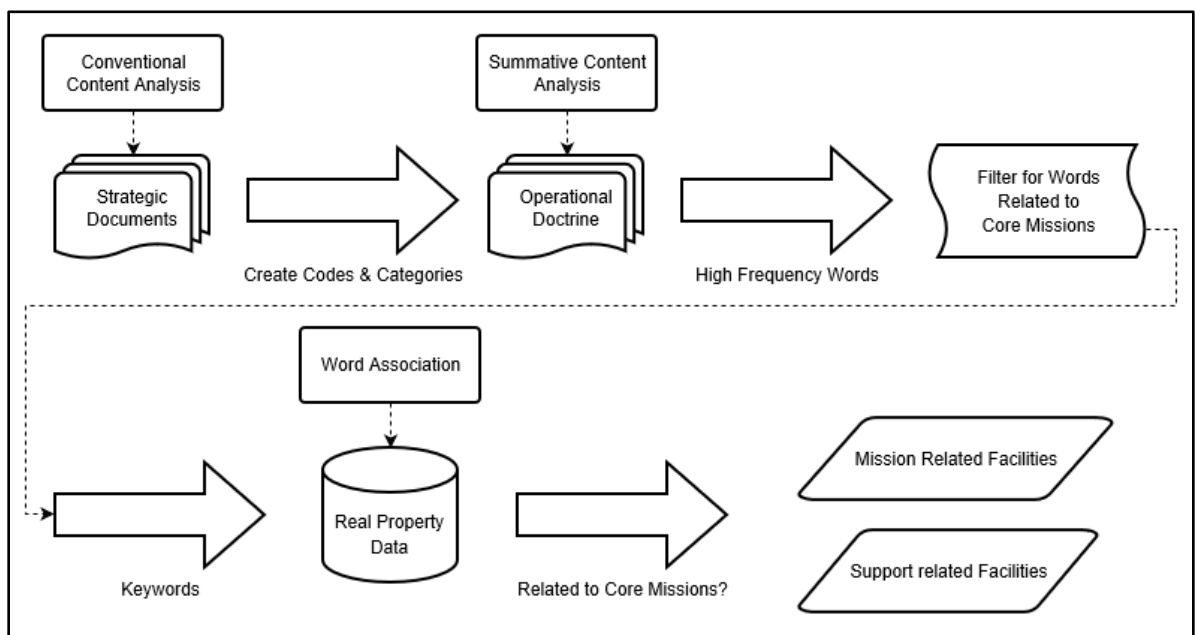


Figure 5: Linear Method Flowchart.

The first piece of data required are strategic documents that describe core mission. The AF has a strategic document set which is comprised of seven documents. All of them were written within the last three years and each one expounds upon a different focus area. The seven documents are: (1) CSAF Focus Area: Strengthening Joint Leaders and Teams; (2) CSAF Focus Area: The Beating Heart of the Air Force... Squadrons!; (3)

Global Vigilance, Global Reach, Global Power for America; (4) The World's Greatest Air Force: Powered by Airmen, Fueled by Innovation; (5) America's Air Force: A Call to the Future; (6) USAF Strategic Master Plan; and (7) Air Force Future Operating Concept.

Not every document is related to AF core missions and so, therefore, not every document is used in the conventional content analysis. The documents that did mention, or have something to do with, core missions are (1) the USAF Strategic Master Plan and (2) Global Vigilance, Global Reach, Global Power for America. An additional document related to Global Vigilance, Global Reach, Global Power for America written as a Senior Leader Perspective from Gen Mark Welsh III is also used. These three documents make up the strategic set utilized to create codes and categories.

The second piece of data used to construct the model are operational doctrine. In total, there are 29 doctrinal annexes which are examined for relevancy to the core missions. After examining their content, 12 were found to be relevant and these are selected for use in the summative content analysis. For use in R Studio, they are converted to *.txt* files and saved as separate documents. From this data, keywords are pulled for use in the model.

The final piece of the model is a combination of AF real property data and DoD real property data. The required subset of AF real property data are a complete list of AF CATCODEs, titles, and descriptions. The required subset of DoD real property data are FAC code titles and descriptions. Both subsets are then combined based on the DoD-defined relationship between FAC codes and CATCODEs. Such a subset of data proved difficult to attain at first. The research combined them from three different sources. A complete list of CATCODEs and FAC code titles and descriptions are found in the FY16

Real Property Categorization System. It contained 919 unique CATCODEs and each one was paired by the DoD to a corresponding FAC code. The second source came from AF Manual (AFMAN) 32-1084. The AFMAN intent is to describe how the AF classifies its real property. The manual contained a complete list of CATCODEs, titles, and descriptions; however, not all CATCODEs had a description. In total, 313 CATCODEs contained a description while 606 did not.

Method Implementation: Conventional Content Analysis

Now that the data has been adequately collected, the following sections describe the conventional content analysis of strategic documents. The first section describes the coding process. It also includes a few examples of how codes were developed which relate to each core mission. The second section describes how codes are turned into categories and concludes with a discussion on how the codes and categories are used throughout the research method. Keep in mind that the goal of developing codes and categories is to gain an objective understanding of the AF core missions.

Coding Strategic Documents

Coding requires the researcher to become immersed in the textual data. The strategic documents are read several times, and very carefully, before beginning the coding process. This ensured that the researcher understood what was being written while also allowing the researcher to focus on the sections of the documents that would provide the most value and insight into AF core missions.

The USAF Strategic Master Plan contains a lot of information. There is some difficulty in determining which sections are related to core missions. Digging deeper,

one finds that bits and pieces related to core missions are sprinkled throughout the document. These pieces are coded according to the proposed research method. Words and phrases are examined closely and chosen only when it relates to a particular core mission. Context is always taken into consideration and preserved as much as possible through the inclusion of additional words or phrases.

Global Vigilance, Global Reach, Global Power for America proves an easier document to code because it is exclusively about the core missions. Although brief, it is laid out in such a way that each core mission has its own section. This makes the coding process much less complicated. Words and phrases are examined closely and chosen only as it relates to a particular core mission. Context is, again, always taken into consideration and is preserved through the inclusion of additional words or phrases as much as possible.

The Senior Leader Perspective is laid out similarly to Global Vigilance, Global Reach, Global Power for America in that each core mission had its own section. Each section was not nearly as long or as detailed the Global Vigilance, Global Reach, Global Power for America sections but they proved to be equally as useful. Words and phrases were examined closely and chosen only as it related to a particular core mission. Context was always taken into consideration and preserved as much as possible through the inclusion of additional words or phrases.

All the codes are combined and separated by core mission. This results in a comprehensive list of codes for each core mission. A sample has been provided in Table 4. In total there are 226 codes created across the three documents for the five AF core missions. For Air and Space Superiority there are 52 codes. For Command and Control

there are 33 codes. For Intelligence, Surveillance, and Reconnaissance there are 47 codes. For Global Mobility there are 53 codes. Finally, for Global Strike there are 41 codes.

Table 4: Sample of Codes for Global Strike Core Mission

Codes for Global Strike		
Deterrence	Promptly (the ability to strike)	Fighters (planes)
Nuclear	Air support (for troops)	Aircraft
Conventional (weapons)	Inserting special operations forces	Modernizing
Precision	Targeting	Conventional (weapons)
Strike	Bombers	Disabling or destroying
Target	Missiles	Enabled by in-flight refueling
Disabling or destroying	Special operations platforms	Updating (technology)

Categories

Once codes are developed, the next step in conventional content analysis is to create categories for the many codes. This is done by examining the words and phrases in context. In a more traditional qualitative analysis, the goal of categories is to help the researcher discover a theme. In this case the themes are the core missions and the codes and categories are developed to reflect each particular theme, or core mission. The goal is to let the documents describe the themes through codes and categories. In this way personal bias and ignorance is removed in explaining core missions.

After reviewing the list several times, categories began to emerge. The first category to emerge was the category of capability. Each core mission had various codes which represented a capability. The category of capability needed to be specific to each core mission and therefore, each core mission contained its own capability category (e.g. ISR capability). The second category to emerge was the category of enabler. This includes the various types of equipment, infrastructure, or tools that are needed to support

a core mission. In the same way, each core mission contained its own enabler category. The third category to emerge was the category of domain. Each core mission operates within a domain, and sometimes multiple domains and the codes reflected these purviews. The final category to emerge was the category of partners. There are several codes that cannot fit into the other categories. These codes all center around the different agencies or personnel that are involved, such as "coalition forces" or "joint".

The codes and there categories are used in selecting operational doctrine to include in the summative content analysis. They are also used as a way of filtering high frequency words and phrases from the summative content analysis. A Sample is provided in Table 5.

Table 5: Example of Codes and Categories.

Codes	Axial Codes	Categories
Deterrence	<i>Product or goal</i>	Global Strike Capability
Nuclear	<i>Tool</i>	Global Strike Enabler
Conventional (weapons)	<i>Tool</i>	Global Strike Enabler
Precision	<i>Goal</i>	Global Strike Capability
Strike	<i>Product or goal</i>	Global Strike Capability
Target	<i>Tool</i>	Global Strike Enabler
Disabling or destroying	<i>Product or goal</i>	Global Strike Capability

Method Implementation: Summative Content Analysis

This section is the second step in the method implementation. It covers the summative content analysis of operational doctrine. The first section covers how operational doctrine is selected. The second section discusses the analysis process, to include data processing through the summative content analysis outputs. The last section describes how keywords were filtered out of the data for use in the model.

Choosing Operational Doctrine

The next step in the research method is to select the different doctrinal annexes to include in the summative content analysis. In total, there are 29 doctrinal annexes within the Curtis E. Lemay Center for Doctrine Development and Education. Each one is reviewed carefully to determine if it should be included, using categories created in the conventional content analysis. If the annex refers to a core mission *capability*, *enabler*, or *domain* it is selected. The final tally includes 12 annexes which can be seen in Table 6.

Table 6: Selected Operational Doctrine.

Doctrine Annexes Related to Core Missions	
Air Mobility Ops	Counterair Ops
Space Ops	Counterland Ops
Global Integrated ISR Ops	Special Ops
Cyberspace Ops	Command and Control
Electronic Warfare Ops	Nuclear Ops
Airspace Control	Targeting

Analysis of Operational Doctrine

Summative content analysis is performed in R Studio using the selected operational doctrine. The data is processed according to the steps laid out in the Methodology chapter. As also referenced in the Methodology chapter, some words are joined intentionally when it is known that the words should be kept together to preserve meaning. A complete list of these joined words can be seen in Table 7.

