SANDIA REPORT
SAND2021-10637
August 2021

Sandia National Laboratories

A Nuclear Security Enterprise Study of High-Reliability Systems, Collaboration, and Data

Terry Josserand

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico
87185 and Livermore,
California 94550

Issued by Sandia National Laboratories, operated for the United States Department of Energy by National Technology & Engineering Solutions of Sandia, LLC.

Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone: (865) 576-8401
Facsimile: (865) 576-5728
E-Mail: reports@osti.gov
Online ordering: http://www.osti.gov/scitech

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5301 Shawnee Rd
Alexandria, VA 22312

Telephone: (800) 553-6847
Facsimile: (703) 605-6900
E-Mail: orders@ntis.gov
Online order: https://classic.ntis.gov/help/order-methods/

ABSTRACT

It may seem simple and trivial, but defining the difference between data and information is contested and has implications that may affect the security of United States interests and even cost lives. For security, data are raw facts or figures without context, while information is the compilation or articulation of data that forms context. Security depends on clarity in the differences between data and information and controlling them.

Control is necessary to ensure that data and information are not inadvertently released to foreign governments, the public, or those without Need-to-Know. A primary concern in the practice of security is the control of data to avoid the inadvertent conversion to sensitive information. The complexity of this concern is further augmented when institutions are part of tightly coupled networks that informally share data and information. Additionally, those that share data as a function of legislative action—and/or formally integrate data and information system infrastructures—may be a higher security risk. This paper will present a case study that utilizes elements of literature from Knowledge Management and networks to tell a story of an issue in security—specifically, controlling the conversion of data to information.

This page left blank

CONTENTS

LIST OF FIGURES

This page left blank

ACRONYMS AND DEFINITIONS

Abbreviation	Definition
CA	Classified Association
CY	Calendar Year
DBMS	Database Management System
DOD	Department of Defense
DOE	Department of Energy
FFRDC	Federally Funded Research and Development Center
FRD	Formerly Restricted Data
GAO	Government Accountability Office
IDC	International Data Corporation
IR	International Relations
IS	Information Systems
KDD	Knowledge Discovery from Data
KM	Knowledge Management
M&O	Management and Operating
NC3	Nuclear Command, Control, and Communications
NDAA	National Defense Authorization Act
NNSA	National Nuclear Security Administration
NPR	Nuclear Posture Review
NSE	Nuclear Security Enterprise
NSI	National Security Information
NTK	Need to Know
NW	Nuclear Weapon
OUO	Official Use Only
PA	Public Administration
PM	Project Management
PMBOK	Project Management Body of Knowledge
RD	Restricted Data
SNL	Sandia National Laboratories
US	United States

This page left blank

TERMS AND DEFINITIONS

Term	Definition
Wicked Problem	Dynamic and complex problems, with no clear definition or solution involving multiple stakeholders in multiple organizations
Tame Problem	Problems that are well defined and easily addressed

This page left blank

INTRODUCTION

It may seem simple and trivial, but defining the difference between *data* and *information* is contested and has implications that may affect the security of United States (US) interests and even cost lives. For security, data are raw facts or figures without context, while information is the compilation or articulation of data that forms context. Security depends on these differences between data and information and how to control them.

Control is necessary to ensure that data and information are not inadvertently released to foreign governments, the public, or those without Need-to-Know (NTK). A primary concern in the practice of security is the control of data to avoid the inadvertent conversion to sensitive information. The complexity of this concern is further augmented when institutions are part of tightly coupled networks that informally share data and information. Additionally, those that share data as a function of legislative action and/or formally integrate data infrastructures may be at high security risk. This paper will present a case study that utilizes elements of literature from Knowledge Management (KM) and networks to tell a story of an issue in security—controlling the conversion of data to information.

Data has many meanings. Data has been described as the new gold, the new oil (Ransbotham 2016), an organizational asset (Lake 2013), the world's most valuable resource (Parkins 2017), even a source of political power (Harari 2017). The current era is one in which data are literally everywhere as technology allows everyone and everything to be continuously connected. As Donald Kettl (2018) describes it, "the world is constantly searching for the next big thing. In government, this is it—the *Data Revolution*. Even if government wanted to ignore the *Data Revolution*, it could not—the data are flooding in from everywhere." This *Data Revolution* era began in the 1990s, but has categorically escalated over the last decade.

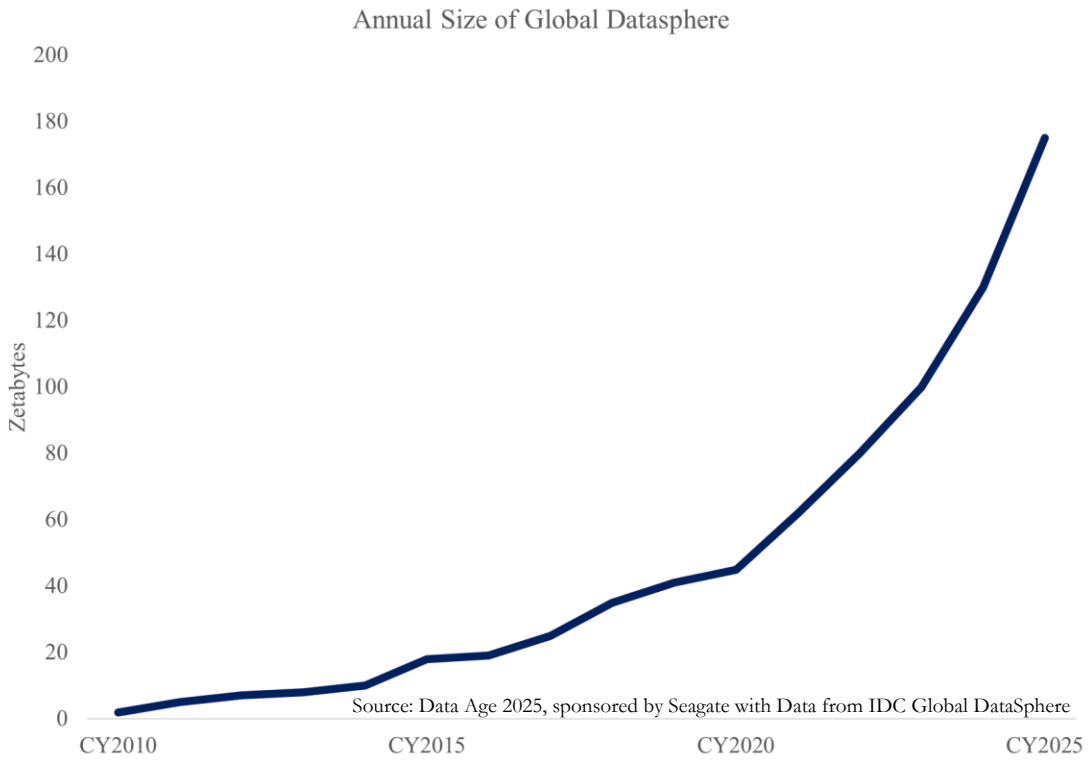

Figure 1. Global Data Growth

One-way to observe this flood of data is by quantifying its global growth since Calendar Year 2010 (CY2010) with projections through CY2025 as illustrated in Figure 1.

