2020 AFCEA ARMY SIGNAL CONFERENCE

MARCH 24-27, 2020 • WATERFORD AT SPRINGFIELD, SPRINGFIELD, VA

SOLUTIONS SHOWCASE

AFCEA Army Signal Conference Solution Reviews Showcase

Welcome to the 2020 AFCEA Army Signal Conference!

AFCEA International is pleased to host this important forum for members of the military. As the U.S. Army's modernization efforts continue to evolve, the service's CIO and other senior leaders have determined their needs with more granularity. Consequently, they have identified several specific requirements for data, including security, information sharing and storage.

The Army continues to strengthen its relationship with technology providers. It is relying on the expertise, experience and best practices vendors bring to the table to address some of their most immediate needs.

Prior to the conference, the Army shared the following areas as opportunities for industry to offer potential solutions to problems the service and joint force must address to continue its progress toward modernization:

- Critical Data Enablers
- Content Discovery and Retrieval/Discovery of Metadata
- Accessibility
- Usability and Understandability
- Trusted Data
- Interoperability
- Security
- Cloud and Data/Application Migration/Enterprise/Hybrid Cloud Strategy

AFCEA International received potential solutions for these problem areas from a range of industry partners and conference participants. Members of the AFCEA Technology Committee reviewed 95 submissions and eight solutions were selected to be presented at the conference.

These abstracts offer a variety of approaches to address some of the Army's toughest problems. They demonstrate how technology and business processes have transformed during the past decade and offer some novel ideas to solving age-old problems. The ultimate goal is to support warfighters with the best possible tools to complete missions safely and successfully.

It has been a pleasure working with the Army, its planning team and industry to bring this conference together. The discussions that occur not only improve future Army operations but also enhance collaboration between the military and industry to combat continuously evolving threats to national security.

Introduction

GUIDANCE: To outpace peer competitors, Army leadership is challenged to build a survivable, unified, end-to-end network that enables leaders to prepare, lead and fight in high-intensity conflict with Unified Action Partners against any adversary from anywhere they choose at any time and to win decisively in all domains and all environments. Data sharing is a critical capability that will enable the global integration of military forces to combat transregional, multidomain, multifunctional threats with the highest security and speed.

Problem Statements

MODERNIZE

Critical Data Enablers (LOE #1.3)

Problem statement: The pace of an evolving threat landscape mandates a change in the culture and operational construct for how we manage data, which is now considered a strategic and critical asset. The Army Cloud Strategy, in conjunction with the Army Data Plan, provides opportunities to deliver dynamic, real-time data that supports defensive and offensive operations through relevant and timely information to decision makers. Data-informed multidomain operations will allow the Army and its joint and multinational partners to defeat every adversary.

A data-aware and data-educated force is needed to execute the Army Data Plan where the right people with the right skillsets ask the right questions to get maximum value from Army data. Talent and culture are key foundational enablers and will directly impact mission outcomes.

The Army must be able to attract, train and retain a talented and experienced workforce imbued with a data-centric culture. This workforce must understand and appreciate the military role for data at the strategic, operational and tactical levels. A lasting cultural change across the Army requires a deliberate, sustained effort emanating from the top of the organization down to the soldier.

Why this is a problem: Today's Army has disparate, isolated data sources, which limits sharing, hinders speed of decision making at echelon (a network infrastructure problem) and prohibits the use of current and emerging cloud capabilities, including evolving artificial intelligence (AI) and machine learning (ML) services and tools. Enhanced use of data can transform the Army's ability to deliver lethal capabilities, augment decision making and increase accountability. In the present state, the Army cannot use data management as a force multiplier, partially because of the lack of skilled, mission-focused and a data-aware and data-educated workforce.

Desired outcome: The Army Data Plan (dated 15 November 2019) applies to all Army data and describes a global, standards-based environment where data and information are visible, accessible, understandable, trusted, interoperable and secure (VAUTIS) throughout the life cycle.

To meet these requirements, the Army will need a robust, integrated approach that serves the warfighter and properly manages resources while implementing protections and security. The Army is looking forward to engaging with commercial vendors who can provide advice on a comprehensive enterprisewide, hybrid data strategy that will help develop a strong data-centric workforce.

MODERNIZE: VISIBLE - Content Discovery and Retrieval/Discovery of Metadata (LOE #1.3)

Problem statement: The U.S. Defense Department Data Tagging Strategy proposed that data must be tagged when it is created, acquired or modified. Data has value only when it is actually used or can be made useful in the future. Delaying the association of metadata results in imprecise descriptions of the data as this descriptive information may be forgotten or lost. Data that is not immediately rendered discoverable or is inaccurately described cannot be relied upon to support analysis or effective decision making.

In order to treat data as a strategic asset, the Army must identify methods of provisioning data services. Leveraging data services provides data-centric functionality to consumers–such as search, create, retrieve, update, delete–and validation against specified criteria. Critical to mission success are methods of advertising data services to allow manipulation, aggregation and transformation of data to make the data more useful by business applications.

Why this is a problem: A foundational goal of the data plan is to identify, tag and register all authoritative data in a way that makes it easily discoverable by users across the enterprise. The Army is still in the process of establishing the governance of its enterprise data so that data security is optimum while at the same time ease of access is timely.

Desired outcome: The Army needs industry support in cutting-edge data management and sharing, with the ability to maintain a strategic and tactical advantage to win over its adversaries at anytime and anywhere in the world.

MODERNIZE: Accessible (LOE #1.3)

Problem statement: The Army must provide all credentialed users access to authoritative and non-authoritative data via commonly supported access methods and shared data services in accordance with law, policy and security controls. The Army requires protected means and mechanisms for all credentialed users to have timely, need-based and authorized access to the right data, including access to security-related metadata that clarifies historical context.

Why this is a problem: The Army has a need to enable authorized users to discover authoritative and non-authoritative data by registering in a shared space. Authoritative data is particularly important because it is, by definition, the most valid, trusted source data that exists. Only by registering data assets, with metadata related to structure and definition consistent with Army standards, can the objectives of data discovery and accessibility be achieved. In addition, the Army will require linking user access to authorize data sets with the user's credentials. The Army requires a holistic approach to data accessibility: authorized users will gain access to authoritative data only through access-approved applications.

Desired outcome: The Army needs to continue to stay on the leading edge of data-centric operations. The service is seeking industry support to improve the accessibility of its critical data. The Army wants industry partners to provide information on technical Authoritative Data Source (ADS) that supports smooth global operations. Every request the application sends to the Application Programming Interface (API) must be authorized. User credentials, applications, data and metadata must all be integrated to ensure seamless operations. The Army must be able to get data down to the decision makers, even when that data is housed or stored on sensitive or classified portions of the environment, so that full correlation tied to the desired mission outcomes can be achieved. The Army needs clear data policies, rules and guidance for facilitating integration at the corporate, business or mission levels and, at a minimum, for metadata-models and data architecture governed or influenced by the enterprise architecture. The Army will lose the ability to gain an operational advantage over adversaries if it cannot deliver the necessary and relevant data to the warfighter regardless of where it is housed. The use of advanced cross-domain solutions will enable those functions throughout Army operations.

MODERNIZE: Understandable (LOE #1.3)

Problem statement: The Army must ensure the data is useable and understandable by authorized consumers, known or unanticipated, through development and use of shared vocabularies, common data standards and documented data dictionaries. The Army has a way to go toward becoming a culture that actively promotes and cultivates a data-aware and data-educated workforce. Across the Army enterprise, the influence and governance of an Army Data Board–creating robust data models and data standards; integrating data; providing data architecture overviews; and identifying data requirements–must be the priority to improve information traceability.

Why this is a problem: According to a published report, the Pentagon failed its first-ever audit in November of fiscal year 2018. The audit report highlighted three issues of note: 1) audit and inventory management; 2) cybersecurity; and, 3) poor data quality within the existing systems. Unstructured data is yet another problem. Examples of unstructured data include share-often text or binary files such as Word documents, PowerPoint presentations, audio files, video files and image files.

It is not uncommon for medium-to-large organizations to have terabytes of structured and unstructured data that they need to manage, backup and potentially recover. This situation presents both a problem for the IT department and a significant, if not always visible, cost to the organization in terms of resources, including storage space, backup space, backup time and staff resources, to manage so much unstructured data.

Desired outcome: The Army needs industry support to improve the usability and information or insights of any data sources. Industry should relate scalability and automation technologies that could augment data capability. The Army is looking for industrial best practices and knowledge of network-based information-sharing that transcends traditional governmental boundaries and coordination to address "the biggest impediment to all-source analysis and to a greater likelihood of connecting the dots: the human or systemic resistance to information sharing." Industry support is also needed to provide technologies or process improvement for data at rest, data in motion and data access across the global enterprise. Recommend industry provide insights into their best practices for metric development and methodologies to assist the Army in determining a process to score or determine data quality in a quantified approach. This quantified approach will enable Army decision-makers to invest in or divest of systems.

MODERNIZE: Trusted (LOE #5.3)

Problem statement: The Army needs to be able to identify and designate authoritative data sources for all data artifacts. Analyses are required to establish traceability and ensure data integrity with timely configuration control and life-cycle management. Data integrity, in turn, requires that data is measured, recorded and protected at the source. Data reliability requires that data is validated and reconciled to establish known pedigree and confidence in the data.

Data protection and integrity remain a top priority for the Army. Data is constantly being used, transmitted and stored at multiple levels. Protecting data must include protection from internal and external threats; protecting data should not degrade operations; and protecting data must include the appropriate security measures and proper handling of security controls of all data from the point of creation through use and destruction. Additionally, the Army has a need to tag or apply attributes to all data. This will allow sensors to identify data movements that fall outside the approved scope, regardless of platform.

Endpoints are one of the Army's most critical avenues of approach or breach, so protection methodologies should include endpoints, because manipulation of data could be just as malicious as theft of data. As required by the Army Data Plan, the Army must ensure data confidentiality, integrity and availability are maintained at all times throughout the data life cycle.

These measures assure data and database products and analyses are accurate and reliable and data-driven decisions are enabled, answering multi-faceted, inferenced or correlated questions. This critical area drives three requirements: the use of strong authentication, a hardened environment and inadvertent disclosures of critical data reduction.

Why this is a problem: Loss of data should almost always be considered worse than initially suspected. The primary reason for this consideration is aggregation. With so much data being used, transmitted and stored, it is nearly impossible to determine what data has already been lost or stolen. An adversary could collect enough "meaningless" data to solve the puzzle of more critical data. The next area of concern is access to data. The Army does not have an effective access control methodology for all levels. The Army does a good job for data at some level of security, while other data is basically overlooked or weakly filtered and protected.

Desired outcome: Fortify the security posture for Army data by reducing the number of vulnerable points through which an adversary could gain access and exploit the service's data. An aggressive accountability and auditing tool is needed to inventory, safeguard and manage data at all levels. The Army needs a methodology to first determine all of the data, then to properly categorize and classify the data, next to tag or apply attributes to data and finally to safeguard the data. This must be an automated process verified by humans and with the assistance of artificial intelligence or machine learning. The protection from loss of secrecy, modification and loss of availability should be in the forefront of Army cyber protection and cyber readiness. The ideal tool should address Test Access Port controls to ensure all malicious avenues of approach are considered.

MODERNIZE: Interoperable (LOE #1.3)

Problem statement: The Army must ensure that all data are useable across multiple systems and applications by stakeholders using non-proprietary, open source, industry or DoD-designated standards. If a system produces data that can only be consumed by its proprietary platforms then we fail to achieve the Army Data Strategy: visible, accessible, understandable, trusted, interoperable and secure (VAUTIS). This highlights the need for open source data platforms and methodologies.

Relationships and dependencies of data will be established and global data standards will be codified. Interoperability relies on shifting from vertical data distribution to horizontal distribution to ensure maximum data sharing such that data is accessible for linking resources together.

Why this is a problem: The Army has data in multiple sources in various degrees of data cleanliness, with uncertain data quality. Poor data management and operations without enterprise oversight result in "dirty data" influencing decisions. Data siloes make the problem worse, distancing the Army from its goal to be a data-driven organization. Army leaders are unable to see, share and act on accurate and quality data. In addition, the identification of siloes and determination of importance to overall information value chain need to be highlighted by the Army Data Board in its governing activities.

The Army requires a standardized format for interoperability. There is a critical need for an open API service, versioning or standardized schema that will allow applications to interconnect. The process of calling and responding to data requests from various authorized entities is often filled with confusion and convolution because of the lack of use of open API services and standardization.

Desired outcome: The Army needs industry support with efficient and effective data modeling and tagging so that the Army can operationalize its data. Open source is a best practice and of tremendous value to the Army, both in the near and long terms. Because APIs and services are complementary, the Army needs industry support in adding automation capabilities and data services, versus tools to refine data as an enterprise asset. The Army data capacity at the enterprise level is too much to conduct assessments without the use of machine learning or artificial intelligence capabilities. Further, the Army needs innovative ways to protect data throughout its life cycle, no matter the network environment. Finally, the Army needs to eliminate duplicative, out-of-date and erroneous data and information policies. The Army requests that industry provides courses of action to execute a current framework that correlates to data interoperability.

MODERNIZE: Secure (LOE #5.3)

Problem statement: The Army's inability to have proactive, internal security at all levels will prove to be detrimental, especially as the Army moves more to the cloud and the Internet of Things. Insider threats pose multifaceted problems for the Army. Internal incidents reported in 2018 were mostly attributed to abuse and malicious intent, with a portion of the internal incidents being unintentional acts. External attacks were less than half of all attacks. The majority of external attacks were accredited to web application, software vulnerabilities and the use of stolen credentials.

Over half of known breaches were the result of insider activity in 2018, according to a Forrester report. The threats can be intentional and unintentional. Both circumstances provide similar devastation, if proactive and reactive, real-time data loss prevention and disaster recovery controls are not employed. The majority of the Army's security controls address external threats to the infrastructure with very few considerations given to internal threats other than policy-based controls.

Why this is a problem: The number of networks, people, devices and workflows combines to produce and receive a lot of data, putting the Army at constant risk. One weak link or unattended risk at any level could cause the entire infrastructure to fall. The Army has a need to improve, build and extend its tactical radius to provide comprehensive cybersecurity for its assets.

Cleared and vetted personnel are able to maliciously manipulate information systems and data, usually without detection. The detection of an insider threat is difficult and, once identified, the distinction between intentional and unintentional can be another challenge. Insiders pose a greater threat because they are closer to the core than external threat agents. Insiders often know the policies and other security controls that are in place to mitigate, detect, deter and prevent attacks against the Army. Many security policies address insider threats, but there is minimal standardization of controls and vetted tools to address the threat.

Desired outcome: There needs to be a Never Trust (Zero Trust), Always Verify framework (Risk Management Framework [RMF]) to ensure the Army's cyber readiness. The Army is interested in a holistic security solution that could help ensure a zero-trust environment through monitoring and trust assurance at every level. The Army is soliciting industry feedback focused on baked-in RMF processes automated into their technologies. The solution should include micro segmentation with the goal to reduce internal incidents or insider threats and the development of a standard tool to identify indicators and defend against insider threats. Additionally, the Army requires assistance with composing tactics, techniques and procedures to make a more security-aware environment.

MODERNIZE: Cloud and Data/Application Migration/Enterprise/Hybrid Cloud Strategy (LOE #1.3)

Problem statement: The Army develops and sustains applications and data in a highly distributed manner and does not have a holistic mechanism to modernize or manage application life cycles. Army applications include both commercial off-the-shelf (COTS) and the government off-the-shelf (GOTS) categories. The focus of the Army migration is mainly on GOTS applications and data. Cloud brings the essential elements of elasticity, resiliency, broad access, efficiency, secure computing platforms, data standardization and compliance tools. The Army has implemented a strategy to modernize and migrate thousands of applications and data to the cloud, but the Army needs assistance with protecting its data through the use of solutions that generate efficiencies through automation.

Why this is a problem: Army application and data owners have been reluctant to migrate en-masse because of technical limitations, funding availability, priorities and perceived risk. Cloud computing capabilities should be adopted in a shared or common approach that is content-service-provider agnostic to prevent limitations in abilities to change vendors in both the near and long term. This model limits the Army's ability to aggregate data for the purpose of artificial intelligence and machine learning.

Desired outcome: The Army needs assistance to establish enterprise cloud and data ecosystems that are artificial intelligence and machine learning-ready and hybrid as well as protect Army data, increase lethality at each echelon (heavy emphasis on processes and architecture data modeling versus cloud services) and generate reinvestment opportunities for modernization. The Army needs to deploy an agile and flexible cloud framework to adapt legacy software to quickly meet changing operational environments, increase readiness and improve cybersecurity. The Army is looking forward to finding out about commercial vendors that can provide advice on a comprehensive enterprise and hybrid cloud strategy that will integrate tactical and non-tactical infrastructure.

Table of Contents

Highlighted abstracts selected as presenters at conference as of March 10.

CRITICAL DATA ENABLERS

CONTENT DISCOVERY

ACCESSIBLE

UNDERSTANDABLE

TRUSTED

INTEROPERABLE

SECURE

CLOUD

CRITICAL DATA ENABLERS

Explainable AI

Kelly Carter, PhD, Data Scientist COL(R), CACI • kelly-marie.carter@caci.com

ABSTRACT

The DOD must move quickly from black-box artificial intelligence (AI) solutions to explainable AI. This presentation will delineate the pitfalls of black-box AI solutions and demonstrate methods to provide explainability in AI and deep learning.

AI that uses deep learning identifies patterns and correlates answers in a way that is not explicitly programed, creating both higher risks and rewards. In 2016, an exposé by ProPublica was done on the Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) algorithm. This algorithm is used by criminal justice officials to determine recidivism rates. This is one of several algorithms used by California, New York, Wisconsin and parts of Florida. The COMPAS algorithm was found to have significant racial bias, which was not correlated to actual recidivism rates. Using data from Broward County, Florida, ProPublica asserted that COMPAS predictions of violent crimes were correct 20 percent of the time, and that blacks were labeled at higher risk almost twice as much as whites. This study was widely criticized, including by the think tank Community Resources for Justice, for misinterpretations of the data and the subject matter. Regardless, the problem of bias and trust in AI outcomes still exists.

The deeper the neural networks, the more layers of analysis, the less insight into how the answer is obtained; adding deep learning compounds the problem. One solution to explainability is a heat map. A heat map shows where the algorithm is focused.

In 2019, it was revealed that an algorithm developed by a Stanford graduate student diagnosed TB correctly 75 percent of the time, compared to doctors in South Africa, at 62 percent. By using a heat map, it was revealed that the algorithm included margin data, i.e. non-image information. For example, by considering markings on the X-ray from radiologist, where the X-ray was taken, and what type of machine took the X-ray, the algorithm boosted the score for positive disease find. The X-rays that the algorithm was trained on had markings from radiologists that indicated a positive find for disease–a simple tic mark. Machines located in hospitals were more apt to find disease than mobile doctor's office machines, so images from these machines were scored more likely to be diseased. This is a significant bias in the algorithm and means the actual intended performance of the algorithm for image-only data is probably much lower.

Explainable AI can be built. Explainability can be created via dashboards, reports and constant model updates. In addition to heat maps, methods for AI explainability include: Maximum Mean Discrepancy (MMD), which shows differences in dataset distributions; Partial Dependency Plot for Individual Conditional Explanation (PDP-ICE), which shows the effect a variable on a model; Accumulated Local Effects

(ALE), which describes how local features influence the prediction; and Interactions-based Method for Explanation (IME), which calculates the contribution of factors. These methods can be applied to black-box AI also. These are just a few examples of ways to build in explainability. Explainability must be expected, demanded and built-in.

BIO: Kelly Carter is a data scientist with CACI and has provided artificial intelligence solutions and research for several USG agencies including the Defense Threat Reduction Agency, the Customs and Border Protection Agency and the Air Force Research Laboratory. Carter is a retired Army Signal Corps colonel living in North Carolina.

Splunk: Partnering with the Army's Data Strategy

Melissa Andrews, Technical Sales Engineer, Splunk Inc. • mandrews@splunk.com

ABSTRACT

Splunk, the industry's leading platform for machine data and used by thousands of customers, allows the collection of human readable data from sensors, systems, applications and hundreds of other data sources, and accommodates its velocity, variety, variability and volume. Splunk can allow enrichment of this data with contextual information from the Army's relational databases and metadata sources. Splunk makes data accessible, usable and valuable across a wide-range of use cases, including cyber security, information technology (IT) operations, supervisory control and data acquisition and industrial control systems (SCADA/ICS), Internet of Things (IoT), business analytics and more. Its flexible, easy-to-use user interface (UI) enables creation of ad hoc or saved reports, alerts, dashboards and visualizations that allow actionable insights into operational data and fully supports mobile access. It supports VAUTIS: visible, accessible, understandable.

Splunk is an open, extensible platform with support for more than 1,600 apps that allow its customers to gain immediate insight and benefit from the platform. Splunk provides ODBC/JDBC connectivity to structured databases, an HTTP Event Collector for sources that are unable to output data to a syslog or file, a complete REST API and software development kits in C#, Python, Java and Javascript. Splunk supports integration with noSQL, big data products such as Hadoop and provides modular input events for data streams that may require specialized programmatic access. Splunk is heavily investing in OSS by building it into its own products, creating seamless integrations with the best of breed OSS, and continuing to maintain the Apache Pulsar project. It supports VAUTIS: interoperable.

Splunk's scalable data fabric platform can provide visibility across multi-domain elements, service branches and technology platforms. Its flexible platform enables distributed search across multiple instances from a central location or distributed locations and leverages native role-based access controls (RBAC) to govern access rights to specific data sets by customer, organizations, stakeholder or any other qualification. It supports VAUTIS: visible, accessible, secure, interoperable.

Security is built into the entire Splunk platform. Splunk uses the latest encryption standards both in-transit and at rest. Its code undergoes rigorous static and dynamic analysis along with application penetration testing, to ensure minimal vulnerabilities. Splunk also includes built-in controls to ensure data integrity and prevent data tampering. Splunk has authorities to operate (ATO) from multiple organizations within the DOD and is certified under Common Criteria. It supports VAUTIS: trusted, secure.

Splunk has a “cloud first” strategy in that all of its products are built to support cloud deployments; hybrid and on-premise deployments are also supported. Splunk’s platform provides built-in and ad hoc machine learning and automation capabilities.

Splunk can provide a context-rich data environment, improve operational awareness, and accelerate the warfighter’s ability to prioritize actions and make data-centric decisions in near real time.

Based on its work with more than 16,000 customers across government, the Department of Defense and commercial industries and as the industry’s leading machine data platform, Splunk is pleased to partner with the Army around creating a comprehensive enterprisewide, hybrid data strategy that will develop a strong data-centric workforce.

BIO: Melissa Andrews has been working with the Department of Defense and technology since 1998. She currently works for Splunk Inc., where she supports U.S. Army customers, helping them understand how they can use Splunk’s industry-leading data platform to support the warfighter. Prior to Splunk, she worked at Oracle Corporation where she provided support to Army customers running large Oracle databases.

Use of AR Contact Lens To Increase Information Speed of Critical Data to Army End Users

Mark Colby, Senior Vice President, Government Business Development, Tectus Corp. •
mark.colby@tectuscorp.com

ABSTRACT

Tectus is developing the world's first augmented reality (AR) contact lens (CL). The technology overlays images, symbols and text on wearers' natural field of vision without obstructing their view and without the need for bulky, heavy, conspicuous goggles or headsets that are socially and physically awkward, interfere with normal vision and generally get in the way of productivity and mobility. These comfortable contact lenses will be safe to wear all day, are inconspicuous and will display as much or as little information as needed.

Since the advent of the smartphone, people have become increasingly dependent upon screens for information, communication and entertainment; but the more people look at screens, the less aware they become of the people, places and things around them. Augmented reality promised to change this paradigm by displaying information in the world itself rather than on a screen; but solutions like Google Glass, Microsoft Hololens, and Magic Leap One just replace the screen in the hand with a screen (or two) on the face. While technologically impressive, few could imagine wearing these headsets all day long. Commercial, consumer and military end users alike will find that AR headsets limit their mobility and interfere with most routine tasks. Unlike those bulky devices, Tectus' AR CL is designed to make augmented reality an invisible and indispensable part of a daily routine.

The Tectus AR contact lens will provide the entire military a best-in-class AR solution that is adapted to the needs of the modern warfighter. The company's technology will allow users to enjoy the benefits provided by head-up displays and AR technology while maintaining their focus on the world around them and without having to wear cumbersome or awkward AR equipment. Beyond AR, Tectus' CL also provides other superhuman capabilities like night vision, facial and object recognition, "see what I see" video and picture sharing. As a platform, Tectus' innovative technology will greatly enhance operational effectiveness and provide the modern warfighter the tools necessary to continue to maintain air, land and sea superiority. This potential has been recognized by AFSOC, USAF PEO ISR/SOF, USSOCOM PEO SOF-Warrior, USSOCOM SOF S&T and Naval Special Warfare.

According to the AFSOC, “The mission impact of this project on the Air Force and the Department of Defense will be transformational for aircrew, ground operators and support personnel in a multitude of activities.”

USSOCOM SOF S&T added, “We believe that technology development under this SBIR Phase II Strategic Increase will eventually contribute to solving a significant SOF operational need as well as a much wider DOD need, including potentially USAF strategic and operational requirements.”

AFLCMC/WIS SOF and Personnel Recovery Division (PEO SOF/ISR) wrote, “Tectus’ AR CL represents not just an improvement but a game-changing warfighter capability,” and “We want the AR CL to seamlessly integrate into existing and future systems so it can become the AF display of the future.”

With the Tectus AR contact lens, the Army could ensure that critical real-time data is visible, accessible, understandable, trusted, interoperable and secure (VAUTIS).

BIO: Mark Colby is a retired naval officer and senior executive with more than 30 years of leadership experience in technology solutions for government and consumer markets. His education includes the United States Naval Academy, a BS in mathematics, Naval Postgraduate School, MS in electrical engineering.

Colby served more than 22 years in the U.S. Navy as Surface Warfare Officer with assignments including nuclear engineering, operations and combat systems, executive officer, strike group maritime operations and commanding officer in two ships. In business, he worked at the senior vice president/president level in sales and operations for several renewable energy, software and FinTech companies, including Envision Solar, Power Analytics, ClearEdge Power, Solar Alliance and Ygrene Energy Fund. He now leads government business development and sales for Tectus Corporation.

Getting a Second Opinion on Your Data

Carlos Cosme, U.S. Army Cyber Command ISSM, ARCYBER • carlos.f.cosme2.civ@mail.mil

ABSTRACT

Creating a data-centric culture begins with trusting gathered data. Currently, the Army relies heavily on tools to gather data to drive processes such as patch management, end-point management, account management and vulnerability scanning without a way to validate that data or to measure the efficiency of those tools. These tools are treated as isolated data sources, often managed by separate teams/people. This practice is often ineffective and leads to knowledge gaps. Further, it hinders the ability for leadership to make well-informed decisions and hinders any move toward cloud/AI/ML capabilities. However, if the Army began comparing the data outputs of these tools against each other or against other tools that gather similar data, that data would become validated and trusted. False negatives and false positives would become identifiable, and missing data might be found. Further, through checking the data with supplementary tools and determining the data with the highest validity, an organization can track the efficiency of the tools it uses as technology to accurately measure return on investment. It can better determine configuration errors in its toolsets that leads to errors in reporting and effectiveness. This method works for a variety of tools and types of tools.

Consider asset management: If Tool A sees 125,000 assets on the network and Tool B sees 145,000 assets on the network and these two tools agree on about 100,000 assets that sit on the network, the 100,000 assets the tools agree upon are the assets of which the organization can be most sure. For the additional 45,000 assets that have not been confirmed by both tools, an incident response process and/or service desk process can be put into action to confirm the existence of those 45,000 assets. Consider these the different levels of data validity that allow an organization to determine the effectiveness of its tool portfolio and on what data it can be most rely. If using more than two tools to check the original tool, the highest level of data validity would be the enterprise configuration assets or elements that all the tools can see and agree on, the moderate level of validity would be the ones that only a few of the tools can see, and the lowest level of validity would be the ones that only one tool can see. At a bare minimum, two tools must be used to confirm enterprise configuration aspects.

By trusting data and knowing the validity of that data, the Army can promote a more data aware and data educated approach to business intelligence in the workforce. Army leaders can better analyze the return on investment of the tools they use and better analyze the productivity and training of the workforce in regard to those tools. Having identifiable metrics of data validity and being able to track growth in those metrics provide an opportunity for the Army to gain confidence in its data, in the tools it uses to gather that data, and in the use of that data.

BIO: Carlos F. Cosme is a native of Bayamon Puerto Rico, and a 1999 graduate of Randolph Macon Academy with a GED. He holds a 2000 25B/74B military computer science certificate school of information technology, Fort Gordon, Georgia.

Cosme's operational tours include information assurance manager, Department of Army (Pacific); security operations manager, Qatar (OIF/OEF); incident response handler, National Reconnaissance Office; service desk manager, FBI Special Technologies and services; network operations manager, Soto Cano Airbase Honduras; cyber SME supporting ARCYBER Technical Warfare Center; security operations program manager, Department of Justice; DMZ program manager, DISA; principal security operations consultant, Hewlett Packard (ArcSight); command ISSM, Cyber Security Division ARCYBER G-6.

Data Democratization: Advancing Data Literacy To Achieve Mission Outcomes

Andrew Churchill, Vice President, Federal, Qlik • andrew.churchill@qlik.com

ABSTRACT

The Army has committed to being a multi-domain force by 2035. Having the networks and IT infrastructure in place to bind the forces together is critical to achieving this. Equally critical is the data moving across the infrastructure and strong data governance policies that increase transparency, improve consistency, eliminate redundancy, enhance security, enable interoperability and increase the speed of decision-making at all echelons. They must also be formed with data democratization in mind.

The Army has, receives and processes vast amounts of data from a wide array of sources that is housed across disparate networks. This data is far less useful if personnel across all levels—from the highest-ranking officers to the newest enlistees to civilian support staff— don't understand where, when and how to access it. Data democratization will enable the Army to get the right data to the right personnel at the right time.