Once the data had been cleaned, R Studio can return accurate frequencies of single words and *n*-grams. Each corpus is turned into a term document matrix (TDM) which contained a list of all the words, or *n*-grams, used across all documents and their

respective frequency by document. Three total TDMs are created: one for single words, one for bigram (two-word phrases), and one for trigram (three-word phrases). R Studio finds over 6,000 single words across the 12 annexes. Not all words are used in all documents. As stated in the Methodology chapter, it is assumed that only those words, or *n*-grams, with high frequencies will reveal the areas of greatest concern.

Table 7: Complete List of Intentionally Joined Words.

Phrase	Joined Word
air force	<i>air_force</i>
command and control	<i>command_and_control</i>
airspace control	<i>airspace_control</i>
strategic attack	<i>strategic_attack</i>
combat support	<i>combat_support</i>
counterair operations	<i>counterair_ops</i>
counterland operations	<i>counterland_ops</i>
cyberspace operations	<i>cyber_ops</i>
electronic warfare	<i>elec_warfare</i>
information operations	<i>info_ops</i>
global integrated intelligence surveillance reconnaissance operations	<i>global_isr_ops</i>
air mobility operations	<i>mobility_ops</i>
nuclear operations	<i>nuclear_ops</i>
space operations	<i>space_ops</i>
special operations	<i>spec_ops</i>
host nation	<i>host_nation</i>

Each TDM is then split into individual annexes and sorted to reveal the top used words and phrases from each annex. The lists is trimmed to only show the top 100 used words and phrases, eliminate words with low counts and presumed irrelevance. This leaves three lists for each annex, a total of 36. Each list is exported from R Studio into Excel (2013) for easy manipulation. A single Excel file is then created for each annex

containing: single words and their counts, bigrams and their counts, and trigrams and their counts. The next section describes how keywords are filtered from these files for use in the model. The programming lines of code used for summative content analysis in R Studio can also be found in Appendix D.

Selecting Keywords

With a complete list of words and phrases for each annex, the method transitions to selecting keywords. Keywords are defined as words with high frequencies that fit under the categories created during the conventional content analysis. Frequency of use is only the starting point for identifying words of potential interest. By using the categories as a measuring stick for each word or phrase, it is ensured that keywords will effectively represented the AF core missions.

For the selection process, single word counts are examined first. If a word fits into a category, bigrams and trigrams are then used to determine context. If the word did not fit into a category based on its context, it is discarded. If the word did fit, the researcher had to determine how to include it as a keyword. In some cases this is straightforward. For example, the word "airlift" would garner inclusion and is exclusive. It does not need another word to preserve context. As another example, the word "air" is found as a word of high frequency but lacks context. However, when joined with another word to create "air defense", a bigram found with high frequency, it becomes meaningful and specific. In cases such as these the bigram is selected as a keyword.

Of the 3600 words and phrases garnering examination, 142 are selected as keywords relating to core missions. Some are duplicates, keywords existing in multiple annexes. For example, many contained the words "command" and "control". For the

sake of the model, the keywords were only included once. After eliminating duplicates, 63 are considered unique. These 63 unique keywords are then included in the model. A complete list of the 63 keywords can be seen in Table 8.

Table 8: Complete List of Selected Keywords.

List of Key Words			
aircraft	information	intel	space control
control	guidance	network	space operation
air support	combat	frequenc	special operation
communicat	radar	satellit	air operation
air control	attack	reconn	coordinat
missile	air defense	technolog	task force
weapon	data	computer	counterair
air traffic control	command control	aeromedic	counterland
aviation	tanker	space system	air interdiction
electro	nuclear	airspace	cyber
launch	planning	data link	disrupt
target	surveil	mobility	jam
airfield	aerial port	sensor	orbit
runway	combat operation	airlift	space capabilit
command	refuel	mass destruction	missile defense
missile tracking	special force	air tasking	

Method Implementation: Model Assembly

The model for identifying core infrastructure contains two parts: (1) real property data, and (2) keywords. Once both are collected, the model assembles them in Excel (2013). While this research method was performed in Excel, a final deliverable model is available through Shiny (Chang, Cheng, Allaire, Xie, & McPherson, 2016), a web application written in R Studio. The model performs a word association by searching through the infrastructure asset descriptions to determine which ones contain keywords. It flags CATCODEs that contain at least one keyword. The model then provides the

results of the analysis by displaying the Basic Categories which contain a majority ($\geq 75\%$) of flagged CATCODEs. All the CATCODEs within these Basic Categories are then assumed to be related to the AF core missions.

Model Results

Out of the 161 Basic Category groupings, 34 receive a score higher than 75%. All the CATCODEs within those Basic Categories are considered to be directly related to the AF core missions. This equates to 21% of the 162 unique Basic Categories. The 34 Basic Categories that were considered core infrastructure assets cover 243 of the 919 CATCODEs, equating to 26%. A list of the 34 Basic Categories considered core and their titles can be seen in Table 9.

Of the 63 keywords used in the model, 48 are found in either the titles or descriptions. That means 15 are not found in any titles or descriptions and added no value to the modeling process. A complete list of the words which were found and the number of times they were found can be seen in Appendix E. In total, keywords are found in the real property data 811 times. The words which are found the most are "aircraft", "control", and "air support" with 116, 90, and 68 flags respectively.

Table 9: Complete List of Basic Categories Found to be Related to Core Missions.

Code	Title
111	Airfield Runways
112	Airfield Taxiways
113	Airfield Aprons
116	Other Airfield Pavements
121	Aircraft Fuel Dispensing Facilities
123	Land Vehicle Fuel Dispensing Facilities
131	Communications Buildings
133	Aviation Navigation and Traffic Aids Buildings
134	Aviation Navigation and Traffic Aids Facilities Other Than Buildings
135	Communication Lines
136	Airfield Pavement Lighting
140	Ground Operational Buildings
141	Airfield Operational Buildings
145	Ground Operational Facilities Other Than Buildings
146	Airfield Operational Facilities Other Than Buildings
163	Moorings
172	Simulation Facilities
174	Impact, Maneuver, and Training Areas
176	Weapons Ranges
177	Team and Unit Ranges
211	Aircraft Maintenance Facilities
212	Guided Missile Maintenance Facilities
215	Weapons and Spares Maintenance Facilities
216	Ammunition, Explosives, and Toxics Maintenance Facilities
217	Electronics and Communications Equipment Maintenance Facilities
221	Aircraft Production Facilities
222	Guided Missile Production Facilities
227	Electronics and Communications Equipment Production Facilities
311	Aircraft Research, Development, Test and Evaluation Facilities
312	Missile and Space Research, Development, Test and Evaluation Facilities
315	Weapons and Weapon Systems Research, Development, Test and Evaluation Facilities
317	Electronic and Communication Equipment Research, Development, Test and Evaluation Facilities
371	Research, Development, Test and Evaluation Range Facilities
690	Administrative Structures Other Than Buildings

Model Verification

The model verification process, as outlined in the Methodology chapter, compares the means of the missions support index (MDI) as a way of verifying that the model works as intended. It is found that the difference in the means of core groupings compared to non-core groupings are statistically different. Additionally, the mean MDI of the Basic Categories selected as core are higher than the mean MDI of the Basic Categories left as non-core. A graph of the results can be seen in Figure 6. This analysis suggests that the model works as intended by the research method.

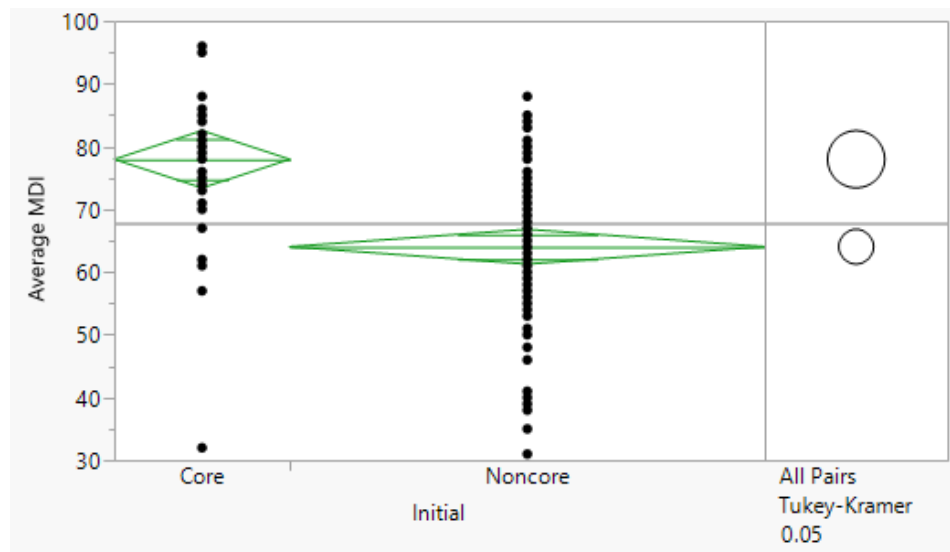


Figure 6: Statistical Analysis Comparing Mean MDI's of Core and Non-core

There are some anomalies that show up during the verification process. An outlier is found in the core infrastructure classification and can be seen clearly in Figure 6. This outlier belongs to the Basic Category 690, *Administrative Structures Other Than Buildings*. Upon further investigation it was found that the unique CATCODEs, while flagged for containing keywords, clearly did not relate to any of the AF core missions.

This Basic Category is an anomaly that contains a low number of CATCODEs, each a false positives.

The highest average MDI within the non-core classification is also investigated to determine if it should have warranted inclusion as core infrastructure. The highest Basic Category within the non-core classification belongs to 155, *Small Craft Berthing*. Upon further investigation it was found that this Basic Category contained one unique AF CATCODE and refers to piers or wharfs where small boats are parked. This Basic Category does not appear relate to any AF core missions and therefore does not garner inclusion. This serves as further confirming that the model works as intended. Other top average MDI's not included as core can be seen in Table 10.

Table 10: Basic Categories Not Selected as Core with High Average MDIs.