Richard Box (2018) describes this *Data Revolution* as an output of a world tightly connected by electronically controlled systems, devices, and programs that may shape theory and practice in the field. The rapid growth in data can partially be attributed to the economies of scale with respect to the cost of data producing technologies and acquisition, storage, and retrieval methods (Ronsenthal and Rosenthal 2012). The growth in data is so fast paced that analytic methods to gain meaningful insights from it cannot keep up. It is projected that of all the data that is generated, 99% of it goes untouched (Burn-Murdoch 2012). This does not lead to organizations deleting untouched or unused data. Instead, data are stored for future use as an asset and a foundational element of organizational KM.

In the KM literature, data are foundational elements to the concept of knowledge. However, the KM literature tends to focus on the concept and forms of knowledge itself with no consensus definition across the disciplines. Knowledge continues to be an ongoing debate as a "multifaceted concept with multilayered meanings (Nonaka 1994)." Within the KM literature, data are often viewed as having a quasi-hierarchical relationship to information and knowledge. Although it is common to see the interchangeable use of the words *data* and *information*. Furthermore, it is not surprising to see the words *information* and *knowledge* used interchangeably. While a consensus definition of knowledge is an ongoing debate, the concepts of data and information may also be up for interpretation. While the distinction between data and information may not be critical in the field of knowledge management, it is in the field and practice of security where the unintended release of information may cause damage to US interests or lives.

"One of the most critical aspects of defining and understanding the meaning of security is to recognize that it is heavily dependent on risk or threat (Herron, Jenkins-Smith, and Silva 2012)." Security programs in the US must identify the risks and threats associated with the release of information to foreign governments, the public, and other entities. The organizations responsible for security programs in the US are also responsible for the control of the data and release of information per policies and laws including Executive Order 13526[1]. Thus the distinction between data and information in the field and practice of security is vital in order to inform the calculation of risk or threat to US interests or lives.

In Public Administration (PA) literature data, information and knowledge come up in network research that is pertinent to security. In PA, the term *network* often focuses on "horizontal coordination mechanisms between actors (mostly organizations) and assumes the outcomes and performance result from interactions between a variety of actors rather than from the actions of one actor alone (Klijn and Koppenjan 2012)." Furthermore, Klijn and Koppenjan describe that trust is often considered one of the core coordination mechanisms and attributes of successful outcomes in networks (2012). Trust has been seen to enhance the possibility that actors within a network will

[1] This order prescribes a uniform system for classifying, safeguarding, and declassifying US national security information, including information relating to defense against transnational terrorism. US democratic principles require that the American people be informed of the activities of their Government. Also, the nation's progress depends on the free flow of information both within the Government and to the American people. Nevertheless, throughout US history, the national defense has required that certain information be maintained in confidence in order to protect US citizens, democratic institutions, homeland security, and interactions with foreign nations. Protecting information critical to US security and demonstrating the nation's commitment to open Government—through accurate and accountable application of classification standards with routine, secure, and effective declassification—are equally important priorities.

share data and information (Lane and Backman 1998). Others view "knowledge as the currency of collaboration (Emerson and Nabatchi 2012)," data sharing as a requisite for success and designate one of the views of network studies as a focus on the information processing and knowledge management capabilities of networks (Rathemeye and Hatmaker 2008). Networks arise for many reasons, one of them being to address *wicked problems* that are typical of security programs. *Tame problems*[2] are well defined and easily addressed whereas *wicked problems* are dynamic and complex, with no clear definition or solution involving multiple stakeholders in multiple organizations (Emerson and Nabatchi 2012).

Herron et al. (2012) states that a dimension of security that deals with wicked problems is nuclear security.

> *Nuclear security encompasses Nuclear Weapons and their development, management, modernization, and uses; nuclear materials and their production, applications, and safeguards; nuclear proliferation and associated implications; and public perceptions of and support for policies relating to each of these aspects of nuclear security.*

Peters (2017) states that wicked problems may be a symptom of another problem. The wicked problems of nuclear security present another wicked problem in the requirement to have clear distinctions and control of data and information both *intra*-organizationally and *inter*-organizationally.

Institutions in any dimension of US security responsible for the control of data and release of information—that may cause damage to interests or lives—must clearly distinguish data from information.These institutions must comprehend the implications of the implementation of policy both *intra-* and *inter*-organizationally. This paper will present a case study that will explore a complex phenomenon that has resulted from the *Data Revolution.*

The case study will lean on elements of literature from KM and networks to tell a story of a wicked problem in security—*controlling the conversion of data to information*. The purpose of this case study is to illustrate and gain a contextual comprehension to inform discussions regarding a specific real-world issue. The paper will begin with a background of the organization and network being utilized for the case study, conclude with a general summary, and posit areas for discussion with contemporary topics for practitioners and academicians.

[2] Examples may include puzzles, algebraic equations, planning for relocation, etc. Often able to be solved utilizing common linear methods.

This page left blank

1. CASE STUDY ORGANIZATION AND NETWORK BACKGROUND

Frederickson (2016) summarizes the varying characterizations of high-reliability organizations and high-reliability systems by their low tolerance to risk-taking due to:

- Catastrophic implications of failure
- Praise for error reporting
- Efficiency over economy
- Tightly coupled physical systems and organizational networks
- Adequate funding, rigid procedures and standards
- High redundancy
- Substantial reliance on expertise

Examples of high-reliability systems include provision of electricity, nuclear power plants, nuclear submarines, and Nuclear Weapons (NW) to name a few. "Owing to the fact that most of our systems are not high reliability, the literature has tended to focus on agencies and systems that are error-tolerant and that have goals that are difficult to measure (March and Olsen 1995)."

High-reliability organizations and high-reliability systems are often the result of wicked problems. Tame problems are well defined and easily addressed whereas wicked problems are dynamic and complex, with no clear definition or solution involving multiple stakeholders in multiple organizations (Emerson and Nabatchi 2012). In the US, nuclear security is a wicked problem that is the responsibility of numerous high-reliability organizations and comprises various high-reliability systems including NWs mentioned earlier.

The Nuclear Security Enterprise (NSE) of the US is one of the networks that is responsible for the wicked problems of nuclear security. The NSE is governed by the National Nuclear Security Administration (NNSA)—a semiautonomous agency within the Department of Energy (DOE)—and is responsible for maintaining the NW stockpile, monitoring and promoting nonproliferation, powering the nuclear Navy, and responding to nuclear and radiological emergencies.[3] The NSE is comprised of seven Management and Operating (M&O) contractors[4] responsible for eight high-reliability organizations that execute these responsibilities.

One of the eight high-reliability organizations is Sandia National Laboratories (SNL). SNL is a Federally Funded Research and Development Center (FFRDC) that collaborates broadly with government agencies, industry, and academic institutions in the strategic areas of NW, energy, and national and global security. Within the NSE, SNL has some of the highest consequence responsibilities as the design agency for approximately 97% of the design of modern NW components, weaponization of the physics package, and overall systems engineering and integration. SNL is responsible for NW systems and components over their entire lifecycle, from original design through final dismantlement and disposal. NW systems have some of the most rigorous technical

[3] From DOE NNSA website https://www.energy.gov/nnsa/about-nnsa last accessed Februray 2021.

[4] M&O contracts are agreements under which the government contracts for the operation, maintenance, or support, on its behalf, of a government-owned or -controlled research, development, special production, or testing establishment wholly or principally devoted to one or more of the major programs of the contracting agency. 48 C.F.R. § 17.601 (2018). The sites that comprise the nuclear security enterprise are the Kansas City National Security Campus in Missouri, Lawrence Livermore National Laboratory in California, Los Alamos National Laboratory in New Mexico, Nevada National Security Site, Pantex Plant in Texas, Sandia National Laboratories primarily in New Mexico, Savannah River Site in South Carolina, and Y-12 National Security Complex in Tennessee.

requirements in defense as they sit dormant for decades, must be immediately available when needed, must always work when authorized by the President of the US and must never detonate otherwise.