And these personnel must be data-literate, meaning they are equipped with the right skills to effectively analyze, visualize and leverage this data. Data literacy is at the heart of data democratization because data is inactionable without it.

The establishment of a data-aware, data-educated force is a cultural movement that must be driven by leadership at the very top and actively adopted throughout the ranks. Leadership must demonstrate its commitment to cultivating a data-skilled workforce through open communication, an investment in training at all levels and a willingness to retool processes. They must exhibit patience and understanding for the tangible learning curves and troubleshooting that comes with learning new technologies. They must also lead by example. They must be willing to get their hands dirty and leverage the tools, reports and solutions produced by personnel to reduce redundant efforts or rework.

The data management and visualization tools that the Army implements must be easy to use, and therefore, easy to train personnel at all levels. They must be trained in how to think critically and strategically about the data so they can properly contextualize it and derive meaningful insights that enable more timely decision-making. Each member of the branch—regardless of role or rank—should be comfortable working with and interrogating their data in natural language without needing to be a data scientist.

The Army must allow personnel at all levels throughout the enterprise to leverage data to conduct their jobs more effectively. This doesn't mean full and open access, rather it means setting appropriate

access based on individual roles within the enterprise. Once users experience how the tools enable them to complete their jobs more efficiently and more accurately, they will become more invested in using and maintaining them.

By instilling a culture of data literacy, Army personnel at all levels will have the added capabilities they need to comprehend, analyze and manage this data.

This session will address the criticality of data literacy for achieving mission outcomes, the educational resources available to advance the data literacy of the force and the ways to inculcate a culture of data literacy across the Army.

BIO: Andrew Churchill leads Qlik's federal go-to-market strategy and customer success initiatives. His leadership responsibilities cover all facets of the business, from coordinating resources across the team, to leading them in opportunity identification and pursuits, to ensuring complete solution delivery. Churchill brings his own unique energy and personality and leverages his 25 years of experience supporting government IT initiatives to drive continued growth. He has worked tirelessly to instill a culture of data literacy across federal agencies with the goal of ensuring that each and every public servant is comfortable analyzing and consuming data. He strives to highlight the opportunity to enhance outcomes and performance with data-driven initiatives.

Accessing the Invisible Data

JP Morgenthal, Chief Technical Officer—Americas, Automation Anywhere •
jp.morgenthal@automationanywhere.com

ABSTRACT

Everyday humans perform hundreds of tasks as part of critical mission processes. The data about what those humans are doing, the task they are performing and information pertinent to the overall process at the point where that task is being executed are lost as they are not captured. Intelligent Process Automation (IPA) allows command to capture and gain insight into those tasks driving greater efficiencies, identifying potential bottlenecks and constraints, and highlighting latencies that have the potential to limit warfighter awareness. Moreover, IPA should be coupled with IoT and other sensor-based efforts to generate a complete picture of data as it flows through the process. This session will present an introduction into IPA and illustrate why it should be a critical tool in the military's toolbox.

BIO: JP Morgenthal is a veteran IT solutions executive and Global CTO. He has been delivering IT services to business leaders for the past 30 years and is a recognized thought-leader in applying emerging technology for business growth and innovation.

Intelligent Process Automation—A Data-Centric and People-Centric Way To Transform How the Army Engages With Data

Mahesh Srinivasan, Chief Technology Officer, OM Group Inc. • mahesh@omgroupinc.us

ABSTRACT

Creating a talented, data-centric culture in the Army can be accomplished by making data easier to engage with and understand. Intelligent Process Automation (IPA) is a next-level data solution.

Converging the abilities of automation with innovative technologies that can "see" and understand data and documents in context, much like humans do, intelligent software bots will radically increase the sophistication of data analytics and decision-making for the Army.

IPA converges multiple advanced technologies into a single smart data processing solution. The technical components include:

- Robotic Process Automation (RPA) – Automate human activities that are manual, rule-based and repetitive, extracting data from varied systems.
- OCR/ICR/CB – Optical Character Recognition: Convert text or print document into a machine-readable format. Intelligent Character Recognition—Convert handwritten text characters into a machine-readable format. Computer Vision: Human-like recognition of user interfaces.
- Artificial Intelligence/Machine Learning (AI/ML) – Identify hidden patterns in knowledge-intensive process and learn from data.
- Business Process Management (BPM) – Manage and improve business processes.
- Predictive Analytics – Analyze data to predict unknown future events.
- Natural Language Processing (NLP) – Understand, interpret and manipulate human language.

The Army can use IPA to solve the following problems:

- VISIBLE: Discover and process a high volume of data automatically.
- ACCESSIBLE: Extract data through automated triggers and on demand from various systems. Speed of data, available in the field, in real-time.

- UNDERSTANDABLE: Automatically present both unstructured and structured data, extracted from various sources, displayed in a way that makes sense.
  - » Categorize data and documents based on AI/ML-based optimization.
  - » Historical data is analyzed over time providing insights and predictions for decision-making.
- TRUSTED: Compare data extracted from multiple applications and an authoritative data source to highlight errors for human review or autocorrect simple errors based on pre-defined business rules.
- INTEROPERABLE: The company's IPA-based automation can help with interoperability across heterogeneous systems and access to all types of data, including mission partners.
- SECURE: The bots operate within their realm of permissions to maintain data security without violating segregation of duty policies.

IPA will conserve Army resources and increase the speed and accuracy of processing and analyzing data. Like a tower crane is a lever for lifting heavy weights, IPA does the heavy lifting of data-centric processes to increase the analytical power of warfighters. Routine, repetitive tasks are handled automatically, while intelligent software bots can quickly deliver more sophisticated information to staff to augment decision-making. IPA leverages the best in automated data-centric and people-centric capabilities.

OM Group's Research and Development team applied RPA and IPA solutions with Agile and DevSecOps methodologies to reduce processing time by 92 percent on a mission-critical project for the Headquarters, Department of Army. The company brings a deep understanding of Army networks and systems, anchored by modern, lean, commercial best practices rooted in CMMI DEV/SVC L3, ISO 9001 and ISO 27001 standards.

BIO: Mahesh Srinivasan is the chief technology officer for OM Group Inc. He has more than 17 years of experience architecting and delivering secure enterprise-class solutions, including intelligent process automation, robotic process automation, cloud solutions, enterprise architecture (EA) and systems integration (SI). Srinivasan is a senior cloud solution architect and client technical advisor with a specialty in bridging business goals and the technologies used to create state-of-the-art solutions. He delivers digital transformation projects, leading large teams in commercial best practices, for federal and commercial clients. PMP, TOGAF 9, ITIL V3, AZURE & AWS Architect.

Trusted and Secure Documents

Matthew Shabat, U.S. Strategy Manager, Glasswall Solutions • mshabat@glasswallsolutions.com

ABSTRACT

Glasswall FileTrust is easily integrated into a comprehensive solution to fulfill the need for safe and trusted documents. Currently embedded in U.S. intelligence community cross-domain solutions, Glasswall's advanced content disarm and reconstruction technology provides deep-file inspection, remediation, sanitization and file regeneration using its patented d-FIRST technology, which is available as a software developer kit (SDK). Glasswall ensures that common business files comply with the published standard for each file type, such as the Microsoft Office file specifications and ISO 32000 for PDFs. The standard is a "known good" state of the file, which in turn eliminates the chance of malware in the files at the structural layer of the file. Glasswall also has the ability to remove active content items at the functional layer of files, which can be exploited by adversaries and include Dynamic Data Exchange, macros, Javascript, embedded files, active URL hyperlinks and Acroforms. Deploying the Glasswall SDK in front of each data center will allow the Army to ensure that files are safe each and every time they are moved or shared. With Glasswall as part of the overall solution, all files will be standardized and secure. The Glasswall solution also allows data to be accessed quickly by analysts for decision making. Data can be sanitized within a fraction of a second and transferred to an analyst without the fear of a malware attack in order to make tactical decisions. For example, if data is confiscated in the field, it can be sent through Glasswall's SDK to sanitize the document prior to an analyst viewing it for intelligence purposes, thereby protecting the analyst's endpoint and preventing the analyst from having to switch to a separate endpoint to view the intelligence.

BIO: Matt Shabat is the U.S. strategy manager for Glasswall Solutions. He served for nearly 10 years in the U.S. Department of Homeland Security's Office of Cybersecurity and Communications, most recently as a cybersecurity strategist and as the director of performance management, and previously as the National Cyber Security Division's chief of staff. While at DHS, Shabat led DHS and interagency implementation of the Cybersecurity Information Sharing Act, collaborated with members of the public and private sectors to increase adoption of security automation and orchestration, supported maturation of the cyber insurance marketplace, developed an operationally relevant approach to measuring the costs of cybersecurity, contributed performance goals to the NIST Cybersecurity Framework, led strategic planning and developed program performance metrics.

Before DHS, Shabat practiced securities, mergers and acquisitions and general corporate law at Mayer Brown LLP in Chicago. He is a Harvard Kennedy School Senior Executive Fellow, earned a JD from the University of Pennsylvania Law School, an MA in security policy studies from the George Washington University's Elliott School of International Affairs, and a BA from Stanford University. He also is ISACA certified in Risk and Information Systems Control, and he is an ISACA Certified Information Security Manager.

Assess and Build Data Skills With Pluralsight

Wes Novack, Systems Architect, Pluralsight • register@wesleytech.com

ABSTRACT

In 2014, Pluralsight was a small engineering organization with just a few teams. As it scaled the organization, it built a new architecture that allowed for the company to decompose its monolithic codebase into bounded contexts using domain driven design.

Each team now owns a vertical slice of product functionality, where they are responsible for the full stack and operation of their bounded contexts, all the way down to the database layer. Other teams are prohibited from reading into databases that they do not own. This allows the owner team to evolve the data, change database schema, and even switch out the database engine completely to a new solution, without affecting other teams.

But how do other teams get access to the data? The company's architecture employs a centralized message bus that was built on top of RabbitMQ, using AMQP and JSON-encoded messages, with a fanout exchange methodology. The team that owns the source of truth for a particular data set publishes updates to RabbitMQ, which enables any number of additional teams to subscribe to those messages to receive updates.

As Pluralsight continued to scale, it found some pain with its RabbitMQ data replication methodology, as it did not retain history. This led the company to implement a Data Vascular System (DVS) within its organization. This system is backed by Apache Kafka, which allows it to publish data entities, adhering to an avro schema, into a topic that is stored in a log format retaining history. The company now has the best of both worlds, where it can use its DVS for bootstrapping the entire history of a dataset, and it can use RabbitMQ for event notification and choreography, where a full data entity is not required.

This talk will take attendees on a tour of the data systems used in Pluralsight's product and give them ideas for implementing their own.

Pluralsight is the enterprise skills platform. With Pluralsight SkillIQ and RoleIQ, customers can benchmark and measure the skills of their technology workers. The company's advanced analytics provide insights to technology leaders, which helps them assess skills across teams and departments. The platform finds and then fills skill gaps in an organization with expert-authored video courses, interactive coding courses and hands-on projects. The company's catalog goes both wide and deep and covers a variety of data technologies, including SQL databases, NoSQL databases, Graph databases, artificial intelligence, machine learning, big data, data lakes, ETL and more.

BIO: Wes Novack is a systems architect at Pluralsight, where he helped form and foster the company's devops culture and practices. His tenure at Pluralsight started with a migration of the company's infrastructure from an on-premises data center into the AWS cloud.

For the past five years, he has enabled experience organization scalability by embedding devops engineers into cross functional product experience teams, and by implementing configuration management, continuous delivery, and infrastructure as code practices. Novack has helped Pluralsight decompose its monolith into a new architecture based on domain-driven design and bounded contexts, enabling loosely coupled applications and product teams.

The company's devops culture, practices, architecture and organizational structure have increased product team flow efficiency, which has resulted in rapid feature delivery and continual iterations and have led to Pluralsight's dominance in the enterprise technology skills market.

Impact of Culture on Data-Centricity

Richard Dillard, Deputy Program Manager, Telesto Group LLC • proposals@telestogroup.com

ABSTRACT

Many factors influence a successful data-centric culture. The operating culture—influenced by the philosophy and espoused values of founders and senior leaders—limits an organization's effectiveness and efficiency with converting data into an information asset for better decision support. Culture can change, and with cultural change comes the requirement to manage change, particularly changes to business processes that drive intended outcomes.

Telesto Group's approach to enabling a data-centric culture utilizes a human-centered view that is: concerned with the real functioning of organizations, including variations inherent in people, process, technology and data; focused on the organization as a system of interacting parts; and based on the successful medical model, i.e., human system of research: diagnosis, prognosis and prescription. This human-centered view provides a lens for identifying and resolving several pain points emerging from human-machine teaming efforts, including net organizational pain commonly caused by the tendency to separate the data from its process context. It also hardwires a formula for effective decision-making (ED), expressed as ED = T x Qt x A, where: T represents shortest sustainable lead-time to collecting data based on its perishable nature, Qt represents higher levels in quality of thinking when converting data to information, and A represents higher levels of acceptance in the outcomes of that conversion.

To really operationalize insights-driven decision-making, the company's innovative alternating data flow value stream drives the emergence of a data culture by attending to four common pain points within the current operating culture and IT architecture. This approach supports the Army Cloud Strategy and the Army Data Plan, as it will enable a global, standards-based environment where data and information are visible, accessible, understandable, trusted, interoperable, and secure (VAUTIS) throughout the life cycle.

- Mission Needs: Applied research into an organization's business goals and processes begins to hardwire the connection between teams and produce the prioritization of efforts before the collection and analysis of data begins. Telesto Group assesses the business process visibility (V) and accessibility (A) requirements, ensuring alignment of data strategy and capture of the best opportunities for improvement.

- Advanced Analytics: Based on a customer's goals and priorities, the company applies higher mathematics and analytical tools for proper diagnosis of the business to uncover vital insights within IT architecture limitations. Practical data science produces better understandability (U).

- Automated Applications: Once business users/managers see the benefit of feasible machine learning and predictive use cases, the data and information becomes trusted (T) and results in a positive prognosis, increasing demand for data science, as well as the need for automation to free up scarce resources.

- Consumable Results: It's been said that the quickest way to the brain is visualization, but because most end users don't find R-code or R-visualization sufficient for consuming interoperable (I) and predicted results on a daily basis, the prescription is to iteratively co-design/develop dashboards with end users that securely (S) enables them to easily consume results in a user-friendly format.

Leading change through the alternating data flow value stream, both human-human and human-machine teaming improves, resulting in a data-centric culture.

BIO: Telesto Group LLC is an integrated solutions provider and SBA Small Business Contractor with offices in West Palm Beach, Florida; O'Fallon, Illinois; and a substantial project presence in the Washington, D.C., area. It is an SAP Partner and has significant experience within multiple Army ERP programs, including LMP, GFEBS and AESIP. Telesto Group holds a prime IDIQ contract supporting the U.S. Army ERP suite and a GSA PSS Schedule as well as a GSA IT-70 contract. In 2018, Telesto Group LLC was awarded a $50 million transportation management system (TMS) prototype at USTRANSCOM through a competitive OTA process, which provided a Software as a Service (SaaS) SAP TMS system and was hosted in the Amazon Web Services (AWS) cloud to USTRANSCOM and DOD. The successful program provided leading-edge thinking for DOD SaaS model implementations. TMS allowed transporters to order, ship and track unit moves through a single system and allows them, port operators and commanders to have end-to-end in-transit visibility of a unit's equipment all the way to the foxhole.

In February 2020, Telesto was selected as the lead systems integrator to produce a prototype for the U.S. Army Enterprise Ammunition Supply Chain System solution. Telesto Group and its subcontractors will support the U.S. Army's mission to create an integrated and automated end-to-end supply chain system for Class V (munitions) to meet joint, operational, accountability, visibility and auditability requirements. Telesto Group brings nearly two decades of working across the DOD to implement sustainable, innovative solutions to reduce the total cost of ownership.

Threat Intelligence Protocol

Aftab Ahmad, Professor, CUNY John Jay College of Criminal Justice • aahmad@jjay.cuny.edu

ABSTRACT

The current suggestions for sharing threat intelligence over the Internet are comprehensive but not flexible. By suggesting a new protocol, the Threat Intelligence Protocol (TIP), CUNY proposes to enhance the STIX/TAXII infrastructure such that the participants in a community of threat intelligence system can dictate their own terms when processing the intelligence instead of letting the cyber threat database proliferate over the years. The proposed protocol has records initiated by participants but stored in a common website. On receiving a TIP signal (TIPs), participants can give a significance level according their risk management profile, update a local threat intelligence database and generate appropriate alerts, as well as raise or lower the level for an imminent threat of a cyber attack.

BIO: Dr. Aftab Ahmad is a tenured professor at the City University of New York, John Jay College of Criminal Justice. His current work includes software reverse engineering, cyber threat intelligence and brain data analysis. Ahmad has authored two books, two book chapters and more than 50 peer-reviewed journal and conference papers.

Leveraging AI/ML To Provide Total Talent Awareness

Jeff Gibson, Managing Partner, Oplign LLC • jeff@oplign.com

ABSTRACT

Oplign is an AI/ML system that automates the alignment of labor demand and labor supply in any organization. The system works for commercial companies, small businesses and government agencies, as well as for platoon, battalion, brigade, corps and any military-sized organizations. The Oplign core has been trained to recognize every labor demand requirement in the global labor market. These are posted every night by every company that has open labor positions. With a fully indexed and encoded map of global labor demand in place and updated every night, Oplign can now ingest any organization's demand for every job, activity, task or collateral duty they have, and automatically map an organization's complete demand to the fully encoded and related market demand.

Once an organization's demand is encoded and related to the core, getting the organization's labor supply to align to its demand takes about 60 seconds for each supply node/individual soldier. A soldier's alignment is done via a mobile interface, which starts by asking for the branch, rank and MOS. With this input, the system then prompts the individual about the top 40 to 50 skills and experiences the organization's "market" wants to know about them specifically. The greater the interaction from the individual at this level, the more precise the AI/ML can be about interrogating and extracting information the organization is targeting. All data in this exchange is data, as there is no place to input text strings or create semantic divergence in the exchange. The supply-demand awareness database created from this continual interaction is a fully structured, fully queryable, fully extensible and pervasive in a manner that allows the organization to always be optimizing its labor supply to its demand.

Oplign's ability to automatically and immediately align and clear labor demand to supply provides real value. Right now, leaders understand the 20th century military labor skills their troops contain because they can see it in their MOS. What they don't understand about their 21st century troops—the ones that never used a rotary phone and could text before they could walk—are all the skills and experiences that are hidden behind their MOS or that have no relation to their MOS. With the increasing number of nontraditional enlistments and the explosion of technical skills required to navigate the commercial market world of today, the enlisted soldiers have deeper, broader and more varied skillsets than at any time in the past. But the commanders are blind to those additive soldier skills.

Simply put, the organization is completely aware of how all its labor supply aligns to all its labor demands, as well as the gaps, and time to close any gap of any supply point to any demand. Oplign will enable the Army to employ "a data-aware and data-educated force to execute the Army Data Plan where the right people with the right skillsets ask the right questions to get maximum value from Army data."

BIO: Jeff Gibson has 25 years of leadership experience spanning across military, Fortune 100 and defense services industry in senior and C-Suite positions. He possesses a strong understanding and has demonstrated competence in leadership demands, operations, organizational change, strategic planning and risk management. Gibson has an MBA from Washington University in St. Louis, and is a retired U.S. Navy SEAL.

Modernize All Data Enablers

Chris Hauter, Federal Account Executive, Alteryx Inc. • chauter@alteryx.com

ABSTRACT

Every organization Alteryx talks to about some element of data-related transformation, whether it be predictive analytics machine learning or AI, wants to focus on the technology. While technology is important, it misses a key component of successful transformation, and that is people. As in many operational areas, the ability to succeed comes down to the training and capability of people, and so it is in data analytics. The issue facing the Army as it endeavors to implement a viable data strategy is that of finding, building and retaining a data literate workforce that can support the needs of a data-agile organization.

Many organizations know that the ability to leverage advanced analytics, including spatial, predictive and machine learning, are critical to addressing their operational challenges. The telling fact is that while 74 percent of organizations want to use advanced analytics only 23 percent really do. Why is that? The main reason is that advanced analytics like predictive analytics is hard. Every organization wishes its data was well structured and organized so that it could be run through a spreadsheet and insights generated, but that is not the reality. Modern data is messy and complex, and organizations need advanced analytics to even begin to understand it.

This need for better analytics is ubiquitous across enterprises whether a supply officer trying to improve the flow of inventory or a staffing organization trying to prevent churn or a procurement team trying to prevent fraud, the need for analytical insight is always there and continuing to grow.

With this increasing need for insight, organizations have focused on hiring the right people to answer these questions: in other words, data scientists. The data scientist is a unique individual at the center of three areas of expertise: domain knowledge, statistics and software engineering, in other words, a unicorn. However, unicorns in this world are extremely rare, hard to find and even harder to retain. As a result, organizations are often forced to make sacrifices, hiring someone who may know math but is unable to code, or has an understanding of software but without the mathematics to back it, and even if the person has the math and the coding background, he or she lacks the domain expertise.

At Alteryx, the approach is based on building and upskilling the existing data resources through self-service analytics. From the analyst to the citizen data scientist and even the classically trained data scientist, Alteryx focuses on enabling the data worker with the ability to build foundational knowledge to find, blend and format data in preparation for analysis. With an intuitive and easy to use interface, Alteryx upskills users to leverage spatial analytics and build predictive models in a code-free, code-friendly environment.

A key factor in building a sustainable culture of analytics is the availability of ongoing learning and skill development. Alteryx has built an engaged online community of users to share knowledge and answer questions, enabling all Alteryx users to develop new skills and deepen existing ones.

BIO: Alteryx is a recognized leader in data science and machine learning through its ability to deliver a self-service analytics end-to-end platform that unifies the analytic experience across the enterprise, enabling organizations to break down data barriers. The Alteryx platform provides the flexibility that business analysts, data scientists and IT need to discover, prep, analyze and operationalize analytic models through a collaborative and governed platform, enabling every data worker, regardless of technical acumen, to become a problem solver. Alteryx enables data workers to find and understand what verified and trusted information is at their disposal, giving them the ability to analyze data from multiple sources to deliver business insights with agility. Alteryx is revolutionizing business operations through easy-to-deploy advanced analytics, including geospatial, predictive and assisted modeling capabilities in a code-free or code-friendly environment. Alteryx helps democratize data across the enterprise so that people within the chain of command can understand their data resources and create actionable insights faster with the highest degree of confidence. Alteryx has built an engaged online community of users to share knowledge and answer questions enabling all Alteryx users to develop new skills and deepen existing ones.

The Data Lake API: No Need To Localize the Data

Ryan Yu, Chief Operating Officer, Sunayu LLC • ryan.yu@sunayu.com

ABSTRACT

The notion of a utopian world where multiple applications and data owners can move to the cloud and aggregate their data is ultimately a fantasy that needs to be erased. Integrating disparate systems and trying to create a standardized data set is a waste of time and resources.

Each application has its own data structure and every effort to create a monolithic data model standard is just trying to shove a square peg through a round hole. The concept of converging into a single data model standard for multiple applications is extremely difficult to execute; this is because a single data structure cannot meet the needs of different types of business logic.

However, this convergence of data does not have to be a physical aggregation; it can be a logical one. Using software, it is far easier to access the data from disparate applications from multiple locations and aggregate it within memory. This allows data scientists and software engineers to view all the data and build new tools and structure the data to meet their needs. Analysts can view all the data, merge it, sort it, dedupe it and effectively mold it to find the metaphorical needle. Building artificial intelligence and machine learning capabilities requires vast amounts of data, and the more data that is provided, the better the training models are. The reality, though, is that the data does not need to be localized; it just needs to be accessible. If it can be accessed by software, providing it to a data scientist, a software engineer or an analyst is easily accomplished. Fundamentally, this is the ideal application for an API layer. This interface would allow data scientists and engineers to make function calls to access data that resides in any location. Once accessed, they have the freedom to do what they do best. Building an API that can access data from anywhere is the easiest and simplest way to provide an end user with aggregated data.

BIO: Ryan Yu has been supporting the DOD for more than 10 years and has supported commercial customers in the software and finance industries for over 15 years. His background is in systems engineering and supporting big data applications and infrastructure. His BS is in information systems from Penn State, and he has a master's in information systems engineering from Johns Hopkins.

Building an Army Data-Centric Culture

Katrina Matthews, Army BD Senior Manager, GDIT • katrina.matthews@gdit.com

ABSTRACT

The Army has recognized the need to establish a data-aware culture that will recognize the imperative for data accuracy, data sharing and the strategic implications of data for the Army as a critical asset. As culture is the "collective conversations of an enterprise," fostering data conversations will help drive data-centric behaviors and belief systems across the Army. To begin, increasing organizational data literacy is key. This is a challenge area as there commonly exists a language barrier between business and information technology functions; however, by improving data literacy, the Army workforce will obtain the knowledge, aptitudes and skills to interpret and act upon data. The enterprise will begin to internalize the importance of its data assets and become more data-quality conscious. Data literacy can be promoted within the Army by:

- Educating the entire workforce on the Army's strategic, operational and tactical data policies is an initial step in building data literacy.
- Further promoting awareness of Army enterprise data management policies, BMA, WMA and EIEMA policies, and mission-owner data management guidelines.
- Encouraging a mindset of data learning as an opportunity for personal growth and career development will engage those who are interested in advancement.

Additionally, to develop a truly data-focused culture, the Army should create a set of data-centric competencies. For example, competencies in data management, information asset management, analytics, data sourcing, use case creation, data science, including machine learning, analytics platforms, and information governance all will lay the ground work for a broad Army organizational data focus. Widely available, focused and iterative training opportunities and self-service tools will be required to build these competencies. Coursework that educates the Army workforce on the strategic criticality of data as an asset should be mandatory for all Army personnel. Additional role-specific training should be available and required for some job functions. Not only should the training focus on specific tool sets that are part of the Army's data ecosystem but also on how to best perform the job function that uses those tools and an understanding of the upstream data sources and downstream data destinations.

Training and education on the Army Enterprise Data Catalog (AEDC) also is required. Phased training by role (data owner, data analyst, data consumer) will create understanding of the importance of data accuracy, methods for data cleansing and appropriate data tagging necessary to maintain the AEDC, as well as resources and tools that are available for performing data analytics. These analysts, armed with an understanding of the AEDC, will be able to find the data assets they need and respond rapidly to data requests from Army leadership to make actionable decisions.

The Army can significantly speed its cultural shift and augment its training and competency development by fostering communities of practice. Engaging cross-domain membership and participation and promoting competency-centered learning events and discussion forums will build an Army-wide sense of community and awareness of data challenges and initiatives, encourage sharing of best practices, and drive innovations by building upon the achievements of others.

BIO: Speaker Bio: Dave Vennergrund is a senior director and Distinguished Technologist at GDIT. Vennergrund has more than 25 years of experience in artificial intelligence, data analytics, data science, IT management and R&D. He led dozens of successful data mining, big data, AI and business intelligence efforts in intelligence, defense, and federal agencies including data lakes for Navy, predictive analytics at EPA, HUD and DOI, improper payment prevention at the IRS, USDA, CMS, VA, DFAS and OPM. Vennergrund has expertise in data analysis, predictive modeling, big data, cloud analytics (AWS, Azure, Google) and the deployment of analytic solutions in mission-critical settings. Vennergrund has built data mining, business intelligence and business analytics centers of excellence, special interest groups, and innovation centers. He is an industry expert in AI/ML, predictive analytics, fraud detection and data science.

Vennergrund earned his BS in computer science at the University of Illinois, and his MS in computer science from Arizona State University, specializing in the application of artificial intelligence to software engineering. Vennergrund has researched and applied computer science, statistical analysis and artificial intelligence methods to a broad range of national missions.

Providing Visible, Accessible, Understandable, Trusted, Interoperable and Secure Data With White Cloud Security Trusted Apps and Data Trust-Listing

Steven Shanklin, Founder and CEO, White Cloud Security Inc. • ziggy@whitecloudsecurity.com

ABSTRACT

The U.S. Army's data ecosystem can be secured and managed by White Cloud Security's Trust Lockdown (TL) Zero Trust App Security and Data Trust-Listing Framework, which enforces app control and data visibility, access, classification and protection at the endpoint regardless of whether the data is shared in the cloud, in a data center or on a specific endpoint. TL monitors and controls the creation, visibility, access and modification of data by users and automata processes across disparate systems including Windows 2000 and XP legacy systems.

One of the central problems with data protection in a cloud ecosystem is controlling which users and process have visibility, access and modification privileges to the data. While network access between endpoints and data sources can be monitored, controlled and secured, traditional file control access mechanisms are difficult to implement and manage at the endpoint processing layer. As with the distribution and management of crypto keys, creating and managing data access privilege lists must be transparent to the disparate systems, applications and users that need to access the data files. Furthermore, the creation and management of the data file access policies must be outside the control of an administrator who has root/supervisory privileges on the endpoints.

TL's cyber-metric handprint file identification technology uniquely identifies each file based upon the file's own data content, eliminates manipulating of data file identification tags and is always unique to each data file's content.

TL is a kernel-level file filter driver on Windows and a Linux Security Module in Linux that communicates with a secure service in the cloud or in a data center appliance that contains the trust-listing policies that determine which software is allowed to run and which data files can be seen, accessed or modified by a software package or component. TL's proven execution control security agent only allows trusted executables, libraries and scripts to run on endpoints according to specific trust pol-

icies for the endpoint or user on that endpoint. The data trust-listing extends the Execution Control Trust-Listing Framework to identify, monitor and control the creation of data files and modifications to them regardless of their location in the data ecosystem or whether the creator or user is a real user or an automata process.

Data attributes include, but are not limited to:

- Cyber-Metric Handprint Identifier unique to each data file or segment
- SHA-1, SHA-256, SHA-512, MD5, CRC32 and the data file or segment's length
- Creation time
- Last modification time
- Blockchain history
- Host identifier
- Host subgroup identifier
- App signature ID
- App compatibility profile list
- IP/MAC address of host
- Data type
- User and user's domain/group
- Classification of data
- Category of data
- Dirty or clean data attribute

TL works with both modern and legacy endpoints from Windows 2000 and Windows Servers 2003 without changes to the legacy endpoints other than installing TL's endpoint agent. It is supported from Redhat/CentOS kernels 3.10 after adding and enabling the Linux Security Module to the Kernel (Ports to Debian and Raspbian in Q2 2020).

