National Security Classification
Changing an Outdated System

A White Paper June 2017
By: Harry Cooper

Current State of National Security Classification

The US government has created and ingested into vast electronic systems potentially 10s and maybe 100s of billions of pages of electronic information and much of that is considered classified. This information is accessed by upwards of 5 million users[1] who have secret or higher clearances. Requirements to re-validate the trustworthiness of those who hold Top Secret clearances every 5 years are impossible to meet, resulting in a backlog of as many as 500,000[2] users. Some of these users may no longer be eligible, but they go undetected.

Millions of cleared users access millions of computer workstations worldwide. They have access to amounts of information far beyond any individual's needs under the belief that only the users themselves can determine what they need to do their job. The reaction to the terrorist attacks of 9/11 was the perception that these attacks might have been prevented if more people had better access to more information that had been collected and was stored on classified government IT systems. As a result, rules that limited each person's access to classified information were rewritten, and the stage was set for access by millions of cleared users to billions of pages of classified information.

[1] 5.1 million Americans have security clearances. That's more than the entire population of Norway; Washington Post, March 24, 2014; by Brian Fung.

[2] Periodic reinvestigation backlog more than doubled in 2015; Federal News Radio; April 1st, 2016; by Nicole Ogrysko.

Information is determined to be classified using standards developed during the Cold War. In the current world of asymmetric warfare, and a world where information is ubiquitous, it is far less clear what information causes identifiable damage to US national security.

The national security classification system in the United States is broken. Developed in 1940 by President Roosevelt, the system was changed by presidents Truman, Eisenhower, Kennedy, Nixon, Carter, Reagan, Clinton and Obama, but many of the initial fundamental concepts remain and the system has not changed or grown to meet the needs of the digital age.

The Current U.S. Classification System

Fundamentally the system consists of three levels of classification and three levels of security clearance. It is predicated on the concept that a small percentage of uniformed military and government civilians would have national security clearances. Of this small percentage with clearances, at least at the confidential level, substantially fewer people would be given secret clearances and even fewer than that would have Top Secret. The system envisioned that perhaps only 10% would be cleared with as few as 2% having Secret and less than 1% with Top Secret.

Four high level concepts make up the current security paradigm;

1) Determining who should have a clearance to access classified information,

2) Determining what information should be classified,

3) Limiting the risk of compromise of classified information by ensuring that access to that information is limited to the access necessary to fulfill the foreign relations, defense and intelligence missions of the United States, and

4) Providing electronic and physical safeguards sufficient to prevent those without authorized access from obtaining sensitive information that would damage the national security of the United States.

The National Security Clearance Process

Today over 60% of the federal workforce has a security clearance. That is a staggering 5.1 million military, federal civilians, and contractors.[3] This number was reported in 2014 and it seems reasonable to assume that the total number and percentage of the eligible workforce has increased. In 2017 access to any of the electronic systems that process classified information requires that every user be cleared. For the SIPRNet[4] secret level system, that carries vital communications to most of the US Military, a Secret clearance is required. For the JWICS[5] network that carries most national intelligence and the imagery that the military relies on for operations, all users must have a Top Secret clearance.

All intelligence elements, the FBI and many military commands require 100% of employees to have a security clearance. Many of these clearances are expressly for access to classified systems even though many of the individuals do not need significant access to classified information. Many of these organizations require all clearances to be at the Top Secret level.

The number of clearances is overwhelming the system that grants and maintains them. Federal News Radio reported that at the end of 2014 there were over 32,000 security clearance reinvestigation cases in backlog.[6] In September of 2016 Federal News Radio reported that "...part of the delay for 2017 is the NBIB plans to award a contract to increase vendor support around investigations to help further address a backlog of security

---

[3] 5.1 million Americans have security clearances. That's more than the entire population of Norway; Washington Post, March 24, 2014; by Brian Fung.