NW systems serve as a deterrent that protects the US homeland and allies abroad (NPR 2018). "The complex nuclear deterrence approach has been the basis of the US nuclear policy since about the 1960s." (Vergun 2020) According to deterrence theory, in order for a deterrent to be effective a state must persuade adversaries of the following:

1. That it has an effective military capability that can inflict unacceptable costs.
2. That the threat of use is credible (Sagan 1994, Powell 2003).

Deterrence theory requires persuasion of adversaries because it assumes actors have incomplete information similar to rational choice theory. Thus, in order for deterrence theory to be effective an adversary must have incomplete information.

The safeguarding of nuclear security information that enables a nuclear deterrence posture is the responsibility of the organizations that own the data. "Owners of data are responsible for determining the sensitivity of information before it is used, processed, or stored on information systems." (DOE 471.2A) For the purposes of this case study, the NSE and SNL are responsible for wicked problems of nuclear security that includes the control of related data and information.

2. CASE STUDY

2.1. Knowledge Management

What is knowledge? There is no consensus definition of knowledge, as it is a "multifaceted concept with multilayered meanings (Nonaka 1994)." However, knowledge is commonly broken out into two distinct forms—explicit and tacit—that are the result of the interpretation of data and information as illustrated in Figure 2.

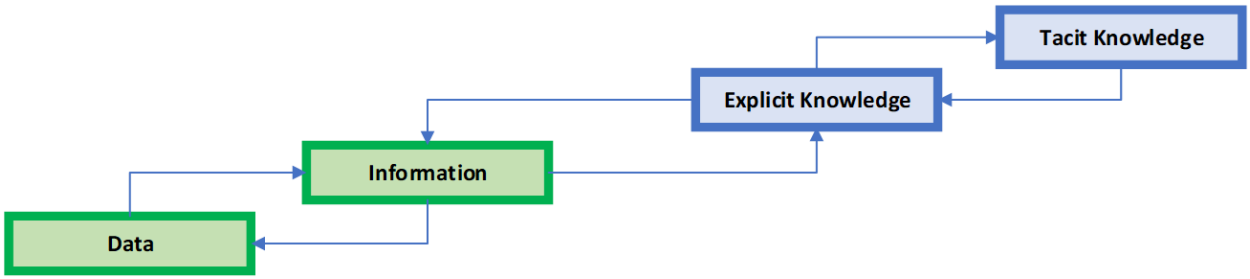

Figure 2. Knowledge Management Hierarchy

Polanyi (2011) describes tacit knowledge as something someone knows but cannot describe. Explicit and tacit knowledge are distinguished through the concepts of *what is known* to the *know-that* respectively. Furthermore, through the concepts of *knowing* or the *know-how* respectively (Cook and Brown 1999; De Marco et al. 2012). The analogy commonly used to distinguish these two types of knowledge is through the story of riding a bike adapted from Polanyi (2011):

> *If a person reads all of the information that exists regarding how to ride a bike and is capable of interpreting it, then they may have explicit knowledge of everything that is known about how to ride a bike. This explicit knowledge of how to ride a bike may improve someone's abilities but, if this person has never put their understanding of what is known into practice—actually riding a bike—they may not be able to ride a bike. In order to ride a bike, tacit knowledge is required.*

This example describes tacit knowledge as existing in the mind of an individual or the *knower*, whereas explicit knowledge is something more tangible. This is why it is often times used interchangeably in the literature with information.

In the management literature, organizations are viewed as a system that processes information and applies knowledge to solve problems (Malhotra 2005). Some view organizations as information-processing machines or knowledge producers, however organizations are more than this—they are entities that create knowledge through action and interaction. Nonaka (1994) goes on to describe the creation and preservation of tacit knowledge through social interactions and practice. The management and processing of knowledge are considered the most important source of an organizations renewable and sustainable competitive advantage (Agrifoglio 2015).

As described previously, high-reliability organizations and systems lean heavily on expertise—*know-that* and *know-how*. This expertise requires both data and information as illustrated in Figure 2—and due to the nature of wicked problems—requires sharing of data and information across organizations in trusted networks to develop forms of explicit and tacit knowledge through socialization.

The *Data Revolution* has brought an unprecedented growth in data that may place strain on all aspects of the KM hierarchy illustrated in Figure 2. For the purposes of the case study, an Information

Systems (IS) research approach from the KM literature is applied to make as clear a distinction as possible between data, information, and knowledge (Agrifoglio 2015).

2.1.1. Data

Utilizing the IS research approach, data refers simply to raw facts and figures (Gallaugher 2011). Data alone are rarely useful due to lack of context. Data and information are often used interchangeably in literature and in the public sector (DOE 471.2A). From an IS research approach, data serve as raw materials to form information. Thus, data are distinguishable from information as illustrated in Figure 2.

Organizations commonly generate, acquire, and manage data in databases. A database is simply a list of data. Organizations often have multiple databases that serve functional purposes, i.e., Human Resources, Finance, Purchasing, Engineering, etc. The list can be organized in unique columns or rows that capture data records or instances. Organizations utilize database management systems (DBMS) to manage all of the databases with varying languages and software to communicate and interact with them. It is common that different functional areas of an organization have a need for a data field that exists in other functional areas databases. As a result, these data fields may be cross-integrated in multiple databases and/or access may be given to cross-functional individuals within an organization.

Figure 3 illustrates the six unique databases that the case study analyst has been granted access to at SNL. It is important to note that several of the data fields exist within multiple databases.

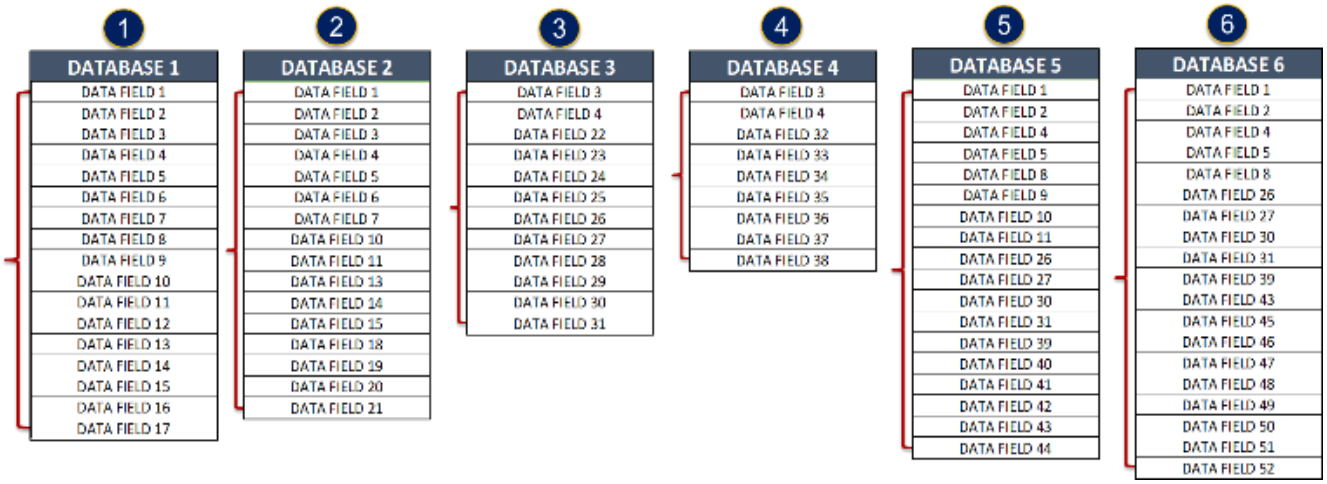

Figure 3. Example Databases

The databases from Figure 3 and the data contained within are what an analyst would utilize to form information as a function of their role within and for the organization.

