NOAA Data Loss Prevention Plan

Office of the Chief Information Officer
Governance and Portfolio Division
August 2016

Background

[01] The protection of sensitive and personal information is more important than ever with electronic communications becoming increasingly prevalent. Safeguarding Personally Identifiable Information (PII) in the possession of the Federal Government and preventing its breach are essential to retaining the trust of the American public[1]. This responsibility is shared by officials accountable for administering operational, privacy, and security programs. PII is any information that, by itself or in combination with other information, may be used to uniquely identify an individual. Within NOAA systems, this primarily can consist of Social Security Numbers (SSN), names, addresses, dates and places of birth, bank account numbers, e-mail addresses, telephone numbers, and passport numbers. The Office of Management and Budget (OMB) and the Department of Commerce (Commerce) released several memoranda to address the issue of safeguarding PII[2].

[02] This plan is intended as a framework for future action that will address user and system specific restrictions, controls, use cases, parameters, and other actions implemented based on the needs of individual systems and mission goals. This plan is intended to satisfy the implementation plan obligations to meet the minimum Privacy DLP Standards within 1 year as outlined in the April 15, 2016 Memorandum entitled “Departmental Privacy Standards for Commerce Data Loss Prevention (DLP) Security Tools”, as well as the corresponding May 3, 2016 data call issued by Commerce.

NOAA Data Overview

[01] NOAA provides the data, science, and information that allow the economy to function effectively and grow sustainably. NOAA helps to ensure a competitive economy by monitoring and predicting changes in the Earth's environment, protecting lives and property, and conserving and managing the nation's coastal and marine resources. NOAA’s data portfolio mirrors the diversity and complexity of its mission … and NOAA is very complex! Our mission and data diversity includes:

- 21,335 Staff (federal, contractor, associate)
- 435 Buildings
- 122 Weather Forecast Offices
- 13 River Forecast Centers
- 1,429 Real-Time Weather Stations
- 17 Satellites
- 8 Buoy Networks: 1042 Stations Deployed
- 13 National Marine Sanctuaries and 1 Marine National Monument
- 286 Data Centers

[1] The definition of PII can be found in the OMB Memorandum M-06-19, July 12, 2006.

[2] See, e.g., Memorandum from David A. Sampson, RE: Safeguarding Personally Identifiable Information, November 6, 2006.

- 94 Federal Information System Management Act Systems
- 33 Exhibit 300 IT Investments

NOAA DLP Strategy Overview

[01] NOAA uses, and will further deploy, a "Defense in Depth" approach to DLP. NOAA will use existing operational controls and privacy enhancing technologies. These include PII identifying solutions, encryption, firewalls, authorized use system access controls, and system audit logs. To further reduce the risk of compromise of sensitive PII in agency communications, NOAA will implement a Data Loss Prevention (DLP) solution set that monitors network communications and prevents sensitive PII from leaving the network, in addition to other sensitive data, as determined when the scope and capability of the solution is determined. Other sensitive data may include law enforcement sensitive data, business identifiable data, or other data sets for which the DLP solution can feasibly be leveraged. Each of these data sets may have one or more data owners, who will classify the information type, as described in the fourth development step below. In addition to these technical controls, NOAA utilizes administrative policies and procedures, as well as privacy training, to further safeguard information privacy and control access to information systems and information assets.

[02] NOAA conducts Privacy Threshold Analyses, (PTA's), and, where applicable, Privacy Impact Assessments (PIAs) on all information systems to ensure privacy implications are addressed when planning, developing, implementing, and operating information technology (IT) systems that maintain information on individuals. NOAA utilizes a PIA template and guidance on conducting PIAs. The NOAA Bureau Chief Privacy Officer (BCPO) collaborates with system owners and IT security professionals to assess existing, new, or proposed programs, systems or applications for privacy risks, and recommends methods to protect individual privacy.

[03] The NOAA DLP solution(s) will be designed to monitor and prevent data from being leaked. NOAA's DLP strategy, however, is to make sure the DLP solution(s) are as efficient and effective as possible. DLP needs to be rationally deployed in order to ensure that false positives do not overwhelm the system and the capacity of NOAA privacy and cyber security managers and staff. DLP tools such as McAfee Security, as powerful as they may be, require careful and organized deployment, otherwise reported incidents may be of little value.