BIO: White Cloud Security was founded by cybersecurity professionals with a proven track record and more than two decades of cybersecurity software development experience in leading-edge host and network intrusion detection, automated remediation and application whitelisting. Its previous companies were acquired by Cisco Systems (Wheelgroup, Psionic), TIS (Haystack Labs) and Lumension Security (Coretrace). Its Trust Lockdown is a "Zero Trust" App Security Framework that verifies the cyber-metric handprint identity of each Executable, Dynamic Code Library and Script every time they try to run. It blocks everything else.

Blur ST - Big Data SIGINT Search

James Kraemer, CEO, Data Intelligence Technologies Inc. • james@dataintelligencetech.com

ABSTRACT

Data Intelligence is the team behind Blur ST, its big data search platform, the Blur ST: enterprise search and discovery platform.

Blur ST was designed in the Intelligence Community from the ground up to be a petabyte scalable, fault-tolerant, search-and-retrieval platform that supports near-real time (NRT) big data indexing and sits natively on top of the Hadoop cluster.

- Search at petabyte scale: Fast indexing, fast search, big data.
- Geospatial mapping: Query geospatial/SIGINT data at scale.
- Native to Hadoop: Put full big data solution to work.
- Record-level security: Baked into the DNA.

BIO: Data Intelligence is an “All Things Big Data” shop specializing in data engineering, data clouds, data science, data security, data visualizations, data analytics, and data search and discovery supporting the U.S. Intelligence Community and Department of Defense customers.

VAUTIS Achieved With Varonis

Jim Evans, Federal Account Executive, Varonis Public Sector • evansjd0@gmail.com

ABSTRACT

The Varonis Data Security Platform (DSP) treats data management, governance and security as priorities in its overall paradigm shift in managing and protecting data. Varonis looks at the unstructured and semi-structured data first, not last, and empowers organizations to augment their data's efficacy and ensure its security. Personnel using Varonis have a greater understanding of their data, and therefore, are empowered to use it as a force multiplier to accomplish their mission no matter where the data may sit across the enterprise; on-prem locally, on separate data silos across the globe or in the cloud. The Varonis DSP is a foundational component in the Army's security posture to achieve Army Data Plan compliance and VAUTIS. Data will be visible and accessible to only those who should have access (least privilege); understandable (the data is sensitive, labeled and relevant); trusted (every file touch, email event, SharePoint folder, and AD touch is monitored and logged); interoperable (trusted and relevant to a given data set); and secured user and entity behavior analytics (UEBA platform monitors all data and alerts based on more than 180 threat models for insider threats, ransomware and malware behaviors, and new and ongoing APTs).

First and foremost, the Varonis DSP understands all metadata associated with unstructured and semi-structured data. This metadata collection allows the Army to quickly understand which user accounts across an enterprise can access specific datasets, identify how those individual users are gaining that access, understand what the users are actively doing with that data, and ensure that the data is secure from both internal and external threats by ensuring no accounts are behaving abnormally. User activity is collected in near-real time and alerts and monitoring are delivered as close to instantaneously as possible.

Varonis provides the capability to automatically identify and tag sensitive files based on common archetypes (classification levels, personally identifiable information, personal health information, just to name a few). This context brings greater understanding to personnel who need to make actionable decisions on how to best protect and secure data. These metadata streams can be used to ensure personnel are able to access their data while also ensuring the chain of custody and preserving data security. Varonis will also automatically identify and quarantine sensitive content to a known-to-be-secure location. In this way, even if spillage occurs, it will be secured until the requisite individuals can act.

Not only does Varonis provide the Army with an unprecedented understanding and risk profile for a critical attack surface, it also provides recommendations for permission changes to proactively help achieve a least permissive model and zero trust. Varonis can commit these changes in a way that does not impact mission requirements and automatically quarantines sensitive data as an incident response mechanism.

The Varonis Data Security Platform contains a full user and entity behavior analytics (UEBA) engine built using semi-supervised machine learning algorithms that define baselines for all users on how they interact with the unstructured and semi-structured data environment and automatically alert and respond to abnormal activity.

BIO: Jim Evans leads the Army team for Varonis Public Sector and has supported the American warfighter for 33 years.

Critical Data Enablement for VAUTIS

Deep Uppal, Vice President, Public Sector Technology Innovation, Information Builders •
deep_uppal@ibi.com

ABSTRACT

The Army Data Plan of November 15, 2019, describes a global, standards-based environment where data is visible, accessible, understandable, trusted, interoperable and secure (VAUTIS) throughout the life cycle. The Army Cloud Strategy in conjunction with VAUTIS has an opportunity to deliver dynamic, real-time data that supports defensive and offensive operations through relevant and timely information to decision makers. To harness the full potential of VAUTIS, the Army requires a multitenant cloud environment to provide secure access through a secure landing zone provisioned to provide federated data access across DISA's CAP while maintaining a zero trust perimeter. Core to the VAUTIS platform would be an integrated platform that provides multiple, concurrent methodologies to federate, abstract, integrate and provide secure, managed, mastered data access to potential tenants in the VAUTIS data platform.

Information Builders' integrated Omni-Gen platform was developed to improve analysis and decision making through automation and machine learning across a wide array of rapidly updating data sources. It was designed for open architecture integration, providing a flexible solution for agnostic-data integration. Native to the platform is an entire library of connectors that facilitates connection to virtually any data source. Hundreds of adaptors allow an easy and secure method to provide that data to any application, tenant or coalition partner. Additionally, the Omni Integration Platform has a native BI component that supports all levels of data analysis, reporting and visualizations. An integrated data science module would allow the Army to provide its analysts, data scientists and business users an easy-to-use interface to create and deploy a variety of data analytic models to include predictive and prescriptive. Finally, the platform can provide federated data analytics: the ability to provide meta-analytics to the VAUTIS environment without the need to develop, maintain or host data repositories, data lakes or data reservoirs.

Through numerous deployments throughout DOD and multiple federal agencies, the Omni-Gen platform can assist the Army with its siloed data assets that reside in several security zones (NIPR, SIPR, Coalition, Five-Eyes, K/J), commercial and government clouds (Azure, AWS) and physical subnets associated with legacy systems. This solution would provide data access without the buildout of additional infrastructure, providing an economical, viable and future-facing integration platform for current and future VAUTIS tenants.

Solution specifics:

- Connect to multiple external data sources in different environments with different associated security classifications to the GovCloud VAUTIS environment.
- Be compliant with Defense Information Systems Agency's Secure Cloud Computing Architecture (SCCA). This includes being able to leverage currently deployed DISN, non-secure Internet protocol router network (NIPRNet).
- Be ready to implement secret Internet protocol router network (SIPRNet) access migration through the secret-joint regional security stacks (S-JRSS), to virtualize tenant access in parallel to VAULT SIPR environment buildout.
- Install and configure Information Builders' Omni-Gen data integration platform to pull meta-analytics without having to store any data locally.
- Demonstrate the ability of the Omni-Gen integration platform to present the data in real time agnostically to tenant applications while adhering to STIG in accordance with RMF.

BIO: Deep Uppal is an innovator, technical problem solver and change agent with a proven track record of defining the technical vision; communicating complex processes; and successfully creating, integrating and deploying next-generation enhanced analytics, network architecture and all associated facets of technology related to state and federal government. Uppal is a member of the U.S. Army Signal Corps.

CONTENT DISCOVERY

Data Governance Best Practices Including Automated Metadata Generation

Thomas Ward, AI Project Leader, IBM Global Chief Data Office, IBM • tomward@us.ibm.com

ABSTRACT

An effective and efficient data catalog supports multiple functionalities:

- Curation tools for annotating and cataloging incoming data properly.
- Enforcement governance policies for data quality and standards compliance.
- Extensions and enrichment to the business glossary.
- Search and exploration of the data in the catalog.

As data catalogs grow in size and complexity, automated AI solutions are critically required to scale these functions. IBM Watson Knowledge Catalog (WKC), powered by IBM Cloud Pak for Data, is a data catalog that tightly integrates with an enterprise data governance platform. Data catalogs can help data citizens easily find, prepare, understand and use the data they need.

Watson Knowledge Catalog (WKC) helps business users quickly discover, curate, categorize and share data assets, data sets, analytical models and their relationships with other members of the organization. It serves as a single source of truth for data engineers, data stewards, data scientists and business analysts to gain self-service access to data they can trust. With data governance, data quality and active policy management, WKC helps an organization protect and govern sensitive data, trace data lineage and manage data lakes.

Automated Metadata Generation (AMG) automates the process of discovering, organizing, and curating data using deep learning technologies. AMG offers suggested metadata labels by looking for patterns in field-level data with technical metadata. AMG enhances speed to access and understand data within WKC. AMG has delivered to IBM's Chief Data Office running on a Cognitive Enterprise Data Platform (CEDP) a 90 percent reduction in cycle time for meta data analysis, resulting in $27 million in productivity savings over the past two years. AMG has dramatically enhanced data quality with regulatory and governance checks.

Metadata is just as important as data. It is the underpinning necessary in today's data era to derive meaningful business insights. Every enterprise struggles with the problem of labeling massive amounts

of data. It's usually a labor-intensive manual process completed by several subject matter experts that can take weeks. With the explosion of data from several technologies, it's critical to have metadata definitions.

Classifying data, such as sensitive data, is a required step to meet regulatory compliance such as GDPR or Government-Owned Entities (GOE). This enables the right course of action in the handling of this data. To address the scale and speed necessary in data labeling, artificial intelligence is a key foundational technology.

This session will describe the main characteristics, components and approaches to building and maintaining a catalog integrated with the data lake. This session will specifically cover:

- The components, characteristics of a catalog.
- The approaches to building and maintaining the business terms in a catalog.
- How the catalog is used to govern the data lake assets.
- How the catalog is used to support self service business insights and other user activities.
- The role of AI/ML in delivering further levels of automation and efficiency to this catalog.
- How the catalog can leverage graph technology and ontologies.

BIO: Thomas Ward is currently an AI project lead within the Chief Data Officer (CDO) organization. In this capacity, he leads the global development and deployment of cloud and AI projects across the supply chain. He has led supply chain cloud projects over the past 10 years. For the past five years, Ward has led the implementation of several Watson applications for the IBM Internal Supply Chain. He has been a featured supply chain conference speaker and published author in several magazines.

Ward is an IBM Academy of Technology Member and one of 25 certified Supply Chain Management Professional–Consultants, globally within IBM. Ward has 30 years of technical leadership experience with IBM in all facets of the supply chain from manufacturing engineering to production supervision and materials management through logistics and procurement.

Ward has a MS in electrical engineering from Rensselaer Polytechnic Institute.

Spectral Hypergraph Analytics for Pattern Discovery and Data Tagging as a Fog-Level Service

James Ezick, Vice President, Engineering, Reservoir Labs • ezick@reservoir.com

ABSTRACT

Spectral Hypergraph Analytics (SHA) provides a novel, unsupervised machine learning approach to decomposing large-volume sensor data into coherent patterns. The approach solves a critical problem in the application of machine learning in data science: how to gain deep insight from the entirety of large-scale unlabeled multidimensional data while supporting, but not requiring, heroic up-front feature engineering. A hypergraph is an extended form of graph where edges can link more than two nodes. SHA treats tabular data as a hypergraph where edges are derived from rows linking elements spanning multiple dimensions of data. Spectral analysis is then a factorization of the resulting adjacency structure into components, with each component capturing a specific set of correlations in the data. For sensor data, a component can typically be interpreted as a pattern of behavior that can be independently analyzed, tagged and stored for later reference. Tagging derived from spectral components can include the recognition of recurrent patterns over time or important co-occurrences between multiple discrete entities across multiple dimensions of data. In this way, the set of components derived from spectral analysis form a concise roadmap of a dataset in that they separate, summarize and weight the dominant activities present in the data.

Fog computing refers to an emerging layered model where services can be deployed in a fog layer between cloud resources and edge devices. For the Army, the fog model provides a way to deploy localized data analytics that sit between edge sensors and cloud repositories. The model is ideal for deploying data preparation and tagging based on SHA in that the computationally intensive factorization routines can be moved off of sensors, but analysis and the association of metadata occurs before data is submitted to the cloud to be made accessible for search and retrieval.

Reservoir Labs has developed an implementation of SHA called ENSIGN. The company shows how ENSIGN supports data decomposition in multiple domains, including successes from the discovery of patterns in human behavior and from threat detection in cybersecurity. It shows how components factored from datasets can be used to add descriptive information to clusters to data as well as provide another way to support immediate forensic analysis and decision making. Reservoir Labs also describes the computational requirements of the approach and makes the case for why a fog deployment model makes sense.

BIO: James Ezick is the vice president, engineering for Reservoir Labs and also serves as the technical area lead for Reservoir's analytics and cybersecurity teams. Since joining Reservoir in 2004, Ezick has developed solutions addressing a broad range of research and commercial challenges in high-performance computing, machine learning, data analytics, cybersecurity, compilers and verification. Ezick has more than 15 years experience managing and executing advanced R&D projects, leading to commercial transitions, patents and research publications. He received his BS in computer science and applied mathematics from the State University of New York at Buffalo in 1997, and MS and PhD degrees in computer science from Cornell University in 2000 and 2004.

AI-Backed Categorization and Tagging With Tamr

Burt Wagner, Senior Solutions Engineer, Tamr • burt.wagner@tamr.com

ABSTRACT

For every organization in existence today, the Army being no exception, the amount of data it handles is vast and growing at an accelerating rate. The ability to effectively identify, categorize and most importantly understand every row of data in every data source at even a small organization quickly outstrips the available resources. Without some means to force multiply domain experts in identifying and tagging data, organizations quickly fall ever further behind. With the vast quantities of data the Army faces, only cutting-edge tools backed by state-of-the-art artificial intelligence and machine learning (AI/ML) can address this problem.

Tamr is a company leading the charge in leveraging such AI/ML tools to address the problem of categorizing data for large organizations to achieve data mastering. To align business need with actual production data, Tamr's unification platform provides clustered views of entities based on data descriptions and their impact in terms of dollar amounts, personnel impacted or any other available metric. Current DOD use cases for data modeling and unification include logistics, supplies and parts, preventive maintenance, personnel, training and readiness, procurement data and intelligence.

Both existing data holdings as well as new data as it is generated, captured or discovered can be placed into meaningful taxonomies of tags. This is done by having human domain experts quickly and easily teach Tamr how to identify and tag rows of data, and then the AI/ML engine runs on highly scalable architectures to address production-size data holdings.

Tamr can help the Army identify, tag and make useful its data holdings in many diverse subject areas, from logistics, training and financial data to personnel data to mission specific capabilities.

BIO: Burt Wagner has been helping national security customers solve data problems for nearly 20 years. He has been a data architect, data engineer and data scientist within the IC for most of that time, working with both relational and NoSQL data solutions to produce meaningful advanced analytics while meeting data security needs for national security customers. Now with Tamr, Wagner is helping customers across the federal government unify data sets across silos and make data-driven decision making possible.

Supply Hub for Operational Predictive Maintenance Analytics (SHOPMAN)

Nikhil Shenoy, President, Colvin Run Networks Inc. • nikhil@colvinrun.net

ABSTRACT

Colvin Run Networks' SHOPMAN analytics platform is powered by MicroStrategy, a leading global business intelligence platform with hundreds of thousands of users worldwide. SHOPMAN supports physical and logical asset management through hundreds of out-of-the-box application program interfaces (APIs) that integrate with a variety of legacy and modern data systems. Platform-agnostic data integration and data wrangling (DI/DW) enables data governance infrastructure with rapid prototyping and scalable, robust management capabilities and ultimately flexible delivery pipelines, including PC or mobile-native (i.e. text messages) or mobile-tailored (i.e. custom apps) to get the right information where it's needed, to who needs it, when they need it and how they want it.

Colvin Run's data scientists perform exploratory data discovery and anomaly detection via SHOPMAN using methods such as dimensionality reduction and outliers analysis on a variety of vectors, and then choose algorithm weights tailored to Army requirements to unearth data of interest from a variety of sources. This means data can be authoritatively identified, clustered and tagged within the same SHOPMAN environment after DI/DW. The key benefit is maximizing data utilization—for example, if data is the new oil, you must still mine it and refine it for value— with seamless sharing using the leading enterprise platform from MicroStrategy, now the largest independent public business intelligence company in the world.

BIO: Nikhil Shenoy is the co-founder and CEO of Colvin Run Networks, an applied data science company that has received Virginia Center of Innovative Technology (CIT) Business Awards in 2017, 2018 and 2019. Colvin Run Networks has customers in commercial industry and defense, including multiple DOD SBIR Awards for blockchain and artificial intelligence/big data analytics applications.

Shenoy is a sitting member of the AFCEA Technology Committee and was featured in AFCEA SIGNAL Magazine in August 2019. He is also a leader with the NDIA's Electronics Committee. Shenoy has more than a decade of experience in technical product development, delivering initiatives for leading companies in a variety of industries, including finance, consumer goods and mobile technology at Goldman Sachs, Procter & Gamble, and Kastle Systems, respectively. Shenoy holds a chemical engineering degree from the Massachusetts Institute of Technology, and an MBA from the University of Chicago's Booth School of Business.

Object Technology for the Army's Data

Scott Rich, Deputy CTO Americas, NetApp • scott.rich@netapp.com

ABSTRACT

Like any physical construction effort, building a new, global data architecture for the Army will require the proper foundation. The ability to deliver modern applications and services that provide leading-edge capabilities to the warfighter will depend on that foundation being able to perform and scale with the needs of a dynamic, distributed and diverse data environment. Deploying a cloud-native architecture based on open protocols will allow the Army to support existing, legacy applications while also being able to provide capabilities to move the services deployed to the soldiers to a more modern, hybrid-cloud solution. Deploying this data-centric architecture using object technologies will provide all of the capabilities the Army is driving to achieve.

Object repositories, which scale to hundreds of petabytes in capacity and billions of objects, can deliver the content discovery and retrieval capabilities to make data available to any authorized user on any device from any location. In order to utilize the public cloud in an efficient manner, the Army must approach its own requirements with a hybrid-cloud thought process—deploying cloud technologies throughout its own enterprise in order to manage data in a single, global namespace—discoverable from anywhere in the enterprise within the Army networks or in public clouds. By deploying object technologies across the enterprise, the Army can manage data through policies defined at the program level, allowing data to transparently move throughout the infrastructure, into and out of the public clouds, across echelons and to the tactical community when needed.

The use of an object foundation for the Army's primary data architecture brings with it the ability to layer COTS capabilities that are found in the public clouds today—connecting identified datasets to cloud-based discovery tools like Elastic, deploying a private, web-based file storage capability, running content and media asset management systems used across commercial enterprises today, or creating new Army specific applications that can transparently move from public cloud to deployed systems.

Object technology is the data capacity technology of the public cloud today and will be across the enterprise in the future. NetApp has more than a decade of history working with the Army to manage data and provides the tools and services to deploy a modern, data-centric object environment to help build future capabilities for the warfighter.

BIO: NetApp is a leader in data management and has over half of the Army's data currently hosted by NetApp data management systems. The company is uniquely positioned to help move the Army to a more data-centric culture and infrastructure in order to leverage innovative technologies from cloud service providers, AI and ML services, and mobile, containerized application solutions.

Data Management Using Metadata To Discover, Search, Distribute, Access and Retain Your Data

Bobby Rountree, Data Intelligence Technical Lead, Hitachi Vantara Federal •

rountree.bobby@gmail.com

ABSTRACT

Object storage is powerful because of the metadata. Metadata is used to describe the content and objects being managed. The government needs to use metadata as a strategic way to enhance its data management solutions. The use of metadata allows the government to search for key data associated with files, determine where and when the content should be distributed, deliver content based on business rules and metadata values embedded in the objects and rely on content metadata when applying retention rules. A strong data management solution in the government will allow it to securely search, access and distribute data from anywhere. A strong data management solution with metadata in the government will allow an organization to turn data into value.

BIO: Bobby Rountree joined Hitachi Vantara in 2016 and is currently data intelligence lead for Hitachi Vantara Federal and CEO of the Dapper Data LLC brand. Throughout his career at Hitachi, Rountree has focused on the content intelligence ecosystem and helping government organizations build a strong data management platform, using products such as Hitachi Content Platform (HCP) object store, HCP Anywhere File-Sync and Share, and Hitachi Content Intelligence Elastic Search. His passion has always been to help people make better decisions with their data. Rountree has been a thought leader and influencer in the knowledge management and data science industry.

Rountree has more than 10 years of experience in data management, programming and data science. He holds a BS in computer science from the Bowie State University, an MS in cyber security from University of Maryland University College and currently is pursuing his doctorate in data science from North Central University.

Regain Control of Your Data

Allen Greene, Account Manager, Veritas Technologies LLC • allen.greene@veritas.com

ABSTRACT

Today, many enterprises maintain complex data landscapes that include a multi-cloud architecture and a diverse storage infrastructure. Because of massive data growth and increasing data privacy country- and vertical-specific regulations, organizations are seeking ways to visualize the data in their infrastructure and reduce risk.

With Information Studio, Veritas addresses this problem, providing IT professionals with a tool to easily identify specific types of details about data and the information it contains and pinpointing areas of risk, waste and potential value.

Information Studio offers clear visibility, targeted analysis and informed action on data, so organizations can confidently address security concerns, upcoming regulations and continued data growth—ultimately improving end-to-end efficiency in data management and regaining control of their data.

Information Studio works by gathering information about data within a company's infrastructure from both native data sources and Veritas NetBackup.

It then allows users to filter information found for specific combinations. For example, users can choose to filter for data older than 2 years containing personally identifiable information (PII) or that includes a specific phrase like "fraud" or "trade confirmation" that needs to be retained for legal hold.

Once Information Studio has identified important information, users can extract reports that enable them to make evidence-driven decisions about the data. For example, users can defensibly delete stale data and better protect personal data, demonstrating a good faith effort to protect PII and regulated data.

Information Studio is offered as a virtual appliance deployed on-premises, which removes concerns about data residency. It was purpose-built to offer flexible licensing and increased control on where data should reside.

BIO: Veritas Information Studio is a comprehensive intelligent data platform that helps organizations address their data-related challenges. With the power to organize data and take informed action, organizations can be confidently prepared to handle security concerns, new regulations and continuous data growth to ultimately regain control of their data.

Intelligent Metadata Management

Michael Anderson, Chief Federal Strategist, Informatica • mianderson@informatica.com

ABSTRACT

To better understand all the information available in an enterprise and unleash its full value, orgnizations need context. Metadata provides this crucial element, allowing them to better understand their data's quality, relevance and value.

Metadata helps organizations discover data, understand data relationships, track how data is used and assess the value and risks associated with its use. As data continues to grow at an explosive rate and become more distributed, these are turning into mission-critical processes—which is why metadata management now plays a central, strategic role in driving digital transformation.

Metadata becomes even more valuable if it is active—overlaid with machine learning, augmented with human knowledge and integrated. It makes the wider data management processes intelligent and dynamic. Active metadata can be the vital foundation of a well-architected data management system, yielding benefits across the entire life cycle of data projects. For example, metadata can highlight missing, incorrect or anomalous data. By tapping into the metadata, systems can automatically correct and enrich the data feeding into a report, avoiding costly errors and enhancing the quality of the analytics to improve decision making.

Informatica Enterprise Data Catalog (EDC)

Informatica's metadata management approach is designed to help enterprises fully harness the value of all their data with active metadata. Informatica EDC allows enterprises to start this journey by tapping into four major categories of metadata:

- Technical: Database schemas, mappings and code, transformations, quality checks.
- Business/Mission: Glossary terms, governance processes, application and business context.
- Operational and infrastructure: Run-time stats, time stamps, volume metrics, log information, system and location information.
- Usage: User ratings, comments, access patterns.

Metadata in these four categories becomes the basis for a common metadata foundation. Informatica EDC uses a rich set of capabilities to create this shared foundation:

- Collect: Scan the metadata from the entire enterprise's data systems across cloud and on-premises, including databases and file systems, integration tools and processes, and analytics and data science tools—with a high level of fidelity.

- Curate: Document the mission and business view of data with glossary terms, concepts, relationships and processes. Augment the collected metadata with this business context. Gather user input in the form of ratings, reviews and certifications to help assess the usefulness of data assets to other users.

- Infer: Apply intelligence to derive relationships not obvious in the collected metadata, including data lineage, data similarity and ranking the most useful data sets for different types of users.

Summary of Informatica EDC key benefits:

- Automatically catalog and classify all types of data across the enterprise using an AI-powered catalog.

- Identify domains and entities with intelligent curation.

- Enrich data assets with governed and crowdsourced annotations.

- Find data assets through powerful Google-like semantic search.

- Discover and understand your data assets with data profiling and quality stats, 360 relationship views and lineage.

- Get a complete picture of your data environment.

- Open APIs to integrate into your environment and expose intelligent metadata anywhere.

BIO: Col. Michael R. Anderson, USA (Ret.), serves as the chief federal strategist at Informatica LLC. In this role, he leads strategic activities for Informatica with a focus on enterprise data management and data security software capabilities supporting the U.S. Department of Defense and U.S. federal agencies.

Know Your Data With Veritas Information Studio

Doug Snyder, Chief Technologist, Veritas Technologies LLC • doug.snyder@veritas.com

ABSTRACT

Today, many organizations maintain complex data environments that include cloud and on-premise storage infrastructures. Because of massive data growth and increasing data privacy regulations, organizations need ways to visualize that data to reduce risk and simplify management.

With Information Studio, Veritas addresses this problem. IT professionals now have a tool to easily identify details about data and the information it contains, pinpointing areas of risk, waste and potential value. Information Studio offers clear visibility, targeted analysis and informed action on data, so organizations can confidently address security concerns, upcoming regulations and continued data growth.

Information Studio works by gathering information about data within a company's infrastructure. It then allows users to filter information found for specific combinations. For example, users can choose to filter for data older than 2 years containing personally identifiable information (PII) or that includes a specific phrase like "Top Secret" or a program name that needs to be retained or deleted. Once Information Studio has identified important information, organizations can make evidence-driven decisions about the data. For example, they can defensibly delete stale data or protect personal data, demonstrating a good faith effort to protect PII and regulated information.

BIO: Doug Snyder is the chief technologist for the U.S. public sector, which includes the U.S. government, Department of Defense, state and local governments as well as health care providers. Snyder is currently in his 13th year at Veritas.

Snyder holds certifications for OpenStack (Certified Architect); CISSP (Certified Information Systems Security Professional) and ITIL v3 (Information Technology Infrastructure Library).

Controlling the Visibility and Discovery of Data With White Cloud Security Data Trust-Listing

Steven Shanklin, Founder and CEO, White Cloud Security Inc. • ziggy@whitecloudsecurity.com

ABSTRACT

White Cloud Security's Trust Lockdown (TL) Data Trust-Listing Framework provides the U.S. Army's data ecosystem with control over data visibility and discovery at the endpoint regardless of whether the data is shared in the cloud, in a data center or on a specific endpoint.

TL tracks each data file based upon a Cyber-Metric Handprint File Identification technology that uniquely identifies each file based upon the file's own data content. This in turn prevents file identity spoofing through manipulating of data file identification tags and is always unique to each data file's content.

TL's Data Trust-Listing Framework automatically tracks the creation, access and modification of data files via a blockchain of the Cyber-Metric Handprints along with the relevant data file attributes. These data file attributes provide data administrators with the ability to trust users and automata to discovery and access and/or modify specific data files or sets of data files defined by type, category or classification.

The relevant data file attributes include but are not limited to:

- Cyber-Metric Handprint Identifier unique to each data file or segment
- SHA-1, SHA-256, SHA-512, MD5, CRC32 and the data file or segment's length
- Creation time
- Last modification time
- Filename/pathname
- Blockchain history
- Host identifier
- Host subgroup identifier
- App signature ID

- App compatibility profile list
- IP/MAC address of host
- Data type
- User and user's domain/group
- Classification of data
- Category of data
- Dirty or clean data attribute

TL's tracking of data creation and modification automates the identifying, tagging and registering all authoritative data in a way that makes it easily discoverable by trusted users and automata across the enterprise ecosystem.

TL's kernel-level control of file visibility, discoverability, access and modification is transparent to the disparate applications and users that need to access the data files and can be managed either directly in the Trust-Listing Framework or via its API. Further, TL controls which apps run on endpoints and which data these apps can access. This is all done at the kernel level and cannot be bypassed even by administrators with root/supervisory privileges at the endpoint or in the network.

TL is a kernel-level file filter driver on Windows and a Linux Security Module in Linux that communicates with a secure service in the cloud or in a data center appliance that contains the trust-listing policies that determine which software is allowed to run and which data files can be seen, accessed or modified by a software package or component. TL's proven execution control security agent only allows trusted executables, libraries and scripts to run on endpoints. Data Trust-Listing extends the Execution Control Trust-Listing Framework to identify, monitor and control the creation of data files and modifications to them.

TL works with both modern and legacy endpoints (from Windows 2000 and Windows Servers 2003) without changes to the legacy endpoints other than installing TL's endpoint agent. It is supported from Redhat/CentOS kernels 3.10 after adding and enabling the Linux Security Module to the Kernel (Ports to Debian and Raspbian in Q2 2020).

BIO: White Cloud Security was founded by cybersecurity professionals with a proven track record and over two decades of cybersecurity software development experience in leading-edge host and network intrusion detection, automated remediation and application whitelisting. Its previous companies were acquired by Cisco Systems (Wheelgroup, Psionic), TIS (Haystack Labs) and Lumension Security (Coretrace). Its Trust Lockdown is a zero trust app security framework that verifies the cyber-metric handprint identity of each executable, Dynamic Code Library and Script every time they try to run. It blocks everything else.