2.1.2. Information

Information is created through the articulation of data—raw facts and figures—with context. Thus, data can be viewed as having a hierarchical relationship to information as information may be reduced into data without context.

How can the hierarchical and inverse view of the relationship between data and information be applied to information and explicit knowledge as previously described? Another perspective of knowledge states that knowledge can and does in fact exist before articulated information and raw data and facts (Tuomi 1999). The arguments of a hierarchical and inverse relationship between data and information is not too widely criticized, especially in the IS literature. It is not until the interpretation of information to knowledge that the literatures are divergent mainly around concepts of the role of human agency or processing that arguably must occur in order for information to become a form of knowledge. (Alavi and Leidner 1998)

Figure 3 is a snapshot of six databases at SNL. These six databases house data. It was previously described that one of the characteristics of the *Data Revolution* is the rapid growth in the volume of data and that SNL deals with wicked problems that may require cross-functional intra-organizational data sharing. This is illustrated in Figure 3 with the repeated data fields in multiple databases. How then are the databases that house data within the organization presented in Figure 3 pertinent to the control of information?

2.1.3. Data Association and Security

The *Data Revolution* has brought about a rise in many analytic techniques to comprehend data through a process of associations. These analytic techniques may result in context that meets the KM literature IS research definition of information. There are multiple definitions and methods for doing similar things across disciplines.

In IS, computer science, and statistics these analytic techniques are often referred to as data mining or the science of extracting useful information from large data sets or databases (Hand, Heikki, and Padhraic 2001). In International Relations (IR), these techniques may be referred to as associations or the process of discovering interesting relationships hidden in large datatsets (Van Puyvelde, Coulthart, and Hossain 2017). In PA, it may be referred to as Knowledge Discovery from Data (KDD), which uses sophisticated statistical or automatic reasoning methods to identify patterns of interesting relationships (Wiig 2000).

Figure 4 describes two pathways that individuals are trained[5] to protect classified or sensitive information regarding security programs. Specific to the NSE, the US DOE Office of Classification is responsible for the classification levels and categories regarding specific program information (US DOE 1991). This paper does not intend to elaborate upon the comprehensive process of classification of sensitive information. Instead, this paper will provide two general methods by which individuals working on security programs are trained to protect sensitive information. Both pathways involve creating a relationship of two pieces of unclassified data that results in an association that may provide sufficient context that has damaging implications if released in an unaccredited environment. The DOE Directive 475.2B (2014) calls this "classification based on association" or a Classified Association (CA).

Figure 4 offers two classic examples:

1. The A + B phenomena where stating that Weapon System 1 is related to Technology 3 is determined to provide sufficient context to be a CA.

[5] Executive Order 13526 has many requirements including that any individual that has access to sensitive information complete the proper background checks for clearance authorization and have been granted information specific Need-to-Know (NTK).

2. The A = B phenomena where stating or inferring that Weapon System 1 is equal to Technology 3, is determined to provide sufficient context to be a CA.

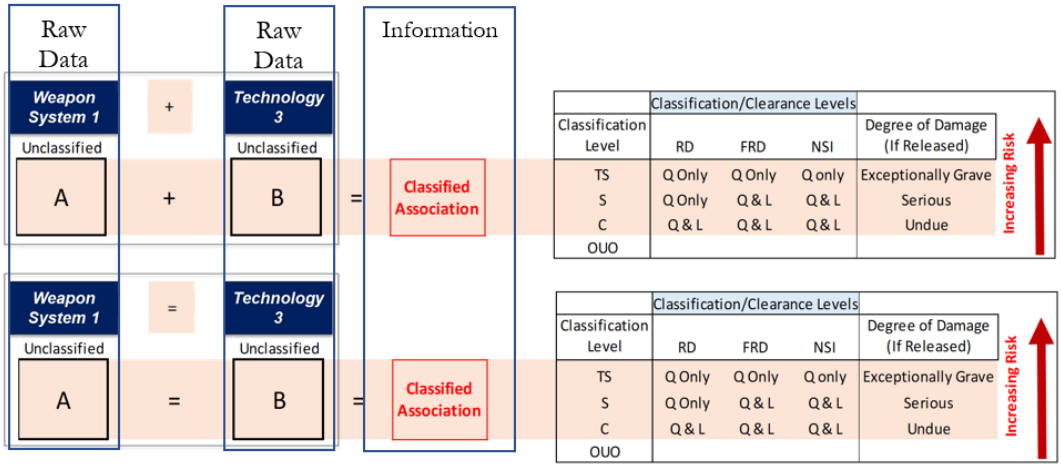

Figure 4. Example of Data in Context from DOE Directive 475.2B

What is important from Figure 4 is that the data are considered unclassified or non sensitive until it is related to other unclassified or non sensitive data resulting in sufficient context to be considered a CA.

2.1.3.1. Tabular Illustration

Utilizing Figure 3, it is possible to align these disparate databases in tabular format to comprehend how many total unique data fields there are as well as how many data field relationships exist between these databases. Figure 5 depicts the tabulation of the databases from Figure 3.

There are 52 unique data fields and a total of 137 unique data field relationships across these six databases. Data Field 3 is circled as an exemplar of the unique data field relationship calculation.

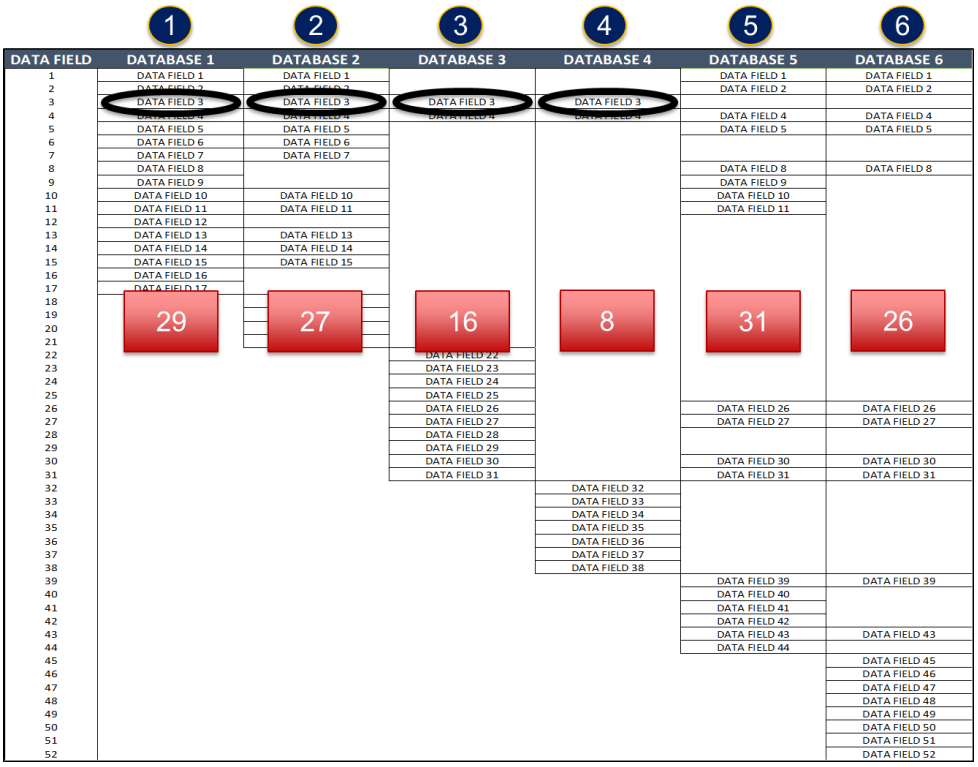

Figure 5. Tabular Illustration of Example Databases

As a calculation for count of unique data field relationships for Database 1, Data Field 3 is repeated in Database 2, Database 3, and Database 4. This equates to three unique data field relationships. Once this calculation process for Database 1 is completed for all unique data fields, the total count of unique data field relationships that Database 1 has to all other databases is 29. Meaning Database 1 has 29 instances where its unique data fields are repeated in the other five databases.

