

United States Marine Corps: File Analysis Project (War Records)

The Challenge

The U.S. Marine Corps had accumulated a massive amount of NIPR (Non-classified Internet Protocol Router Network) and SIPR (Secret Internet Protocol Router Network) data after ten years of operating in Afghanistan. The data was unstructured and potentially contained vital information covering the Marine Corps' presence and legacy in the region. The data consisted of emails, record message traffic, Microsoft Office files, PDF documents, images, personally identifiable information (PII) and collections of hour-by-hour, day-by-day records of events, incidents, actions and outcomes.

Solution

After extensive vetting, the Marine Corps selected Active Navigation's file analysis software to analyze, sort and organize this massive corpus of data. A Standard Operating Procedure (SOP) document was created to outline the data analysis and migration process in 5 steps:

Discover	Cleanse	Organize	Evaluate	Migrate
<ul style="list-style-type: none"> Conduct an initial skim of the data Perform initial analysis and data aggregation 	<ul style="list-style-type: none"> Remove redundant, obsolete, and trivial (ROT) data Quarantine sensitive information Detect data spillage 	<ul style="list-style-type: none"> Classify content using metadata tags 	<ul style="list-style-type: none"> Review rule sets Modify workflows as needed 	<ul style="list-style-type: none"> Transfer data to temporary/permanent storage locations

Results

- Over 100 terabytes of data were indexed, resulting in millions of collected files
- 77% of the analyzed data was ROT
- Millions of documents were categorized using the Department of Navy (DON) Standard Subject Identification Code (SSIC) retention schedule
- Tens of thousands of sensitive data files were identified within unclassified data collections

Business Benefits

- Quarantined Sensitive Data:** PII, medical, and other spillage information were identified and properly handled
- Achieved Records Management Compliance:** The original Afghanistan war record data was unstructured and out of compliance with **DoD Instruction 5015.02**
- Gained Operational Insight:** After developing best practices, the Marine Corps is expanding the project to additional units to gain insight into the operational history and evaluate training project plans (TPPs) and doctrine

By deploying a file analysis software to automate the processing of their wartime data, the Marine Corps has been able to confidently manage the ever-increasing growth of data and continue to meet their mission requirements.

Real World Application

- Intel Targeting Reports: A collection of records that outline who, how, when, and where our forces conducted missions
- Improvised Explosive Device (IED) Case Files: A collection of records that contain descriptions, images, and plans regarding various IED attacks for lessons learned



Americas: Reston, VA, USA +1 571 346 7607

sales@activenavigation.com

www.activenavigation.com

@ActiveNav