Army Data Tagging for Strategic Value

Katrina Matthews, Army BD Senior Manager, GDIT • katrina.matthews@gdit.com

ABSTRACT

There are two major elements to this effort. Data tagging and building a metadata catalog of the massive amounts of Army data is key if the Army's data is to be useable for strategic analysis by Army leaders and organizations. Tagging creates groupings to organize the data along technical, business, operational and security dimensions. Manually tagging metadata for every data source in the Army is not practical, too time consuming and costly. The Army needs to utilize automation to do tagging, augmented by machine learning (ML) so the cataloging gets smarter over time. Some user intervention is always necessary, but this method is more practical, more sustainable and less labor intensive.

Tools have been developed for just this purpose. Data hosted in the cloud, such as AWS GovCloud, can make use of Glue crawlers that will read and store the metadata in the catalog. Informatica, Waterline Data and Alation are COTS tools that perform similar functionality. These tools utilize crawlers and can process data in Hadoop clusters and relational databases and even inferring schema based on field content, when no header is available. This is where artificial intelligence and machine learning are particularly valuable. By using tags already assigned to data structures, ML methods could learn how to tag new data, treating the task as a typical ML classification task. The Army's functional analysts play a key part in accepting or rejecting inferred tags, which teaches ML software to make better recommendations in the future.

As the Army begins to systematically tag and extract metadata during ingestion of data, the target result is the second major element, the creation of an Army Enterprise Data Catalog (AEDC). The catalog is the mechanism for providing full data awareness to organizations within the Army, enabling data analysts to find and filter the right data for the intended purpose. There are numerous enterprise-level products available to build the catalog. AWS Glue, Collibra, Talend, Informatica Enterprise Data Catalog and Alation are all viable options, and each has its own feature sets for the Army to consider. One significant benefit of a data catalog is that users can be aware that data exists without having access to it. Access can still be controlled through the Army's governance model and revoked when no longer needed.

To accomplish the objectives above, the Army might form cross-functional tiger teams, overseen by the appropriate Army directorate. These teams must be comprised of three key roles. The first is personnel with strong functional knowledge who can recognize the data and its potential uses. They can assign quality and usefulness scores to the data assets to increase the trust in their use. The second is data engineers who are experts in using the described information above, as well as in sound data

science principles. The third is the translator, an individual who knows the functional value of the data and enough about data science to broker clear understanding between the functional SMEs and the data scientists.

BIO: Speaker Bio: Dave Vennergrund is a senior director and Distinguished Technologist at GDIT. Vennergrund has more than 25 years of artificial intelligence, data analytics, data science, IT management, and R&D. He led dozens of successful data mining, big data AI, and business intelligence efforts in intelligence, defense and federal agencies including data lakes for Navy, predictive analytics at EPA, HUD and DOI, improper payment prevention at the IRS, USDA, CMS, VA, DFAS and OPM. Vennergrund has expertise in data analysis, predictive modeling, big data, cloud analytics (AWS, Azure, Google) and the deployment of analytic solutions in mission critical settings. Vennergrund has built data mining, business intelligence and business analytics centers of excellence, special interest groups, and innovation centers. Vennergrund is an industry expert in AI/ML, predictive analytics, fraud detection and data science.

Vennergrund earned his BS in computer science at the University of Illinois, and his MS in computer science from Arizona State University, specializing in the application of artificial intelligence to software engineering. Vennergrund has researched and applied computer science, statistical analysis and artificial intelligence methods to a broad range of national missions.

Knowledge Management at Echelon

Gregory Wallsten, Consultant, U.S. Army Retired • gregorywallsten@gmail.com

ABSTRACT

Correctly tagging data for wide distribution and timely access across multiple echelons is a very challenging task and currently not effective across the joint force because of the large volumes of information and the large numbers of participants. The capability of the network expanded and massive amounts of data was easily saved for sharing, which created a new problem of information overload. Most knowledge management (KM) systems are designed to work for a select few who use their information daily, but because their KM is not standardized, engines are not optimizing results. This is problematic as it is time consuming and sometimes not practical to collaborate with multiple stakeholders. Additionally, data management is critical for the entire force as we routinely conduct joint/multinational operations together as one team. While planning, staffs have an increasingly difficult task of developing the best options for commanders. However, the staff will drive on with the information they have in a time-constrained environment, even if it is short of what is available.

The Army must better organize data to enable superior mission command. This is best achieved through a collective KM system executed through KM teams at all echelons. Assign this responsibility to those who are certified with this critical process and establish measures of effectiveness. From ATP 6-01.1 (1-1), “Knowledge management (KM) is the process of enabling knowledge flow to enhance shared understanding, learning, and decision making.” To further strengthen the process, attach an industry representative to the unit’s MTOE/TDA to provide industry support where it will be most effective.

The KM cell should be charged and held accountable for enabling this knowledge flow. Knowledge managers should become more proactive and continuously look for techniques to improve their systems holistically. The KM team needs to enforce its SOPs, but it should also aid the entire team to ensure tagging and naming conventions are universal across the enterprise. The KM team must work together with the Signal (Communications) sections at all echelons above the company level to ensure ease of access for those who need it and support network security, which is another topic.

The knowledge management teams must ensure proper tagging of all reports from inside their command through direct action by monitoring what is published on the portal/shared drive. They should assist subordinate staffs with KM training and improved SOPs. When the knowledge management team works with highly trained staffs, the KM process will become institutionalized and self-healing as expectations rise and sloppy KM practices are addressed at lower levels. The best knowledge managers can share their lessons learned with the community as they refine their TTPs and SOPs. The industry representatives can reach back to their organizations for solutions greatly enhancing

modernization efforts. As this area continues to evolve, the Army will be better prepared to face new challenges that arise. The KM team will oversee and control its unit's knowledge requirements while synchronizing with the entire enterprise for optimal data management.

BIO: Lt. Col. Gregory Wallsten,USA (Ret.), served 24 years as an Armor officer with four combat deployments. His last assignment was with the U.S. Army Joint Modernization Command where he served for 20 months as the G-3 Chief of Operations. He was CHOPs for the Network Integration Evaluation (NIE) 18.2; and the Joint Warfighting Assessment 18 and 19.

An Authoritative Solution for Enterprise Information Management (U.S. Army)

Deep Uppal, Vice President, Public Sector Technology Innovation, Information Builders •
deep_uppal@ibi.com

ABSTRACT

The foundational goal of the U.S. Army data plan is to identify, tag and register all authoritative data in a way that makes it easily discoverable by users across the enterprise. The Army is in the process of establishing the governance of its enterprise data so that data security is optimum. The Army requires an enterprise information management (EIM) plan to support a master data management approach to data governance and stewardship. As the DOD Data Tagging Strategy proposed that data must be tagged when it is created, acquired or modified, the service would require a tool that provides an integrated offering that can support EIM activity.

The Army requires a system that automatically samples a wide variety of sources, including ETL, unstructured, relational, series and virtualized sources to harvest and append metadata for each associated dataset. The result: a dynamic index that would track and improve efficiencies within the Army's data structure and provide a consistent means of data governance. Finally, the Army would need a meta-data catalog to provide a common process and create a common data fabric to access BI-Tools and associated AI enriched analytics. Because of security concerns, the Army would require the creation of said catalog and associated mapping without the true movement of data and only though the ingestion, creation and appending of meta-events.

Information Builders' (IB) Omni-Gen (Omni) master metadata management integration and mastering platform enables rapid, model-driven implementation of master data management (MDM), data cleansing and data integration projects. It provides a unified environment through which the Army can quickly and easily define integration, meta-data quality, match/merge, remediation and unification plans. Omni-Gen delivers multidomain enterprise data governance through the Omni Governance Console (OGC). Built-in authorization for each domain ensures that each user has role-specific capabilities. Users can configure customizable workflows to provide automatic alerts when data quality threats require intervention. The 360 Viewer provides a complete web-based view of golden records in a data mastering environment.

IB's platform can harvest metadata and store it in a centralized repository. Each meta-operation results in a complete catalog of meta-events, transformations, dimensions and ownership details related to that data. Smart algorithms enable modeling and indexing of all metadata types to quickly locate and provide context to all cross connections. Additionally, a federated search application provides a smart

engine using hundreds of crawlers that searches all metadata and associated artifacts within seconds. Finally, this catalog contains a dynamic visual repository, a map that provides full-data lineage of data and affiliated systems as it flows through multi-vendor, external and internal systems.

Solution specifics:

- Data Automation: the automation of tagging and profile datasets, including the automated review of metadata profiles through AI filters and heuristic rulesets.
- Scalability: IB's Omni Platform allows for scalability and caching of metadata artifacts for quick mapping, updating and visualization. As the flow of data increases, Omni can continue to process and update catalog entries without resorting to resource-intensive operations.
- Flexibility: IB's platform functions equally well in the cloud, on premise or through any variation of hybrid offerings.
- Tool Integration: the creation of API's and an emergent metadata consumption layer for easy connections back to the associated database for ease of BI-reporting

BIO: Deep Uppal is an innovator, technical problem solver and change agent with a proven track record of defining the technical vision; communicating complex processes; and successfully creating, integrating and deploying next-generation enhanced analytics, network architecture and all associated facets of technology related to state and federal government. Uppal is a member of the U.S. Army Signal Corps.

Identifying and Tagging Sensitive and Classified Data Sets

Jim Evans, Federal Account Executive, Varonis Public Sector • jevans@varonis.com

ABSTRACT

The Varonis' Data Security Platform treats data governance and management as a priority. The mission of the Data Security Platform is to aid enterprise organizations in protecting their data from misuse and abuse. Using the platform, relevant personnel can gather, correlate, understand and respond to any possible metadata associated with any unstructured or semi-structured monitored platform. Within the Data Security Platform, the Data Classification Engine allows enterprises to automatically discover and tag data based on the content of individual files. Discovering this sensitive content is made easy: hundreds of built-in patterns looking for commonly understood sensitive data sets, such as U.S. personally identifying information (PII), personal health information (PHI/HIPAA), General Data Protection Regulation (EU GDPR), or California Consumer Privacy Act (CCPA). Users can also create rules to scan for customer-specific sensitive data and file sets. Examples of how these custom rules are commonly leveraged to discover and tag data include classified documents (spillage), Freedom of Information Act Requests (FOIA) and many more. These custom rules can also be used to identify if documents are classified appropriately, essentially determining when document content may be more sensitive than currently-applied classifiers. Varonis' classification capabilities include the application of specific tags that can be leveraged by other industry partners, taking an integrated approach to data loss prevention.

One major focus of Varonis is to aid enterprise customers in identifying their sensitive data; however, this is only an initial step. Varonis also seeks to proactively aid customers in securing any discovered sensitive data. To accomplish this mission, the software has automated tools in place to help organizations make decisions on how to best protect their sensitive data. Rules can be created and enforced such that nonsecure sensitive data can be automatically moved into select secure repositories. A common use case involves spillage of classified documents onto NIPR: the document in question can be quarantined to a secure location automatically for review by G-2 personnel. The Varonis Data Security Platform can help deliver a least permissive model; delivery of a least privilege framework includes support for "the need to know." The platform informs security staff where individuals or groups may have access that is not required by their job role or duties. This process limits risk by facilitating the removal of those individuals that have access to sensitive datasets but do not need it. Secondly, Varonis provides organizations with an efficient and secure enterprisewide eDiscovery solution: DatAnswers. This module considers all metadata, including discovery rules tagged within the Data

Security Platform and allows users with the requisite permissions to search for data no matter where it may be stored. Varonis will automatically enumerate search results based on users to ensure they only see data to which they have access. This allows users to get the most up-to-date information regarding active mission requirements, while also preserving the need to know by not allowing users to search for data they shouldn't be able to access.

BIO: Jim Evans leads the Army team at Varonis and has supported the American warfighter for 33 years.

Enabling Unsampled Network Visibility on 100/200/400G Links

Scott Rey, Director, NetQuest • srey@netquestcorp.com

ABSTRACT

As network traffic continues to grow at a staggering rate, maintaining effective network visibility is increasingly challenging. Further, the expanding rate of network-based attacks is mandating full unsampled network visibility. Security teams require real-time pervasive visibility to detect and overcome attacks in a rapidly evolving threat landscape. However, we are fast approaching a compute technology tipping point in which network traffic growth is overwhelming today's security monitoring technology.

Using real-time packet capture to monitor and analyze security threats across high-capacity 100Gbps network links is impractical and costly. Even network flow visibility solutions struggle to monitor these high-speed links using traditional multi-core processors and horizontal scaling, and processing limitations prevent them from extracting metadata with sufficient fidelity to yield effective decision making. Moreover, emerging 400Gbps technology will soon be raising the bar again. New cutting-edge technology must be applied to high-speed network links to enable discovery, identification and tagging of every network flow in real-time, and enrichment of each flow with critical metadata.

NetQuest's OMX3200 Optical Monitoring Exchange (OMX) leverages state-of-the-art field-programmable gate array (FPGA) technology in a monitoring architecture that performs line-rate packet processing free of multi-core CPU and PCIe bus capacity constraints. This technology converts high volumes of packet traffic into an authoritative data source that is easily discoverable to support analysis and critical decision making.

The OMX supports industry standard IPFIX flow metadata generation, and its innovative use of FPGA technology maximizes high-capacity traffic processing and metadata generation per RU of rack space. The high-capacity unsampled flow metadata it generates is invaluable for the security monitoring of large-scale networks. The compact modular OMX3200 1RU chassis supports up to four parallel packet processing engines capable of processing up to 16 x 100Gbps links with 1:1 unsampled IPFIX flow metadata generation.

The OMX tags data at the source to create high-fidelity structured data that is efficient to transport, store and feed to AI/ML tools for analysis in support of defensive and offensive operations. The association of metadata immediately at the network source dramatically improves efficiency and scalability across the entire security monitoring infrastructure. The solution also integrates NetQuest's OTN/SDH signal discovery and targeting for integrated WAN monitoring and cyber intelligence applications.

The OMX solution can also combine application-specific information from network packets with stateful session-level awareness to provide a richer context of network activity. This includes configurable levels of metadata fidelity and payload signature detection based on application type or criteria describing the network endpoints involved. Further, despite the efficiencies of leveraging enriched flow metadata for analysis, the OMX also supports deeper packet-level analysis by supporting the targeting of specific packet flows for forwarding to packet-based security tools. When targeting packet streams, the OMX provides advanced packet preprocessing to offload security tools and increase their capacity.

The OMX cutting-edge technology enables effective 100/200/400G network traffic visibility in a scalable architecture that provides a strategic and tactical network information advantage.

BIO: Scott Rey has more than 25 years of IC and DOD networking and cyber experience. NetQuest has been a longstanding and trusted supplier of cyber appliances to government agencies.

ACCESSIBLE

Super-Bots Are Here To Save the Data

Keith Nelson, Global Head, Public Sector, Automation Anywhere •

keith.nelson@automationanywhere.com

ABSTRACT

In the Defense Department and throughout government, the quantity of data is not a concern. Data lakes abound with a variety of structured, unstructured and semistructured data in varying formats, styles, classification levels and recency. By providing cutting-edge robotic process automation (RPA) enabled with artificial intelligence, aka "bots," Automation Anywhere is committed to assist the Army with its data interoperability issues.

Founded in 2003 and headquartered in San Jose, California, Automation Anywhere is the global automation leader with 1.7 million bots currently deployed throughout 3,500 global customers of all sizes. Automation Anywhere can liberate soldiers and civilians from the mundane, repetitive tasks they must do. It allows employees to use their intellect and creativity to solve higher order business challenges.

The company envisions a world where Army staff are working side by side with RPA bots to successfully identify, organize and extract data in a way that helps promote national defense. Automation allows higher productivity, a reduction in cost, a reduction in error rate and ultimately increased scale for companies. Organizations can automate the things that are rote and routine and allow people to focus on the things that require creativity and intellect.

Using a combination of traditional RPA and artificial intelligence elements, like unstructured data processing and natural language understanding, AAI's machine learning-powered systems can crunch through tasks that normally take hundreds of thousands of hours to perform.

The Office of Management and Budget (OMB) has initiated a goal for government to move from "low value to high value work" using RPA and similar technologies, which will free up thousands of hours that were previously dedicated to manual tasks like data entry, formatting spreadsheets, cutting-and-pasting data from legacy IT systems and similar administrative work.

The Army is already seeing the benefits of Automation Anywhere' s RPA software, as referenced in a recent article in *Army AL&T* magazine. By deploying Automation Anywhere, 7,000 Army acquisitions officers may each regain up to 13 days per year that can be more strategically dedicated to negotiations, market research and contracts analysis.

Based on a recent survey of more than 10,000 office workers spanning nearly a dozen countries, average public sector workers spend 3.51 hours per day on manual, repetitive computer tasks that aren't part of their primary job and are ripe for human error.

The research showed that nearly half of workers surveyed find digital administration boring (47 percent) and a poor use of their skills (48 percent), while the majority say it gets in the way of doing their main job (51 percent) and reduces their overall productivity (64 percent). Over half (52 percent) of millennial respondents felt that they could be more productive if they had less administrative tasks to complete.

At the very top of the hated task heap is general data entry—manually inputting data into a computer or other devices—followed closely by managing email traffic and filing digital documents—such as documents, spreadsheets, images or PDFs—into the correct digital folder.

BIO: Keith Nelson brings 18 years of public service in the federal government to his role of technology evangelist at Automation Anywhere. His service as chief information officer, chief financial officer and chief human capital officer at three Cabinet agencies has given him insights into how cutting-edge commercial technology can be applied to solve many of government's long-standing challenges.

Achieving a Data-First Ecosystem

Jim Evans, Federal Account Executive, Varonis Public Sector • jevans@varonis.com

ABSTRACT

The Varonis Data Security Platform (DSP) prioritizes understanding all metadata associated with all unstructured and semistructured data. This metadata collection enables a near-instantaneous understanding of risk factors, such as: Which users across an enterprise have the ability to access specific sets of data? How do those individual users have that access? What is their history of using or accessing that data? Is this data secure from both internal and external threats? The Varonis Data Security Platform also provides the capability to automatically identify and tag sensitive files based on common archetypes such as classification levels, personally identifiable information (PII) and personal health information (PHI). Context developed from the metadata about the data brings greater understanding to personnel who need to make actionable decisions on how to best protect and secure data. These metadata streams can be used to ensure personnel are able to access their data while also ensuring the chain of custody and preserving data security.

The first step the Army must take to achieve the desired outcome and alignment to the Army Data Plan is to understand all data (what and where the data is; who is accessing and what are they doing), identify sensitivity and label the data, achieve and maintain a least privilege security posture, and monitor and alert for all suspect user behavior. The Varonis DSP provides the most comprehensive and quickest method to achieve VAUTIS.

As data is understood and protected, the Varonis DSP enables assignment of data owners or knowledge managers to manage their data enterprisewide. Users who require access to data sets "owned" by these data owners quickly and easily request access through the user interface. Upon request, data owners are immediately notified of a user requesting access to their data and are provided all relevant information regarding the user, including who they are, what data they are requesting access to, why they are requesting access, and user and entity behavior analytics (UEBA) and ML-backed recommendations on if they should have access. These data owners, understanding the implications of granting this user access to their data, make the appropriate decision on whether this requester should have access. Once approved, permission changes will take place granting the user access to the data for the time period specified. This greatly speeds up sharing of data and limits confusion. All actions within the software are audited, automated rules can be enforced, and entitlement reviews are fast and simple.

The Varonis DSP provides organizations with an efficient and secure enterprisewide eDiscovery capability. This module considers all metadata, including discovery rules tagged within the DSP, and allows users with the requisite permissions to search for data on-prem or in the cloud. Varonis will automatically enumerate search results based on the users to ensure they only see data to which

they have access. This allows users to get the most up-to-date information regarding active mission requirements, while also preserving the need to know by not allowing users to search for data they shouldn't be able to access.

BIO: Jim Evans leads the Army team at Varonis Public Sector and has supported the Army for 33 years.

Data-Centric Operations at the Tactical Edge: Moving Data Between Two Different Security Domains at the Speed of the 21st Century Mission

Mario Soto, Solutions Architect, General Dynamics Mission Systems • mario.soto@gd-ms.com

ABSTRACT

A critical component of data-centric operations is the movement of data between classification levels at the location where the data is needed: in the tactical environment. The ability to embed a trusted, rugged, secure and low-SWAP device in operational environments is critical to timely access to mission and situational awareness data. One approach to doing this could be General Dynamics' TACDS tactical cross-domain solution, which enables information sharing across different security domains in tactical vehicles, aircraft and dismounted soldier systems. TACDS enables automated and secure transfer of data down to the decision makers, regardless of the classification of the source network. This way, warfighters can achieve full correlation of data tied to desired mission outcomes at the speed the mission needs.

TACDS provides a low-cost, small size, weight, and power (SWaP), rugged, tamper-resistant cross-domain solution ideal for inclusion in almost any vehicle, mobile shelter, ground sensor system, aircraft or UAV. TACDS works by executing programmable rule sets that filter information such as messages, allowing individual messages or data fields within them to be selectively passed, blocked or changed. This method ensures data security on both networks and automates away the need for time-consuming man-in-the-middle screening of message exchanges. TACDS uses an application programming interface (API) to ensure the seamless operation and integration of user credentials, applications, data and metadata.

TACDS is an advanced cross-domain solution that was recently assessed compliant by a National Cross Domain Strategy and Management Office (NCDSMO) certified lab for NCDSMO's Phase 1 Raise the Bar (RTB) requirements. The National Security Agency's RTB initiative establishes security guidelines and requirements for cross-domain solutions deployed by the U.S. government to protect National Security Systems data.

This presentation would review the features and benefits of TACDS, including how deploying it at the tactical edge will help solve the Army's need to enable authorized users to discover authoritative

and non-authoritative data in shared spaces. It will also show how TACDS holistically enables the Army's mission of linking user access to authorized data sets with the user's credentials. Finally, it will demonstrate how TACDS enables the automated and secure sharing of data between two different security domains at the speed required by the missions of the 21st century.

BIO: Mario Soto has more than 30 years of experience in the DOD community, first as an Army sergeant for over six years and for the past 24 years as an engineer, manager and architect at General Dynamics. His skills include C4ISR, satellite operations, Solarwinds, Cisco Systems products and information assurance, and he holds two master's degrees from Boston University in project management and computer information systems. Soto serves currently as a solutions architect for General Dynamics' TACDS tactical cross-domain solution, as well as the executive officer for his local American Legion Post.

Governing Critical Data

Michael Anderson, Chief Federal Strategist, Informatica • mianderson@informatica.com

ABSTRACT

All Informatica technologies are built to be modular, integrated and highly interoperable.

Credentialed user access to authoritative and non-authoritative data: Achievable via a combination of components from Informatica's Intelligent Data Platform.

Dynamic Data Masking (DDM): DDM is a data security product operating between an application and a database to prevent unauthorized access to sensitive information. DDM intercepts requests sent to the database and applies data masking rules to the request to mask the data before it is sent back to the application. DDM will mask or prevent access to sensitive data stored in production and non-production databases. Administrators create rules to specify the database requests to intercept and the masking actions to apply. DDM monitors incoming database requests from the application. DDM applies the data masking rules to the database request before it sends it to the database. The database processes the modified request as normal and returns masked results to DDM. DDM then sends the results to the application.

DDM can also mask data for specific types of database requests or to restrict access to data from certain organizational groups. For example, a created rule to apply a masking function to social security numbers (SSN) when the database request comes from a support team member. When the database sends the data back to the application, the support team member sees the masked numbers instead of the real SSN.

Enterprise Data Catalog (EDC): In addition to data discovery, with metadata management and cataloging capabilities within Informatica EDC, users can have different roles. As a catalog administrator, a person can specify access permissions on resources for specific users and user groups. The type of access permissions depends on the specific security and privilege requirements in the enterprise. For example, in finance offices, apart from the data steward who validates the integrity, consistency and quality of the data, no one in the institution must be able to view the details of the data sources that store confidential soldier financial details. Identification of data sources that store soldier details by unauthorized personnel might lead to hacking of the data sources and leaking of confidential information.

Automation and application of data policies, rules, and guidance across strategic/operational/tactical levels:

Data Governance (DG): Informatica's Enterprise DG solution combines Informatica Axon DG, EDC and Data Quality. The combination enables data discovery, understanding what needs to be governed, and

ensures application of appropriate business policies. As a result, data is governed appropriately and is monitored for quality, ensuring trusted data and support for compliance requirements. The solution scans and catalogs data assets throughout the organization, both on-premises and in the cloud. It enriches what organizations know about their data while automating data quality checking and other routine data management tasks. Finally, it delivers intelligent recommendations and suggestions for more efficient data mission value.

BIO: Col. Michael R. Anderson, USA (Ret.), serves as the chief federal strategist at Informatica LLC. In this role, he leads strategic activities for Informatica with a focus on enterprise data management and data security software capabilities supporting the U.S. Department of Defense and U.S. federal agencies.

Panic-Proof Identity Authentication

Hitoshi Kokumai, President, Mnemonic Security Inc. • kokumai@mneme.co.jp

ABSTRACT

Most desirable would be an identity authentication measure that is practicable when we have lost cards/tokens, injured and panicked in a chaotic situation. For it, organizations need to rely on the deployment of credentials that can stand stress and panic. Making use of autobiographic memories, especially episodic image memory, would make it feasible; images of toys, dolls, dogs and cats, for example, that people and their children used to love for years would jump into a person's eye even when placed in heavy pressure and caught in severe panic.

What is practicable in an extreme environment can be practiced in an everyday environment, though the reverse is not true. Such an authentication system that copes with the panicked people can be operated for all the everyday applications, too, as a stand-alone authenticator, as a factor of multi-factor schemes and as the master password of ID federation schemes.

This is not a hypothesis. This solution has a seven-year history of the trouble-free military use in the field in one of the OECD countries.

BIO: Advocate of "Identity Assurance by Our Own Volition and Memory." Member of Kantara Initiative, which is itself a member of AFCEA. Promoting the concept of expanded password system that accepts images as well as texts, which is intended to be a legitimate successor to the time-honored seals, autographs and text-only password systems.

UNDERSTANDABLE

Data Driven Insights: Getting the Most out of Your Data

Michael Anderson, Chief Federal Strategist, Informatica • mianderson@informatica.com

ABSTRACT

The U.S. Army requires quality data delivered quickly to users and leaders in order to drive insights across multiple data sources to ensure mission success. Whether structured or unstructured data, data at rest or data in motion, access is required across the enterprise. Manual and hand-coding approaches simply cannot scale to both manage the Army's hybrid architectures and support the many mission needs. A shared source of governed and trusted data that can serve the needs of an entire organization is necessary.

The Informatica Intelligent Data Platform (IDP) fills that demand, delivering data at scale for all users and use cases. The industry's most complete solution is also the most modular, built on a microservice architecture to ensure automated delivery of data for self-service access by people, applications and machines. Data migration and integration, data quality, data governance, data catalog, data protection, master data management, data engineering and cloud data services are all available through Informatica.

Any data: The IDP collects data from even the most fragmented sources across complex hybrid enterprises. It connects:

- Any data type—Structured, semi-structured and unstructured data
- Any integration pattern—Batch, real-time and streaming or API
- Any metadata—Application-, pattern- and product-aware
- Any source—Databases, data warehouses, applications, big data systems, IoT, social media and so on
- Any location—On-premises, cloud, hybrid and big data

And it transforms that data into trusted, secure, governed, accessible, timely and actionable intelligence, enabling the intelligent digital transformation of organizations with the thorniest data challenges.

Most importantly, the Army can use Informatica's IDP to grow and evolve at its own speed. Because it is modular, it can start with a single Informatica product or solution and add data management capabilities as needs require.

Artificial intelligence built-in: The CLAIRE engine is the industry's first metadata-driven artificial intelligence to power data-driven disruption. It leverages the IDP's enterprisewide metadata management and adds the intelligence to make intelligent recommendations, to automate the development and monitoring of data management processes, and to adapt to changes from within and outside the enterprise. CLAIRE drives the intelligence of all the data management capabilities in the IDP.

The CLAIRE engine will boost productivity across the platform by:

- Accelerating data delivery: CLAIRE automatically recognizes data types and data entities. It can also group data, relate data and intelligently tag it for faster understanding.
- Accelerating business self-service: CLAIRE can provide intelligent recommendations to help business analysts and data scientists. It can suggest other relevant data sets to consider in their situation. It can also automatically relate business terms to technical data, making data more immediately understandable and usable.

CLAIRE provides the metadata-driven intelligence that is used across the entire Intelligent Data Platform.

BIO: Col. Michael R. Anderson, USA (Ret.), serves as the chief federal strategist at Informatica LLC. In this role, he leads strategic activities for Informatica with a focus on enterprise data management and data security software capabilities supporting the U.S. Department of Defense and U.S. federal agencies.

Democratize Your Data

Chris Hauter, Federal Account Executive, Alteryx Inc. • chauter@alteryx.com

ABSTRACT

Does this scenario sound familiar? A generalized question from an operational group is put forward, but before any analysis is done, a database expert from the business intelligence team must follow up with the operational group to better understand the context and narrow the scope of the question. The person then must find relevant data assets, which could be extremely time consuming, and communicate with the operational group the description and context of data found. Add to the process follow-up questions and access requests for relevant data and this is a very inefficient analytics process.

Typically, most data analysts spend more time searching for data assets than they do on the analysis itself. Too much time spent on the front end translates to inefficiencies and delays in answering critical business questions and lost time in implementing the actions that can improve operations and outcomes. Many times, others in the organization may have already collected the same data or performed a similar analysis, but others simply don't know it—and have no way of finding it—so as a result there is a duplication of effort. Alteryx works to change this dynamic and elevate how data teams discover, prioritize and analyze all relevant and trusted information with their organization. Alteryx provides a data exploration platform for the enterprise that empowers data resources to find, manage, understand and collaborate on the data that resides within the organization.

With Alteryx, organizations can combine data cataloging with human experience to make better decisions. Alteryx infuses the power of metadata with human insight to document the types of information the data contains, where the information comes from, who is using it and how it is used. Through familiar social interactions, organizations can share and utilize organizational tribal knowledge. Alteryx makes discovering everything in the analytic process—data, analytic apps, workflows, macros, visualizations and dashboards—easier. By being able to seamlessly share and identify trusted information assets, along with insight into how they are being used and their lineage, organizations can make more impactful operational decisions.