These data field relationships enable an analyst to go both horizontally and vertically through these databases in order to generate information for the organization. Within SNL, when asked a question, the analyst's role is to form data-driven information from this raw material and to maintain and produce new knowledge. The volume of data within these databases continues to grow as a function of the *Data Revolution*, along with the sharing of unique data fields, the growth of additional databases, and advancement of analytical methodologies.

Another way to view the tabular illustration is through a network lens.

2.1.3.2. Network Illustration

It has been described that data are simply raw facts and figures without context and information is data in context (Boisot and Canals 2004). It was previously described that in order to protect sensitive information, security professionals are trained to avoid relating raw data in such a way that may give it context or reveal a CA.

2.1.3.2.1. Analyst Network

When asked a question, the SNL analyst that has access to these six unique databases utilizes them to develop a response that is informative and data-driven. The analyst is seeking to add context to the data to form information through the utility of the 137 unique data field relationships. Figure 6

illustrates the database network that the SNL analyst utilizes with the green lines representing the existence of shared data fields and thus a database relationship.

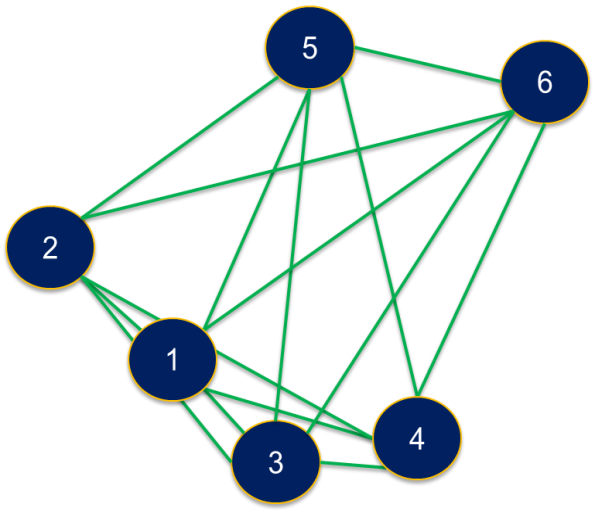

Figure 6. Analysts Network

Figure 6 describes the snapshot of the network of databases that the SNL analyst utilizes; it does not describe the volume of data or the analytic techniques. This is simply meant to illustrate that the databases that house data are related to one another. If the volume of data within these databases follows the behavior of the global growth of data depicted in Figure 1 then this analyst may well be inundated with data.

The individual that has access to the network of databases depicted in Figure 6 is required to comprehend and control at all times how their use of data to generate information may result in a CA from Figure 4.

2.1.3.2.2. Intra-Organizational Network

At an organizational level, there may be a large count of disparate databases. This may be a result of cross-functional or even cross-program needs. Figure 7 illustrates a hypothetical snapshot of all of the databases that exist within the SNL organization.

The green solid lines represents the known relationships of the data fields between the databases to which the case study analyst has access . The red dotted lines represent the unknown relationships of the data fields between the databases that the case study analyst does not have access to and possibly does not know exist.

From the IS research and KM perspective, these databases are the core of an organizations information-processing capability. It is through these databases that individuals within the organization communicate with one another. Individuals communicate by taking raw data compilations to form contextual information.

From a security perspective, it is the responsibility of the institution to ensure that the data in the intra-organizational databases is controlled to the extent that it does not result in a CA. That is, to

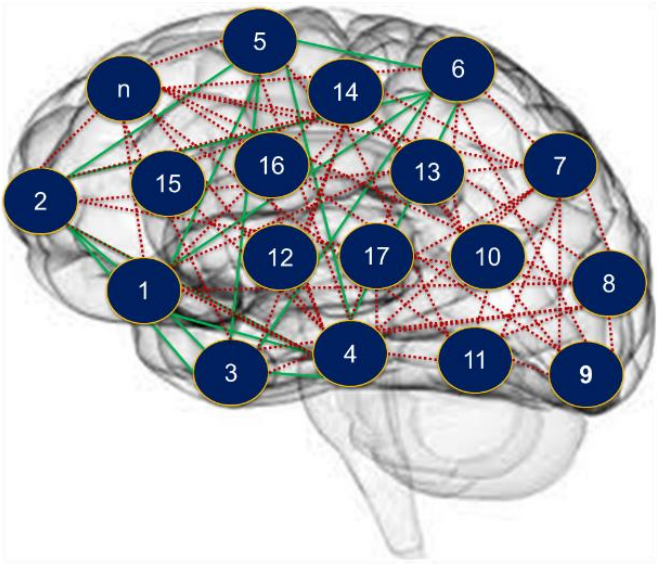

Figure 7. Intra-Organizational Network

determine the sensitivity of information that exists on the intra-organizational network and ensuring it is accredited to sustain the approved level of information at any given point in time (DOE 471.2A). The complexity facing security professionals in the public sector responsible for controlling data in context is daunting as the volume and count of databases continues to increase.

When wicked problems stimulate the organization, various individuals fire like synapses through the various database connections in order to respond. As described previously, the NSE includes SNL and eight other organizations responsible for the wicked problems of the NSE.

2.1.3.2.3. Inter-Organizational Network

The wicked problems of the NSE require the organizations within the network to collaborate. As previously described, is has been suggested that knowledge is the currency of collaboration and data sharing is a requirement for success (Emerson and Nabatchi 2012).

Rathemeye and Hatmaker (2008) describe one of the perspectives of network studies as a focus on the information processing and knowledge management capabilities of networks. It may be argued that data sharing is not central to the success of networks, but it would be difficult to argue it is not important.

US government practices and legislation are both recommending and requiring forms of technological coupling for data and information sharing purposes both within and across networks. The Government Accountability Office (GAO) emphasized the importance of effective management and oversight of the contracts, projects, and programs that support NSE's mission, which are dependent upon the availability of reliable, enterprise-wide cost information (GAO 2019). Section 3113 of the National Defense Authorization Act (NDAA)[6] of 2017 requires the NNSA to implement a common financial reporting system for the NSE. Furthermore, advancements in the

[6] This section of the NDAA authorizes appropriations for fiscal year 2020 for military activities of the Department of Defense, for military construction, and for defense activities of the Department of Energy, to prescribe military personnel strengths for such fiscal year, and for other purposes.

practice of Project Management (PM)[7] have led to requirements for common-integrated software implementation for high-reliability system efforts within the NSE (Bradshaw and Julian 2014, Bowers 2014). Over the last five years within the NDAA, the use of the word *data* has almost grown by a factor of three as illustrated in

Figure 8.