[04] In response to the OMB M-07-16, Safeguarding Against and Responding to the Breach of Personally Identifiable Information, and NIST SP 800-122, at 2.1, the NOAA Governance and Portfolio Division is leading the DLP initiative to promote secure practices in electronic communications (e-mails and Internet access) on the NOAA network to protect Controlled Unclassified Information (CUI) data. Taking a phased approach, the DLP initiative may include plans to implement specific solutions, for example, McAfee's Data Loss Prevention (DLP) commercial off-the-shelf software solution that is capable of identifying and tracking a number of NOAA's defined PII datasets. The NOAA DLP solution(s) will be designed to give NOAA an enterprise view into where it's most sensitive data are stored, who has access to the data, and where and by whom the data are sent outside the NOAA network. By using this information, NOAA can spot broken business processes and reduce the overall risk of exposure. The DLP solution(s) will take a data-centric approach to security, in which policies can be developed around the content that should be protected and then deployed across multiple data states or functionalities, such as identifying, monitoring, and preventing.

DLP Development Steps

[01] A multi-layered approach will be applied to prevent data leakage for all routes and states. Data is classified under one of several schemes like data in motion, data in use, and at rest; or by data in-store, in-use and in-transit.

- Data in motion: Data that needs to be protected when in transit including HTTP/S, S/FTP/S, IM, P2P, SMTP.
- Data in use: Data that resides on end user workstation and needs to be protected from being leaked through removable media devices like USB, DVD, CD's.
- Data at rest: Data that resides on local storage media or server storage.

[02] Each of these layered types of data will be considered during deployment to maximize the prevention of data leakage. A multi-step approach to deployment, shown below, will be used as well. These steps are discussed in detail below.

1. Define policies
2. Identify sensitive data
3. Determine information flows
4. Identify data owners
5. Identify deployment scenarios
6. Plan DLP operations
7. Deploy DLP product(s)

Define Policies

[01] NOAA will build policies to protect the sensitive data. Every policy will consist of some rules, such as to protect credit card numbers, PII, and social security numbers, if such policies are not already in place. If there is a requirement for NOAA to protect sensitive information and a DLP product such as McAfee DLP does not support it out of the box, then NOAA will create rules using regular expressions (regex). It should be noted that DLP policies at this stage will be defined and not applied.

[02] Those policies will reflect the internal controls relate to management's plans, methods, and procedures used to meet their mission, goals, and objectives. Internal controls include the processes and procedures for planning, organizing, directing, and controlling program operations. They include the systems for measuring, reporting, and monitoring program performance. We determined that the following internal controls were relevant to our objective: Commerce Directives; and OMB, White House, and National Institute for Standards and Technology (NIST) guidelines. Each of these sources provide a framework for implementing an automatic tool to monitor transfers of PII and for developing, or implementing, a commercial off-the-shelf product. We are evaluating these controls against an enterprise life cycle approach, and by reviewing enterprise life cycle commercial off-the-shelf artifacts and documents supporting the procurement, budget, and expenses for the DLP solution.

[03] Policies will take into consideration existing guidelines and recommendations as well as other factors, such as impact and dependency for other systems, also needed to be considered when implementing a DLP solution. The National Institute for Standards and Technology Special Publication 800-122, Guide to Protecting the Confidentiality of Personally Identifiable Information[3][1], recommends that agencies implement automated tools, such as a network data leakage prevention tool, to monitor transfers of PII and to monitor inbound and outbound communications for unauthorized activities. In addition, the Government Accountability Office's Standards for Internal Control in the Federal Government[4][2] provides that application controls should be designed to help ensure completeness, accuracy, authorization, and validity of all transactions during application processing. Controls should be installed as an application interfaces with other systems to ensure that all inputs are received and are valid and that outputs are correct and properly distributed.

Identify Sensitive Data

[01] NOAA will identify all the confidential, restricted, and highly restricted data across the whole organization and across the three categories, i.e. for data in-transit, in-store and in-use. In identifying the sensitive data, NOAA will define the scope within which the DLP Solution will function. Each data set analyzed will be considered as to whether or not leveraging the DLP product would be an efficient use of resources, whether the data is non-sensitive, or whether the DLP would be an effective tool in further securing the data. DLP products work with signatures to identify any restricted data when it is crossing boundaries. To identify the critical data and develop its signatures, there is a term in DLP products known as fingerprinting. Data is stored in various forms at various locations in an organization and it requires identifying and fingerprinting. Various products come with a discovery engine which crawl all searchable data in a given data store, index it and make it accessible through an intuitive interface which allows quick searching on data to find its sensitivity and ownership details.