Code	Title	Average MDI
155	Small Craft Berthing	88
125	POL Pipelines	85
126	Other Liquid Fuel and Dispensing Facilities	85
151	Piers and Wharfs	85
310	Research, Development, Test and Evaluation Science Facilities	84
316	Ammunition, Explosives, And Toxics Research, Development, Test and Evaluation Facilities	84
318	Propulsion Research, Development, Test and Evaluation Facilities	84
321	Technical Services Research, Development, Test and Evaluation Facilities	84
124	Operating Fuel Storage Facilities	83
149	Operational Support Facilities Other Than Buildings	81
319	Miscellaneous Item and Equipment Research, Development, Test and Evaluation Facilities	81
721	Enlisted Unaccompanied Personnel Housing	81
154	Sea Walls, Bulkheads, and Quay Walls	80

Summary

The model built to identify infrastructure assets related to AF core missions identified 34 Basic Categories containing 243 of the 919 unique CATCODEs. The model

is successful in separating infrastructure assets into core and non-core and, therefore, is a first step in distinguishing between mission and support facilities. While this research method chose keywords based on qualitative techniques, other methods or selecting keywords can be used. An application which performs the word association is available for other research. Additionally, it can be used by senior leaders to further investigate CATCODEs and their relation to the AF core missions.

V. Conclusions and Recommendations

Chapter Overview

The model developed in this research provides results which identify groupings of infrastructure assets that directly relate to Air Force core missions. This chapter is meant to be a critical analysis of the results and a discussion about several relevant topics that stem from the research. It also includes a discussion on the implications of such results. The first section provides a summary of the results and a discussion about the study. The next section covers the limitations followed by the significance of the study. Finally, recommendations and future research are presented to continue in building upon the work presented in this paper.

Conclusions of Research

The model produced results showing 34 out of the 161 Basic Categories as being directly related to the AF core missions of air and space superiority; intelligence, surveillance, and reconnaissance; rapid global mobility; global strike; and command and control. The model suggests that only 21% of the Basic Categories directly support core missions. The 34 Basic Categories cover 243 unique CATCODEs out of the 919 used by the AF, approximately 26%. The number of unique CATCODEs does not equate directly to the number of facilities and utilities using those CATCODEs. Multiple facilities and utilities can share the same CATCODE.

This research effort was meant to be a starting point for identifying core infrastructure - facilities and utilities - directly related to AF core missions. While many facility descriptions may be easy to distinguish between mission and support, some are

more difficult. The model was meant to identify groups of infrastructure in an objective manner void of personal bias. The model performed as intended and was verified using MDI.

During model verification, it was found that a high MDI does not always correlate to core missions. Medical facilities are an excellent example of this phenomenon. While they have a high MDI and are very important to the health and welfare of Airmen, they were found not to be directly related to the core missions. Model verification also showed that the model was not perfect. For a complete and accurate list, each infrastructure group should be reviewed carefully to ensure that no Basic Category was selected that clearly does not belong. This should be done as objectively as possible since the point of this model was to arrive at groupings without personal biases.

The researcher recognizes that there can be many emotions involved when talking to commanders about their organizational facilities relevance to mission. While many arguments can be made for certain groups, care should always be taken to ensure a class of infrastructure designated as essential to core missions remains true to the intent laid out in the Strategic Master Plan. If the class becomes overly inclusive it will cease to serve the purpose of differentiating infrastructure as support or mission essential.

Discussion

This discussion section is meant to go into more detail about the designation of core and non-core. There are several things that are worth noting that stem from the qualitative analysis of both the strategic documents and the operational doctrine. All of them are discussed in this first section. The second section discusses the researcher's

thoughts on the method and results. What follows is meant to add additional insight into how the results of the modeling process should be viewed.

Defining Core and Non-core

This research creates two distinct classifications of infrastructure. The first is designated "*core*" and contains infrastructure directly related to the AF core missions. The second, designated "*non-core*", contains all infrastructure that does not directly relate to the core missions. This section discusses the characteristics of the classifications and what the researcher intends for them to mean for the AF.

The Strategic Master Plan discusses the need to distinguish between mission infrastructure (core) and support infrastructure (non-core) for strategic purposes. The distinction is necessary to give senior AF leaders the ability to make better decisions related to infrastructure maintenance, consolidation, demolition, and new construction. While core infrastructure is important to the AF core missions, non-core infrastructure remains important in its own right.

If core infrastructure can be said to directly support core missions, non-core can be said to indirectly support those same missions. Many core facilities and utilities could not exist or function without its non-core partners. The non-core designation does not mean that the AF can universally get rid of or stop funding them. Many of them directly support Airmen. As an example, one of the capabilities of global mobility is aeromedical evacuations. While medical facilities, like clinics, do not directly take part in transporting troops (a core mission of the AF), they do care for them upon arrival.

AF senior leaders universally understand that Airmen are key, and that taking care of those Airmen will always be an important part of the AF. This is an important

characteristic to keep in mind when talking about core and non-core infrastructure. The AF should continue to fund non-core infrastructure in support of Airmen and many other functional areas of the AF.

The intent of the designation is to help senior leaders make difficult funding decisions related to asset management in this fiscally uncertain environment. It is not intended to suggest that funding to a particular group should increase or cease. Condition of an asset should continue to be a driver in determining which pieces of infrastructure need maintenance and repair.

Thoughts on the Results

During the verification stage, the researcher noticed several facilities or utilities that were designated as core but had very low MDI's. Upon further investigation the researcher discovered that some of these belonged to Basic Categories that could easily be dismissed as support infrastructure. One such infrastructure group was 690, the Basic Category of *Administrative Structures other than Buildings*. The FAC descriptions associated with this set of infrastructure clearly does not relate to AF core missions, even though they were flagged as containing keywords. This is an example of a false positive, a basic Category that the model selected but clearly does not relate to the core missions. The results by Facility Class can be seen in Table 11.

Table 11: Results Shown by Facility Class.

Code	Title	Basic Categories	Core Basic Categories	Percentage
1	Operation & Training	44	20	45%
2	Maintenance & Production	18	8	44%
3	Research, Development, Test, and Evaluation	14	5	36%
4	Supply	13	0	0%
5	Hospital & Medical	4	0	0%
6	Administrative	3	1	33%
7	Housing & Community	26	0	0%
8	Utility & Ground Improvements	32	0	0%
9	Land	7	0	0%

The current model assumes that a Basic Category is considered core when a majority of the unique CATCODEs within it contains at least one keyword. For the research, a majority was defined as 75%. This score can be adjusted based on leadership and their belief about what percentage a Basic Category should have to be considered core. A list of Basic Categories that missed the majority but were higher than 50% can be seen in Appendix F.

The model used in this research utilized only words found through summative content analysis. During the conventional content analysis, the researcher noticed two sub-themes that were absent from the summative content analysis. These two themes are (1) the importance of training Airmen and (2) the importance of developing advanced technology. Multiple documents (USAF, 2014, 2015a, 2015c, 2015d, 2015e) referenced these things directly and, therefore, it might be useful to include such associated facilities as mission.

Limitations of the Research

This research effort does have some limitations. The first is that of the documents chosen for summative content analysis. They all appear to be written by different sources

and vary substantially in length and depth of information. If an important topic was only covered briefly within one of them, it may not have shown up during the summative content analysis.

The second limitation is the descriptions associated to each CATCODE. Some are more descriptive than others. More detail is better for a model like this. While every FAC code had an official description, only 313 CATCODEs contained an AF description leaving 606 without one. If the AF had descriptions for these CATCODEs the model output might be different.

The third limitation is strategic value of the documents used in the research. The documents used for this study were meant for public consumption and therefore remained fairly generic. They were all considered operational doctrine and they contained almost no tactical information. While such information resides in the AF's tactics, techniques, and procedures (TTP) manuals, they are classified.

Another limitation is the context of keywords found in CATCODE and FAC descriptions. The model did not investigate the context of keywords found in the descriptions to verify it met the intent. Keywords were selected in such a way as to avoid contextual errors as much as possible. For example, the word "space" was not used even though many of the documents referred to it in the context of outer space. It was assumed that very few descriptions would contain keywords out of context and that they would not affect the selection of a substantial amount of Basic Categories as core.

Significance of the Research

The significance of this research lies in the proof of concept of the model. The literature review showed that no method existed for identifying infrastructure related to organizational objectives which lead to the creation of a model to do so. The research verified that the model developed in the research is useful in identifying facilities that are directly related to the AF core missions. The AF can take the model and input keywords related to core missions to formulate a list of mission critical facilities. Further, the model can help the AF create the distinction between mission and support infrastructure as called for in the Strategic Mater Plan.

The classifications of core and non-core created by this model can be used to help leaders make better, strategic decisions. Currently, the AF's Integrated Priority List (IPL) compares all projects across the enterprise to one another. Other than points given out arbitrarily by AF Major Commands, there is no way to determine the strategic value of a project. This model and the classifications it produces can help senior leaders make difficult funding decisions in this fiscally uncertain environment.

While the model is intended to identify Basic Categories related to AF core missions, it could be used to investigate which unique CATCODEs are related to core missions. Since the model uses keywords to identify CATECODEs, some of this work has already been done. Considerable work would still be required in reviewing each flagged description to determine if it in fact did relate to core missions. This is because some of the keywords may appear in a description in a different context than intended.

Recommendations for Action

The AF should develop its own list of keywords (related to core missions) and use them in the model to definitively determine which infrastructure assets are directly related to AF organizational objectives. In this way the AF would have the ability to objectively identify a particular facility as mission or support, thus complying with the Strategic Master Plan. In short, the AF should use the model – developed and verified through this research – to separate infrastructure assets into core and non-core groups

While the AF has taken considerable steps to improve the way it does asset management, it should continue to evolve and be perfected as additional information and data becomes available. The designation of core or non-core is additional information that is currently not available and, therefore, should be included within the asset management program. The point of the consolidated IPL was to get rid of multiple project lists and give leadership the ability to compare projects across the AF on one list. While this is smart conceptually, the current way the AF measures importance (MDI) does always work as intended; nor does it quantify risk related to the core missions. The list, therefore, does not always provide accurate comparisons.

The researcher believes the additional information provided by the model allows for the creation of two lists. The first list should contain only infrastructure that directly supports core missions, and the second for non-core infrastructure. In this way leadership would no longer be equating a dormitory to a space launch facility as the current method does. Moreover, the two lists can be sorted, creating accurate comparisons between like classifications (core and non-core). This enables accurate comparisons, such as an ISR facility to a space launch facility and a dining facility to a dormitory. Accurate

comparisons of mission to mission and support to support results in a higher degree of confidence that the right decisions are made.

As referenced above in the Discussion section, the researcher is not suggesting that funding within a particular classification should change. The two lists created should receive appropriately proportioned funding as dictated by senior leaders. This ensures that the AF is successful in carrying out its core missions while also taking care of Airmen and other AF capabilities and functions.