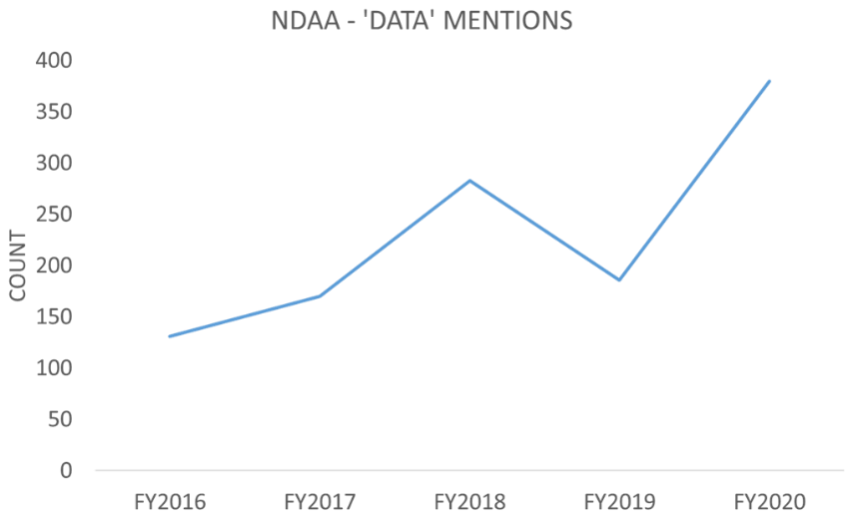

Figure 8. NDAA and Data

PM practices evolve and the economies of scale of data producing technologies and acquisition, storage, and retrieval methods continue to move in a cost-beneficial direction. With each new advancement, tightly coupled organizational networks responding to wicked problems are technologically becoming more coupled (Ronsenthal and Rosenthal 2012). In the context of this study, technological coupling is characterized by the inter-organizational integration of data and information systems architectures. This technological coupling is not necessarily a product of natural intra-network actors making decisions in their self-interest, but are also a product of legislation.

Figure 9 illustrates the organizations[8] that comprise the NSE and is meant to represent a quasi-hypothetical state of inter-organizational integration of data and information systems architectures.

[7] According to the Project Management Institute, a project is a temporary endeavor that has a defined beginning and end in time. Project management is the application of knowledge, skills, tools, and techniques to project activities to meet the project requirements. The elements of the practice reside in the Project Management Body of Knowledge (PMBOK).
[8] The sites that comprise the nuclear security enterprise are the Kansas City National Security Campus in Missouri, Lawrence Livermore National Laboratory in California, Los Alamos National Laboratory in New Mexico, Nevada National Security Site, Pantex Plant in Texas, Sandia National Laboratories primarily in New Mexico, Savannah River Site in South Carolina, and Y-12 National Security Complex in Tennessee.

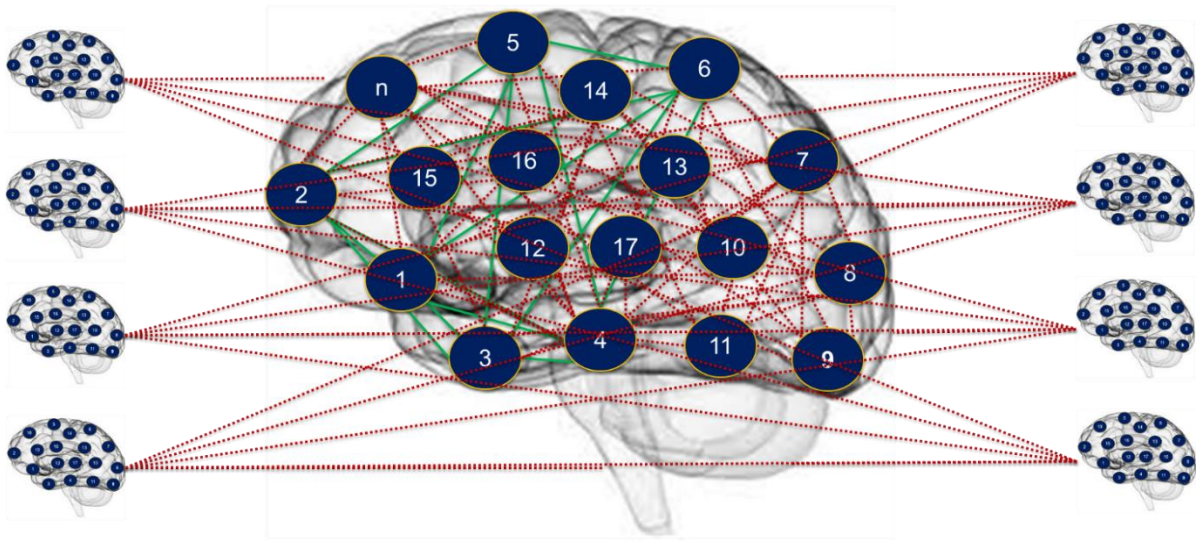

Figure 9. Inter-Organizational Network

Figure 7 illustrated the complexity of comprehending intra-organizational relationships of data fields across databases in order to control sensitive information. Figure 9 exacerbates the concept with the inter-organizational integration of data and information system architectures across the tightly coupled NSE network.

The inter-organizational complexity of Figure 9 becomes even more extreme as other institutions that are more loosely or indirectly coupled to the NSE network are included. These institutions may include universities, other FFRDCs, government agencies, suppliers, etc. Furthermore, there is a wildcard in that all of the individuals within the NSE organizations also have access to open source data and data systems outside of the network that is not elaborated on in this case study.

This page left blank

3. DISCUSSION

Most of the government reports on national security related data issues focus on cyber security breaches that look in hindsight to assess damage. More often than not, the entity that was responsible for maintaining the data did not comprehend what it was they were controlling from a data in context perspective.

Cyber security is a delay mechanism to stop those that do not have proper access authority. The primary concern of cyber security is focused on building near-impenetrable cyber fences, not about comprehending the data within. Data breaches are commonplace within the largest public and private institutions that are responsible for multiple dimensions of sensitive information. These government reports would be good to review as exemplars of how data grows faster than technology, describing government risks, cyber challenges, and incidences which include the Equifax breach, and the state of emergency health information systems (GAO-20-123, GAO-20-631, GAO-20-691, GAO-19-105, GAO-18-622, GAO-18-622, GAO-18-210).

The impact of the growth of data and technology highlights concerns in the study of contract theory and supply chain management. More recently in December 2020, multiple US government agencies were part of an adversarial hack. According to the New York Times, these government agencies include the Pentagon, intelligence agencies, nuclear labs, and Fortune 500 companies (Newman 2020). The cyber security of multiple government agencies was partially relying on a third-party contractor. Most of the news wants to point to the third party contractor and the wide spread use of its software within the supply chain. However, the crux of the problem is the conversion of data to information. Once multiple data are compiled, they may result in context that has dangerous consequences to US interests and lives. These government agencies are now scrambling to comprehend what the data thief may be able to comprehend from the facts and figures they gathered. In this process, they may discover that the data that was acquired should not have been integrated in the first place. The integration of these agencies may be the result of the conflict that has been mostly highlighted in the intelligence literature—Need-to-Know versus Need-to-Share.