[4] Secret Internet Protocol Router Network; www.dlhra.mil/perserec/osg/s1class/siprnet.htm

[5] Joint Worldwide Intelligence Communications System; https://en.wikipedia.org/wiki/Joint_Worldwide_Intelligence_Communications_System

[6] Periodic reinvestigation backlog more than doubled in 2015; Federal News Radio; April 1st, 2016; by Nicole Ogrysko.

clearance cases that has ballooned to more than 500,000 over the last two years".[7]

The need for security clearances is on the rise each year and the ability to complete initial and periodic reinvestigations is falling further behind. The way background investigations are conducted is also much the same as it has been for decades. Agents talk to neighbors who often know very little about those who live in the neighborhood. Checks with local law enforcement and schools are also complicated by an increasingly mobile society. Financial checks were added in the 90s in response to identification of cold war spies whose motive was money.

Gone are the days when Americans are born, work, and die in the same community. Social media is potentially a much more efficient way to understand the character of people who must be trusted with secrets, yet the USG has only just begun to view this medium to investigate those who must be trusted.

It is also not clear in the 21st century if the standards for determining trustworthiness are still valid. The standards largely rely on factors that do not predict those who might commit treason out of their own deep-seated beliefs. Famous recent leakers Manning and Snowden do not follow the models designed during the Cold War. Neither of them was recruited by a foreign intelligence service either with pressure to expose their lifestyle secrets or for money. It is far less clear now what causes Americans to turn against their country. This is a powerful signal that the secrecy paradigm, with respect to vetting those who handle secrets, largely unchanged since the 1960s, simply doesn't work anymore.

[7] **Cost for security clearances expected to increase in 2017, 2018;** Federal News Radio; September 7th, 2016; by Jason Miller.

Classified Information

The second critical factor of the current paradigm is guidelines and processes by which information is identified as sensitive and categorized as classified national security information. There is a common perception outside of government that too much information is classified that shouldn't be. While the actual standards for deciding what is classified plays a role, so does the unparalleled growth of electronic information.

The growth in the volume of classified information in the federal government is virtually out of control and is incalculable. We know that the amount of classified information that existed up through 1975 was approximately 1 billion pages[8]. This count was completed when President Clinton ordered, in 1995[9], that all records dated up to 1975 be declassified. That count included only non-electronic records as there was no means then to assess the volume of electronic records that had been created.

In 2015 the General Counsel for the National Archives[10] indicated that when President Obama leaves office in 2017 NARA expects that more than 1 billion pages of electronic documents (including email messages) will be moved to his Presidential Library. This single electronic collection from one 8-year presidential term is equivalent to all classified information held by NARA prior to 1975.

Electronic information that now dates back to the late 70s, began skyrocketing in volume in the early 90s. Electronic information, in 2017, is measured in units of measure that were not even words in the 70s; "petabytes".

---

[8] **ISOO Annual Report to the President for FY 2000**; Steven Garfinkle, Director ISOO; Dated September 17, 2001

[9] **EO 12958**; Classified National Security Information; President William Clinton, 14 April 1995

[10] **Discussion with Mr. Gary Stern**: NARA GC, at the 1st Symposium on Electronic Declassification held at the Institute for Defense Analysis.

Classified holdings of the major US Intelligence agencies today are measured in petabytes (PB). Although not confirmed as accurate, Wikipedia reports that a single data center constructed by NSA was planned to hold between 3 and 12 exabytes[11] of data.[12] An exabyte (XB) is 1000 Petabytes. It is widely understood that holdings within intelligence agencies today are measured in petabytes with reasonable expectations that agencies would have between 2 and 20 petabytes of data as part of their internal electronic systems that process classified information. All data on a classified system is presumed classified until reviewed and approved for transmission to an unclassified system or release to the public.

The table below provides information on how many pages of textural material can be expected from a range of storage sizes.