Recommendations for Future Research

This research effort utilized the qualitative tools of conventional and summative content analysis. Another study could be done using the same methodology but with different documents. While the researcher identified Core Function Support Plans (CFSPs) as a potential candidate, they are classified. They remain an excellent source with one CFSP for each AF core mission. Still another study could use different qualitative techniques. A survey could be developed and sent to senior leaders queering them for keywords, which then could be used in the model. Both could serve to validate the results found in this study.

Another study could be done to identify which core infrastructure belongs to each AF core mission. The distinction would give senior leaders the ability to compare the condition between the core missions. Such an effort might show that a particular core mission requires more attention than another. Strategic decisions related to funding would contain less uncertainty.

While the point of this study was to identify facilities and utilities, no utility groups were identified in the model. While facilities are important, many cannot function without utilities. A future study could use linear segmentation to identify critical utility sections which connect to core facilities.

In a similar fashion, the researcher found that training is an important part of sustaining core missions. While there are many training facilities in the AF, not all are related to core missions. Another researcher could investigate the different training facilities and identify those that directly support core missions.

This model was intended to identify Basic Categories related to AF core missions. Another research effort could go deeper to investigate which unique CATCODEs are related to core missions. Since the model uses keywords to identify CATCODEs, some of this work has already been done. Additional steps would be required to ensure that keywords found in CATCODE descriptions were used in context so as to not include a CATCODE that does not relate to core missions. A researcher may consider a different method all together which would provide the desired results.

Summary

This research succeeded in identify infrastructure groups directly related to the AF core missions of air and space superiority; intelligence, surveillance, and reconnaissance; rapid global mobility; global strike; and command and control. The model produced results showing which Basic Categories should be considered core infrastructure. The classifications can be used to help senior leaders make better-informed decisions related to infrastructure maintenance, consolidation, demolition, and

construction. Additionally, the study lays the foundation for future research and model development. Building on these methods, the Air Force can better understand how its real property relates to core missions.

Appendix A: Category Groups

Code	Title
11	Airfield Pavements
12	Liquid Fueling and Dispensing Facilities
13	Communications, Navigation Aids and Airfield Lighting
14	Land Operational Facilities
15	Waterfront Operational Facilities
16	Harbor and Coastal Operational Facilities
17	Training Facilities
21	Maintenance Facilities
22	Production Facilities
31	RDT&E Buildings
37	RDT&E Range Facilities
39	RDT&E Facilities Other Than Buildings
41	Liquid Storage - Fuel and Nonpropellants
42	Ammunition Storage
43	Cold Storage
44	Covered Storage
45	Open Storage
51	Medical Centers and Hospitals
53	Medical and Medical Support Facilities
54	Dental Clinics
55	Dispensaries and Clinics
61	Administrative Buildings
62	Underground Administrative Structures
69	Administrative Structures Other Than Buildings
71	Family Housing
72	Unaccompanied Personnel Housing
73	Personnel Support and Services Facilities
74	Indoor Morale, Welfare, and Recreation Facilities
75	Outdoor Morale, Welfare, and Recreation Facilities
76	Museums And Memorials
81	Electrical Power
82	Heat and Refrigeration (Air Conditioning)
83	Sewage and Waste
84	Water
85	Roads and Other Pavements
86	Railroad Facilities
87	Ground Improvement Structures
88	Water Resources & Environment
89	Miscellaneous Utilities
91	Land
92	Other Rights

Appendix B: Basic Categories

Code	Title
111	Airfield Runways
112	Airfield Taxiways
113	Airfield Aprons
116	Other Airfield Pavements
121	Aircraft Fuel Dispensing Facilities
122	Marine Fuel Dispensing Facilities
123	Land Vehicle Fuel Dispensing Facilities
124	Operating Fuel Storage Facilities
125	POL Pipelines
126	Other Liquid Fuel and Dispensing Facilities
131	Communications Buildings
132	Communications Facilities Other Than Buildings
133	Aviation Navigation and Traffic Aids Buildings
134	Aviation Navigation and Traffic Aids Facilities Other Than Buildings
135	Communication Lines
136	Airfield Pavement Lighting
140	Ground Operational Buildings
141	Airfield Operational Buildings
144	Operational Support Buildings
145	Ground Operational Facilities Other Than Buildings
146	Airfield Operational Facilities Other Than Buildings
149	Operational Support Facilities Other Than Buildings
151	Piers and Wharfs
154	Sea Walls, Bulkheads, and Quay Walls
159	Other Waterfront Operational Facilities
163	Moorings
164	Marine Improvements
171	Training Buildings
172	Simulation Facilities
173	Training Support Facilities
174	Impact, Maneuver, and Training Areas
175	Small Arms Ranges
176	Weapons Ranges
177	Team and Unit Ranges
178	Explosives and Flame Ranges
179	Training Facilities Other Than Buildings
211	Aircraft Maintenance Facilities
212	Guided Missile Maintenance Facilities
213	Ships and Spares Maintenance Facilities
214	Tank and Automotive Maintenance Facilities
215	Weapons and Spares Maintenance Facilities
216	Ammunition, Explosives, and Toxics Maintenance Facilities
217	Electronics and Communications Equipment Maintenance Facilities

218	Miscellaneous Items and Equipment Maintenance Facilities
219	Installation Repair and Operation Maintenance Facilities
221	Aircraft Production Facilities
222	Guided Missile Production Facilities
226	Ammunition, Explosives, and Toxics Production Facilities
227	Electronics and Communications Equipment Production Facilities
228	Miscellaneous Items and Equipment Production Facilities
229	Installation Maintenance and Repair Production Facilities
310	Research, Development, Test and Evaluation Science Facilities
311	Aircraft Research, Development, Test and Evaluation Facilities
312	Missile and Space Research, Development, Test and Evaluation Facilities
315	Weapons and Weapon Systems Research, Development, Test and Evaluation Facilities
316	Ammunition, Explosives, And Toxics Research, Development, Test and Evaluation Facilities
317	Electronic and Communication Equipment Research, Development, Test and Evaluation Facilities
318	Propulsion Research, Development, Test and Evaluation Facilities
319	Miscellaneous Item and Equipment Research, Development, Test and Evaluation Facilities
321	Technical Services Research, Development, Test and Evaluation Facilities
371	Research, Development, Test and Evaluation Range Facilities
390	Research, Development, Test and Evaluation Facilities Other Than Buildings
411	Bulk Liquid Fuel Storage
422	Installation and Ready Issue Ammunition Storage
423	Liquid Propellant Ammunition Storage
432	Installation and Ready Issue Cold Storage
441	Depot and Arsenal Covered Storage
442	Installation and Organization Covered Storage
451	Depot Open Storage
452	Installation and Organization Open Storage
510	Medical Centers and Hospitals
530	Medical and Medical Support Facilities
540	Dental Clinics
550	Dispensaries and Clinics
610	Administrative Buildings
690	Administrative Structures Other Than Buildings
711	Family Housing Dwellings
713	Family Housing Trailer Sites
714	Detached Family Housing Facilities
721	Enlisted Unaccompanied Personnel Housing
722	Unaccompanied Personnel Housing Mess Facilities
723	Detached Unaccompanied Personnel Housing Facilities
724	Officer Unaccompanied Personnel Housing Facilities
725	Emergency Unaccompanied Personnel Housing Facilities
731	Safety, Discipline, and Rehabilitation Facilities

738	Miscellaneous Personnel Support and Services Facilities
744	Miscellaneous Indoor Morale, Welfare and Recreation Facilities
751	Outdoor Recreation Facilities
754	Miscellaneous Outdoor Morale, Welfare and Recreation Facilities
760	Museums and Memorials
811	Electric Power Source
812	Electric Power Transmission and Distribution Lines
813	Electric Power Substations and Switching Stations
821	Heat Source
822	Heat Transmission and Distribution Lines
823	Heat Gas Source
824	Heat Gas Transmission
826	Refrigeration (Air Conditioning) Source
827	Chilled Water (Air Conditioning) Transmission and Distribution
831	Sewage and Industrial Waste Treatment and Disposal
832	Sewage and Industrial Waste Collection Lines
833	Refuse and Garbage Facilities
841	Potable Water Supply, Treatment, and Storage
842	Potable Water Distribution System
843	Fire Protection Water Facilities
844	Nonpotable Water Supply and Storage
845	Nonpotable Water Distribution System
851	Roads, Bridges, and Tunnels
852	Sidewalks and Other Pavements
860	Railroad Tracks
871	Grounds Drainage
872	Grounds Fencing, Gates, and Guard Towers
891	Miscellaneous Utilities - Square Feet
892	Miscellaneous Utilities - Each
911	Land Purchase, Condemnation, Donation, or Transfer
912	Public Domain Withdraw
913	License or Permit
914	Public Land of Territories and Possessions
921	Easements
922	In-Lease
923	Foreign Rights

Appendix C: Complete List of Intentionally Joined Words

Phrase	Joined Word
air force	<i>air_force</i>
command and control	<i>command_and_control</i>
airspace control	<i>airspace_control</i>
strategic attack	<i>strategic_attack</i>
combat support	<i>combat_support</i>
counterair operations	<i>counterair_ops</i>
counterland operations	<i>counterland_ops</i>
cyberspace operations	<i>cyber_ops</i>
electronic warfare	<i>elec_warfare</i>
information operations	<i>info_ops</i>
global integrated intelligence surveillance reconnaissance operations	<i>global_isr_ops</i>
air mobility operations	<i>mobility_ops</i>
nuclear operations	<i>nuclear_ops</i>
space operations	<i>space_ops</i>
special operations	<i>spec_ops</i>
host nation	<i>host_nation</i>