Intelligence research explores important topics related to the control of what is typically described as Need-to-Know versus Need-to-Share—experiencing a major roadblock in the revelations of Edward Snowden. These constructs may be useful in the continued exploration of the impacts on control of data in the *Data Revolution*. The literature describes policies, procedures, and technologies that link people, systems, and information from government agencies. It describes technical and policy barriers and the concept of "connecting the dots" for intelligence work. What this literature does not focus on is what the agencies being described comprehend of their raw data and what sharing it may mean from a classification or data in context perspective depending on the collaborators access to specific databases (Best 2011).

Governance has a wide range of breadth and ambiguity in definition as Lynn, Heinrich, and Hill (2001) describe:

> *Governance generally refers to the means for achieving direction, control, and coordination of wholly or partially autonomous individuals or organizational units on behalf of interests to which they jointly contribute.*

Like the ambiguity of governance in the PA literature, the concept of data governance is also fraught with ambiguous definitions. Nielsen did a comprehensive review of data governance literature and for the purposes of research defined it as "companywide processes that specify decision-making rights and responsibilities aligned with organizational goals to encourage desirable behavior in the

treatment of data as an organizational asset (2017)." These literatures would be good starting points to comprehend the crosscutting fields exploring how to improve governance specifically to data within organizations.

There is extensive literature on the concept of big data, but from a control of sensitive information or a public sector national security context, the literature is lacking. According to Van Pyvelde (2017), scholars have tended to focus on issues of privacy and liberties.

> *The absence of a comprehensive study on big data and national security decision-making is problematic because it limits researchers' ability to consider the implications of the big data 'revolution' in the field of security.*

Weber and Khademian described the potential for networks to govern complex wicked problems. Examining networks "as an alternative to the limitations of hierarchical and fragmented systems in public policy development and delivery and as a more democratic means of developing public policy (2008)." Their article has some good ideas and frameworks that may be beneficial to comprehending data as they examine the transfer of knowledge within networks.

Within contemporary PA literature, a concept pertinent to the field of security—personal privacy—has been gaining salience. With contemporary issues such as the Cambridge Analytica scandal that sought to influence political elections, this literature classifies the emergence of algorithms in the public and private sectors as a wicked problem (Andrews 2017).

Eric Schlosser (2014) wrote a book about the extreme consequences of even the smallest risks becoming realized related to high-reliability NW systems. Schlosser's book is a valuable source as it describes a more systematic view of the breadth of NW systems that includes other high-reliability systems including the Nuclear Command, Control, and Communications (NC3).

4. SUMMARY

Tame problems are well defined and easily addressed whereas wicked problems are dynamic and complex, with no clear definition or solution involving multiple stakeholders in multiple organizations (Emerson and Nabatchi 2012). The case study utilizes elements of literature from KM and networks to tell a story of a wicked problem in security—controlling the conversion of data to information. The purpose of this case study is to illustrate and gain a contextual comprehension of the wicked problem to inform discussions.

Organizations—especially high-reliability organizations responsible for high-reliability systems—rely on expertise and the creation and maintenance of both explicit and tacit knowledge. As a result, these organizations must preserve and grant access to raw data and information systems. Due to the nature of the scope of responsibility, these organizations must share this raw data and information intra-organizationally across functional areas as well as inter-organizationally especially within responsible networks.

The wicked problem of controlling the conversion of data is specific to sensitive information within an organization and across a network responsible for its control. As the growth of data and the technological means of acquiring, storing, retrieving, analyzing, and sharing it become simpler and more cost effective, the complexities that organizations face in controlling data to avoid context by association are not trivial. Furthermore, the pivot towards inter-organziational integration of data and information systems between network participants exacerbates these complexities.

So where does this case study go from here? It is often stated that technology moves faster than policy. The *Data Revolution* has shown that data are growing faster than the technological and methodological means to comprehend it. While integrating data and information systems both intra-organizationally and inter-organizationally across tightly coupled networks is driven by good intentions and perceived efficiencies, it will require innovation from the security community to deal with the wicked problem—controlling the conversion of data to information.

This page left blank

REFERENCES

[1] Alavi, Maryam, and Dorothy E. Leidner. 1998. Knowledge management and knowledge management systems conceptual foundations and an agenda for research. Fontainebleau: INSEAD.

[2] Agrifoglio, Rocco. 2015. Knowledge Preservation Through Community of Practice: Theoretical Issues and Empirical Evidence. https://doi.org/10.1007/978-3-319-22234-9.

[3] Andrews, Leighton. 2019. "Public administration, public leadership and the construction of public value in the age of the algorithm and 'big data'". Public Administration. 97 (2): 296-310.

[4] Best, Richard A. 2011. Intelligence information: need-to-know vs. need-to-share. http://ezproxy.library.yorku.ca/login?url=https://www.heinonline.org/HOL/Page?handle=hein.crs/crsmthaaalm0001&id=1&size=2&collection=congrec&index=alpha/I_crs.

[5] Boisot, M., Canals, A. 2004. Data, information and knowledge: Have we got it right?, In: Journal of Evolutionary Economics, vol. 14, no. 1, pp. 43–67.

[6] Burn-Murdoch, John. "Study: less than 1% of the world's data is analysed, over 80% is unprotected." The Guardian. 2012. [Last Accessed December 2020] https://www.theguardian.com/news/datablog/2012/dec/19/big-data-study-digital-universe-global-volume

[7] Bowers, JS. 2014. Program Integration in a Nuclear Weapons Product Realization Environment. SAND2014-20242C.

[8] Box, Richard C. 2018. *Essential History for Public Administration.* Melvin & Leigh Publishers, Irvine, CA.

[9] Bradshaw, Rick and Julian, Cotye. 2014. B61-12 Life Extension Program Project Controls System Overview.

[10] Chen CC, Medlin BD, Shaw RS. 2008. A cross-cultural investigation of situational information security awareness programs. Information Management & Computer Security; 16:360-376.

[11] Cook, S. D. N., and J. S. Brown. 1999. "Bridging Epistemologies: The Generative Dance Between Organizational Knowledge and Organizational Knowing". ORGANIZATION SCIENCE. 10 (4): 381-400.

[12] Dawes, Sharon S., Anthony M. Cresswell, and Theresa A. Pardo. 2009. "From "Need to Know" to "Need to Share": Tangled Problems, Information Boundaries, and the Building of Public Sector Knowledge Networks". Public Administration Review. 69 (3): 392-402.

[13] De Marco, Marco. 2012. Information systems crossroads for organization, management, accounting and engineering ; ItAIS: The Italian Association for Information Systems. Heidelberg: Physica-Verlag.

[14] Eggers, William D. 2016. "Government's cyber challenge: Protecting sensitive data for the public good." Deliotte Review Issue 19. [Last Accessed December 2020] https://www2.deloitte.com/us/en/insights/deloitte-review/issue-19/protecting-sensitive-data-government-cybersecurity.html

[15] Emerson, Kirk, and Tina Nabatchi. 2015. Collaborative governance regimes. Washington, D.C: Georgetown University Press.

[16] Frederickson, H. George, Kevin B. Smith, Christopher W. Larimer, and Michael J. Licari. 2018. The public administration theory primer.

[17] Gallaugher, John. 2011. Information Systems: A Manager's Guide to Harnessing Technology. Flatworld Knowledge, Irvington, NY.

[18] Hand, David J., Heikki Mannila, and Padhraic Smyth. 2001. Principles of data mining. Cambridge, Mass: MIT Press.