File Sizes in Pages					
Size	**Word**	**Excel**	**Powerpoint**	**Email**	**Average**
GB	64,000	166,000	15,500	100,000	86,875
TB	64,000,000	166,000,000	15,500,000	100,000,000	86,875,000
PB	64,000,000,000	166,000,000,000	15,500,000,000	100,000,000,000	86,875,000,000
XB	64 Trillion Pages	166 Trillion Pages	15.5 Trillion Pages	100 Trillion Pages	86.9 Trillion Pages

Figure 1 "File Sizes in Pages"

Systems such as SIPRNet and JWICS are world-wide classified networks that produce more than a billion email messages alone each year[13]. Information collected by intelligence agencies that must be classified to protect the sources and methods is measured in petabytes. Each petabyte of textural data is about **86 billion pages**.[14]

[11] Exabyte is a unit of information equal to one quintillion (1018) bytes, or one billion gigabytes.

[12] https://en.wikipedia.org/wiki/Utah_Data_Center

[13] A guess based on knowledge that a single agency produces over 500 million email messages on its classified system each year.

[14] How Many Pages in a Gigabyte; Discovery Series Fact Sheet; undated, Lexus Nexus

Reinventing a Broken System

SIPRnet was accessed by Private Manning and was the source of more than half a million documents published by WikiLeaks. Edward Snowden, a contractor for NSA, had access to the vast NSA network at the Top Secret level. Snowden is accused of stealing a large number of classified information, possibly as much as 1.7 million documents.[15]

The link between vetting and classification levels is a systemic failure. For access to Top Secret (TS) information individuals need a Top Secret Clearance. In the digital age this means that access to an IT system that is approved for Top Secret information, all users must have this ultimate clearance. No distinction is made between users who only rarely need TS information and those who access it every day. Everyone on the system must have clearances for all the information. The clearance system is predicated on the concept that the higher the security clearance the more trustworthy the individual. It certainly seems likely that no such link exists, else there would be much greater compromise of Secret information than Top Secret. Access to millions of pages of Secret information requires the same minimal vetting as access to a single page.

For the lesser secret clearance, individuals must be free of serious convictions and not be listed in a database of known criminals or radicals intent on doing harm to the US. These checks are virtually the same as for everyone who enters government service or military service, police officers, doctors, lawyers, first responders, and many other positions in our society. The vetting done for these positions should be considered adequate for occasional access to any classified information needed in the performance of their jobs and even routine access to information that does not rise to what is now called Top Secret.

Those individuals who access the most sensitive information should have a greater inspection done on their background and beliefs. Strong emphasis should be placed on review of their social media posts, likes,

[15] WikiPedia; https://en.wikipedia.org/wiki/Edward_Snowden

and connections. We have software now available to analyze a person's social media life and determine if that person leans toward those who believe that our nation is not acting in the best interest of its citizens.

More significant than background investigations or "certifications" that a person is able to access information at a specific level, is the need for ongoing audits of each trusted person's official access to information and an ongoing review of their social ties. Technology exists today capable of detecting normal behavior patterns for individuals who access information. Any sudden deviation from that "normal pattern" should trigger a review. Individuals in trusted positions should get accustomed to being reviewed often, especially when their job duties change.

Better safeguards on information must also be in place. Systems administrators and software developers who work in the most sensitive systems often have access to all data in those systems. Data is rarely encrypted at rest meaning that these administrators and developers have access to the content of the data files. This is a fundamental flaw in our system that could be fixed by changing the priority from vetting individuals to protecting the information that they may access.

One of the core principles of the classified information system from WW-II through the 90s was "need-to-know." This concept embodied the idea that just because a person held a clearance, they did not necessarily need access to all information classified at that level. Everyone who held a security clearance was compelled by this principle to validate that someone else who asked for access to information they held actually needed the information. Widespread belief that this responsibility became a source of power for those who held information and caused them to not share that information led to the demise of this tool. Under the 2009 Obama Executive Order "need-to-know"[16] was transformed to

[16] EO 13526 §6.1 (dd) "Need-to-know" means a determination within the executive branch in accordance with directives issued pursuant to this order that a prospective recipient requires access to specific classified information in order to perform or assist in a lawful and authorized governmental function.

make need to know a presumption. Agency heads are further compelled to ensure that classified information is available "to the maximum extent possible to [cleared individuals]."[17]

The creation of massive electronic repositories filled with sensitive information, the need for millions of employees to access these systems, and the shift in policy to permit the widest possible access to information, created a perfect storm that is leading to a catastrophic failure of the classification system.