Appendix D: R Code for Summative Content Analysis

```
#Load the packages used in this method
library("tm")
library("slam")
library("SnowballC")
library("RColorBrewer")
library("ggplot2")
library("wordcloud")
library("biclust")
library("igraph")
library("fpc")
library("Rcampdf")
library(data.table)
library(cluster)
library(xlsx)

#set environment
doctrine <- file.path("J:/AFIT ENV GEM 17M/Buyer, Joseph/THESIS/Data/Strategic
Docs for Text Mining/Text files")
docs <- Corpus(DirSource(doctrine))
J <- 100

#PREPROCESS THE DATA/DOCUMENTS
#remove punctuation, numbers, capital letters, common words, and specific words
docs <- tm_map(docs, removePunctuation)
docs <- tm_map(docs, removeNumbers)
docs <- tm_map(docs, tolower)
docs <- tm_map(docs, removeWords, stopwords("english"))
docs <- tm_map(docs, removeWords, c("annex","catalog","author"))

#in some cases there may be additional punctuations that need to be removed
for(j in seq(docs))
{
  docs[[j]] <- gsub("&", " ", docs[[j]])
}

#save a docs for use in ngrams
docs2 <- docs

#combining words that should stay together
for (j in seq(docs))
{
  docs[[j]] <- gsub("air force", "air_force", docs[[j]])
  docs[[j]] <- gsub("command and control", "command_and_control", docs[[j]])
}
```

```

docs[[j]] <- gsub("airspace control", "airspace_control", docs[[j]])
docs[[j]] <- gsub("strategic attack", "strategic_attack", docs[[j]])
docs[[j]] <- gsub("combat support", "combat_support", docs[[j]])
docs[[j]] <- gsub("counterair operations", "counterair_ops", docs[[j]])
docs[[j]] <- gsub("counterland operations", "counterland_ops", docs[[j]])
docs[[j]] <- gsub("cyberspace operations", "cyber_ops", docs[[j]])
docs[[j]] <- gsub("electronic warfare", "elec_warfare", docs[[j]])
docs[[j]] <- gsub("information operations", "info_ops", docs[[j]])
docs[[j]] <- gsub("global integrated intelligence surveillance reconnaissance
operations", "global_isr_ops", docs[[j]])
docs[[j]] <- gsub("air mobility operations", "mobility_ops", docs[[j]])
docs[[j]] <- gsub("nuclear operations", "nuclear_ops", docs[[j]])
docs[[j]] <- gsub("space operations", "space_ops", docs[[j]])
docs[[j]] <- gsub("special operations", "spec_ops", docs[[j]])
docs[[j]] <- gsub("host nation", "host_nation", docs[[j]])
}

```

```

#tell R to treat docs as a text document
docs <- tm_map(docs, PlainTextDocument)
docs2 <- tm_map(docs, PlainTextDocument)
docs2 <- tm_map(docs, stripWhitespace)

```

```

#remove endings and whitespace
docs <- tm_map(docs, stemDocument)
docs <- tm_map(docs, stripWhitespace)

```

#STAGE THE DATA/DOCUMENTS

```

#create a document term matrix (document as row and terms as column) and transposed
matrix (document as column and terms as rows)
tdm <- as.matrix(TermDocumentMatrix(docs,control=list(wordLengths=c(2, 35))))
tdm <- as.data.frame(tdm)
tdm <- as.data.table(tdm,TRUE)
names(tdm) <- c("words",dir(doctrine))

```

#Split data table into documents

```

c2_tdm <- tdm[, .(words,C2_Annex.txt)]
isr_tdm <- tdm[, .(words,ISR_Annex.txt)]
mobility_tdm <- tdm[, .(words,Mobility_Annex.txt)]
nuclear_tdm <- tdm[, .(words,Nuclear_Annex.txt)]
airspace_tdm <- tdm[, .(words,Airspace_Annex.txt)]
counterland_tdm <- tdm[, .(words,Counterland_Annex.txt)]
counterair_tdm <- tdm[, .(words,Counterair_Annex.txt)]
cyber_tdm <- tdm[, .(words,Cyber_Annex.txt)]
electronic_tdm <- tdm[, .(words,Electronic_Annex.txt)]

```

```

spec.ops_tdm <- tdm[, .(words,SpecOps_Annex.txt)]
space_tdm <- tdm[, .(words,Space_Annex.txt)]
targeting_tdm <- tdm[, .(words,Targeting_Annex.txt)]

#decending order to show top used words
c2_tdm <- c2_tdm[order(-C2_Annex.txt)]
isr_tdm <- isr_tdm[order(-ISR_Annex.txt)]
mobility_tdm <- mobility_tdm[order(-Mobility_Annex.txt)]
nuclear_tdm <- nuclear_tdm[order(-Nuclear_Annex.txt)]
airspace_tdm <- airspace_tdm[order(-Airspace_Annex.txt)]
counterland_tdm <- counterland_tdm[order(-Counterland_Annex.txt)]
counterair_tdm <- counterair_tdm[order(-Counterair_Annex.txt)]
cyber_tdm <- cyber_tdm[order(-Cyber_Annex.txt)]
electronic_tdm <- electronic_tdm[order(-Electronic_Annex.txt)]
spec.ops_tdm <- spec.ops_tdm[order(-SpecOps_Annex.txt)]
space_tdm <- space_tdm[order(-Space_Annex.txt)]
targeting_tdm <- targeting_tdm[order(-Targeting_Annex.txt)]

#Keep only the top 100 key words
c2_tdm <- c2_tdm[1:J]
isr_tdm <- isr_tdm[1:J]
mobility_tdm <- mobility_tdm[1:J]
nuclear_tdm <- nuclear_tdm[1:J]
airspace_tdm <- airspace_tdm[1:J]
counterland_tdm <- counterland_tdm[1:J]
counterair_tdm <- counterair_tdm[1:J]
cyber_tdm <- cyber_tdm[1:J]
electronic_tdm <- electronic_tdm[1:J]
spec.ops_tdm <- spec.ops_tdm[1:J]
space_tdm <- space_tdm[1:J]
targeting_tdm <- targeting_tdm[1:J]

#save all the files to excel for editing
write.xlsx(c2_tdm, "J:/AFIT ENV GEM 17M/Buyer, Joseph/THESIS/Data/Strategic
Docs for Text Mining/Excel Files/C2_Annex.xlsx")
write.xlsx(isr_tdm, "J:/AFIT ENV GEM 17M/Buyer, Joseph/THESIS/Data/Strategic
Docs for Text Mining/Excel Files/ISR_Annex.xlsx")
write.xlsx(mobility_tdm, "J:/AFIT ENV GEM 17M/Buyer,
Joseph/THESIS/Data/Strategic Docs for Text Mining/Excel Files/Mobility_Annex.xlsx")
write.xlsx(nuclear_tdm, "J:/AFIT ENV GEM 17M/Buyer, Joseph/THESIS/Data/Strategic
Docs for Text Mining/Excel Files/Nuclear_Annex.xlsx")
write.xlsx(airspace_tdm, "J:/AFIT ENV GEM 17M/Buyer,
Joseph/THESIS/Data/Strategic Docs for Text Mining/Excel Files/Airspace_Annex.xlsx")

```

```

write.xlsx(counterland_tdm, "J:/AFIT ENV GEM 17M/Buyer,
Joseph/THESIS/Data/Strategic Docs for Text Mining/Excel
Files/Counterland_Annex.xlsx")
write.xlsx(counterair_tdm, "J:/AFIT ENV GEM 17M/Buyer,
Joseph/THESIS/Data/Strategic Docs for Text Mining/Excel
Files/Counterair_Annex.xlsx")
write.xlsx(cyber_tdm, "J:/AFIT ENV GEM 17M/Buyer, Joseph/THESIS/Data/Strategic
Docs for Text Mining/Excel Files/Cyber_Annex.xlsx")
write.xlsx(electronic_tdm, "J:/AFIT ENV GEM 17M/Buyer,
Joseph/THESIS/Data/Strategic Docs for Text Mining/Excel
Files/Electronic_Annex.xlsx")
write.xlsx(spec.ops_tdm, "J:/AFIT ENV GEM 17M/Buyer,
Joseph/THESIS/Data/Strategic Docs for Text Mining/Excel Files/SpecOps_Annex.xlsx")
write.xlsx(space_tdm, "J:/AFIT ENV GEM 17M/Buyer, Joseph/THESIS/Data/Strategic
Docs for Text Mining/Excel Files/Space_Annex.xlsx")
write.xlsx(targeting_tdm, "J:/AFIT ENV GEM 17M/Buyer,
Joseph/THESIS/Data/Strategic Docs for Text Mining/Excel
Files/Targeting_Annex.xlsx")