[19] Harari, Yuval Noah, and Derek Perkins. 2017. *Sapiens*. HarperCollins. http://api.overdrive.com/v1/collections/v1L2BaQAAAJcBAAA1M/products/807484e0-6f0d-42a6-b976-0546afed02dd.

[20] Herron, Kerry, Jenkins-Smith, Hank, and Silva, Carol. 2012. Sandia National Laboratories, United States. *US Public Perspectives on Security*. Washington, D.C.: United States. National Nuclear Security Administration.

[21] Kettl, Donald F. 2018. "The Big Deal about Big Data." National Academy of Public Administration. [Last Accessed December 2020]

[22] https://www.napawash.org/standing-panel-blog/the-big-deal-about-big-data

[23] Klijn, E-H. (Erik-Hans), and Koppenjan, J.F.M. (Joop). 2012. Governance network theory: Past, present and future. Policy and Politics Vol. 40 No. 4, Pp. 587-606. http://repub.eur.nl/pub/74946.

[24] Lake, Peter, and Paul Crowther. 2013. "Data, an Organisational Asset". Sheffield Hallam University, Sheffield, UK

[25] Lane, C. and R. Bachman (eds.) 1998. Trust within and between organizations; conceptual issues and empirical applications. Oxford: Oxford University Press

[26] Lynn, Laurence E., Carolyn J. Heinrich, and Carolyn J. Hill. 2001. Improving governance: a new logic for empirical research. Washington, Dc: Georgetown University Press.

[27] Malhotra, Yogesh. 2005. "Integrating knowledge management technologies in organizational business processes: getting real time enterprises to deliver real business performance". Journal of Knowledge Management. 9 (1): 7-28.

[28] March, James G., and Johan P. Olsen. 1995. Democratic governance. New York: Free Press.

[29] March, James G., and Johan P. Olsen. 2010. Rediscovering Institutions. Riverside: Free Press. https://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=1974423.

[30] Milward, H. B., and Keith G. Provan. 2001. "Governing the hollow state". Peace Research Abstracts. 38 (2).

[31] Newman, Lily Hay. 2020. "No One Knows How Deep Russia's Hacking Rampage Goes." New York Times. https://www.nytimes.com/2020/12/14/us/politics/russia-hack-nsa-homeland-security-pentagon.html

[32] Nonaka, Ikujiro. 1994. "A Dynamic Theory of Organizational Knowledge Creation". Organization Science. 5 (1): 14-37.

[33] Parkins D. 2017. "Regulating the internet giants: The world's most valuable resource is no longer oil, but data". Economist (United Kingdom). 413 (9035).

[34] Peters, B. G. 2017. What is so wicked about wicked problems? A conceptual analysis and a research program. Policy and Society, 36(3), 385–396.

[35] Polanyi, Michael. 2011. The tacit dimension. Chicago, Ill: University of Chicago Press.

[36] Powell, R. 2003. Nuclear deterrence theory, nuclear proliferation, and national missile defense. International Security, 27(4), 86-118. 2. NPR 2018

[37] Ransbotham, S., Kiron, D., Prentice, P. K. 2016. Beyond the Hype: The Hard Work Behind Analytics Success, In: MIT Sloan Management Review, vol. 57, no. 3, pp. 1–16.

[38] Rethemeyer, R. K., & Hatmaker, D. M. 2008. Network Management Reconsidered: An Inquiry into Management of Network Structures in Public Sector Service Provision. Journal of Public Administration Research and Theory, 18(4), 617-646.

[39] Rosenthal, D. S. and Daniel C. Rosenthal. 2012. "The Economics of Long-Term Digital Storage." Stanford University Libraries, Stanford, CA.

[40] Sagan, S. D. 1994. The perils of proliferation: Organization theory, deterrence theory, and the spread of nuclear weapons. International Security, 18(4), 66-107.

[41] Sandia National Laboratories. Nuclear Weapons. [Last Accessed December 12, 2020] https://www.sandia.gov/missions/nuclear_weapons/index.html

[42] Schlosser, Eric. 2014. *Command and Control Nuclear Weapons, the Damascus Accident, and the Illusion of Safety*. Large Print Pr.

[43] Scholl, Margit. 2018. *Information Security Awareness in Public Administrations*. https://torl.biblioboard.com/content/8e3fd2c1-4e33-4715-a5bd-f9a1c00c959d?organizationId=1f7368e7-f10b-49a1-8ced-2d9476279974

[44] Tuomi, Ilkka. 1999. "Data Is More than Knowledge: Implications of the Reversed Knowledge Hierarchy for Knowledge Management and Organizational Memory". Journal of Management Information Systems. 16 (3): 103-117.

[45] United States. 2018. Nuclear Posture Review. https://purl.fdlp.gov/GPO/gpo88923.

[46] U.S. Department of Energy. December 1991. "Identification of Classified Information," Office of Classification, Chap. IV, Part B

[47] U.S. Department of Energy. 1997. "Information Security Program." DOE 471.2A

[48] U.S. Department of Energy (DOE) National Nuclear Security Administration (NNSA). [Last Accessed December 12, 2020] https://www.energy.gov/nnsa/locations

[49] U.S. Government Accountability Office. *Information Technology Federal Agencies and OMB Need to Continue to Improve Management and Cybersecurity.* GAO-20-691

[50] U.S. Government Accountability Office. *Critical Infrastructure Protection.* GAO-20-631

[51] U.S. Government Accountability Office. *Cybersecurity Selected Federal Agencies Need to Coordinate on Requirements and Assessments of States*. GAO-20-123

[52] U.S. Government Accountability Office. *Information Security Agencies Need to Improve Implementation of Federal Approach to Securing Systems and Protecting against Intrusions.* GAO-19-105

[53] U.S. Government Accountability Office. *High Risk Series Urgent Actions Are Needed to Address Cybersecurity Challenges Facing the Nation.* GAO-18-622

[54] U.S. Government Accountability Office. *Data Protection Actions Taken by Equifax and Federal Agencies in Response to the 2017 Breach.* GAO-18-559

[55] U.S. Government Accountability Office. *Electronic Health Information CMS Oversight of Medicare Beneficiary Data Security Needs Improvement.* GAO-18-210

[56] U.S. Government Accountability Office. *National Nuclear Security Administration Additional Actions Needed to Collect Common Financial Data.* GAO-19-101

[57] Van Puyvelde, Damien, Stephen Coulthart, and M. Shahriar Hossain. 2017. "Beyond the buzzword: big data and national security decision-making". International Affairs. 93 (6): 1397-1416.

[58] Vergun, David. "DOD Official Outlines US Nuclear Deterrence Strategy." Defense.gov. 2020. https://www.defense.gov/Explore/News/Article/Article/2334600/dod-official-outlines-us-nuclear-deterrence-strategy/

[59] Weber, Edward P., and Anne M. Khademian. 2008. "Wicked Problems, Knowledge Challenges, and Collaborative Capacity Builders in Network Settings". Public Administration Review. 68 (2): 334-349.

[60] Wiig, K.M. 2002. "Knowledge management in public administration", Journal of Knowledge Management, Vol. 6 No. 3, pp. 224-239.

[61] Wilson, M. and Hash, J. 2003. Building an Information Technology Security Awareness and Training Program, Computer Security Division, Information Technology Laboratory, National Institute of Standards and Technology, Gaithersburg, MD.

DISTRIBUTION

Name	Org.	Sandia Email Address
Technical Library	01911	sanddocs@sandia.gov

Sandia National Laboratories

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.