Determine Information Flows

[01] It is very important for an organization to identify their information flow. NOAA OCIO will prepare a questionnaire to identify and extract all the useful information. A sample questionnaire would address, at a minimum, the following three issues:

- What is a standard data flow, and what should be the source and destination of the identified data?
- What are all the egress points present in the network?
- What processes are in place to govern the informational flow?

[3] National Institute of Standards and Technology, NIST SP 800-122, Guide to Protecting the Confidentiality of Personally Identifiable Information (PII) (April 2010).

[4] Government Accountability Office (formerly known as the General Accounting Office), GAO/AIMD-00-21.3.1, Internal Control: Standards for Internal Control in the Federal Government (Nov. 1999).

Identify Data Owners

[01] Identification of the NOAA staff and line office owners of data is also an important step in the planning strategy of DLP, so a list will be prepared by OCIO of whom to send the notifications to in case any sensitive data is lost. NOAA OCIO will distribute an assessment to identify the owners of each of the different sensitive data elements across the organization. The data owners also will be responsible for classifying the information types[5]. Many types of data will have multiple owners, governed by separate line and staff office policies for the collection and use of that data, depending on mission needs. The assessment will attempt to identify each offices ownership, collection, storage, and transmission of sensitive data so that when an incident occurs, the incident is properly triaged, escalated where necessary, reported[6], and the DLP processes and application are modified and tuned as necessary.

Identify Deployment Scenarios

[01] The following questions arise in identifying potential Deployment Scenarios. Each of these must be addressed prior to agency-wide deployment of a mature DLP solution.

1. Will the Initial Deployment be applied to all of the traffic of data in use, or in motion, or at rest?
2. Alternatively, should NOAA deploy the DLP appliance by copying the network traffic and analyzing it at a different port before deploying it directly to the data states of the network traffic?
3. Should the deployment occur in high availability mode or should we configure in bypass mode?
4. How will the setup of endpoints with the DLP manager occur?
5. How do we maintaining integrity between communication ports and firewalls?
6. How do we ensure proper configuration of a crawling agent?

[02] As discussed above, sensitive data falls under three categories, i.e. data in motion, data at rest and data in use. After identifying the sensitive data and defining policies, NOAA will prepare for the deployment of DLP product(s). DLP deployment scenario of all three categories include :

- Data in motion: Data that needs to be protected when in transit, i.e. data on the wire. This includes channels like HTTP/S, S/FTP/S, IM, P2P, SMTP etc. NOAA will install the DLP protector appliance or software so it is not directly inline with the traffic. This is prudent to start with a minimally invasive method by not putting the appliances inline, to prevent a huge number of false positives or a network outage if the inline device fails. The NOAA approach will be to deploy DLP appliances or software in a span port first,

[5] See, NIST SP 800-60.

[6] Reporting here is referring to both internal reporting to the Office that owns the information, the Bureau Chief Privacy Officer, and N-CIRT as necessary, as well as external notifications (such as Privacy Incident reporting to DOC) and external reporting to OMB. Organizations report annually on specific privacy and security activities in their annual FISMA reports to OMB. The most recent memorandum is OMB M-10-15, FY 2010 Reporting Instructions for the Federal Information Security Management Act and Agency Privacy Management, available at http://www.whitehouse.gov/omb/assets/memoranda_2010/m10-15.pdf

and then after the DLP strategy is mature, then put into inline mode. In order to mitigate the second risk, NOAA may deploy two options: first, deploy DLP in High Availability mode, and second, configure the inline DLP product in bypass mode, which will enable the traffic to bypass the inline DLP product in case the DLP product is down.

- Data in Use: Data that resides on the end user workstation and needs to be protected from being leaked through removable media devices like USB, DVD, CDs, etc. will fall under this category. In Data in Use, an agent may be installed in every NOAA endpoint device like laptop, desktop, etc. which is loaded with policies and is managed by a centralized DLP management server. Agents would be distributed on the endpoints via pushing strategies like SMS, GPO, etc.
- Data in Store: Data that resides on file servers and DBs and needs to be monitored from being getting leaked will fall under this category. All NOAA data that resides in storage servers or devices would crawled using a DLP crawling agent. After crawling, data is fingerprinted to see if any unstructured data is present or not.

Plan DLP Operations

[01] NOAA will need to split the DLP operations into three phases: a triaging phase, a reporting and escalation phase, and a tuning phase. The security operation's team will monitor the alerts fired or triggered by the policies set up in the DLP product. N-CIRT will fine tune the policies as a result of some mis-configurations earlier or due to eventual policy or guidance changes and apply the changes to the DLP product. NOAA will need to identify the staffing, budget, training, and other resource demands that each phase of the DLP Operations will require, and determine the capabilities in effectively carrying out each phase with the available resources.