Alteryx enables them to create a collaborative data governance platform that harvests metadata from a variety of sources to create a centralized data catalog. Data users can find relevant assets via search functionality and data lineage, and then better understand them with features such as asset descriptions, field linkage, glossary terms and ownership information. Users can also interact by adding comments, requesting access, and certifying and sharing assets.

Best practices for such an approach include creating a centralized data catalog, where users can find relevant assets via search functionality, explore data lineage and better understand data assets with available descriptions, field linkage, glossary terms and ownership information. Additionally, us-

ers need to have an ability to interact and collaborate by adding comments, requesting access and certifying, rating and sharing assets. This enables a significant growth in tribal knowledge. And finally, the critical nature of the metadata load itself is vital to the success of any data governance project. The platform needs to have up-to-date information that's valuable for users. It is important to have a clear understanding up front on who would be using the platform and what information would be most valuable for them to discover during their analyses.

By creating an organizational standard that defines how employees find, share and collaborate on data, organizations can reduce the time spent searching for impactful data and assets and avoid those with poor ratings. They can also easily implement and enforce data governance policies across the organization.

BIO: Alteryx is a recognized leader in data science and machine learning through the ability to deliver a self-service analytics end-to-end platform that unifies the analytic experience across the enterprise, enabling organizations to breakdown data barriers. The Alteryx platform provides the flexibility that business analysts, data scientists and IT need to discover, prep, analyze and operationalize analytic models through a collaborative and governed platform enabling every data worker, regardless of technical acumen, to become a problem solver. Alteryx enables data workers to find and understand what verified and trusted information is at their disposal, giving them the ability to analyze data from multiple sources to deliver business insights with agility. Alteryx is revolutionizing business operations through easy to deploy advanced analytics including geospatial, predictive and assisted modeling capabilities in a code free or code friendly environment. Alteryx helps democratize data across the enterprise so that everyone within the chain of command can understand their data resources and create actionable insights faster with the highest degree of confidence. Alteryx has built an engaged online community of users to share knowledge and answer questions enabling all Alteryx users to develop new skills and deepen existing ones.

Autonomous Cyber Threat Sharing as Prototype for Army Data Model

Mark Maglin, Vice President, DOD Cyber Security Services, ECS Federal •
mark.maglin@ecstech.com

ABSTRACT

In multidomain warfare, especially cyber, there are two primary uses for data: instantaneous sharing of indicators to protect systems and soldiers in a real-time tactical environment; and longer term strategic analysis, including trend analysis, historical artifacts and AI. This requires a hybrid data strategy with complementary dissemination methods once the data is generated on how it is tagged, stored, analyzed and transported. Any system must rely on a common data schema, an open data exchange layer and open APIs that enable automation.

ECS Federal delivers the Army Endpoint Security Solution (AESS) as a managed service that is currently deployed to nearly 700,000 endpoints in all theaters on NIPRnet and SIPRnet. AESS is a platform of tightly integrated tools that provides a data architecture with real-time sharing of newly discovered threat indicators across an enterprise fabric for real-time protections and a threat intelligence platform to ingest, correlate, de-duplicate, score and enrich threat data to share in near-real time across the Army Enterprise and with other participating agencies outside of Army. AESS also provides standard Army and DISA feeds, including to big data platforms such as GABRIEL NIMBUS. AESS uses a connected community of analysts and users to enrich data that is immediately available within the community. Th AESS data model can be scaled and expanded to multidomain information operations, including all types of data.

The company's hybrid solution consists of three methods to store and share data. (1) real-time tactical data is sent to subscribers (human or machines) for instantaneous protections; (2) data is stored and indexed regionally for rapid operational analysis and query; and (3) strategic data is shipped to existing Army and DISA feeds and big data platforms such as GABRIEL NIMBUS. This open data approach ensures that datasets from all sources can be used without additional overhead to make it actionable or informational.

Tactical. AESS generates real-time threat data on the endpoint using behavior heuristics, AI and sandboxing of suspicious payloads. After adjudication, that data is shared via an open source data exchange layer (DXL) that uses a publish/subscribe method. Other endpoints and network devices such as advanced firewalls have immediate access machine-to-machine of the new threat data.

Operational. Data is indexed and stored regionally using open source Elastic for quick response for operational and tactical environments including DDIL. Regional Elastic nodes roll up to a master node to allow global search and analytics. Elastic is based on a distributed index search that allows quicker response times and avoids the high cost and overhead of shipping all the data to a central repository. For DDIL, local commanders have access to their data for alerts and dashboards without need to access enterprise servers. AESS includes a threat intelligence platform (TIP) using ThreatQuotient that ingests feeds, correlates with existing threats, de-duplicates data and risk scores according to mission needs. The TIP is a collaborative user environment where threat analysis can be shared in real time across enterprise increasing reaction times and eliminating duplicative efforts.

BIO: 24-year Navy C4ISR, NSA cyber threat data, currently delivering Army Endpoint Security Solution.

Usability Brief: A Solution To Support Data Management, Quality and Sharing for the U.S. Army

Deep Uppal, Vice President, Public Sector Technology Innovation, Information Builders •
deep_uppal@ibi.com

ABSTRACT

The U.S. Army's Data Plan establishes a methodology to provide data that is usable, understandable and wrapped with the intent of data governance. To help guide this policy, an Army Data Board–creating robust data models, data standards, integrating data, providing data architecture overviews and identifying data requirements–must be the priority to improve information traceability.

Information Builders (IB) proposes a solution that will be derived from IB's established reputation to research the individual needs of user groups within the U.S. Army, coalition users and external contractors to orchestrate a solution that speaks to the functionalities needed today while outlaying infrastructure and capabilities for the future. IB looks at consistent product evolution, future iterative design needs and feedback-driven updating as actual system features through the use of automated processes to provide stewardship feedback, data resolution and outlay a foundation for data sharing, application updates and/or new functionalities to be deployed in the source systems. IB's technology and novel approach to technological implementation allows for easy access to legacy systems without any disruption to established infrastructure. IB's solution will be able to build new experiences, provide enhanced capabilities and ensure the Army Data Plan vision for data sharing, management and microservices, to include use of shared vocabularies, common data standards and documented data dictionaries, all while providing the U.S. Army with support to a multi-tenant ecosystem with an integrated, comprehensive solution set.

A phased approach to system design, implementation and consistent updating will allow the implementation team to provide the value-based functionality that answers today's problems while working in the wider state context. In addition to its flexible technology stack, IB's proposed solution will enable the U.S. Army to realize the benefits of a one stop shop for DOD members, while knowing that their environment meets critical requirements. In the larger context, Information Builders supports a deeper integration of cloud-complementary solutions and capabilities across the Army Chief Data Office portfolio. Tapping the combined power of cloud auto scaling and the ability of Information Builders' technology to expand its hardware usage on demand, the Army can improve surge capacity without requiring major investments in hardware or administrator support.

Solution specifics:

- Data Automation: The automation of tagging and profile datasets, including the automated review of metadata profiles through AI filters and heuristic rulesets.
- Compliance: Provides out-of-the-box industry compliance for ISO, SOC 3, PCI, ATO and other standards, as well as auditing and certifications of data centers, policies and procedures. Additionally, the solution leverages AWS support for DISA compliance around esoteric networks such as SIPR, NIPR and coalition.
- Security and Logging: Supports DISA and US-AG-specific security requirements with approved capabilities such as identity and access management roles and security groups, the configuration of anti-malware protection and an intrusion prevention system.
- IB proposes a governance module that is multidomain and supports domain security, role-based authorization, single sign-on (SAML) and LDAP authentication. The platform can also support ADFS and EFSS solution sets.

BIO: Deep Uppal is an innovator, technical problem solver and change agent with a proven track record of defining the technical vision; communicating complex processes; and successfully creating, integrating and deploying next-generation enhanced analytics, network architecture and all associated facets of technology related to state and federal government. Uppal is a member of the Signal Army Corps.

Improving Data Quality and Data Sharing

Jim Evans, Federal Account Executive, Varonis Public Sector • jevans@varonis.com

ABSTRACT

The Varonis Data Security Platform treats data management, governance and security as priorities in its overall paradigm shift in managing and protecting data. Varonis looks at the unstructured and semi-structured data first, not last, and empowers organizations to augment their data's efficacy and ensure its security no matter where the data may sit across the enterprise; on-prem locally, on separate data silos across the globe, or in the cloud.

First and foremost, the Varonis Data Security Platform understands all metadata associated with unstructured and semi-structured data including, but not limited to, often-shared files such as Word documents, PowerPoint presentations, PDFs and audio-visual files. This metadata collection allows users of the software to quickly understand which user accounts across an enterprise can access specific datasets, identify how those individual users are gaining that access, understand what the users are actively doing with that data, as well as ensure that the data is secure from both internal and external threats.

Varonis also has the tools in place to automatically identify and tag sensitive files based on common archetypes such as classification levels, personally identifiable information (PII), personal health information (PHI). This context brings greater understanding to personnel who need to make actionable decisions on how to best protect and secure data. These metadata streams can be used to ensure personnel are able to access their data while also ensuring the chain of custody and preserving data security.

Not only does the Varonis Data Security Platform give users an unprecedented understanding of and risk profile for their unstructured and semi-structured environments, it also provides recommendations for permission changes to help achieve a least permissive model. Further, Varonis can automatically commit these changes in a way that does not impact mission needs and can automatically quarantine unsecured sensitive data as an incident response mechanism. This automation capability plays a key role in the Varonis approach to remediation and data security.

The Varonis Data Security Platform contains a full user and entity behavior analytics (UEBA) engine built using machine learning (ML) algorithms that define baselines for all users on how they interact with the data environment. The Varonis DSP understands what security groups users are in and who are their peers in those security groups and compares individuals' behavior to their previous behavior as well as that of their peers. In doing these comparisons, Varonis will alert and respond automatically

to malicious behavior on an organization's unstructured data, augmenting the enterprise's cybersecurity posture. Varonis will automatically trigger on any of more than 180 threat models rules that, when broken, signify a user is behaving abnormally in a specific behavioral pattern that Varonis can trace back to a malicious activity. Security personnel can investigate these alerts on data access and respond accordingly.

BIO: Jim Evans leads the Army team at Varonis Public Sector and has supported the American warfighter for 33 years.

Understanding Data

Richard Graham, Chief Executive Officer, CodeMettle • richard@codemettle.com

ABSTRACT

CodeMettle's ConOptic software is an open, flexible and nonproprietary data and decision management tool that by design normalizes disparate data, tags with meta-data at runtime and shares across echelons.

CodeMettle offers an authoritative federated data repository for network operations data. It orchestrates and exchanges data between stovepipe systems, tools and the enterprise network into a distributed repository. Thusly, overall data quality is improved, while providing structure and organization to the data. CodeMettle's repository platform does this by compressing the data, and then schedules and distributes it down from the global repository to local echelons and to needed devices. Bi-directional data flows close the feedback loop between directives and actual performance improving future decision making. The software architecture is designed to operate within the constraints of decreased bandwidth and SWaP as well as reduces dependency on third-party vendors. The overall standardization of data enables warfighters at-echelon and across formations with systematically more accurate, intuitive, digestible information in a secure environment. This greatly reduces redundant errors and reduces fundamental dependencies on MS Office, human action and stovepiped legacy systems by centralizing shareable universal data.

BIO: Richard Graham is the cofounder of CodeMettle LLC, a software business servicing the DOD and commercial enterprise clients. CodeMettle's software provides massively scalable network automation and service management for the most complex networks. The distributed and scalable data management platform enables enterprises to analyze, organize, consolidate and visualize complex processes and operations. Some of CodeMettle's notable clients using it in critical networks include: DOD and federal government, U.S. Army, U.S. Marine Corps, U.S. Air Force, State Department, FAA and FEMA. Commercial: AT&T/DirecTV, SiriusXM and Bell Canada.

Before forming CodeMettle, Graham was the founder and CEO of ILC, a developer of network management software primarily for satellite-related networks. Graham built ILC into the largest developer and producer of network management software for satellite systems in the world.

Graham began his career as an electrical engineer from Georgia Tech. Upon graduation, he entered defense R&D for radar systems. He ultimately became an expert at developing radar environmental simulators for the RF and millimeterwave radar systems. Graham was the chief architect for multiple key radar environment simulators employed by the U.S. Navy and Air Force.

SQL Server/Relativity - eDiscovery Software Solutions

Nirupama Hewawasam, President, SamanMali Consulting LLC • nirupama@saman-mali.com

ABSTRACT

Litigations firms globally face a similar set of problems illustrated in Modernizing Understandable (LOE #1.3). A large amount of litigation support documents are often unstructured data. Confidential documents are often accessed exclusively by the attorneys and are redacted before caseworkers of different levels would have access. Security, fast access and retention are critical for operational continuity. The Army could adopt a similar approach to how global litigation firms solve this problem using an e-discovery solution with customization needed specifically for the Army.

Because of familiarity with the Relativity e-discovery application with the backend of a MS SQL server farm to handle terabytes of data, this example is applicable. The application is designed to provide both secure and fast access to the unstructured data stored. Clients can access data via the Relativity application globally from inside the VPN. Clients have the ability to get quick and controlled access to the data they want. Specific functionalities that the Army might find useful include:

- Solution for audit and inventory management
  - » There are trace and audit trails.
- Solution for cybersecurity
  - » The application often resides inside a VPN behind the firewall. Database servers are secure with specific ports needed to communicate between web, application and database server groups. Documents can have different access levels based on the user groups, with high, low or no access to content. It also has cloud integration capacity, although most commonly used in VPN for privacy and security reasons. The solution will have to be customized for the needs of the Army network and security infrastructure.
- Solution for poor data quality unstructured date
  - » OCR technology: Optical character recognition allows recognizing text inside images, scanned documents and image files. Extracted texts are searchable and indexed.
  - » Redacting capability: Content can be redacted and stored in the database. Only high-privileged users can have access to the original content.
  - » Investigating Audio and Video Files: Has the capacity to transcribe and analyze unstructured multimedia files (audio, video) using the third-party integrated tool aiWare that uses artificial intelligence (AI).

» Social Media and Enterprise Chat Data: Has the capacity to review and analyze social media and chat sources using an integrated third-party tool (RTK.Message), which allows users to extract, filter and analyze social media data (Facebook, Twitter, Enterprise chat data, Bloomberg Chat and Microsoft Lync).

» Emails: Has email threading and clustering capacity that helps organize and identify patterns; with third-party integrated tools (NexLP's Story Engine) allows mapping out relationships, emotions in a story.

BIO: SamanMali Consulting LLC is a SQL server database solution provider in database administration, development and data analysis. President Nirupama Hewawasam has more than 15 years of industry experience in SQL server database management with high profile financial, federal and litigation clients.

TRUSTED

Integrity Verification Through Timed Ledger Stamps

Nisha Panwar, Assistant Professor, Augusta University • npanwar@augusta.edu

ABSTRACT

Advances in computation power embedded in smaller chips has enabled large-scale data generation through collective sensing by IoT devices, sensors and wearables. However, the advances in storage access still have a long way to go as compared to the processing power of these tiny IoT devices. Therefore, the sensing workflow terminates into a third-party storage service provider. In addition, the scale at which these sensors report observations is a time-series that allows zooming-in to the data as fine as required by certain application.

Since the continuous stream of sensor observations has the potential to reveal user-activity, preferences and changes over time, it is certainly privacy-threatening to the owner of these devices and sensors. Therefore, a correct implementation of well-known trust-but-verify paradigm in these data driven settings is highly important.

One perspective is to explore this trust-but-verify paradigm as a tunable—configurable as suitable for the application—slider to find the balance between the two extremes, i.e., right-to-own versus right-to-audit, to leverage the trust in computing as well as in the storage on any public platform.

One method is a ledger-based timestamping approach that combines the trustworthiness of a central solution with the scalability of a de-centralized solution. In particular, a blockchain-based timestamping solution can leverage these central authorities to maintain a public ledger of timestamps. The verifiable ledger enables the integrity check on the data as well as the meta-data. Every time a central authority generates a signature on a unique timestamp, it must be published in the subsequent block of the public ledger. Therefore, these signed timestamps can be verified by anyone whenever the corresponding block is published on the main chain.

The blockchain-based timestamping solution offers a public ledger that records the sequence of time-stamped transaction logs in a shared database model. The pool of timestamped transactions require an additional mechanism, i.e., mining, to fairly select the transaction logs (blocks) and add to the public ledger. Each block contains the selected transaction logs with the integrity proofs. The blocks are further chained through hash pointers to guarantee the immutability across all previous blocks.

There is a computation cost that a mining node pays to offer a sequence of timestamped transactions to be appended to the existing ledger. In addition, the peer nodes must agree on the replicated state

of the ledger, i.e., block validation, which avoids the inconsistent set of transactions to appear on replicated ledgers. This is also termed as fork-consistency or the double-spending attack where the same asset is used as input for multiple transactions to appear in the same ledger. In such a scenario, the fork-consistency requirement guarantees that eventually only one of the transactions will be valid and appended to the ledger, and the other transaction will be invalidated. This vision of a verifiable timestamping ledger-based approach has the advantage that records can be preserved and verified as far back on the timeline as required by any application.

BIO: In 2020, Nisha Panwar became an assistant professor at the School of Computer and Cyber Sciences at Augusta University. She did her post doctorate work (2016-2019) at Department of Computer Science, University of California Irvine. She received her PhD at Department of Computer Science, Ben-Gurion University of the Negev, Israel in 2016. During her post doctorate work, she proposed solutions to ensure verifiability in smart spaces. In addition, she has worked on subliminal privacy aspects in smart homes data generation and device interaction. To continue the work on privacy policies and GDPR compliance in modern digital world, she proposed solutions for verifiable deletion of data from an untrusted cloud. During her PhD, she has worked on wireless communication security in smart connected vehicles. She aims to continue working in access control and post-facto reconstruction of the facts through lifelogs in this digitally self-quantified world. Her research interests include security and privacy issues in smart spaces, verifiable computing, Internet of Things (IoT), smart homes and smart vehicles.

The Holy Grail of Encryption: Securing Data in Use

Brandon Sellers, DOD Account Manager, Enveil • brandon@enveil.com

ABSTRACT

Operations and intelligence analysis methods involving external datasets can be very revealing. This exposure not only includes attribution of the search—who is performing it—but also the content of the search—what you are searching for—that may include sensitive indicators and/or classified selectors that would be extremely damaging to national security if exposed.

Enveil's ZeroReveal solution allows data to be securely processed while remaining in the untrusted domain, extending the boundary of trusted compute. With Enveil, Army units can perform secure searches, watchlisting and analytics using sensitive/classified indicators (JWICS/SIPRNet) against PAI or other less sensitive data on untrusted systems without compromising mission objectives. This unmatched capability enables secure and efficient data sharing, collaboration, reporting and alerting across multiple classification levels to significantly reduce operational risk and accelerate the timeline for turning raw data into actionable intelligence. By eliminating the need to exfiltrate data for processing, Enveil also saves valuable time and resources in areas of limited connectivity.

Enveil ZeroReveal is the first and only certified solution for performing operations from classified domains against sources on lower classification domains. Using homomorphic encryption techniques at levels of scale that have not been previously achieved, Enveil ensures the content of the search itself as well as the interests and intent of the person performing the search are never revealed. Enveil scales linearly across one or multiple domains, allowing aggregated data to remain in the untrusted or less trusted environment while sensitive operations such as search, watchlisting and analytics are securely performed.

Enveil's core technology was developed, deployed and operationalized inside of the U.S. National Security Agency and the technology has been implemented at scale in the most sensitive of environments. Carrying NIAP and CSfC certifications, Enveil ZeroReveal is a lightweight, API-based proxy-layer software system ready for deployment via existing third-party integrations for immediate mission impact. By decoupling from the storage technology layer, Enveil is able to sit above the data, requiring no changes to the underlying environment.

The Enveil ZeroReveal solution can integrate with any type of storage technology, data format or computational platform, including commodity hardware, cloud computing and small form factor. This deliberate engineering approach fully supports interoperability, flexibility and adaptability of Army Signal Corps sensor and data assets, providing Army commanders with decisive battlefield insights.

Building off the technology's broad applicability in the commercial market, Enveil enables secure data sharing across security levels and sources while ensuring the contents of operation and its corresponding results remain encrypted. As part of Enveil's Department of Defense (DOD) outreach during Small Business Innovation Research (SBIR) efforts, Enveil confirmed the significant mission demand for secure search, watchlisting and analytics across numerous DOD mission sets, with specific interest in the areas of information, surveillance, and reconnaissance (ISR); open source intelligence (OSINT) and publicly available information (PAI); cyber operations; tactical edge; and insider threat. As a result, Enveil was awarded a Phase II contract under SBIR 19.2, sponsored by the Air Force PEO Presidential and Executive Airlift.

BIO: Brandon Sellers leads DOD customer engagement at Enveil, leveraging his diverse skillset honed during 20 years as an officer in the U.S. Navy. He has served as a combat-experienced fighter pilot, Senate aide, intelligence operations officer, and accredited diplomat and now helps DOD and IC customers achieve Trusted Compute in Untrusted Locations. In addition to three overseas tours in Asia, his career included assignments with the Marine Corps, Army, State Department and the IC. Sellers joined Enveil from Gartner, where he served as a strategic partner to the Intelligence Community.

Data Integrity

Matthew Shabat, U.S. Strategy Manager, Glasswall Solutions • mshabat@glasswallsolutions.com

ABSTRACT

The U.S. intelligence community funded Glasswall to add a security tagging feature to its Glasswall FileTrust software. Tags can be placed within files to track and audit file movement and user handling. Glasswall contributed to a larger data solution. Glasswall can also add additional security measures against insider threats by sanitizing documents in transit while adding and/or gathering tag information. By processing the files before and after each endpoint, Glasswall can be used to ensure that tag information is inserted and/or gathered and the files are remediated and sanitized, thus protecting endpoints throughout the process. It potentially can identify if the document was accessed while on a previous endpoint depending on the changes in the tags and what was modified while on that endpoint. Having Glasswall processing all files in the cloud prevents endpoints from being compromised even if the previous endpoint had already compromised the file. While this tagging feature was not yet available, Glasswall FileTrust was included as a reference architecture solution in the UK National Cybersecurity Center of Excellence's Data Integrity project.

BIO: Matt Shabat is the U.S. strategy manager for Glasswall Solutions. He served for nearly 10 years in the U.S. Department of Homeland Security's Office of Cybersecurity and Communications, most recently as a cybersecurity strategist and as the director of performance management, and previously as the National Cyber Security Division's chief of staff. While at DHS, Shabat led DHS and interagency implementation of the Cybersecurity Information Sharing Act, collaborated with members of the public and private sectors to increase adoption of security automation and orchestration, supported maturation of the cyber insurance marketplace, developed an operationally relevant approach to measuring the costs of cybersecurity, contributed performance goals to the NIST Cybersecurity Framework, led strategic planning and developed program performance metrics.

Before DHS, Shabat practiced securities, mergers and acquisitions, and general corporate law at Mayer Brown LLP in Chicago. He is a Harvard Kennedy School Senior Executive Fellow, earned a JD from the University of Pennsylvania Law School, an MA in security policy studies from the George Washington University's Elliott School of International Affairs, and a BA from Stanford University. He also is ISACA Certified in Risk and Information Systems Control, and he is an ISACA Certified Information Security Manager.

Ensuring the Data Validity of Your Tool Portfolio

Carlos Cosme, U.S. Army Cyber Command ISSM, ARCYBER • carlos.f.cosme2.civ@mail.mil

ABSTRACT

To fortify the Army's security posture and prevent data loss on the Army networks, it is essential to have a clear picture of what the Army networks look like. The Army does not have a clear picture of what its network comprises or how it is configured. It is impossible to reduce the attack surface to prevent adversary access until the Army can confidently identify with certainty all vectors of approach to the networks. This can be achieved by utilizing multiple data sets from multiple network tools and cross referencing those data sets to eradicate any false positives/negatives to provide clarity on what the networks comprise. By using multiple data sets from multiple tools to cross reference, the picture of the Army networks is validated. Any assets that one data set discovers that another data set does not can go through the incident response process and personnel can discover misconfigurations, issues with life cycling, unauthorized assets on the network and more. A clear picture of the Army network provides resounding impacts for cybersecurity. With a clear idea of the assets on the Army networks, assets that have previously gone unseen can be appropriately removed or patched to prevent critical vulnerabilities to the network and leading to data loss.

Verifying the picture of the Army network can be done by simple SIEM rules that automate the process of pulling data sets from multiple network tools and flagging inconsistencies in those data sets. Those assets that are not verified by all data sets can then go through an active discovery process and can be reconfigured, life-cycled or the tool that cannot identify that asset can be fixed to see or remove those assets on Army network mapping tools. Artificial intelligence/ML tools can take the data from the script scans and begin the active discovery/incident response process in order to provide real-time data loss prevention.

Launder tool sets data against known assets and known IP spaces to identify tool sets administration effectiveness and false positive/negative enumeration. This provides a multi-tool validation that results in highly accurate business intelligence.

BIO: Carlos F. Cosme is a native of Bayamon, Puerto Rico. He is a 1999 graduate of Randolph Macon Academy with a GED. He holds a 2000 25B/74B military computer science certificate school of information technology, Fort Gordon, Georgia. Cosme's operational tours include information assurance manager, Department of Army (Pacific); security operations manager, Qatar (OIF/OEF); incident response handler, National Reconnaissance Office; service desk manager, FBI Special Technologies and Services; network operations manager, Soto Cano

Airbase Honduras; cyber SME supporting ARCYBER Technical Warfare Center; security operations program manager, Department of Justice; DMZ program manager, DISA; principal security operations consultant, Hewlett Packard (ArcSight); Command ISSM, Cyber Security Division ARCYBER G-6.

Trusting Your Data: The Key to the Data-First Ecosystem

Jim Evans, Federal Account Executive, Varonis Public Sector • jevans@varonis.com

ABSTRACT

Trust in the Army's data begins and ends with the absolute ability to visualize, log and understand every file or SharePoint folder touch, every email event, and every AD authentication and change, and then mapping all of this against user activity. Monitoring and alerting must first be based on where the data lives whether on-prem or in the cloud—not an endpoint or other perimeter device, which will always leave the Army chasing the threat and questioning trust of its data.

The Varonis Data Security Platform provides a full user and entity behavior analytics (UEBA) capability built upon semi-supervised machine learning (ML) algorithms that define baselines for all users (human or machine) on how they interact with the data environment. The Varonis DSP understands what security groups users are in, who are their peers are, what types of data individuals typically access (classified documents, PII, PHI), and what they typically do with that data and compares every individuals' behavior to that of themselves, as well as their peers. In doing these comparisons, Varonis automatically triggers on any of more than 180 threat models for insider and external threats, ransomware or malware, and new or existing APTs that when broken individually or in conjunction, signify a user behaving abnormally in a specific behavioral pattern Varonis can trace back to malicious activity, which immediately enhances the Army's security posture. High-fidelity, highly contextual alerts enable security personnel to quickly and easily conduct forensics and determine appropriate playbook response.

The foundation for the Varonis DSP is understanding and continual monitoring of all metadata associated with unstructured and semi-structured data. This metadata collection allows the Army to quickly understand where and what data is, which user accounts can access specific datasets, identify how those individual users are gaining that access, understand what the users are actively doing with that data, and ensure that the data is secure from both internal and external threats.

Varonis provides the capability via a robust, easily customizable data classification engine to automatically identify and tag sensitive files based on common archetypes (classification levels, personally identifiable information (PII), personal health information (PHI)). This context brings greater understanding to data owners who need to make actionable decisions on how to best protect and secure data. The various metadata streams can be used to ensure personnel are able to access their data while also ensuring the chain of custody and preserving data security.

Varonis provides the Army an unprecedented understanding of, and risk profile for, their unstructured and semi-structured environments. Varonis provides recommendations and modeling for permission changes to achieve and maintain a least permissive posture. Varonis can automatically commit permission changes in a way that does not impact mission needs, and can automatically quarantine unsecured sensitive data as an incident response mechanism. Automated capabilities play a key role in the Varonis approach to remediation and data security.

The Varonis DSP would be a foundational component in the Army's security posture to achieve VAUTIS.

BIO: Jim Evans leads the Army team for Varonis Public Sector and has supported the American warfighter for 33 years.

Data Protection Delivered Through Enterprise Data Management

Michael Anderson, Chief Federal Strategist, Informatica • mianderson@informatica.com

ABSTRACT

The U.S. Army requires quality data delivered quickly to users and leaders to drive insights across multiple data sources to ensure mission success. Whether structured or unstructured data, data at rest or data in motion, access is required across the enterprise. Manual and hand-coding approaches simply cannot scale to both manage the Army's hybrid architectures and support the many mission needs. A shared source of governed and trusted data that can serve the needs of an entire organization is necessary.

The Informatica Intelligent Data Platform (IDP) fills that demand, delivering data at scale for all users and use cases. The industry's most complete solution is also the most modular, built on a microservice architecture to ensure automated delivery of data for self-service access by people, applications and machines. Data migration and integration, data quality, data governance, data catalog, data protection, master data management, data engineering and cloud data services are all available through Informatica.

Any data: The IDP collects data from even the most fragmented sources across complex hybrid enterprises. It connects:

- Any data type—Structured, semi-structured and unstructured data
- Any integration pattern—Batch, real-time and streaming or API
- Any metadata—Application-, pattern- and product-aware
- Any source—Databases, data warehouses, applications, big data systems, IoT, social media and so on
- Any location—On-premises, cloud, hybrid and big data

And it transforms that data into trusted, secure, governed, accessible, timely and actionable intelligence, enabling the intelligent digital transformation of organizations with the thorniest data challenges.

Most importantly, the Army can use Informatica's IDP to grow and evolve at its own speed. Because it is modular, it can start with a single Informatica product or solution and add data management capabilities as needs require.

Artificial Intelligence Built-In: The CLAIRE engine is the industry's first metadata-driven artificial intelligence to power data-driven disruption. It leverages the IDP's enterprisewide metadata management and adds the intelligence to make intelligent recommendations, to automate the development and monitoring of data management processes, and to adapt to changes from within and outside the enterprise. CLAIRE drives the intelligence of all the data management capabilities in the IDP.