A New Approach

Changing the massive classification system in the United States is no easy task. Only a clean break with the current system will work; the time has long since passed when incremental change could lead to a dramatic new approach. Sustaining the 1940s model in the digital age is expensive, ineffective, and itself a risk to the security of the United States.

A new classification model can be characterized as follows:

1) Redefine Damage to National Security: The threshold condition that must be met to classify information is that it would damage national security if exposed. The standards for the kinds of information that meets this test must have greater precision and must be crafted in a way that is clear, understandable, and readily apparent to everyone,

2) Change the Vetting Process: The way we approve people for access to sensitive information is out of date, ineffective and costly. We must continue to make a judgment about an individual's suitability for being trusted with our nation's secrets, but we must do this in a manner consistent with our culture, our changing times. We should

[17] EO 13526 § 4.2 Distribution Controls. (a) The head of each agency shall establish procedures in accordance with applicable law and consistent with directives issued pursuant to this order to ensure that classified information is accessible to the maximum extent possible by individuals who meet the criteria set forth in section 4.1(a) of this order.

identify those jobs that require good character, non-criminal behavior, and professional stability. The tests already applied for many jobs including entry into the federal workforce, police, law, medicine and first responders should be deemed reasonably effective at identifying those who should not be trusted. We also need to de-link the vetting from the information. When a trusted individual needs limited access to exceptionally sensitive information that access should be granted. We no longer have the luxury of assigning clearances that match the levels and then telling an individual cleared at one level that he or she can't have access to information at a higher level that might save her life or the lives of others, and

3) Transform Access to Information: Giving any one person access to more information than he or she can read in a lifetime is not necessary. Recent high-profile leakers have leaked more information than they could personally have read or even known the content of. Their access was not limited to what they needed by anyone's definition of need. Steps must be taken to limit access in the following ways:

    a. Identify Personal Need: Everyone with access to classified information should have a personal profile that describes his or her professional need for information. Academic preparation, job duties, and roll in government mission must be included in the profile. Information must be identified as meeting a range of information needs and these needs must be identified as responsive to specific individual profiles.

    b. Limit Access to Volumes of Information: Each person with access to information who queries information systems containing millions or billions of pages should never be given access to more information than that person can read (such as Private Manning was). Users must be required to refine queries until a manageable amount of information is the result. People can't "need" more information than they can read in a

year or a lifetime. Electronic tools exist to help refine searches, make searches more precise, and – more importantly – to exclude information that exceeds the needs of the individual.

c. Improve System Security: Sensitive information simply must be encrypted at rest. Systems administrators (like Snowden) should not have access to the text of information that resides within a system unless they have a specific need for the information. Our information infrastructure must be improved to ensure that vast quantities of sensitive information can be protected even with millions of users querying that information and seeking answers to mission-relevant questions.

Conclusion:

Nothing short of a complete overhaul will fix classification. We must reevaluate what makes information sensitive, we must change the way we decide who should have access to classified information, and above all we must take steps to ensure individuals get all the information they need, but are not allowed to access more information than they can ever use.

No new paradigm can ever fix the problems we have created in our legacy information. A new approach will be expensive and require that all government employees be trained in the new system. Funding now devoted to personnel security must be refocused on better systems design to control what information each user receives and uses.

The author is a retired CIA official who has over 30 years of experience in managing national security classification programs. He participated in the drafting of the Presidential Executive Orders and federal regulations related to classification and secrecy. He is also the author of numerous white papers on national security classification. This manuscript was approved by CIA for publication in accordance with the author's secrecy agreement and CIA regulations.