```

```
#BIGRAMS
```

```
#create bigrams from docs
```

```
BigramTokenizer <-
```

```
function(x)
```

```
  unlist(lapply(ngrams(words(x), 2), paste, collapse = " "), use.names = FALSE)
```

```
tdm.bi <- as.matrix(TermDocumentMatrix(docs2, control = list(tokenize =
BigramTokenizer)))
```

```
tdm.bi <- as.data.frame(tdm.bi)
```

```
tdm.bi <- as.data.table(tdm.bi, TRUE)
```

```
names(tdm.bi) <- c("words", dir(doctrine))
```

```
#BIGRAM - Split data table into documents
```

```
c2_tdm.bi <- tdm.bi[, .(words, C2_Annex.txt)]
```

```
isr_tdm.bi <- tdm.bi[, .(words, ISR_Annex.txt)]
```

```
mobility_tdm.bi <- tdm.bi[, .(words, Mobility_Annex.txt)]
```

```
nuclear_tdm.bi <- tdm.bi[, .(words, Nuclear_Annex.txt)]
```

```
airspace_tdm.bi <- tdm.bi[, .(words, Airspace_Annex.txt)]
```

```
counterland_tdm.bi <- tdm.bi[, .(words, Counterland_Annex.txt)]
```

```
counterair_tdm.bi <- tdm.bi[, .(words, Counterair_Annex.txt)]
```

```
cyber_tdm.bi <- tdm.bi[, .(words, Cyber_Annex.txt)]
```

```
electronic_tdm.bi <- tdm.bi[, .(words, Electronic_Annex.txt)]
```

```
spec.ops_tdm.bi <- tdm.bi[, .(words, SpecOps_Annex.txt)]
```

```
space_tdm.bi <- tdm.bi[, .(words, Space_Annex.txt)]
```

```
targeting_tdm.bi <- tdm.bi[, .(words, Targeting_Annex.txt)]
```

```
#BIGRAM - decending order to show top used words
c2_tdm.bi <- c2_tdm.bi[order(-C2_Annex.txt)]
isr_tdm.bi <- isr_tdm.bi[order(-ISR_Annex.txt)]
mobility_tdm.bi <- mobility_tdm.bi[order(-Mobility_Annex.txt)]
nuclear_tdm.bi <- nuclear_tdm.bi[order(-Nuclear_Annex.txt)]
airspace_tdm.bi <- airspace_tdm.bi[order(-Airspace_Annex.txt)]
counterland_tdm.bi <- counterland_tdm.bi[order(-Counterland_Annex.txt)]
counterair_tdm.bi <- counterair_tdm.bi[order(-Counterair_Annex.txt)]
cyber_tdm.bi <- cyber_tdm.bi[order(-Cyber_Annex.txt)]
electronic_tdm.bi <- electronic_tdm.bi[order(-Electronic_Annex.txt)]
spec.ops_tdm.bi <- spec.ops_tdm.bi[order(-SpecOps_Annex.txt)]
space_tdm.bi <- space_tdm.bi[order(-Space_Annex.txt)]
targeting_tdm.bi <- targeting_tdm.bi[order(-Targeting_Annex.txt)]
```

```
#BIGRAM - Keep only the top 40 key words
```

```
c2_tdm.bi <- c2_tdm.bi[1:J]
isr_tdm.bi <- isr_tdm.bi[1:J]
mobility_tdm.bi <- mobility_tdm.bi[1:J]
nuclear_tdm.bi <- nuclear_tdm.bi[1:J]
airspace_tdm.bi <- airspace_tdm.bi[1:J]
counterland_tdm.bi <- counterland_tdm.bi[1:J]
counterair_tdm.bi <- counterair_tdm.bi[1:J]
cyber_tdm.bi <- cyber_tdm.bi[1:J]
electronic_tdm.bi <- electronic_tdm.bi[1:J]
spec.ops_tdm.bi <- spec.ops_tdm.bi[1:J]
space_tdm.bi <- space_tdm.bi[1:J]
targeting_tdm.bi <- targeting_tdm.bi[1:J]
```

```
#BIGRAM - save all the files to excel for editing
```

```
write.xlsx(c2_tdm.bi, "J:/AFIT ENV GEM 17M/Buyer, Joseph/THESIS/Data/Strategic
Docs for Text Mining/Excel Files/C2_Annex_bi.xlsx")
write.xlsx(isr_tdm.bi, "J:/AFIT ENV GEM 17M/Buyer, Joseph/THESIS/Data/Strategic
Docs for Text Mining/Excel Files/ISR_Annex_bi.xlsx")
write.xlsx(mobility_tdm.bi, "J:/AFIT ENV GEM 17M/Buyer,
Joseph/THESIS/Data/Strategic Docs for Text Mining/Excel
Files/Mobility_Annex_bi.xlsx")
write.xlsx(nuclear_tdm.bi, "J:/AFIT ENV GEM 17M/Buyer,
Joseph/THESIS/Data/Strategic Docs for Text Mining/Excel
Files/Nuclear_Annex_bi.xlsx")
write.xlsx(airspace_tdm.bi, "J:/AFIT ENV GEM 17M/Buyer,
Joseph/THESIS/Data/Strategic Docs for Text Mining/Excel
Files/Airspace_Annex_bi.xlsx")
write.xlsx(counterland_tdm.bi, "J:/AFIT ENV GEM 17M/Buyer,
Joseph/THESIS/Data/Strategic Docs for Text Mining/Excel
Files/Counterland_Annex_bi.xlsx")
```

```

write.xlsx(counterair_tdm.bi, "J:/AFIT ENV GEM 17M/Buyer,
Joseph/THESIS/Data/Strategic Docs for Text Mining/Excel
Files/Counterair_Annex_bi.xlsx")
write.xlsx(cyber_tdm.bi, "J:/AFIT ENV GEM 17M/Buyer,
Joseph/THESIS/Data/Strategic Docs for Text Mining/Excel Files/Cyber_Annex_bi.xlsx")
write.xlsx(electronic_tdm.bi, "J:/AFIT ENV GEM 17M/Buyer,
Joseph/THESIS/Data/Strategic Docs for Text Mining/Excel
Files/Electronic_Annex_bi.xlsx")
write.xlsx(spec.ops_tdm.bi, "J:/AFIT ENV GEM 17M/Buyer,
Joseph/THESIS/Data/Strategic Docs for Text Mining/Excel
Files/SpecOps_Annex_bi.xlsx")
write.xlsx(space_tdm.bi, "J:/AFIT ENV GEM 17M/Buyer,
Joseph/THESIS/Data/Strategic Docs for Text Mining/Excel Files/Space_Annex_bi.xlsx")
write.xlsx(targeting_tdm.bi, "J:/AFIT ENV GEM 17M/Buyer,
Joseph/THESIS/Data/Strategic Docs for Text Mining/Excel
Files/Targeting_Annex_bi.xlsx")

```

```
#TRIGRAMS
```

```
#create trigrams from docs
```

```
TrigramTokenizer <-
```

```
  function(x)
```

```
    unlist(lapply(ngrams(words(x), 3), paste, collapse = " "), use.names = FALSE)
```

```
tdm.tri <- as.matrix(TermDocumentMatrix(docs2, control = list(tokenize =
TrigramTokenizer)))
```

```
tdm.tri <- as.data.frame(tdm.tri)
```

```
tdm.tri <- as.data.table(tdm.tri, TRUE)
```

```
names(tdm.tri) <- c("words", dir(doctrine))
```

```
#TRIGRAM - Split data table into documents
```

```
c2_tdm.tri <- tdm.tri[, .(words, C2_Annex.txt)]
```

```
isr_tdm.tri <- tdm.tri[, .(words, ISR_Annex.txt)]
```

```
mobility_tdm.tri <- tdm.tri[, .(words, Mobility_Annex.txt)]
```

```
nuclear_tdm.tri <- tdm.tri[, .(words, Nuclear_Annex.txt)]
```

```
airspace_tdm.tri <- tdm.tri[, .(words, Airspace_Annex.txt)]
```

```
counterland_tdm.tri <- tdm.tri[, .(words, Counterland_Annex.txt)]
```

```
counterair_tdm.tri <- tdm.tri[, .(words, Counterair_Annex.txt)]
```

```
cyber_tdm.tri <- tdm.tri[, .(words, Cyber_Annex.txt)]
```

```
electronic_tdm.tri <- tdm.tri[, .(words, Electronic_Annex.txt)]
```

```
spec.ops_tdm.tri <- tdm.tri[, .(words, SpecOps_Annex.txt)]
```

```
space_tdm.tri <- tdm.tri[, .(words, Space_Annex.txt)]
```

```
targeting_tdm.tri <- tdm.tri[, .(words, Targeting_Annex.txt)]
```

```
#TRIGRAM - decending order to show top used words
```

```
c2_tdm.tri <- c2_tdm.tri[order(-C2_Annex.txt)]
```

```

isr_tdm.tri <- isr_tdm.tri[order(-ISR_Annex.txt)]
mobility_tdm.tri <- mobility_tdm.tri[order(-Mobility_Annex.txt)]
nuclear_tdm.tri <- nuclear_tdm.tri[order(-Nuclear_Annex.txt)]
airspace_tdm.tri <- airspace_tdm.tri[order(-Airspace_Annex.txt)]
counterland_tdm.tri <- counterland_tdm.tri[order(-Counterland_Annex.txt)]
counterair_tdm.tri <- counterair_tdm.tri[order(-Counterair_Annex.txt)]
cyber_tdm.tri <- cyber_tdm.tri[order(-Cyber_Annex.txt)]
electronic_tdm.tri <- electronic_tdm.tri[order(-Electronic_Annex.txt)]
spec.ops_tdm.tri <- spec.ops_tdm.tri[order(-SpecOps_Annex.txt)]
space_tdm.tri <- space_tdm.tri[order(-Space_Annex.txt)]
targeting_tdm.tri <- targeting_tdm.tri[order(-Targeting_Annex.txt)]

#TRIGRAM - Keep only the top 40 key words
c2_tdm.tri <- c2_tdm.tri[1:J]
isr_tdm.tri <- isr_tdm.tri[1:J]
mobility_tdm.tri <- mobility_tdm.tri[1:J]
nuclear_tdm.tri <- nuclear_tdm.tri[1:J]
airspace_tdm.tri <- airspace_tdm.tri[1:J]
counterland_tdm.tri <- counterland_tdm.tri[1:J]
counterair_tdm.tri <- counterair_tdm.tri[1:J]
cyber_tdm.tri <- cyber_tdm.tri[1:J]
electronic_tdm.tri <- electronic_tdm.tri[1:J]
spec.ops_tdm.tri <- spec.ops_tdm.tri[1:J]
space_tdm.tri <- space_tdm.tri[1:J]
targeting_tdm.tri <- targeting_tdm.tri[1:J]

#TRIGRAM - save all the files to excel for editing
write.xlsx(c2_tdm.tri, "J:/AFIT ENV GEM 17M/Buyer, Joseph/THESIS/Data/Strategic
Docs for Text Mining/Excel Files/C2_Annex_tri.xlsx")
write.xlsx(isr_tdm.tri, "J:/AFIT ENV GEM 17M/Buyer, Joseph/THESIS/Data/Strategic
Docs for Text Mining/Excel Files/ISR_Annex_tri.xlsx")
write.xlsx(mobility_tdm.tri, "J:/AFIT ENV GEM 17M/Buyer,
Joseph/THESIS/Data/Strategic Docs for Text Mining/Excel
Files/Mobility_Annex_tri.xlsx")
write.xlsx(nuclear_tdm.tri, "J:/AFIT ENV GEM 17M/Buyer,
Joseph/THESIS/Data/Strategic Docs for Text Mining/Excel
Files/Nuclear_Annex_tri.xlsx")
write.xlsx(airspace_tdm.tri, "J:/AFIT ENV GEM 17M/Buyer,
Joseph/THESIS/Data/Strategic Docs for Text Mining/Excel
Files/Airspace_Annex_tri.xlsx")
write.xlsx(counterland_tdm.tri, "J:/AFIT ENV GEM 17M/Buyer,
Joseph/THESIS/Data/Strategic Docs for Text Mining/Excel
Files/Counterland_Annex_tri.xlsx")

```

```
write.xlsx(counterair_tdm.tri, "J:/AFIT ENV GEM 17M/Buyer,
Joseph/THESIS/Data/Strategic Docs for Text Mining/Excel
Files/Counterair_Annex_tri.xlsx")
write.xlsx(cyber_tdm.tri, "J:/AFIT ENV GEM 17M/Buyer,
Joseph/THESIS/Data/Strategic Docs for Text Mining/Excel
Files/Cyber_Annex_tri.xlsx")
write.xlsx(electronic_tdm.tri, "J:/AFIT ENV GEM 17M/Buyer,
Joseph/THESIS/Data/Strategic Docs for Text Mining/Excel
Files/Electronic_Annex_tri.xlsx")
write.xlsx(spec.ops_tdm.tri, "J:/AFIT ENV GEM 17M/Buyer,
Joseph/THESIS/Data/Strategic Docs for Text Mining/Excel
Files/SpecOps_Annex_tri.xlsx")
write.xlsx(space_tdm.tri, "J:/AFIT ENV GEM 17M/Buyer,
Joseph/THESIS/Data/Strategic Docs for Text Mining/Excel
Files/Space_Annex_tri.xlsx")
write.xlsx(targeting_tdm.tri, "J:/AFIT ENV GEM 17M/Buyer,
Joseph/THESIS/Data/Strategic Docs for Text Mining/Excel
Files/Targeting_Annex_tri.xlsx")
```