The CLAIRE engine will boost productivity across the platform by:

- Accelerating data delivery: CLAIRE automatically recognizes data types and data entities. It can also group data, relate data and intelligently tag it for faster understanding.
- Accelerating business self-service: CLAIRE can provide intelligent recommendations to help business analysts and data scientists. It can suggest other relevant data sets to consider in their situation. It can also automatically relate business terms to technical data, making data more immediately understandable and usable.

CLAIRE provides the metadata-driven intelligence that is used across the entire Intelligent Data Platform.

Big Data Security Solutions: Informatica's intelligent data security solution for big data includes the Informatica Secure@Source and its data masking solutions. Secure@Source automates the process of discovering, analyzing and visualizing sensitive data so security, compliance and privacy teams can rapidly understand sensitive data risk and apply appropriate controls and policies to the data. It replaces costly, time-consuming manual data audits with precise processes that define, discover and analyze sensitive data for risk. Informatica's data masking solutions let organizations de-identify and desensitize data for reporting, analytics and mission-critical applications to counter privacy and regulatory concerns. Beyond encryption, data masking secures data in use by the organization for customer and operational applications.

BIO: Col. Michael R. Anderson, USA (Ret.), serves as the chief federal strategist at Informatica LLC. In this role, he leads strategic activities for Informatica with a focus on enterprise data management and data security software capabilities supporting the U.S. Department of Defense and U.S. federal agencies.

Data Protection and Integrity

Rick Bueno, CEO and Founder, Cyber Reliant Corp. • rabueno@cyberreliant.com

ABSTRACT

Cyber Reliant Data Protection is able to fortify the security posture for Army data by protecting the data at all levels directly regardless of the number of vulnerable points through which an adversary could gain access and exploit the Army's data. Furthermore, Cyber Reliant Data protection will protect the Army's data regardless of the attacker successfully breaching the network via one of the vulnerable points. Cyber Reliant ensures the protection of the Army's data regardless of a network breach. It is a completely new paradigm in data protection.

Built and designed to protect the Army's most sensitive data in the most austere cyber environment, Cyber Reliant applies data protection directly to the data itself. Traditional security relies on protecting the perimeter and on data encryption. Perimeter-based data protection strategies are important but are not enough as the perimeter can be, has been and will continue to be breached. Encryption is also an important data protection strategy, but it is not enough as encryption can also be broken with relative ease.

The key to successful data protection is to apply the data protection to the data itself in a manner so difficult and complex that not even the most sophisticated attackers would know how to break into it. Cyber Reliant products were designed and built by offensive information operations engineers who designed a data protection product to counter the most sophisticated state-level sponsored attacks.

Cyber Reliant incorporates not just one technique but a series of innovative and specialized techniques to counter any offensive information operations attempt to data exploitation and exfiltration. Cyber Reliant has implemented advanced key management, disassociation, encryption, shredding, embedding and dispersion techniques, which create a framework of data protection that has been adjudicated as information theoretic secure by the NSA. Information theoretic secure implies a quantum resistance data protection methodology.

BIO: Rick Bueno, president and CEO of Cyber Reliant, is a U.S. veteran, entrepreneur and experienced visionary with a deep 35-year history of developing and executing cybersecurity strategic initiatives and solutions in the commercial, defense, intelligence and special operations community. Prior to founding Cyber Reliant in 2010, Bueno served as the National Security Agency Information Assurance Directorate AD Afghan mission manager/NATO Special Operations Forces for the ISR Task Force. In this role, he worked closely with NSA, USDI, CENTCOM and others to develop secure communications strategies in support of NATO and Special Forces missions. Prior to his role at ISR TF, Bueno served as the chief strategic architect for the Director of National Intelligence (DNI). As a member of the director's action group, he was responsible for early establishment of the Intelligence Community's policies and processes for the information-sharing environment and other activities within the DNI CIO scope.

Protecting Critical Data

Matthew Jones, Cybersecurity Sales Specialist, Cisco Systems Inc. • matjone2@cisco.com

ABSTRACT

Cisco's integrated security solutions work together to deliver effective network security, incident response and heightened IT productivity through automation. The company understands application behaviors, automates microsegmentation and uses advanced security analytics to speed detection. The solutions outsmart emerging threats with machine learning and behavioral modeling. They know who is on the network and what they are doing using telemetry from the network infrastructure. They detect advanced threats and respond to them quickly. The company protects critical data with smarter network segmentation.

Cisco can identify, segment and monitor a data source and see who accessed it. It can automatically tag data sources, such as servers, and these tags can be used to dynamically identify data from those sources, enforce policies, monitor flows and verify intended policies are being met. Cisco's integrated security solutions work together to make dynamic, data-driven decisions that enforce least privilege. These decisions reduce the threat surface area, which in turn reduces discourses. Additionally, continuous monitoring, using AI/ML, allows users to detect deviations and respond faster, reducing damage done in the event of a breach. In the event of data loss, the solutions keep a record of the activity so an organization can scroll back to further examine what occurred and scope what was lost.

BIO: Matthew Jones is a cybersecurity sales specialist at Cisco Systems Inc. with 20 years of experience working with the DOD.

INTEROPERABLE

The Network as a Weapon System

Richard Graham, Chief Executive Officer, CodeMettle • richard@codemettle.com

ABSTRACT

CodeMettle's capabilities help transform the network from a sustainment function in support of maneuver to a critical, proactive enabler to every part of a commander's scheme of maneuver. The shared understanding of the network situation as part of the overall tactical situation and greater mission set is the deliverable hallmark of CodeMettle's software in an easy-to-use, holistic, actionable common operating picture (COP). CodeMettle's software ConOptic is an open, flexible and nonproprietary data and decision management tool that by design normalizes disparate data, tags with meta-data at runtime and shares across different formation types at-echelons.

CodeMettle's products integrate the DOD's network, services and operational data with mission context all in a browser-based, dashboard-driven software platform. This approach bridges the gap between network data and operational data. The company's off-the-shelf software solutions are actively utilized in the DOD to bridge S6/G6 NetOps functions to the greater S3/G3 Operations Process for the greater purpose over units' mission sets. CodeMettle's data management software supports the DOD data strategy in providing context to reduce process complexity and the cognitive load required for multi-faceted decision-making during missions and operations. Its central data management, orchestration and network management, and unifying NetOps capabilities increase warfighters' proactivity from the tactical edge to the strategic enterprise.

CodeMettle will be showcasing its latest developments in unifying strategic and tactical networks for the 2nd and 160th Signal Brigades in support of European Defender '20 and real-world operations in Southwest Asia, respectively. The company will also demonstrate planning products, an enterprise SAR/GAR automation system being developed for the Air Force, and the latest updates to the ITN TRIK box management product.

CodeMettle's data normalization and orchestration solutions reduce network complexity for end users at-echelon. It formalizes innate data integration and platform infrastructure to stabilize and manage networks, effectively raising the standard of overall network literacy across warfighting functions. This reduces cognitive load and training burden while simultaneously empowering the warfighter at-echelon. The decreased dependency on FSRs and reduced burden on high-density MOSs and network technicians increases freedom of maneuver in MDMP and the Army operations process. CodeMettle's API-based interface, provisioning workflows and automated processes deliver a comprehensive, intuitive visualization through actionable common operating pictures (COPs) and dashboards. This enables planners and decision makers to proactively interpret and manage networks with direct impacts to a unit's mission set.

These products streamline MDMP and the Army operations process thereby increasing a unit's overall operational tempo. CodeMettle accomplishes this through automated collation and curation of federated multi-echelon data to planners and decision makers to visualize missions, schedules, decision support documents, priorities and risk. CodeMettle's automated end-to-end mission management system simplifies planning, executing and monitoring of this data. CodeMettle's data management software supports the DOD data strategy in providing context to reduce process complexity and the cognitive load required for multifaceted decision making during missions and operations.

BIO: Richard Graham is the cofounder of CodeMettle LLC, a software business servicing the DOD and commercial enterprise clients. CodeMettle's software provides massively scalable network automation and service management for the most complex networks. The distributed and scalable data management platform enables enterprises to analyze, organize, consolidate and visualize complex processes and operations. Some of CodeMettle's notable clients using it in critical networks include: DOD and federal government, U.S. Army, U.S. Marine Corps, U.S. Air Force, State Department, FAA and FEMA. Commercial: AT&T/DirecTV, SiriusXM and Bell Canada.

Before forming CodeMettle, Graham was the founder and CEO of ILC, a developer of network management software primarily for satellite-related networks. Graham built ILC into the largest developer and producer of network management software for satellite systems in the world.

Graham began his career as an electrical engineer from Georgia Tech. Upon graduation, he entered defense R&D for radar systems. He ultimately became an expert at developing radar environmental simulators for the RF and millimeterwave radar systems. Graham wasthe chief architect for multiple key radar environment simulators employed by the U.S. Navy and Air Force.

Helix: Secure Collaboration With Industry and Academia

Claire Cuccio, President and CEO, SNVC, LC • claire.cuccio@snvc.com

ABSTRACT

The DOD struggles with a platform to collaborate freely with industry and academia. The requirement for CAC access credentials often eliminates the ability of a platform such as the DOD Enterprise Portal or milSuite to work as a true collaboration space. Organizations in academia and industry who collaborate regularly with the DOD usually have one person with CAC access who downloads the information from a government collaboration site, emails the files around to their reachback personnel who edit it and email it back. Then the contractor with CAC access reposts the document with suggested edits. The bottom line is email is used as the collaboration platform, not the platform itself, and the data is exposed many times while in transit. Additionally, many DOD organizations block collaboration websites such as Google Docs. The DOD requires a secure collaboration platform accessible by DOD, industry and academia.

At SNVC, the company developed its own internal collaboration environment called helix, which allows personnel to collaborate with industry or academic partners on projects. Helix is a user environment for Microsoft Sharepoint that makes Sharepoint easy to use for managers. Sharepoint can be challenging to use, and it requires an administrator with specific skills. DOD will not need IT personnel or industry FSOs to manage the environment. In helix, a manager can easily create a site, upload and download documents, and control site, functional and document permissions, which allows teams to collaborate. All team members can task and track progress without using email. Helix maintains an audit trail so users can track who did what and when. SNVC uses helix internally to develop proposals with partners, onboard employees, as an employee portal and to manage a nonprofit fundraiser. For example, if SNVC is bidding on an opportunity, and it has an industry partner, the proposal manager can grant an outside partner access to the folder or just access to specific files. The company segregates that server so the rest of SNVC's data is not accessible to the outside person.

Benefits to DOD include uses in the DODIN and active directory for secure DOD access. Outsiders can VPN in using a different access control method such as username/password plus text access code to a cellphone.

The helix environment for internal DOD users could be isolated by a firewall with a read-only active directory instance in the cloud. External users would have to request access by registering for an account with Sharepoint via helix. The request is approved by an authorized agent in the DOD and supervised by a DOD internal manager when an external user is approved into the environment. The external user could see nothing until permission to projects and documents is granted and then can

only view authorized projects or documents. Data is not orphaned in the cloud when a project ends or a person leaves a project. The DOD keeps the data and can remove a departing person's access.

BIO: Claire Cuccio, PhD, is a retired U.S. Army Signal Corps colonel. SNVC has been providing network operations, project management, cybersecurity and policy support to the DOD/Army for over 20 years.

Operationalizing the Army's Data

Jim Evans, Federal Account Executive, Varonis Public Sector • jevans@varonis.com

ABSTRACT

The Varonis Data Security Platform first and foremost understands all metadata associated with all unstructured and semi-structured data. This metadata collection enables a near-instantaneous understanding of risk factors, such as:

- Which users across an enterprise can access specific sets of data?
- How do those individual users have that access?
- What is their history of using or accessing that data?
- Is this data secure from both internal and external threats?
- What is the sensitivity of the data?

This context brings greater understanding to personnel who need to make actionable decisions on how to best protect and secure data. The Varonis Data Security Platform contains related solutions that leverage this metadata that it already collects, in specific ways, to aid in the sharing and data access across an enterprise.

First, DataPrivilege allows enterprises to assign data to manage their unstructured data enterprise-wide. Enterprise users who require access to data sets owned by these data owners simply request access. Upon request, data owners are immediately notified that a new user is requesting access to their data but are also given all relevant information regarding the user, such as who they are, what data they are requesting access to, why they are requesting access, and most importantly user and entity behavior analytics (UEBA) and machine learning-backed recommendations on if they should have access. These data owners can make the decision about whether this requester should or should not have access. Once approved, permission changes will take place granting the user access to the data for the time period specified. This greatly speeds up the sharing of data and limits confusion as users will have their requests sent to the appropriate individuals immediately.

Second, Varonis provides organizations with an efficient and secure enterprisewide eDiscovery solution: DatAnswers. This module considers all metadata, including discovery rules tagged within the Data Security Platform, and allows users with the requisite permissions to search for data no matter where it may be stored. Varonis will automatically enumerate search results specific to the user to ensure they only see data to which they have access. This allows users to get the most up-to-date information regarding active mission requirements, while also preserving the need to know by not allowing users to search for data they shouldn't be able to access.

Finally, once data is searched and access is granted, all actions on data are audited by Varonis. This ensures that all access is tracked and correlated into a user's behavioral profile. The user and entity behavioral analytics profile for each user allows Varonis to ensures that users are acting normally. Once an alert on suspicious activity is triggered, automatic action can take place and investigations can be conducted that provide a full contextual view to describe the entire attack scenario. Using all the solutions within the Data Security Platform ensures that the entire life cycle of data is protected. Data is searched for, shared and accessed all in a secure and audited way so that data integrity is ensured.

BIO: Jim Evans leads the Army team at Varonis Public Sector and has supported the American warfighter for 33 years.

AI-Backed Data Interoperability

Burt Wagner, Senior Solutions Engineer, Tamr • burt.wagner@tamr.com

ABSTRACT

It is not uncommon today for organizations to have data silos across projects and departments, and it is not trivial to bring these silos together into a single unified data model for sharing data across the organization. Having multiple, disparate copies of data in inconsistent formats with no standards of definition also makes producing meaningful analytics impossible. A person can typically map a few separate data silos into a unified data structure using standard rules-based tools today, but the complexity of those rules typically grows exponentially with the addition of each additional data source. Simply put, existing tools in place simply do not scale to the problem at hand.

Tamr has created a human-led artificial intelligence and machine learning (AI/ML) solution to analyze existing data silos, create a unified data model and transform data. As new data sets are added, Tamr uses its AI/ML architecture to recommend mappings into the unified model, as well as de-dupe and resolve the underlying data. Organizations can automatically map columns that score higher than their predefined confidence factor, so a human only needs to address columns that the ML model is unsure about. As each additional data set is added to the unified data model, the ML model gets smarter, thus reducing the amount of total time that a human must perform manual column mappings. Learn how Tamr can help the Army unify its data to establish global standards, facilitating system interoperability and allowing meaningful analytics.

BIO: Burt Wagner has been solving data problems for national security customers for nearly 20 years. He has worked as a data architect, data engineer and data scientist for elements within the IC for most of that time, creating unified data models and producing meaningful advanced analytics for national security problems. Now with Tamr, Wagner is helping customers across the federal government use cutting-edge machine learning and AI to address data unification and work across data silos.

Automate Your Data Life Cycle With Object Technology

Scott Rich, Deputy Chief Technology Officer Americas, NetApp • scott.rich@netapp.com

ABSTRACT

NetApp StorageGRID Object Repository technology includes an Integrated Lifecycle Manager (ILM) that allows for data to be managed by policy not administrator actions. Data location, protection scheme, lifespan, replication and access can be managed through policies defined by the various Army programs rather than individual administrators or automation scripts. The ILM acts on extensible object metadata, so policies can be tailored to fit program requirements. Security data can be held locally for a short window and moved to a DCO archive for analysis and long-term storage. Imagery can be moved to COCOM Intelligence Brigades based on geocoordinates added to the object metadata. Collected data can remain in theater for a set period of time, replicated to CONUS for overwatch, but automatically removed from the theater assets to conserve capacity, all while still remaining discoverable across the global namespace.

StorageGRID Object software also provides the REST-based API's so that additional functionality and data management can be performed through tools and applications developed to run on top of this object foundation. Integration with public clouds to tier data for archiving, access native cloud functions like AI/ML algorithms or data analytics. Through StorageGRID capabilities, the Army can manage, protect and secure data with minimal administrator interaction, leveraging the leading-edge capabilities being developed in the public cloud as well as Army-specific applications in the future.

BIO: As a leader in data management and with over half of the Army's data currently hosted by NetApp data management systems, the company is uniquely positioned to help move the Army to a more data-centric culture and infrastructure in order to leverage innovative technologies from cloud service providers, AI and ML services, and mobile, containerized application solutions.

Unlocking the Power of Data With Multi-Cloud Analytics

Geoff Tudor, Senior Vice President, Panzura/Vizion.ai • nbabayan@aequussg.com

ABSTRACT

Panzura's Vizion.AI platform delivers a single, unified vision into data across the enterprise and provides powerful search, analysis, recovery and control of multi-cloud data for greater productivity, operational intelligence, improved security and reduced storage costs.

Indexing and search functions are powered by Vizion.ai's breakthrough hyper-scale, multi-cloud data engine, helping companies quickly query and locate data across multiple clouds. Running on this hyper-scale data engine is Elasticsearch, a multi-tenant full-text search engine with schema-free NoSQL document store. The combined solution enables customers to scale from one record to billions with immediate search results across an unlimited number of Panzura's record types and unlimited number of data repositories. The platform service also provides a powerful analytics engine that addresses a broad range of use cases. With integrated machine learning, Vizion.ai provides cloud cost analytics based on historical usage trends to estimate cloud savings based on moving to lower-cost tiers of storage. With audit data search, forensic discovery can be used to track user behavior and actions across multiple files, folders and clouds.

BIO: Geoff Tudor is senior vice president of Panzura's Vizion.ai. Tudor has more than 22 years of experience in storage, broadband and networking. Previously, he was chief cloud strategist at Hewlett Packard Enterprise, where he led CxO engagements for *Fortune* 100 private cloud opportunities, resulting in 10 times growth to over $1 billion in revenues. Before that, he was cofounder and launched award-winning products in cloud storage at Nirvanix (acquired by Oracle), backup and recovery at GNS (acquired by Symantec), and gigabit ethernet last-mile networking with Advent Networks and Tellaire (acquired by MRV Communications). Tudor holds an MBA from the University of Texas at Austin and a BA from Tulane University. He holds patents in satellite communications.

Modernized Data Management Strategy for the Army

Ash Banerjee, Principal, The Brite Group Inc. • ash@thebritegroup.com

ABSTRACT

By developing an overall enterprise data management vision, goals and plan, the Army will be well positioned to generate, prioritize and launch high-impact initiatives that help expand enterprise capabilities. The Brite Group approach recommends proceeding to use lessons learned to inform assessment and planning for data strategy, infrastructure, service model, project governance and change management approaches.

The company recommends developing an enterprise data management strategy with a focus on next-generation data sharing practices that can rely on a multi-tiered architecture capable of sharing data between all services and divisions at all classification levels and maintaining the data standards to maximize efficiency in sharing. Incorporating enterprise standards toward activities such as data modeling and metadata management will unlock further capabilities. Shifting to a horizontal data distribution model is key to maximizing interoperability; the Army's investment in automation capabilities and data services will further enhance to the Army's data competencies. After an initial search and discovery phase, the installment of data management best practices, such as adopting a standard enterprise taxonomy for elements in critical data sources, will be vital to the success of the objectives.

The Brite Group understands the challenges with dirty data stemming from the Army having data in multiple sources that are not integrated and of questionable quality. A data quality management plan needs to be incorporated as a part of the overall data management strategy and investments to improve accessibility (e.g. API development) of the various data sources to the central data architecture team and other components should be prioritized. The process of publishing and subscribing to data requests from various authorized entities needs to be a trusted and reliable drill.

The Brite Group's approach to developing a comprehensive data strategy and roadmap are based on the following four pillars:

- People: Data analytics talent includes staff resources with data science skills, data analytics contract/project management and agile expertise; opportunity to upskill existing staff to fill some of these needs.
- Capability: Required core competencies will include expertise in policy/customer needs, research question development, data analytics project and contract management and analytics execution.

- Structure: A centrally coordinated hub-and-spoke model strikes an appropriate balance between empowering individual Army components and strengthening an enterprise data management strategy to drive enterprise projects and capabilities.

- Process: The Army's enterprise data management strategy definition and execution oversight should be led by a steering committee composed of a senior champion and cross-functional group of senior leaders.

An assessment of the state of the Army's data assets and efforts in progress needs to be initiated first. Based on the company's initial assessment and experience with the Army, it believes there will be significant opportunities to build capabilities across each of the key levers for accomplishing the Army's desired outcome.

The TBG data capabilities maturity assessment framework can be tailored to meet the Army's specific challenges.

BIO: Ash Banerjee is a seasoned IT professional with more than 20 years of experience in data management, data architecture and data analytics programs in the federal space. He leads The Brite Group, a small business firm that has been providing exceptional service to federal agencies in the field of data analytics. The firm is a corporate AFCEA Small Business member for several years. Banerjee also has been a member of the AFCEA Small Business Committee for the past two years, providing active support to the broader community for various causes.

21st Century, Unified Approach to Data Analytics

Harold Heriford, President/CEO, Solutions4Less Inc. • robert.heriford@solutions4less.tech

ABSTRACT

Incorta is a unified data analytics platform for warfighters and decision makers and provides an option to have an advantage over adversaries.

Current 1990s-style data warehouse/data analytics platforms are slow, costly and labor-intensive. They use a process of data modeling and extract-transform-load (ETL) to place a subset of real-time production data into a data warehouse. A smaller subset is loaded into data marts, an attempt to accelerate retrieval. The ETL applications are expensive and require skilled individuals to ensure data is properly formatted and linked ("joins"). A query requiring five or more joins creates a noticeable increase in response time. A traditional data warehouse takes months to years to build and implement, with presumptions of a decision makers requirements. Some questions a decision maker asks cannot be answered because the data was not loaded. Changing situational requirements demands a new sets of information. It takes months to reconfigure the system to provide answers. Up to 80 percent of these warehouses fail. Traditional warehouses typically refresh data every 24 hours, risking information being overcome by events.

Another costly, complicated component is the user interface, business intelligence/data intelligence application. Decision makers rely on skilled information technology personnel to act on their request for vital dashboards and reports. The request is placed in a queue, with results taking weeks. Traditional data warehouses don't scale easily and are performance challenged.

Incorta's innovative and emerging technology is designed for today's needs in the age of big data and fast analytics. Incorta's Direct Data Mapping (DDM) provides 100 percent access to an organization's data with 100 percent fidelity. Data is refreshed every 5 minutes, providing real-time information. The extract-transform-load process and data modeling are not needed, and their time, labor and costs disappear. Row-level security from source system date is preserved, enforcing consistent security policy across users, user-roles, sessions and applications. Encryption with detailed audit logs enforce governance.

Incorta uses Apache Parquet in place of a data warehouse. Parquet offers many advantages for data management, including reduced file size through compression, performance gains with parallel processing efficiencies and the ability to store complex nested data structures and break away from relational model constraints. Advanced analytics capabilities include predictive modeling and machine learning, using Apache Python-ML for data cleansing and Spark-AI for predictive analytics. Easy-

to-use data visualization and analytics features are built in. Incorta is on-premises or cloud-based. On-premises deployment uses low-cost commodity hardware. Data can be accessed through desktop or mobile devices.

Typical Incorta implementation takes 12 to 18 weeks. Dashboard and reports are delivered in sub-seconds. Incorta's 10x data compression provides a smaller storage footprint. The user interface is simple and designed for the average user. Current users report higher return on investment and significantly lower total cost of ownership.

Incorta is the solution to the Army's data concerns and issues.

BIO: Harold Heriford is a retired U.S. Air Force combat veteran with more than 18 years at Oracle as a data warehouse architect for over 10 years and in pre-sales and sales for 7 years. He is the president/CEO of Solutions4Less Inc., a certified Service-Disabled Veteran-Owned Small Business. (SDVOSB)

Data Catalog to Data Lineage: Data Storytelling

Pragyansmita Nayak, Chief Data Scientist, Hitachi Vantara Federal •

pragyan.nayak@hitachivantarafederal.com

ABSTRACT

Data is undoubtedly recognized today as a valued asset. Data fusion and feature engineering maximize the utility of a datasets where the sum is greater than the parts, almost like how diamonds formed from a couple of carbon atoms under intense heat and pressure but in much less time!

The level of effort spent in identifying these key metrics and nuggets of information are invariably wasted since they are not usable downstream due to various factors not only limited to knowledge sharing, training and documentation. Data governance processes help streamline this process of data acquisition all the way through advanced analytics for knowledge and economic value creation. Data-centric organizations such as the U.S. Army can thus leverage their data assets for data-driven decisions and effective self-service analytics for enhanced tactical, operational and strategic intelligence.

A solution architecture that can work seamlessly with the existing data sources and yet remains extensible for future needs via open standards is crucial to meet this critical need of today. Data governance encompasses a suite of tools and processes that enables quick insights into the data flow at the organization and associated sub-organizations. Data catalog and data lineage formalizes the data architecture and forms the solid backbone that supports the advanced analytics use-cases related to machine learning.

BIO: Pragyansmita Nayak is chief data scientist at Hitachi Vantara Federal (HVF), a wholly owned subsidiary of Hitachi Vantara. She has more than 20 years of experience in software applications and data science-related research and development. She holds a doctorate in computational sciences and informatics from George Mason University (GMU) and BS degree in computer science from Birla Institute of Technology and Science (BITS), Pilani, India. Her doctorate thesis focused on the application of machine learning techniques such as Bayesian networks for redshift estimation. She is the founder of the NoVA Deep Learning meetup and Washington DC Pentaho User Group (PUG).

Flattening Procurement Data Structures To Increase Data Velocity via Web API Service

Mason Beninger, Technical Program Manager, nGAP Incorporated • mason@ngap.com

ABSTRACT

Compress what constitutes a “Contract Line Item,” its properties and its relations to the rest of the program life cycle. The abstract envisions the creation of a government-only API utilizing the procurement data standard. The API would digitize the contract file into an interactive model, a model where properties can be updated in real time by other business applications. If a contract is modeled as an object, it allows machine learning algorithms greater access to relevant datasets.

BIO: Mason Beninger is a technical program manager at nGAP Incorporated. He holds a BA in political science from Temple University. His technical trade is in software usability testing and test architecture in web applications. He is currently pursuing his master’s degree in cybersecurity. Beninger researches and applies contract life cycle data management and automation approaches for DOD applications, in particular, contract change management on large or complex weapons platforms or programs, such as ships or multiyear maintenance programs.

nGAP develops acquisition management software and researches utilizing API endpoints in conjunction with machine learning to reduce human error, manage large amounts of contract change transactions and contract closeout process optimization within the context of the FAR/DFAR.

SECURE

Zero Trust Architecture

Mackenize Morris, Cyber Security Lead, Savannah River National Laboratory •
mackenize.morris@srs.gov

ABSTRACT

The integration of new communication standards like 5G facilitate the paradigm shift toward the Internet of Things (IoT) and transformation of data flow. Low power, high bandwidth microdevices in IoT are designed to be as cheap as possible inversely proportional to the security of such devices. Compromise of any node on the network opens up the mesh to adversary and malicious behavior heightening the risk of insider capabilities. Implementing zero trust architecture is the needed solution. Zero trust is a behavior and design choice rather than a single product solution. No vendor solution captures and sells a zero trust environment. Rather, it is defined in an array of controls laid out in a well-crafted risk management framework.

The principles of zero trust focus on strengthening the C-I-A triad for internal devices; separating from the old mantra of perimeter security. Traditionally, enterprise networks have taken a bilateral approach to implicit and explicit trust. The network implicitly trusts everything internal while explicitly distrusting all outside sources. By taking away implicit trust relationships on the internal network, a zero trust architecture eliminates the ability for insiders to have free reign of the domain. Segmenting each user and each device into their own internal subdomain allows for detailed anomalous detection of their regular behavior and network footprint. Deviation from their individual norm is detectable and preventable by monitoring the typical device and user interaction captured from network traffic data.

Removing the trust relationship eliminates the concept of insider threat. Once fully implemented, a zero trust network has no insiders or outsiders, just users. Whitelisting, machine learning and regular group policy will strip away the ability of insiders to cause harm outside their micro domain. With each successful implementation of zero trust concepts, the mesh of networked devices becomes more resilient to attack from outside or inside.

BIO: Mackenize Morris moved into cybersecurity after working as a process control engineer, realizing that his career before him was not the right choice. Morris built and led the industrial control systems cybersecurity team within the Savannah River Site Tritium processing facility for three years. During this time, he built out the risk management framework and system security plan for the operational technology systems on site. Morris now works in research and development of new cybersecurity techniques and solutions for critical infrastructure at Savannah River National Laboratory. Morris holds numerous cyber security certifications including: GSEC, GCIH, GMON, GICSP, GRID, GREM, GDSA, GPEN, CISSP and Security+.

Intelligence Observation From the Russian Threat Landscape

Christian Rencken, Intelligence Advisor/Subject Matter Expert, CrowdStrike •

lindsey.hunter@crowdstrike.com

ABSTRACT

Over the last decade, Russian cyber threat actors have continued to display their highly sophisticated and technical capabilities affecting not just their Ukrainian neighbors but also those around the world. Connected closely to geopolitics, Bears (CrowdStrike's name for Russian nation-state affiliated actors) have worked closely and competently to achieve nation-state goals. These ambitions are fulfilled through stealing intellectual property, influencing elections around the globe, and even initiating attacks that affect the kinetic world. Given this activity, this presentation will cover three aspects of cyber intelligence:

- Strategic-level threats: Strategic intelligence shows how global events, foreign policies and other long-term local and international movements can potentially impact the cyber security of an organization.
- Operational-level threats: In the same way that poker players study each other's quirks so they can predict their opponents' next move, cybersecurity professionals study their adversaries. Behind every attack is a "who," "why" and "how."
- Tactical-level threats: Tactical intelligence is focused on the immediate future, is technical in nature, and identifies simple indicators of compromise (IOCs).

At the conclusion of this presentation, attendees should walk away with an understanding of the primary Russian threat actors tracked by CrowdStrike Intelligence; what their motivations are; and a glimpse into the future operations of the Russian state.