Deploy DLP Product(s)

[01] Deployment of security components is of no use if they cannot be monitored, and a DLP product is no exception. Below is an overview of what a DLP operation of an organization can be. First of all, the DLP product needs to be created with the right set of policies on the identified data among data at rest, in motion or in transit categories. The DLP operations can be separated into three phases, namely: the triaging phase, the reporting and escalation phase, and the tuning phase. These will need to be modified depending on the nature of the incident identified in the triaging phase for referral to N-CIRT and for DOC notification, as necessary. The triaging phase, incident reporting and escalation, as well as the any parameter modifications and tuning will be carried out in accordance with existing PII/BII Breach Response and Notification Plan.

Conclusion

[01] NOAA will employ a Defense in Depth approach to DLP. NOAA's DLP solution(s) need to minimize deployment and operating costs. As an off-the-shelf product, the McAfee Total Protection, or a similar product solution would potentially be an additional tool within the DID approach to effectively protect PII and BII data wherever it may be.

[02] NOAA has maintained a high awareness of data security, and is vigilant in protecting the sensitive information located within its systems.   These Data Loss Prevention measures will enhance the security of NOAA information systems and maintain the highest level of compliance with all regulatory and guidance documents that govern Data Loss Prevention at the agency[7].

Definitions

Business Identifiable Information (BII) – Information that is defined in the Freedom of Information Act (FOIA) as "trade secrets and commercial or financial information obtained from a person and privileged or confidential." Commercial or financial information is considered confidential if disclosure is likely to cause substantial harm to the competitive position of the person from whom the information was obtained.

Personally Identifiable Information (PII) – Information that can be used to distinguish or trace an individual's identity, such as name, Social Security number (SSN), biometric records, etc., alone, or when combined with other personal or identifying information, which is linked or linkable to a specific individual, such as date and place of birth, mother's maiden name, etc. [OMB M-07-16].

Sensitive Personally Identifiable Information (Sensitive PII) – Sensitive PII is defined as PII which, when disclosed, could result in harm to the individual whose name or identity is linked to the information. Further, in determining what PII is sensitive, the context in which the PII is used must be considered. For example, a list of people subscribing to a government newsletter is not sensitive PII; a list of people receiving treatment for substance abuse is sensitive PII. As well as context, the association of two or more non-sensitive PII elements may result in sensitive PII. For instance, the name of an individual would be sensitive when grouped with place and date of birth and/or mother's maiden name, but each of these elements would not be sensitive independent of one another. For the purpose of determining which PII may be electronically transmitted, the following types of PII are considered sensitive when they are associated with an individual. Secure methods must be employed in transmitting this data when associated with an individual:

- Place of birth
- Date of birth
- Mother's maiden name
- Biometric information
- Medical information, except brief references to absences from work
- Personal financial information
- Credit card or purchase card account numbers
- Passport numbers

[7] NIST SP 800-53A, Recommended Security Controls for Federal Information Systems, establishes common criteria for assessing the effectiveness of security controls in federal information systems. Organizations use the recommended assessment procedures from NIST SP 800-53A to develop their own assessment procedures.

- Potentially sensitive employment information, e.g., personnel ratings, disciplinary actions, and result of background investigations
- Criminal history
- Any information that may stigmatize or adversely affect an individual.

This list is not exhaustive, and other data may be sensitive depending on specific circumstances. Social Security Numbers (SSNs), including truncated SSNs that include only the last four digits, are sensitive regardless of whether they are associated with an individual. If it is determined that such transmission is required, then secure methods must be employed. [DOC Electronic Transmission of PII Policy].

Controlled Unclassified Information (CUI) – Information that requires safeguarding or dissemination controls pursuant to and consistent with applicable law, regulations, and government-wide policies but is not classified under Executive Order 13526 or the Atomic Energy Act, as amended.

Signed this ____ day of ____________, 2016.

GOLDSTEIN.ZACHARY.G.1228698985

Digitally signed by
GOLDSTEIN.ZACHARY.G.1228698985
DN: c=US, o=U.S. Government, ou=DoD, ou=PKI, ou=OTHER, cn=GOLDSTEIN.ZACHARY.G.1228698985
Date: 2016.08.30 15:28:33 -04'00'

Zachary Goldstein, NOAA CIO