Appendix E: Keyword Counts Found in FAC/CATCODE Descriptions

Key Words	Counts	Key Words	Counts	Key Words	Counts	Key Words	Counts
aircraft	116	guidance	19	frequenc	3	air operation	0
control	90	combat	17	satellit	3	coordinat	0
air support	68	radar	15	reconn	2	task force	0
communicat	53	attack	10	technolog	2	counterair	0
air control	48	air defense	9	computer	2	counterland	0
missile	43	data	7	aeromedic	2	air interdiction	0
weapon	39	command control	7	space system	2	cyber	0
air traffic control	33	tanker	5	airspace	1	disrupt	0
aviation	26	nuclear	5	data link	1	jam	0
electro	25	planning	4	mobility	1	orbit	0
launch	25	surveil	4	sensor	1	space capabilit	0
target	22	aerial port	4	airlift	1	missile defense	0
airfield	22	combat operation	3	mass destruction	1	missile tracking	0
runway	20	refuel	3	space control	1	special force	0
command	19	intel	3	space operation	1	air tasking	0
information	19	network	3	special operation	1		

Appendix F: Basic Categories Not Selected and >50%

Code	Title	Score
149	Operational Support Facilities Other Than Buildings	0.74
132	Communications Facilities Other Than Buildings	0.67
178	Explosives and Flame Ranges	0.67
872	Grounds Fencing, Gates, and Guard Towers	0.67
219	Installation Repair and Operation Maintenance Facilities	0.6
813	Electric Power Substations and Switching Stations	0.6
214	Tank and Automotive Maintenance Facilities	0.57
124	Operating Fuel Storage Facilities	0.5
126	Other Liquid Fuel and Dispensing Facilities	0.5
173	Training Support Facilities	0.5
175	Small Arms Ranges	0.5
319	Miscellaneous Item and Equipment Research, Development, Test and Evaluation Facilities	0.5
390	Research, Development, Test and Evaluation Facilities Other Than Buildings	0.5
751	Outdoor Recreation Facilities	0.5

Appendix G: R Code for Shiny Application Model

```
---
title: "Identifying Core AF Infrastructure by Keywords"
author: Joseph Buyer
output:
  flexdashboard::flex_dashboard:
    orientation: rows
    vertical_layout: fill
    theme: spacelab
runtime: shiny
---

```{r setup, include=FALSE}
#Read in the different packages used in the model
library(flexdashboard)
library(shiny)
library(gdata)
library(dplyr)
library(readr)
library(scales)
library(DT)

#read in data
FC <- readr::read_csv("J:/AFIT ENV GEM 17M/Buyer, Joseph/THESIS/R Model/FAC
and CAT.csv")
FC <- mutate_each(FC, funs(toupper))
basic.cats <- readr::read_csv("J:/AFIT ENV GEM 17M/Buyer, Joseph/THESIS/R
Model/Basic Categories.csv")
cat.groups <- readr::read_csv("J:/AFIT ENV GEM 17M/Buyer, Joseph/THESIS/R
Model/Category Group.csv")
fac.classes <- readr::read_csv("J:/AFIT ENV GEM 17M/Buyer, Joseph/THESIS/R
Model/Facility Class.csv")

...

Inputs {.sidebar}

```{r}
#Create and define the different inputs and buttons used in the app
selectInput(inputId = "class", label = "Choose how to display results",
            choices = c("Facility Class (1 digit)","Category Group (2 digit)","Basic Category
(3 digit)"),selected = "Basic Category (3 digit)")
```

```

hr()

textInput(inputId = "keyWord", label = "Keyword", placeholder = "Type keywords
here")

actionButton(inputId = "action", label = "Add Word")
hr()
div(style="display:inline-block", actionButton(inputId = "reset", label = "Clear"))

rv <- reactiveValues(words = data.frame(NULL))

# When the Submit button is clicked, save the input form data
observeEvent(input$action, {

  `if`(is.null(rv$words),
    rv$words <- data.frame(input$keyWord, stringsAsFactors = F),
    rv$words <- data.frame(c(rv$words,input$keyWord), stringsAsFactors = F))

})

observeEvent(input$reset, {
  rv$words <- data.frame(NULL)
})

...

Row {data-width=650} {.tabset .tabset-fade}
-----

### CATCODEs and FAC Codes with Titles and Descriptions

```{r}
#Code chunk that displays table showing real property data
DT::renderDataTable({

keyword2 <- paste(toupper(as.vector(unlist(rv$words))), collapse = '|')

FC3 <- unlist(lapply(X = 1:nrow(FC[,3]), FUN = function(x) grepl(keyword2, FC[x,3])))
FC4 <- unlist(lapply(X = 1:nrow(FC[,3]), FUN = function(x) grepl(keyword2, FC[x,4])))

either <- as.numeric(FC3==1) + as.numeric(FC4==1)

DT::datatable(FC[which(either > 0),], options = list(searchHighlight = TRUE), rownames
= F, fillContainer = T)
})

```

...

### ### Model Results

```{r}

#Code chunk that displays the model results by user-chosen grouping
DT::renderDataTable({

keyword2 <- paste(toupper(as.vector(unlist(rv\$words)))), collapse = '|')

FC3 <- unlist(lapply(X = 1:nrow(FC[,3]), FUN = function(x) grepl(keyword2, FC[x,3])))

FC4 <- unlist(lapply(X = 1:nrow(FC[,3]), FUN = function(x) grepl(keyword2, FC[x,4])))

either <- as.numeric(FC3==1) + as.numeric(FC4==1)

basic.cats <- subset(basic.cats, basic.cats[,3] > 0)

cat.groups <- subset(cat.groups, cat.groups[,3] > 0)

fac.classes <- subset(fac.classes, fac.classes[,3] > 0)

for(i in 1:nrow(basic.cats))

{

basic.cats[i,4] <- sum(substr(data.frame(FC[which(either >
0,)]),1,3)%in%basic.cats[i,1])

basic.cats[i,6] <- percent(as.numeric(basic.cats[i,4])/as.numeric(basic.cats[i,3]))

}

for(i in 1:nrow(cat.groups))

{

cat.groups[i,4] <- sum(substr(data.frame(FC[which(either >
0,)]),1,2)%in%cat.groups[i,1])

cat.groups[i,6] <- percent(as.numeric(cat.groups[i,4])/as.numeric(cat.groups[i,3]))

}

for(i in 1:nrow(fac.classes))

{

fac.classes[i,4] <- sum(substr(data.frame(FC[which(either >
0,)]),1,1)%in%fac.classes[i,1])

fac.classes[i,6] <- percent(as.numeric(fac.classes[i,4])/as.numeric(fac.classes[i,3]))

}

basic.cats\$`Core Msn Key Words Count` = NULL

cat.groups\$`Core Msn Key Words Count` = NULL

fac.classes\$`Core Msn Key Words Count` = NULL

```

results <- switch(as.character(input$class),
  "Facility Class (1 digit)" = fac.classes,
  "Category Group (2 digit)" = cat.groups,
  "Basic Category (3 digit)" = basic.cats)

DT::datatable(results,
  rownames = FALSE,
  options = list(columnDefs = list(list(className = 'dt-center', targets = "_all")),
    fillContainer = T
  ) %>% DT::formatStyle(
    "Percent Considered Core",
    color = 'red',
    fontWeight = 'bold'
  )
})
```


Row {data-height=250}

Selected Keywords

``{r}

Table that shows all previous user input keywords

(update with current response when Submit is clicked)

DT::renderDataTable({

DT::datatable(data.frame(unlist(rv$words), stringsAsFactors = F),

rownames = F, colnames = 'Used Words',

fillContainer = T,

caption = htmltools::tags$caption(

"This table shows all the ', htmltools::em('keywords'), ' that are in use'

)

)

)

)