BIO: Before joining CrowdStrike, Christian Rencken had four different roles through his time at a previous cyber intelligence company. Most recently he was on the strategic intelligence team, which assigned him three main priorities: researching and writing RFIs, writing daily and weekly analyses of the cyber threat landscape, and contributing to the weekly cybersecurity podcast where he and his team would discuss cyber trends observed over the week and give commentary on perceived future threats. His job before strategic intelligence was a cyber intelligence analyst at both the entry and senior role. In these roles, he monitored his clients' digital footprint by crawling the open, deep and dark web to look for threats. Rencken has worked to improve the security postures of companies from *Fortune* 500 to SMB across multiple different verticals.

Federal Zero Trust Architecture Case-Study

Drew Epperson, Director, Federal Solution Architecture, Palo Alto Networks •

depperson@paloaltonetworks.com

ABSTRACT

Zero Trust is a strategic initiative that helps prevent successful data breaches by eliminating the concept (emotion) of trust from an organization's network architecture. Rooted in the principle of "never trust, always verify," zero trust is designed to protect modern digital environments by leveraging network segmentation, preventing lateral movement, providing Layer 7 threat prevention, and simplifying granular user-access control.

Palo Alto Networks hired the founder of Zero Trust (John Kindervag) in 2016 with the intention of bringing the theoretical benefits of zero trust into an operational capability. Over the past several years, Palo Alto Networks has conducted numerous zero trust workshops, production prototypes, and policy documents for federal clients. This session will provide real-world use cases, operational guidance, highlight unique federal challenges, and provide a zero-trust maturity model that any customer can leverage.

The session will provide participants with a true case study of operational zero trust deployments. In today's market, there is an abundance of marketing around zero trust, but a lack of meaningful guidance and lessons learned. The company's approach is to provide transparent feedback and recommendations for those who are planning their zero trust journey. Palo Alto Networks is currently engaged in several DOD zero trust prototypes that can provide insight to organizations and individuals who are interested in the topic.

Overview of key topics covered:

- Why the federal government faces unique challenges regarding zero trust and what they are.
- Lessons learned from operational experience covering topics such as: methodology matters, data and application categorization, "inside-out vs. outside-in" designs, identity considerations, and maturity models.
- Real-world reference architectures from federal customers.
- Policy and organizational alignment for success.

BIO: Drew Epperson is the chief solution architect for Palo Alto Networks Federal. In this role, Epperson leads a team of technical leaders who support customers ranging from operations staff to the General Officer/C-level as a trusted advisor. Epperson has designed and implemented critical cyber solutions incorporating network, endpoint, analytics and the cloud to classified/unclassified environments across the entire federal government. Epperson also leads Palo Alto Networks innovation initiatives with a focus on solving hard problems customers face in meeting their mission. Currently, Epperson also serves on the NSTAC subcommittee for Software Defined Networks. Before joining Palo Alto Networks, he served as the chief technical strategist at McAfee Federal. Epperson's background in hands-on project manager, sales engineer and field CTO roles provide unique abilities in solution development. He holds a BS in IT operations and an MS of information assurance and cybersecurity with an emphasis in cyber analytics from Penn State University.

A Comprehensive Zero Trust Security Approach

Matthew Jones, Cybersecurity Sales Specialist, Cisco Systems Inc. • matjone2@cisco.com

ABSTRACT

Zero trust is a comprehensive approach to securing all access across networks, applications and environments. This approach helps secure access from users, end-user devices, APIs, IoT, microservices, containers and more. It protects the workforce, workloads and workplace. Assume zero trust when someone or something requests access to work assets. Trustworthiness must be verified before granting access.

The Cisco Zero Trust security framework helps prevent unauthorized access, contain breaches and reduce the risk of an attacker's lateral movement through the network.

Cisco Zero Trust enables organizations to:

- Consistently enforce policy-based controls.
- Gain visibility into users, devices, components and more across the entire environment.
- Get detailed logs, reports and alerts that can help better detect and respond to threats.

Zero Trust for the Workforce

Cisco Duo helps protect users and their devices against stolen credentials, phishing and other identity-based attacks. It verifies users' identities and establishes device trust before granting access to applications.

Zero Trust for Workload

Secures hybrid, multicloud workloads and contains lateral movement with application segmentation from Cisco Tetration. Get complete visibility and determine the dependencies within databases and applications.

Zero Trust for Workplace

Cisco Software-Defined Access (SD-Access) helps gain insight into users and devices; identifies threats and maintains control over all connections across the network, including Internet of Things (IoT) devices.

BIO: Matthew Jones is a cybersecurity sales specialist at Cisco Systems Inc. with 20 years of experience working with the DOD.

Zero Trust for Army's Internet of Things

Jim Taylor, Chief Technology Officer, Onclave Networks Inc. • jtaylor@onclave.net

ABSTRACT

The U.S. Army is a complex enterprise consisting of many "things." These things can be networks, people, devices and workflows. All these things have various levels of security associated with them, and their behaviors vary depending on current activity. The question is how to generate a security framework that encompasses and protects this growing Internet of Things (IoT).

In answer to this question, Onclave Networks Inc. developed Secure IoT, a zero trust secure virtual segmented network and communications platform designed to protect an enterprise. The solution delivers enterprisewide protection for all IoT, operational technology (OT) and even information technology (IT) regardless of age, operating system or protocol. In conformance with the NIST Risk Management Framework (RMF), it provides trusted, secure and protected networks that use their own root of trust, are cryptographically separated from other networks running on the same infrastructure and are continuously monitored for changes inside the secure networks.

Secure IoT eliminates the OT attack surface, making it invisible and inaccessible and prevents crossover and man-in-the-middle breaches that compromise IT assets. This is achieved using Onclave's Trusted Communications Framework (TCF) and is enforced in the Secure IoT platform. TCF applies zero trust through the Secure IoT Orchestrator's topology and policy settings. Using Orchestrator, trusted operators build a policy of trusted devices and services that allows them to establish secure segmented tunnels over existing networks called enclaves. This removes the attack surface of any IoT and OT devices hidden inside the enclaves. Trust is established between the Onclave Gateways and Bridges using trust keys that are created by a unique process. The public key of each device is stored in Onclave's Secure Blockchain. Initial communications to establish the trust is fully encrypted for both the transport and content. Once a trust is established, the participating devices create individual Layer 2 over Layer 3 secure tunnels over the existing network to create a secure and protected segmented transport with limited or no change to the current network infrastructure.

Secure IoT employs unique hardware and specially designed software and follows the Risk Management Framework. The hardware creates cryptographically separated networks on the same wires and is more secure than applying VPNs, firewalls and policies using standard IT tools. Its software and firmware manage secure communications and perform automated continuous monitoring, making everything inside inaccessible to hackers and unauthorized users.

The Secure IoT platform life cycle maps directly to the NIST RMF and involves segmentation and access management resulting in reduced overhead. Onclave developed Secure IoT based on enforcement of Zero Trust employing RMF Steps 1-6: Categorize, Select, Implement, Assess, Authorize and Monitor.

Secure IoT's Orchestrator defines the need to have trusted identities that include people, services and devices. The establishment of trusted individuals is enforced by an organization using Secure IoT by incorporating its policies and its policy engine into the platform. Onclave enforces this policy by preventing any unauthorized individuals from having access to secure enclaves.

BIO: James Taylor, CTO, has more than 30 years of experience working in the computer industry where he led the development of numerous emerging technologies into successful commercial endeavors. Most recently, Taylor led the design and development of Onclave's Secure IoT platform centered on his authored patent—Dynamic Cipher Key Management. The inimitable tactic Taylor implemented into his design of the Secure IoT platform has fostered new commercial opportunities focused on building Secure Virtual Segmented Networks over existing LANs and WANs. Taylor initiated a comprehensive analysis of the industry and market, creating the value proposition, and a close identification of the customer. His experience from his other business endeavors helped him guide and build the business operations, policies, legal structure of the business, marketing and growth plans, financials and the team needed to make Secure IoT a success. Earlier in his career while at Unisys, Taylor developed the ES7000 Portfolio. He supported marketing of solutions, including development of marketing plans, sales force training and staff development. This program was a major success for Unisys. Taylor also developed, implemented and managed the business profile, business mode, and processes for Unisys' new US Microsoft Metro Practices. Practices launched in Los Angeles, San Francisco, Minneapolis, Chicago, New York and Boston.

Insider Threat Prevention

George Kamis, Chief Technology Officer, Forcepoint • kamis@forcepoint.com

ABSTRACT

Forcepoint is depended upon by distributed enterprises and government agencies around the world to connect and protect their highly sensitive environments. Forcepoint's human point system combined with its strength in cross-domain security enables agencies to take a game-changing, risk-adaptive protection approach to ensure comprehensive security and visibility across these highly sensitive environments.

At the forefront of adaptive security is behavior-centric analytics. By fusing data from traditional security systems and output from data loss prevention with that of other agency sources (e.g., SIEM data, HR, travel logs, email and chat communication), agencies get a more informed contextual picture of behavior to quickly identify anomalies. Using this context, analytics adapts policies automatically based on changes in risk levels, providing Risk-Adaptive Protection. Risk-Adaptive Protection automatically responds to risk and adapts policies down to an individual user level—controlling data across government environments.

Game Changer, enabled by Forcepoint Dynamic Data (DDP), a next-generation data loss prevention (DLP) and behavioral analytics (BA) product, delivers proactive, risk-based protection. Its human-centric model provides behavioral context to identify anomalies through analytics, which ingests and fuses data from traditional security systems, output from DLP and user activity monitoring (UAM) and internal agency sources. DDP utilizes context and risk scores provided by behavioral analytics and dynamically enhances UAM and DLP data collection and/or invokes enforcement to adapt policies automatically down to the user. Dynamic protection is used to control data and access or alert at all classification levels on-premises, endpoints and the cloud, enabling agencies to secure data at rest, in-motion and in-use across IT environments.

Endpoint monitoring occurs through both UAM and DLP, with each classification level capable of monitoring the same types of information and metrics as the other classification levels. Each classification level keeps data, control and monitoring local to the respective enclave. Enclaves can utilize the built-in NiFi traffic flow mechanism within the behavior analytics platform to forward ingested data through the High Speed Guard architecture up to a single enclave that serves as the unification point for all user risk level data. Only data that is positively vetted can traverse through the NGFW/High Speed Guard architecture.

Risk-adaptive protection dynamically applies monitoring and enforcement controls to protect data based on the calculated behavioral risk level of users and value of data accessed. This allows security organizations to better understand risky behavior and automate policies, dramatically reducing the

quantity of alerts requiring investigation. Each user has a unique and dynamic risk level. Risk levels are driven up and down based on changes in behavior. Risk levels drive different outcomes. Security adapts to risk levels as they fluctuate.

Benefits:

- Insider threat protection.
- Automated risk-adaptive learning, adapting policies and enforcement down to individual user.
- Micro segmentation to reduce internal insider threats.
- Early detection through behavior indicators to defend against insider threats.
- Professional services provide comprehensive program set up to include composing tactics, techniques and procedures (TTPs) to make a more security aware environment.

BIO: George Kamis is the chief technology officer, Forcepoint Federal. He is also the chairperson for the Forcepoint CTI Council, which is made of subject matter experts from each product and capability area within Forcepoint. In this position, Kamis works closely with information assurance industry leaders, government executives and the Forcepoint executive management team to help guide Forcepoint's long-term technology strategy and keeps it aligned with industry requirements. By leveraging his wealth of experience in cyber and cross-domain solutions, Kamis has helped lead Forcepoint to the forefront of cross-domain and cybersecurity solutions in the public sector. Prior to his role as CTO, Kamis served as vice president of engineering at Raytheon Trusted Computer Solutions for more than 10 years. As part of those duties, he ran both the professional services and development organizations. Kamis has over 27 years of operational and management expertise in secure systems engineering. Prior to joining Raytheon Trusted Computer Solutions, he worked at the U.S. Naval Research Laboratory's, Center for High Assurance Computer Systems where he developed cross-domain and multilevel secure systems and for the Navy. Kamis holds a BS in electrical engineering from West Virginia University and holds active memberships in both the Institute of Electrical and Electronics Engineers (IEEE) and AFCEA International.

Protection Beyond the Perimeter

Brent Hansen, Chief Technology Officer, Thales Trusted Cyber Technologies •
mary.shiflett@thalestct.com

ABSTRACT

Thales Trusted Cyber Technologies' (TCT) mission is to protect the U.S. federal government's most vital data from the core to the cloud to the field. TCT provides an extensive portfolio to secure data at rest, data in use and data in motion with authentication, access controls, encryption and cryptographic key management. Its capabilities include making data accessible, trusted and secure to meet compliance with the Army Data Strategy, map to recommendations in the Defense Information Board (DIB) Zero Trust Architecture and NIST SP 800-53 security and privacy controls. In order to make data secure, the company addresses enabling controls to meet laws, regulations and guidance; protect PII; secure proprietary information; and make data available and prevent data breaches. Recommendation 1.4 of DIB ZTA calls for encryption of data at rest and data in transit.

To secure the modern hybrid, multi-cloud architecture, TCT eliminates traditional data siloes by providing secure multi-tenancy with the establishment of security domains with clear separation of duties. This approach provides a much stronger security posture over traditional methods such as self-encrypting drives or other types of native encryption tools. The policy-based granular access controls create guardpoints around files, folders or shares enforced by encryption to clearly establish who, when, what and where data is accessed. This information can then be sent SIEM tools to enhance overall cyber situational awareness.

The Army network cannot rely on network and endpoint security alone but rather needs defense in depth for multi-level security for the contingency of assuming network compromise, making it imperative to secure data around granular policy enforced by encryption. Encrypting sensitive data is a fundamental part of any cybersecurity strategy. However, the cryptographic keys used to encrypt and decrypt data are often overlooked. Cryptographic keys are the keys to the kingdom. Large volumes of encrypted data yield copious amounts of cryptographic keys, which need to be managed, stored and secured. If these keys are compromised, attackers can gain access to encrypted information. To reduce enterprise complexity, key management has to provide a secure key management life cycle within a FIPS (140-2 L3) boundary to support data on premise and in the cloud rather than relying on disparate point solutions.

TCT's technology protects sensitive information by controlling access to data with strong encryption controls and centralized key management to protect critical information both at rest and in transit to ensure data integrity across data fabrics in a multidomain operational environment. The company's goal is to defend data by gaining a clear understanding of traffic flows across the various fabrics and establishing a trusted security baseline. Any deviations outside of the identified baseline would

be perceived as a potential threat, and therefore access to encryption keys to decrypt data would be denied. In addition to providing strong cryptographic controls, TCT provides a policy-based data discovery product that expands upon its data security platform to discovering and securing sensitive data dynamically. This methodology ensures the integrity of mission critical data, without disruption in a real-time environment.

BIO: Brent Hansen is Thales TCT's chief technology officer. Hansen leads Thales TCT's sales engineering organization and spends the majority of his time evangelizing and strategizing on data-centric approaches for federal agencies looking to avert being the next victim of a data breach.

Hansen brings more than 19 years of IT experience in data and enterprise architecture, data warehousing, big data and business intelligence. He is an industry expert in data encryption and tokenization. He leads teams that architect security strategies to secure and protect sensitive data for both federal government and large commercial enterprises across the globe.

Securing the Internet of Things

Chris Rouland, CEO, Phosphorus Cybersecurity Inc. • chris@phosphorus.io

ABSTRACT

Phosphorus Cybersecurity Enterprise is a solution for managing the firmware and security patches, credentials and endpoint configuration of embedded devices (IoT/BoT/ICS). Historically, this type of security is delivered to PCs and servers with agents; however with the billions of IoT devices from thousands of manufacturers, agent-based technology is not possible. Phosphorus has developed an agentless, completely automated platform for IoT security that acts as a radical force multiplier for security tasks and vulnerability remediation that have either been done by hand or, in most cases, not at all. The impact of this technology is full visibility into the customers embedded device world, and the ability to secure devices in an automated fashion that have most likely never been secured at all. By securing the "things" in the environment, Phosphorus helps create a safer operating environment and mitigates insider threat by bringing devices into security compliance.

BIO: Chris Rouland is co-founder and CEO of Phosphorus Cybersecurity Inc. A 25-year veteran of the information security industry, Rouland is a renowned leader in cybersecurity innovation and disruption. In his career, Rouland has founded and led several multi-million dollar companies including Bastille, the first company to enable enterprise security teams to assess and mitigate the risk associated with the growing Internet of Things, and Endgame. Rouland stood up the X-Force for Internet Security Systems and later became CTO.

Comply-to-Connect (C2C): Enabler of Zero Trust for the DODIN

Dean Hullings, Global Defense Solutions Strategist, Forescout Technologies •

dean.hullings@forescout.com

ABSTRACT

As the DOD rolls out the Comply-to-Connect (C2C) cybersecurity program, it is not too early to consider how the components of C2C will support the department's zero trust strategy. Recent guidance from the National Institute of Standards and Technology (NIST), Draft Special Publication (SP) 800-207, Zero Trust Architecture, makes major strides toward establishing a common understanding of what of zero trust is. Among SP 800-207's most important conclusions is that zero trust is not a binary end state, but rather a journey more effectively measured through the perspective of a maturity model framework. Zero trust is not achievable through the implementation of a single technology or product. It is, instead, a strategic approach that requires a comprehensive re-evaluation of how an organization leverages processes and technologies to meet its security objectives. Whether the goal is to implement zero trust in enterprise, cloud, IT as a Service or hybrid environments, the basic security objectives often remain the same. And it is a core responsibility of the cybersecurity industry to partner to deliver truly integrated solutions that deliver consistent security across all environments and not to encumber soldiers with unwieldy integrations.

NIST SP 800-207 highlights the specific components of zero trust, many of which–not coincidentally–are shared with C2C, especially asset discovery, application-based policy enforcement, security information and event management, and identity and access management. This session will explore the specific capabilities Forescout is providing to the C2C framework that, in concert with other technologies, enables a more interoperable, comprehensive and automated zero trust security vision across DOD information networks.

BIO: With 30 years of experience in the Information Technology industry, Dean Hullings provides strategic recommendations and guidance to the Forescout Public Sector account management team, connecting engagements and initiatives to maximize team productivity. Before joining Forescout, Hullings spent 26 years in the U.S. Air Force as a communications and cyber officer, serving in various leadership positions for Air Force and joint commands. Hullings has a BS in computer and information science from the University of Delaware and three master's degrees in public administration, military operations, and national security and strategic studies.

Insider Threat Prevention

Rick Bueno, CEO and Founder, Cyber Reliant Corp. • rabueno@cyberreliant.com

ABSTRACT

Cyber Reliant data protection is able to provide the Army with proactive, internal data security at all levels, especially as Army moves more to cloud and IoT. The Cyber Reliant data protection products will provide the Army with Never Trust (zero trust), Always Verify framework (RMF) to ensure the Army's data protection. Cyber Reliant prevents insider threat by differentiating system access from data access. Only the authorized data owner has access to the data and with that is closely monitored to report when who, what, when and where the data was accessed. This combined with other auditing and monitoring capabilities will provide a chain of use for all data elements.

Cyber Reliant also prevents the exfiltration of data altogether and provides aggressive monitoring of any data moves. Cyber Reliant is a completely new paradigm in data protection. This capability was built and designed from funding received by U.S. Army Intelligence and USSOCOM to protect the Army's most sensitive operational data in the most austere cyber environment. Cyber Reliant applies data protection directly to the data itself. Traditional security relies on protecting the perimeter and on data encryption. Perimeter-based data protection strategies are important but are not enough as the perimeter can be, has been and will continue to be breached. Encryption is also an important data protection strategy, but it is not enough as encryption can also be broken with relative ease.

The key to successful data protection is to apply the data protection to the data itself in a manner so difficult and complex that not even the most sophisticated attackers would know how to break it. Cyber Reliant products were designed and built by offensive information operations engineers who designed a data protection product to counter the most sophisticated state-level sponsored attacks. Cyber Reliant incorporates not just one technique but a series of innovative and specialized techniques to counter any offensive information operations attempt to data exploitation and exfiltration. Cyber Reliant has implemented advanced key management, disassociation, encryption, shredding, embedding and dispersion techniques, which create a framework of data protection that has been adjudicated as Information Theoretic Secure by the NSA. Information Theoretic Secure implies a quantum resistance data protection methodology.

BIO: Rick Bueno, president and CEO of Cyber Reliant, is a U.S. veteran, entrepreneur and experienced visionary with a deep 35-year history of developing and executing cybersecurity strategic initiatives and solutions in the commercial, defense, intelligence and special operations community. Prior to founding Cyber Reliant in 2010, Bueno served as the National Security Agency Information Assurance Directorate AD Afghan Mission Manager/NATO Special Operations Forces for the ISR Task Force. In this role he worked closely with NSA, USDI, CENTCOM

and others to develop secure communications strategies in support of NATO and Special Forces missions. Prior to his role at ISR TF, Bueno served as the chief strategic architect for the Director of National Intelligence (DNI). As a member of the director's action group, he was responsible for early establishment of the Intelligence Community's policies and processes for the information sharing environment and other activities within the DNI CIO scope.

Security of Data and Endpoints with White Cloud Security Trusted Apps and Data Trust-Listing

Steven Shanklin, Founder and CEO, White Cloud Security Inc. • ziggy@whitecloudsecurity.com

ABSTRACT

The U.S. Army's Data ecosystem can be secured by White Cloud Security's Trust Lockdown (TL) zero trust app security and data trust-listing framework, which enforces zero trust app control and data visibility, access, classification and protection at the endpoint regardless of whether the data is shared in the cloud, in a data center or on a specific endpoint, and monitors and controls the creation, visibility, access and modification of data by users and automata processes.

Beside controlling which apps a user can run and data they can access, a data security system requires more than a data tagging system to prevent unauthorized data access or manipulation. For example, a root user in Linux can modify the SELinux tags used to control file execution and access. Therefore, the creation and management of the file execution and data file access policies must be outside the control of an administrator who has root/supervisory privileges on the endpoint.

TL's cyber-metric handprint file identification technology uniquely identifies each file based upon the file's own data content, prevents file identity spoofing that can be done via SHAttered Attack or manipulating of executable and data file identification tags and is always unique to each executable and data file's content.

TL also includes three significant and unique layers of protection against malicious security administration insiders to prevent Edward Snowden types of internal security breaches. Lone Wolf Detection and Deterrence informs other admins of policy changes; Lone Wolf Remediation using 1-click distrust to remove an admin immediately from an trust-listing inheritance tree; and Lone Wolf Prevention, which uses a 2-man+ rule that requires at least two or more admins to apply the same app trust or data trust policy.

Data trust-listing prevents malicious insiders (even those with root privileges) from accessing or manipulating data they are unauthorized to access and tracks which files have been modified via a blockchain of the cyber-metric handprints along with the relevant data file attributes. TL will also prevent users with root privileges from running apps and scripts that are untrusted or unauthorized for use.

TL is a kernel-level file filter driver on Windows and a Linux Security Module in Linux that communicates with a secure service in the cloud or in a data center appliance, which contains the trust-listing policies that determine which software is allowed to run and which data files can be seen, accessed or modified by a software package or component. TL's proven execution control security agent only allows trusted executables, libraries and scripts to run on endpoints. The company's data trust-listing extends its execution control trust-listing framework to identify, monitor and control the creation of data files and modifications to them.

TL works with both modern and legacy endpoints (from Windows 2000 and Windows Servers 2003) without changes to the legacy endpoints other than installing TL's endpoint agent. It is supported from Redhat/CentOS kernels 3.10 after adding and enabling the Linux Security Module to the Kernel (Ports to Debian and Raspbian in Q2 2020).

BIO: White Cloud Security was founded by cybersecurity professionals with a proven track record and more than two decades of cybersecurity software development experience in leading-edge host and network intrusion detection, automated remediation and application whitelisting. Its previous companies were acquired by Cisco Systems (Wheelgroup, Psionic), TIS (Haystack Labs) and Lumension Security (Coretrace). Its Trust Lockdown is a zero trust app security framework that verifies the cyber-metric handprint identity of each executable, dynamic code library and script every time they try to run. It blocks everything else.

Securing Army Data Assets Against Insider Threats

Katrina Matthews, Army BD Senior Manager, GDIT • katrina.matthews@gdit.com

ABSTRACT

Employing data analytics to complement comprehensive monitoring for suspicious activity within an information technology system is key to detecting insider threats and creating a "Never Trust (zero trust), Always Verify" climate. Each Army program has some level of logging on servers and networks, and transactional logging within systems. This generates an enormous amount of data. For this reason, comprehensive examination of that data isn't possible with human labor alone. Fortunately, many Army programs have the means to automate examination of this data through their investment in Splunk, which is designed to identify insider threats and generate alerts when insider threat fingerprints appear in the data. Most Splunk implementations fall short of delivering full value as they are inadequately configured and utilized.

To successfully harness the power of Splunk, it is absolutely essential to combine the knowledge of cybersecurity SMEs who understand threat tactics, techniques and procedures (TTPs) and what to look for in the data, data scientists who know how to process and organize the data, and Splunk developers who know how to configure Splunk. The combination of these three roles is the only way to fully harness Splunk's full capability for which the Army is paying. Patterns within the logs need to be identified, examined and flagged as suspicious. These patterns are then programmed back into Splunk as machine learning models that get smarter as additional scenarios are identified and added. Most Splunk implementations only scratch the surface of this capability as it is left to cybersecurity personnel to implement Splunk, without the help of data scientists and Splunk developers to complete the implementation loop. Once data science is used to identify the TTPs on a single program, the Splunk configuration can be shared across other programs, and by this sharing, the programs can rapidly strengthen their defenses against insider threat. Even though the log types may differ, programs can share the signatures to seek.

Splunk has built-in capabilities for compliance, risks analysis (RMF) and remediation, and GDIT has specifically expanded the capability of Splunk to increase the options available to the operators to perform their daily duties in addition to the reporting capabilities for decision makers. Splunk identifies the individual systems reporting and providing status, allowing the operator to easily determine the source of anomalies or changes in the environment that are not authorized, not approved by change management or no longer in compliance.

Splunk is a powerful tool that is available in cloud and on-prem platforms. Building a workforce with the right skills (cyber, data analytics and Splunk SME) will increase the Army's ability to rapidly detect and halt insider threats before serious harm occurs. Ensuring that TTPs are leveraged widely is another key to making this solution effective. The Army can facilitate this by hosting a monthly or bi-weekly workshop or consortium for the cyber SMEs, Splunk SMEs and data scientists to discuss and distribute their latest discoveries, and by compiling a catalog of Splunk-enabled TTPs for use by programs.

BIO: Dave Vennergrund is a senior director and Distinguished Technologist at GDIT. Vennergrund has more than 25 years of artificial intelligence, data analytics, data science, IT management and R&D. He led dozens of successful data mining, big data, AI and business intelligence efforts in intelligence, defense, and federal agencies including data lakes for Navy, predictive analytics at EPA, HUD and DOI, improper payment prevention at the IRS, USDA, CMS, VA, DFAS and OPM. Vennergrund has expertise in data analysis, predictive modeling, big data, cloud analytics (AWS, Azure, Google) and the deployment of analytic solutions in mission critical settings. Vennergrund has built data mining, business intelligence and business analytics centers of excellence, special interest groups and innovation centers. Vennergrund is an industry expert in AI/ML, predictive analytics, fraud detection and data science.

Vennergrund earned his BS in computer science at the University of Illinois, and his MS in computer science from Arizona State University, specializing in the application of artificial intelligence to software engineering. Vennergrund has researched and applied computer science, statistical analysis, and artificial intelligence methods to a broad range of national missions.

Cyber Readiness via Zero Trust and RMF

Jim Evans, Federal Account Executive, Varonis Public Sector • jevans@varonis.com

ABSTRACT

Varonis builds products around the concepts of the Risk Management Framework and controls described in NIST 800-53, risk mitigation, zero trust and risk management in general. This common strategy underscores both RMF as well as the company's operational journey, which acts as both a playbook and CONOPS for the Army utilizing the platform. By first understanding where all data resides, sensitive or otherwise, Varonis can provide a complete mapping and interactive interface for RMF compliance and zero trust. Granular understanding of access-based or directory service-based risk to data is a key component when determining appropriate security controls and policies. Once policies are determined through the RMF, Varonis provides capabilities to track these control sets and families appropriately.

Varonis has always viewed insider threats as the most problematic. This is because most attacks are executed by, or through, individuals otherwise approved for network access. Thanks to the collection of all relevant data and entity metadata, including real-time access monitoring, Varonis DSP provides true user behavior analytics. All activity is logged for alerting, analytical and forensic purposes. This ensures that any manipulation of data can be traced back to individual users with precise date stamps and activity descriptions. Every entity within the environment is monitored and a baseline of activity is created. Activity that is contrary to this established threshold of regular behavior will result in alerts of varying severity. Building on this baseline and threshold approach via semi-supervised machine learning (ML), Varonis DSP provides more than 180 pre-defined threat models. These tested and proven models include potential threats originating at the perimeter, such as communication by a compromised entity to a command and control server. Improper activity, such as repeated failed access attempts, credential escalation or ransomware-style actions are all logged, alerted on and delivered to analysts with complete and detailed forensic analysis. This allows for faster responses to true threats, as well as confidence to deploy more restrained responses in the event of non-malicious mistakes committed by users.

Ultimately, Varonis seeks to achieve 'data maturity', i.e. a complete zero trust architecture via remediation of folder-level permissions access. Only Varonis DSP can execute such remediation in either a step-wise commit or fully-automated fashion, thereby saving time and resources. Once the effectiveness of security controls provided by native operating systems have been leveraged to their fullest, any additional measures can be tracked and alerted on using Varonis.

The metadata collection, auditing and UEBA capabilities within Varonis combines for a solution that is primed to detect external threats as readily as it detects internal ones. With its complete understanding of all entities, both users and service accounts, and their behavior, Varonis DSP can leverage dozens of relevant threat models to detect when accounts may be compromised or when legitimate users are acting maliciously. Further, Varonis DSP leverages that same breadth of metadata to detect risky behavior originating outside the network. Credential escalation, brute force and DNS tunneling are just a few attack methods that the Data Security Platform can detect and alert using its unique insights.