97


```

## Bibliography

- AFCEC. (2015). *FY17-18 AFCAMP – Overview*.
- Antleman, A., Dempsey, J., & Brodt, B. (2008). Mission Dependency Index - A Metric for Determining Infrastructure Criticality. *Infrastructure Reporting and Asset Management*, 141–146. Retrieved from <http://ascelibrary.org/doi/pdf/10.1061/9780784409589>
- ASCE. (2013). *Report Card for America's Infrastructure*. American Society fo Civil Engineers. <https://doi.org/doi:10.1061/9780784478837>
- Berelson, B. (1952). Content analysis in communication research.
- Cavanagh, S. (1997). Content analysis: concepts, methods and applications. *Nurse Researcher*, 4(3), 5–13. <https://doi.org/10.7748/nr1997.04.4.3.5.c5869>
- Chang, W., Cheng, J., Allaire, J. J., Xie, Y., & McPherson, J. (2016). shiny: Web Application Framework for R. Retrieved from <https://cran.r-project.org/package=shiny>
- Civil Engineering Research Foundation. (1996). *Level of investment study: Facility and infrastructure maintenance and repair*. Washington D.C.
- Elsawah, H., Guerrero, M., & Moselhi, O. (2014). Decision Support Model for Integrated Intervention Plans of Municipal Infrastructure, 1039–1050.
- Faber, M. H., & Stewart, M. G. (2003). Risk assessment for civil engineering facilities: Critical overview and discussion. *Reliability Engineering and System Safety*, 80(2), 173–184. [https://doi.org/10.1016/S0951-8320\(03\)00027-9](https://doi.org/10.1016/S0951-8320(03)00027-9)
- Feinerer, I., & Hornik, K. (2015). tm: Text Mining Package. Retrieved from <https://cran.r-project.org/package=tm>
- GAO. (1996). Content Analysis: A Methodology for Structuring and Analyzing Written Material.
- GAO. (2008a). *Federal Real Property: Government's Fiscal Exposure from Repair and Maintenance Backlogs Is Unclear*.
- GAO. (2008b). Physical Infrastructure: Challenges and Investment Options for the Nation's Infrastructure.
- GAO. (2015). *Transportation Infrastructure: Information on Bridge Conditions Bridge*.
- GSA. (2009). *FY 2008 Federal Real Property Report*.

- GSA. (2013). *FY 2012 Federal Real Property Report*.
- Hastak, M., & Baim, E. J. (2015). Risk Factors Affectiong Management and Maintenance Cost of Urban Infrustructure, *20*(6), 14–15.  
<https://doi.org/10.1227/01.NEU.0000249266.26322.25>
- Hemme, K. (2015). Critical Infrastructure Protection: Maintenance is National Security. *Journal of Strategic Security*, *8*(Supplement: Eleventh Annual IAFIE Conference), 25–39. <https://doi.org/10.5038/1944-0472.8.3s.1471>
- Holsti, O. R. (1969). Content analysis for the social sciences and humanities.
- Hsieh, H.-F., & Shannon, S. E. (2005). Three Approaches to Qualitative Content Analysis. *Qualitative Health Research*, *15*(9), 1277–1288.  
<https://doi.org/10.1177/1049732305276687>
- Hubbard, D. W. (2014). *How to Meassure Anything: Finding the Value of “Intangibles” in Business* (3rd ed.). John Wiley & Sons, Inc.
- Hughes, J. W. (1991). Accounting and Reporting for Physical assets by Governments. AMER ACCOUNTING ASSOC 5717 BESSIE DR, SARASOTA, FL 34233.
- IFMA. (2009). *Strategic Facility Planning: A White Paper*. International Facility Management Association. <https://doi.org/10.1108/00438029910279394>
- Infrastructure Asset Management Exchange. (2015). *Infrastructure Asset Management Report: From Inspiration to Practical Application*.
- ISO. (2014). ISO 55000:2014(E) - Asset Management - Overview, principles and terminology, 2014. [https://doi.org/10.1007/978-3-7908-2720-0\\_12](https://doi.org/10.1007/978-3-7908-2720-0_12)
- Krippendorff, K. (1980). Content analysis: an introduction to its methodology.
- Kujawski, E., & Miller, G. (2009). The Mission Dependency Index: Fallacies and Misuses. *INCOSE International Symposium 2009*, 1565–1580.  
<https://doi.org/10.1002/j.2334-5837.2009.tb01035.x>
- Marlow, D. R., Beale, D. J., & Mashford, J. S. (2012). Risk-based prioritization and its application to inspection of valves in the water sector. *Reliability Engineering and System Safety*, *100*, 67–74. <https://doi.org/10.1016/j.res.2011.12.014>
- Merriam, S. B. (2009). *Qualitative research: a guide to design and implementation*. The Jossey-Bass higher and adult education series. San Francisco : Jossey-Bass,.
- Nichols, M. J. (2015). *A Delphi Study using Value-Focused Thinking for United States Air Force Mission Dependency Index Values*. Air Force Institute of Technology.

- NRC. (1996). *Budgeting for facilities maintenance and repair activities*, Standing Committee on Operations and Maintenance. <https://doi.org/10.17226/9226>
- NRC. (2008). *Core Competencies for Federal Facilities Asset Management Through 2020: Transformational Strategies*. <https://doi.org/10.1103/PhysRevB.81.041203>
- NRC. (2012). *Predicting Outcomes from Investments in Maintenance and Repair for Federal Facilities*.
- Oostermeijer, K., & Boss, J. (2010). *Best Practices in Asset Management Implementation*.
- Payne, T. (2000). Facilities management a strategy for success.
- Provost, F., & Fawcett, T. (2013). Data Science and its Relationship to Big Data and Data-Driven Decision Making. *Data Science and Big Data*, 1(1), 51–59. <https://doi.org/10.1089/big.2013.1508>
- R Core Team. (2016). R: A Language and Environment for Statistical Computing. Vienna, Austria. Retrieved from <https://www.r-project.org/>
- Salman, B., & Salem, O. (2012). Risk Assessment of Wastewater Collection Lines Using Failure Models and Criticality Ratings. *Journal of Pipeline Systems Engineering and Practice*, 3(3), 68–76. [https://doi.org/10.1061/\(ASCE\)PS.1949-1204.0000100](https://doi.org/10.1061/(ASCE)PS.1949-1204.0000100)
- Savin-Baden, M., & Major, C. H. (2013). *Qualitative Research: The essential guide to theory and practice*. Routledge.
- Stemler, S. (2001). An Overview of Content Analysis. *Practical Assessment, Research & Evaluation*. <https://doi.org/10.1362/146934703771910080>
- Strauss, A., Corbin, J., & others. (1990). *Basics of qualitative research* (Vol. 15). Newbury Park, CA: Sage.
- The Facilities Society. (2014). Strategies for Facilities Management. Retrieved from <http://www.facilities.ac.uk/j/cpd/62-facility-management/118-strategies-for-facility-management>
- The Institute of Asset Management. (n.d.). PAS-55.
- The White House, & Bush, G. W. (2004). *Executive Order #13327 - Federal Real Property Asset Management*.
- The White House, & Obama, B. (2013). *Presidential Policy Directive 21 - Critical Infrastructure Security and Resilience*. Retrieved from <https://www.whitehouse.gov/the-press-office/2013/02/12/presidential-policy-directive-critical-infrastructure-security-and-resil>

- USACE. (2011). *Program Management Plan for Asset Management – Draft*.
- USACE, Grussing, M. N., Gunderson, S., Canfield, M., Falconer, E., Antelman, A., & Hunter, S. L. (2010). *Development of the Army Facility Mission Dependency Index for Infrastructure Asset Management*.
- USAF. (2012a). *Program Action Directive (PAD) 12-03*.
- USAF. (2012b). *Programming Plan (P-Plan) for Implementation of Enterprise-Wide Civil Engineer Transformation (PAD 12-03), Volume 1 (Vol. 1)*.
- USAF. (2014). *America's Air Force: A Call to the Future*.
- USAF. (2015a). *Global Vigilance, Global Reach, Global Power for America*.
- USAF. (2015b). *Levels of Doctrine (Vol. 1)*.
- USAF. (2015c). *Strategic Posture Annex to the USAF Strategic Master Plan*.
- USAF. (2015d). *The World's Greatest Air Force: A Vision for the United States Air Force*. <https://doi.org/10.1017/CBO9781107415324.004>
- USAF. (2015e). *USAF Strategic Master Plan*.
- Vanier, D. J. (2001). Why Industry Needs Asset Management Tools. *Journal of Computing in Civil Engineering*, 15(1), 35–43.
- Water Research Foundation. (2012). *Condition Assessment of Water Main Appurtenances [Project # 4188]*.
- Weber, R. P. (1990). *Basic content analysis*. Sage.
- Welsh, G. M. A. (2014). *Senior Leaders Perspective*.

## **Vita**

Captain Joseph C. Buyer graduated from Broadneck Senior High School in Annapolis, Maryland. He entered undergraduate studies at the Pennsylvania State University (Penn State) in State College, Pennsylvania where he graduated with a Bachelor of Architectural Engineering focusing on mechanical systems in 2012. He was commissioned as an officer through the Detachment 720 AFROTC at the Penn State upon graduation.

His first assignment was to Goodfellow AFB in San Angelo, Texas in September of 2012. There he was assigned to the 17<sup>th</sup> Civil Engineer Squadron where he served as a civil engineer officer. While there he served as a project manager, an element chief, and a deputy flight chief, all while honing his leadership skills. In September of 2015, he entered the Graduate School of Engineering and Management, Air Force Institute of Technology in Dayton, Ohio. Upon graduation, he will be assigned to the 820<sup>th</sup> RED HORSE Squadron at Nellis AFB, Nevada.

REPORT DOCUMENTATION PAGE				Form Approved OMB No. 074-0188	
The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of the collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.   <b>PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.</b>					
1. REPORT DATE (DD-MM-YYYY) 23-03-2017		2. REPORT TYPE Master's Thesis		3. DATES COVERED (From – To) September 2015 – March 2017	
TITLE AND SUBTITLE  Refining Air Force Asset Management Strategy: Connecting Air Force Infrastructure to Core Missions				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)  Buyer, Joseph C., Captain, USAF				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(S) Air Force Institute of Technology Graduate School of Engineering and Management (AFIT/ENY) 2950 Hobson Way, Building 640 WPAFB OH 45433-8865				8. PERFORMING ORGANIZATION REPORT NUMBER  AFIT-ENV-MS-17-M-117	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) AGENCY: Air Force Civil Engineering Center ADDRESS: 2261 Hughes Ave, JBSA Lackland, TX 78236 PHONE and EMAIL: N/A ATTN: Capt Matthew Nichols (no sponsor enter: Intentionally left blank)				10. SPONSOR/MONITOR'S ACRONYM(S) AFCEC	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT DISTRIBUTION STATEMENT A. Approved for public release; Distribution unlimited.					
13. SUPPLEMENTARY NOTES This material is declared a work of the U.S. Government and is not subject to copyright protection in the United States.					
14. ABSTRACT At this time, Air Force organizational objectives are not tied to the Air Force asset management program. Additionally, the Air Force is unable to objectively differentiate between mission and support facilities as the Air Force Strategic Master Plan requires. While difficult, there is value in adjusting Air Force asset management strategic objectives by connecting Air Force investments in infrastructure to organizational objectives. Such a connection would allow senior leaders to more confidently make difficult decision related to infrastructure in this fiscally uncertain environment. This research focuses on improving Air Force asset management strategic objectives; specifically, it investigates infrastructure assets and Air Force core missions to determine if linkages can be made between them. Qualitative and quantitative skills are used to identify keywords from Air Force strategic documents and build a model which identifies infrastructure groupings related to Air Force core missions. The model breaks infrastructure assets into two groups, (1) core – directly related to Air Force core missions – and (2) non-core. This method of identifying infrastructure related to core missions can be used by the Air Force to differentiate between mission and support infrastructure while also reducing the amount of uncertainty related to asset management project funding.					
15. SUBJECT TERMS Asset management; mission infrastructure; mission facilities; core mission; qualitative research;					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT  UU	18. NUMBER OF PAGES  102	19a. NAME OF RESPONSIBLE PERSON Lt Col Vhance Valencia, AFIT/ENV
a. REPORT  U	b. ABSTRACT  U	c. THIS PAGE  U			19b. TELEPHONE NUMBER (Include area code) (937) 255-3636, ext 4826 vhance.valencia@us.af.mil