BIO: Jim Evans leads the Army team at Varonis Public Sector and has supported the American warfighter for 33 years.

CLOUD

Operational Intelligence in the Cloud

Temika Cage, Solutions Engineer, Splunk • tcage@splunk.com

ABSTRACT

Splunk Enterprise is the leading platform for real-time operational intelligence. It takes the machine data generated by IT systems and technology infrastructure—whether it's physical, virtual or in the cloud—and turns it into valuable insights. Splunk Enterprise is geared for deploying within the cloud, as well as across hybrid environments—with a mixture of on-prem and cloud infrastructure.

For organizations looking for the full-feature set of Splunk Enterprise delivered as a cloud service, Splunk Cloud is an alternative or addition to Splunk Enterprise. Splunk Cloud is an AWS-based service that delivers all of the functionality of Splunk Enterprise with the flexibility of Software as a Service (SaaS). Using Splunk Cloud, users can search, analyze and visualize data from applications and devices across the entire environment.

Splunk Cloud meets the FedRAMP security standards, and helps U.S. federal agencies and their partners drive confident decisions and decisive actions at mission speeds. Now, agencies can ingest data once—in real-time—and use that same data to address a variety of challenges across various programs and initiatives spanning security and IT operations, as well as modernization and mission objectives.

Splunk Cloud is SOC 2 Type 2 and ISO 27001 certified. Splunk Cloud PCI and Splunk Cloud HIPAA are available. Additionally, Splunk Cloud is FedRAMP Authorized by the General Services Administration FedRAMP Program Management Office at the moderate impact level and also meets U.S. Persons requirements under ITAR. Splunk provides dedicated cloud environments for each customer as well as encryption in-transit and encryption at rest.

Collect and index any machine data—whether it's from physical, virtual, cloud environments or legacy software. Splunk is vendor agnostic and enables users to search, monitor and analyze any machine data to discover powerful insights across multiple use cases like security, IT operations, application delivery, industrial data and IoT. Additionally, with the power of machine learning baked in, users can make faster, more informed decisions across the organization.

Splunk offers granular visibility and unprecedented real-time insights for federal government officials who want to modernize and transform their agencies, so they can confidently adopt new paradigms, migrate to new technologies, and monitor their performance and availability. With Splunk, agencies can embrace modernization initiatives, improve efficiencies, enhance security and deliver superior citizen experiences. Using the Splunk platform, government agencies strengthen their future and

ensure success by extending citizen and cyber safety, delivering service excellence and embracing innovations responsibly.

Splunk can easily help with mapping events like tasks and code submissions in the build process and see problems as they're happening and build a little data science behind it to predict issues. Application life cycle analytics offer a platform approach that bridges team silos and provides insight into each phase.

BIO: Temika Cage is a solutions engineer and Splunk Cloud SME for Splunk's Army team.

Army Enterprise Cloud Strategy

Katrina Matthews, Army BD Senior Manager, GDIT • katrina.matthews@gdit.com

ABSTRACT

The Army can achieve short-term wins in the ability to use its data by utilizing a phased approach to analytics in the cloud. While the Army deals with obstacles to migrating its applications to the cloud, a hybrid cloud data lake is an ideal way to allow data in different hosting platforms to be utilized together as one unit. Using a secure, cloud native data analytics (SCNDA) reference architecture to establish an on-prem/off-prem data lake is an achievable first step to aggregating the Army's data and harnessing the power of AI and ML to draw mission-critical value from the data.

The SCNDA is a framework and methodology that is cloud provider agnostic and can be instantiated in any cloud, such as those from AWS, Microsoft Azure, Google Cloud Platform, and can even support on-premise and hybrid solutions. It aligns with the Army Data Strategy and can incorporate the Army's existing data tools such as Hive and Denodo, as well as other open-source and cloud-native pay-per-use analytics tools, crawlers and governance tools. Data stored in AWS, Azure, MilCloud clouds, or in big data platforms such as Hadoop HDFS, can be incorporated into the data lake using a logical data lake model, or alternately, the data can be sent to the data lake periodically and time stamped. This model allows the Army to leverage existing visualization tools, such as Tableau, Palantir, Microsoft ASP.Net. It also allows the Army to take advantage of powerful cloud native tools such as AWS QuickSight and Kibana or tools native to Azure and Google Cloud.

The Army can set up a strong governance model right from the beginning to ensure the security of the data. If using AWS GovCloud, LDAC's customers can access the data and utilize analytic products employing the numerous cloud-native analytical tools in a lower-cost model that does not require the need for large upfront licensing costs, nor will it incur the transactional costs of taking data out of the cloud. This model has a relatively small cost-per-use. Cloud native storage and compute is considerably more cost effective than traditional data center, if architected appropriately to use only the resources needed at the time they are required.

The longer-term win for the Army happens as applications migrate to the cloud. Re-hosting that includes refactoring to take advantage of the new technologies available in the cloud do not necessarily bind applications to a single cloud platform. Many tools and constructs are available on AWS, Azure, Google Cloud, and can in some cases be instantiated in on-premise hosting. The correct platform should be considered on a case-by-case basis for each application with several factors in mind. Some of the benefits of cloud-hosting are the elimination of hardware and support labor costs and the availability of modern container-based solutions that improve speed, ease of deployment and security of applications.

BIO: Speaker Bio: Dave Vennergrund is a senior director and Distinguished Technologist at GDIT. Vennergrund has more than 25 years of artificial intelligence (AI), data analytics, data science, IT management, and R&D. He's led dozens of successful data mining, big data, AI, and business intelligence efforts in intelligence, defense, and federal agencies, including data lakes for the Navy, predictive analytics at EPA, HUD and DOI, improper payment prevention at the IRS, USDA, CMS, VA, DFAS and OPM. Vennergrund has expertise in data analysis, predictive modeling, big data, cloud analytics (AWS, Azure, Google) and the deployment of analytic solutions in mission-critical settings. Vennergrund has built data mining, business intelligence and business analytics Centers of Excellence, special interest groups, and innovation centers. Vennergrund is an industry expert in AI/ML, predictive analytics, fraud detection and data science.

Vennergrund earned his BS in computer science at the University of Illinois, and his MS in computer science from Arizona State University, specializing in the application of artificial intelligence to software engineering. Vennergrund has researched and applied computer science, statistical analysis and artificial intelligence methods to a broad range of national missions.

Safe Migration To, From and Between Government Clouds With Veritas InfoScale

Doug Snyder, Chief Technologist, Veritas Technologies LLC • doug.snyder@veritas.com

ABSTRACT

Enterprise IT leaders face a demand for unprecedented levels of IT service continuity from their customers, both internal and external. The rise of cloud and virtualization, consumerization of IT, globalization of business and relentless growth means an increasing need for always-on modes of operations coupled with reduced tolerance for downtime. Veritas InfoScale Enterprise draws on a long heritage of world-class availability and storage management solutions from Veritas to provide reliable migrations and operations across physical, virtual and cloud infrastructures.

InfoScale Enterprise ensures critical IT environments stay up and running. Highlights include:

- Migration to Hybrid Cloud—Enables data migration for workloads to private or public clouds such as AWS or Azure, increasing platform flexibility.
- Sub-minute recovery for critical applications and databases—Enables highly available critical applications and databases such as Oracle RAC environments by leveraging easy-to-use cluster and storage management capabilities.
- Recovery across any distance—Recover locally (HA), in a metro region (through campus clusters) and globally (through wide-area DR), with wide support for software- or hardware-based replication across physical, virtual and cloud ecosystems. Test recovery readiness with Fire Drill with minimal disruption to production environments.
- Fast and intelligent recovery—Instantaneous failure detection and recovery with intelligent monitoring and optimal failover target selection.
- Availability across virtual, physical and multicloud—Such as VM rebootfree failover (without impacting VMware® vMotion™ and DRS) for VMware and failover across metro sites without data loss for Microsoft Hyper-V and failover within and across cloud data centers such as AWS and Azure.

BIO: Doug Snyder is the chief technologist for U.S. Public Sector, which includes the U.S. government, Department of Defense, state and local government, as well as healthcare providers. Snyder is currently in his 13th year at Veritas.

Snyder holds certifications for OpenStack (Certified Architect); CISSP (Certified Information Systems Security Professional) and ITIL v3 (Information Technology Infrastructure Library).

DNS as the Foundation of Hybrid Cloud Management

Ben Ball, Director of Strategy, BlueCat Networks • bball@bluecatnetworks.com

ABSTRACT

The cloud plays to the natural strengths of the Army's decentralized network environment, offering standardized services and innovative new capabilities on demand. At the same time, the cloud also has the potential to exacerbate the weaknesses of that decentralized approach, further balkanizing a network that needs standardized data for automation, security, and optimized performance.

Many network teams experience cloud remorse when they try to manage core network infrastructure across a complex web of hybrid environments. Ensuring the reliability of data pathways, maintaining visibility into user activity, and securing the enterprise can be a real challenge when DevOps and cloud teams are able to build or tear down infrastructure at will.

The Domain Name System (DNS), Dynamic Host Configuration Protocol (DHCP) and Internet Protocol Address Management (IPAM) are collectively referred to as DDI. DDI seems like an unlikely candidate to solve these challenges. Most network administrators see DDI as legacy infrastructure—something to be managed and controlled, not something to be leveraged. Yet as cloud deployments unfold, network teams quickly find that DDI is the linchpin of their success or failure.

The Army wants to manage, automate and secure its cloud environments without slowing down innovation. To do so, its decentralized, ad hoc DDI infrastructure must give way to a purpose-built, single source of truth that promotes consistent operations, increased automation and network security.

From the perspective of automation, a standardized approach to DDI would enable the Army to deploy capabilities to the warfighter faster while streamlining traditional network operations. In this proposed architecture, developers would be able to consume network resources on a self-service basis, while network teams would be able to focus on optimizing network infrastructure instead of drowning in a sea of service requests. This would result in more efficient and agile utilization of cloud resources, empowering DevOps teams to take full advantage of the cloud.

From the perspective of security, a unified DDI infrastructure would provide unparalleled visibility into activity in cloud environments. Network and security teams would be able to track and monitor cloud resources as they are spun up and spun down. This improved understanding of what is active in the cloud would help to ease concerns around cyber risk without the need for heavy-handed security tactics.

This presentation will focus on the strategic opportunity to build the Army's cloud environment on a solid foundation of a standardized, flexible, resilient and secure DDI infrastructure. It will reveal how, by consolidating control of pathways for data and compute flowing through hybrid environments through automation, the Army will eliminate siloed services and empower their hybrid environments to operate at scale. It will demonstrate how integrated DDI will benefit the entire organization by laying the groundwork for automation, decreasing the cost of cloud management, and locking down critical data.

BIO: Ben Ball is the director of strategy at BlueCat Networks. A former foreign service officer and homeland security official, Ball creates compelling marketing material that spreads the word about the critical role of the Domain Name System (DNS) in network management and security.

Bringing Critical Mission Data to the Warfighter

Nathaniel Wells, Director, Cloud and Federal Alliances, Panzura • nbabayan@aequussg.com

ABSTRACT

Unstructured file storage is the fastest growing category of data due to duplication of files, the explosive growth of machine-generated data such as application log files, machine learning output, IoT telemetry or sensor data, 4K video and 3D imaging. Additionally, increasingly distributed workforce generating massive amounts of data from multiple platforms has forced enterprises to consider new storage paradigms such as cloud storage.

Yet, adopting the cloud as a storage tier comes with the complexity of integrating with existing IT environments, ensuring data security, and managing data in a multi/hybrid cloud world.

This is when Panzura's Freedom answers the requirement of a modern, globally-distributed enterprise accelerating its secure digital transformation. Built from the ground up to power enterprise-scale deployments in the cloud and allow collaborative global teams to seamlessly work with challenging file sizes while maintaining military-grade security, Panzura delivers unprecedented performance and scale across all unstructured data workloads. Panzura provides a multicloud file services platform that enables global block level deduplication, high performance tiered NAS, global file collaboration, active archiving, backup, and disaster recovery across all locations. Panzura Freedom architecture enables the enterprise to consolidate their unstructured data while eliminating islands of storage. The end result is the warfighter having low-latency access to critical mission data anywhere in the world while saving the government billions on IT storage spend.

BIO: An Army veteran, Nathaniel Wells leads Panzura's federal business. His objective is to provide government customers with capabilities that not only modernize federal IT but also contribute to overall mission.

Distributed Application and Data Management and Modernization Across Army Enterprise Ecosystems

Tom Culpepper, Enterprise Architect, IBM • tculpep@us.ibm.com

ABSTRACT

This talk will focus on hybrid cloud strategies and solutions for interoperability, workload portability and flexibility of open source software to support Army tactical and non-tactical IT environments and provide a roadmap to hybrid cloud and edge computing. The hybrid cloud abstracts everything above and below the operating system, including every environment and every application to provide consistent interaction without retooling, retraining, splitting management or sacrificing security. Discussion points will address how the hybrid cloud connects multiple computers, consolidates IT resources, scales out and quickly provisions new resources and moves workloads between environments. It also has a single, unified management tool and orchestrates processes. As part of the hybrid cloud we will present edge computing capabilities that bring consistency to applications and operations extending their workloads out to different physical locations. Discussion will include edge computing solutions that use automated provisioning, management, and orchestration to simplify operations and establish a common infrastructure across compute, storage and network requirements to support fast set up/tear down time. To minimize operational complexity in austere locations with limited IT staffing we will discuss the benefits of a centrally managed edge computing solution to allow the Army to: (1) centralize where they can, and (2) distribute when they must. We will chat about edge computing solutions that provide the capabilities to support easier maintenance of equipment with rapid and continuous integration, including automating and managing infrastructure from the current and future tactical and non-tactical infrastructures to the remote edge sites; provisioning, updating and maintaining firmware and software across the infrastructure; supporting hybrid workloads (e.g. virtual machines, containers, applications and microservices), and how to continually operate with reduced capabilities, even when Internet connectivity is not reliable. The foundation, virtual machines and high-performance computing workloads for scalability and consistent deployment models, containerized workloads, minimizing footprints, storage, messaging and communication, telecommunications operations at the edge and a common data fabric will be addressed. Discussion will lay out an architecture that combines cloud and edge computing to deliver capabilities that bring the Army's tactical and non-tactical environments into advanced, state-of-the-art and proven technologies. The architecture and technology underpinnings will highlight efficiencies of operations; provide the warfighter with real-time and near real-time data in connected and disconnected environments; leverage community-driven software that promotes stability, security and innovation; and provides a common data fabric that can be used to connect,

replicate and share data across high-availability and limited bandwidth environments, with additional capabilities for artificial intelligence (AI) and machine learning (ML) to assist in predictive and prescriptive insights not only for the environment but to assist the warfighter in making optimum decisions.

BIO: Tom Culpepper is presently serving IBM as an enterprise architect for the U.S. Army Account. He provides advice to clients to help them achieve their enterprise cloud strategy by rationalizing their application portfolio on how best to move their applications to cloud environments (public, private, hybrid) through transformations to cloud service provider services (IaaS, PaaS, and SaaS). As they move and build into the cloud, he collaborates with clients to adopt the best approaches for migrating and modernizing their applications, while helping them establish the right mix of management and monitoring capabilities to manage and support the maintenance and sustainment of their applications and the supporting infrastructure. Culpepper has supported *Fortune* 500 companies, including IBM, 3M Health Information Systems, Infinity Insurance and SunGard in their enterprise products and service offerings. Currently, he is supporting PEO EIS ALTESS in their Application Migration and Modernization to the Cloud and has supported PEO DHMS, AMC CIO/CTO/G6, the Office of the Deputy Chief of Staff CIO/CTO/G6 HQ AMC and LOGSA. He is a mentor and subject matter expert for blockchain, cloud, hybrid cloud and AI/ML.

Applying Data Trust-Listing to a Cloud and Data/Application Migration/ Enterprise/Hybrid Cloud Strategy

Steven Shanklin, Founder and CEO, White Cloud Security Inc. • ziggy@whitecloudsecurity.com

ABSTRACT

The U.S. Army's Modernize Cloud Strategy would be supported by White Cloud Security's Trust Lockdown (TL) Data Trust-Listing, which enforces data visibility, access, classification and protection at the endpoint, regardless of whether the data is shared in the cloud, in a data center or on a specific endpoint, and monitors and controls the creation, visibility, access and modification of data by users and automata such as AI and ML processes.

One of the central problems with data protection in a cloud ecosystem is controlling which users and process have visibility, access and modification privileges to the data. While network access between endpoints and data sources can be monitored, controlled and secured, traditional file control access mechanisms are difficult to implement and manage at the endpoint processing layer. As with the distribution and management of crypto keys, creating and managing data access privilege lists must be transparent to the disparate applications and users that need to access the data files. Furthermore, the creation and management of the data file access policies must be outside the control of an administrator who has root/supervisory privileges on the endpoints.

TL's Cyber-Metric Handprint File identification technology uniquely identifies each file based upon the file's own data content and prevents spoofing or manipulating of data file identification tags. It is always unique to each data file's, or file segment's, content.

TL is a kernel-level file filter driver on Windows and a Linux Security Module in Linux that communicates with a secure service in the cloud or in a data center appliance that contains the trust-listing policies that determine which software is allowed to run and which data files can be seen, accessed or modified by a software package or component. TL's proven Execution Control security agent only allows trusted executables, libraries and scripts to run on endpoints according to specific trust policies for the endpoint or user on that endpoint. The Data Trust-Listing extends the Execution Control Trust-Listing Framework to identify, monitor and control the creation of data files and modifications to them regardless of their location in the data ecosystem or whether the creator or user is a real user or an automata process.

Data attributes include, but are not limited to:

- Cyber-Metric Handprint Identifier unique to each data file or segment
- SHA-1, SHA-256, SHA-512, MD5, CRC32 and the data file or segment's length
- Creation time
- Last modification time
- Blockchain history
- Host identifier
- Host subgroup identifier
- App signature ID
- App compatibility profile list
- IP/MAC address of host
- Data type
- User and user's domain/group
- Classification of data
- Category of data
- Dirty or clean data attribute

TL works with both modern and legacy endpoints (from Windows 2000 and Windows Servers 2003) without changes to the legacy endpoints other than installing TL's endpoint agent. It is supported from Redhat/CentOS kernels 3.10 after adding and enabling the Linux Security Module to the Kernel (Ports to Debian and Raspbian in Q2 2020).

BIO: White Cloud Security was founded by cybersecurity professionals with a proven track record and over two decades of cybersecurity software development experience in leading-edge host and network intrusion detection, automated remediation and application whitelisting. White Cloud Security's previous companies were acquired by Cisco Systems (Wheelgroup, Psionic), TIS (Haystack Labs) and Lumension Security (Coretrace). White Cloud Security's Trust Lockdown is a zero trust app security framework that verifies the cyber-metric handprint identity of each executable, dynamic code library and script every time they try to run. It blocks everything else.

Embracing DevSecOps: A Changing Security Landscape for the U.S. Government

Derek Weeks, Vice President, Sonatype • dpratt@sonatype.com

ABSTRACT

The United States is facing a growing cybersecurity threat. While there is increasing awareness of, and conversations around, the need for a coordinated strategy to prevent, identify and respond to threats stemming from unmanaged software supply chains, too little is actively being changed.

For instance, one of the biggest threats comes from the contractors paid to support the federal government that are supposed to be helping protect its sophisticated systems. Too often these companies are inadvertently introducing vulnerabilities into the supply chain. This is due, at least in part, to long-held emphasis on cost and overall performance, rather than security protocols. While the former are important, as cyber security threats multiply daily, the short-term benefit of awarding contracts to the cheapest contractor may have profound long-term effects on national security.

Newly reintroduced legislation by Senator Warner on the need to build security into the Internet of Things (IoT) landscape, and the highly discussed report Deliver Uncompromised from The MITRE Corporation provide a roadmap for where the United States should be headed. This legislation also offers an opportunity for savvy contractors and agencies to get ahead by prioritizing security in their development process now.

The Deliver Uncompromised report urges the DOD, and by extension the entire U.S. government, to "lead by example and use its purchasing power and regulatory authority to move companies to work with DOD to enhance security through addressing threat, vulnerabilities and consequences of its capabilities and adapt to dynamic, constantly changing threats." Similarly, the Internet of Things Cybersecurity Improvement Act draft legislation outlines that a final bill would "include four separate areas: secure development, identity management, patching, and configuration management. Under the language in the bill, vendors selling IoT devices to federal agencies will have to meet the NIST standards for those areas."

What does that mean exactly? It means industry is on the cusp of an incredible shift in how proposals will be evaluated by the federal government. In fact, the DOD set a Q4 2018 goal of incorporating security into its awards criteria. It also means that both government software developers and external vendors will have to build security into development or risk losing their job.

One of the largest continued areas of mismanagement within the software supply chain continues to be open source components. They form the foundation of so many mission-critical applications in government but often have unknown pedigree, risky software licenses and known security vulnerabilities. These vulnerabilities especially represent a large and fast-expanding attack surface for adversaries.

How can users keep up? As part of this change, all contractors and government software developers will need to think critically and not only ask themselves "does the code have vulnerabilities," but "could it have vulnerabilities," and "how do we know either way?"

Sonatype vice president and DevOps advocate Derek Weeks argues that with the right tools and embedded security across the entire development process, many of the issues discussed above can be easily addressed, leaving the software supply chain secure so mindshare can be left for other critical national security issues.

He will offer tips such as:

- How to reduce attack surface area from OSS components by just knowing what's in the software.
- Why introducing security checkpoints can reduce rework and risk.
- How to easily get started with software supply chain security to ensure users are "delivering uncompromised."

BIO: Derek E. Weeks is a huge advocate of applying proven supply chain management principles into DevOps practices to improve efficiencies and sustain long-lasting competitive advantages. He currently serves as vice president and DevOps advocate at Sonatype, creators of the Nexus repository manager and the global leader in solutions for software supply chain automation.

Weeks is also the co-founder of All Day DevOps, an online community of 40,000 IT professionals, and the lead researcher behind the annual State of the Software Supply Chain report for the DevOps industry.

In 2018, Weeks was recognized by DevOps.com as the "Best DevOps Evangelist" for his work in the community.

Multi-Cloud Workload Security and Visibility

Martin Isaksen, Senior Architect, Cisco • marisaks@cisco.com

ABSTRACT

A new security mechanism is needed to monitor and record all ports, processes and protocol sessions in the application for its entire life cycle regardless of where the workload is migrated. Compressing the metadata and sending it to a big data repository will provide a security manifest of every session that tried to connect to the workload in an automatic and agile fashion across clouds with no user interaction required. With this full visibility, zero trust security automation can help organizations with cloud migrations by automating the creation of network security policies across diverse compute environments. Updates to global workload security policies can be updated in one click, leveraging business intent policy tags such as mission-level, geo-location or any other mission abstraction. Integrating security into a big data architecture provides the underlying infrastructure for unsupervised deep learning using clustering analytics to find indicators of compromise (IoC) across workloads in real time as well as long-term historical data.

BIO: Martin Isaksen has more than 30 years of DOD experience in applications, networking and software solutions. He is currently the senior architect for Cisco's federal defense business. In his prior role as Microsoft Federal DOD CTO, he led the Microsoft technology strategy for unified communications, hybrid cloud, business productivity, and mobility. Prior to that role, Isaksen worked at Nortel for 10 years as DOD chief architect, supporting several mission-critical systems for the federal government, including the NASA Space Shuttle IP Multicast Network and the first DOD-certified VoIP solution at Ft. Huachuca, Arizona. Prior to that, Isaksen worked in the federal government for 11 years across DOD and civilian agencies as a branch chief and architect, where he helped publish the first DOD websites at the Defense Technical Information Center. He holds a BS in computer science from the University of Maryland University College.

Protecting Your Data in Their Cloud

Brent Hansen, Chief Technology Officer, Thales Trusted Cyber Technologies •
mary.shiflett@thalestct.com

ABSTRACT

Thales Trusted Cyber Technologies' (TCT) mission is to protect the U.S. federal government's most vital data from the core to the cloud to the field. TCT provides an extensive portfolio to secure data at rest, data in use and data in motion in hybrid, multicloud environments. The company's capabilities include making data accessible, trusted and secure to meet compliance with the Army Data Strategy, map to recommendations in the DIB Zero Trust Architecture and NIST SP 800-53 security and privacy controls.

Many infrastructure, platform and software-as-a-service providers offer data-at-rest encryption capabilities with encryption keys managed by the service provider. Meanwhile, many industry or internal data protection mandates, as well as industry best practices as defined by the Cloud Security Alliance, require that keys be stored and managed remotely from the cloud service provider and the associated encryption operations. Providers can fulfill these requirements by offering Bring Your Own Key (BYOK) services to enable customer control of the keys used to encrypt their data. Customer key control allows for the separation, creation, ownership and control, including revocation, of encryption keys or tenant secrets used to create them.

Leveraging cloud provider BYOK API's, the CipherTrust Cloud Key Manager (CCKM) reduces key management complexity and operational costs by giving customers life-cycle control of encryption keys with centralized management and visibility. Take control of cloud encryption keys to leverage the value of Bring Your Own Key services with full life-cycle cloud encryption key management; comply with the most stringent data protection mandates with up to FIPS 140–2 Level 3 validated key origination and storage; gain higher IT efficiency with centralized key management across multiple cloud environments, automated key rotation and key expiration management.

The requirement to protect sensitive data across Infrastructure, Platform and Software-as-a-Service (IaaS, PaaS and SaaS) has resulted in broader cloud provider encryption offerings. Meanwhile, the Cloud Security Alliance and industry analysts state that encryption keys should be held by customers. The challenges of holding keys grow with up to hundreds of master keys per subscription to be secured and managed across multiple clouds. There is also the imperative of knowing how, when and by whom encryption keys are used. CCKM provides comprehensive key life-cycle management to fulfill requirements for safe, comprehensive key management across multiple clouds.

With the requirement for key security mechanisms such as safe storage of cloud backup keys, CCKM acts as a key escrow for supported clouds and allows for full key metadata control both during upload and for keys in use. CCKM offers multiple capabilities in support of enhanced IT efficiency: centralized

key management, automated key rotation and federated login. CCKM simplifies the need to hold and manage encryption keys for cloud services, a critical solution for fulfilling industry and organizational data protection mandates. Additional Thales TCT multicloud security products, including Bring Your Own Advanced Encryption, all with centralized, FIPS validated key management, enable individuals to encrypt and control cloud storage to reduce the chance of sensitive data being leaked.

BIO: Brent Hansen is Thales TCT's chief technology officer. Hansen leads Thales TCT's sales engineering organization and spends the majority of his time evangelizing and strategizing on data-centric approaches for federal agencies looking to avert being the next victim of a data breach.

Hansen brings more than 19 years of IT experience in data and enterprise architecture, data warehousing, big data, and business intelligence. He is an industry expert in data encryption and tokenization. He leads teams that architect security strategies to secure and protect sensitive data for both federal government and large commercial enterprises across the globe.

Visible, Accessible, Understandable, Trusted, Interoperable and Secure (VAUTIS) Data Einstein's Way

Christopher Gunderson, Adaptive Acquisition Architect, Frontier Technology Inc. •

cgunderson@fti-net.com

ABSTRACT

Joint Vision (JV) 2020, published in 2000, predicted "globally interconnected, end-to-end … information capabilities, associated processes, and people to … provide information on demand to warfighters, policy makers..." Policy mandates to make "need-to-share" (NtS) co-equal to "need-to-know" (NtK) soon followed. Yet, the Army's latest data strategy demonstrates that JV2020 has not materialized, and that risk-embracing security policies have not catalyzed the intended behavior. Einstein would suggest that finally succeeding will require fundamentally departing from failed paradigms. For example:

- Focus government R&D on virtual security services to interface with COTS cloud and AI.
- Employ risk-reward analysis—by humans and machines—to establish N-t-S policies for each data element.
- Employ Model-Based Systems Engineering and virtual technology to build N-t-S cross-domain-service architecture.
- Employ semantic technology and Model-Driven Architecture to compose open systems from legacy stovepipes.
- Adjust compliance requirements to include demonstration of N-t-S decision logic.
- Incentivize value over compliance in acquisition strategies for information systems.
- Employ Social Network Analysis (SNA) to discover and empower innately innovative human resources.
- Engage successful innovation coaches from outside the defense sector to:
  - » Mentor-the-mentors.
  - » Write the doctrine for a datacentric workforce.

FTI has leveraged hundreds of Small Business Innovative Research grants to evolve a defense decision support framework that delivers shovel-ready designs aligned with defense acquisition compliance

requirements. FTI partners with expert engineering firms in a virtuous cycle of build-a-little/analyze-a-little/adjust fire.The company will leverage recent DOD work re: adaptive acquisition; tactical, software-defined, open-systems; cloud-enabled, authoritative data warehouses; software-defined cross domain solutions; organizational change; cyber sensor nets. The company's approach applies equally well to building tactical and non-tactical infrastructure. FTI proposes:

- Human-System Engineering. Discover, nurture and incentivize innate innovators.
- Apply SNA to identify hidden heroes.
- Mentor the mentors.
- Incentivize risk-taking.

Adaptive acquisition. Measure and incentivize value; treat compliance as a boundary condition. Employ:

- Value-centric measures.
- Automated, metrics-based market research and trade-off analysis.
- Adaptive contracting (e.g. "Other Transactions," SBIR phase III).
- Cyber-authorize Agile + DevSecOPS.

Virtually Assured Cloud Information Sharing Services (VACISS). Focus defense research on virtual security services to embed in COTS cloud. Convince authorities to accept assured policy-based logical separation as criteria for authorization.

- Leverage cloud's virtual machine paradigm to provision logical separation.
- Apply high-assurance standards (e.g. ARINC 653) to harden virtual separation.
- Assure that authentication and authorization are independent.
- Develop machine-readable, high-assurance, N-t-K and N-t-S policy language based on identity, dynamic attributes and dynamic conditions.
- Compose N-t-K/N-t-S policy enforcement process from independent, high-assurance services for authentication and authorization.
- Embed cybersecurity threat-tracking into system-performance-monitoring services.
- Work with cross domain authorizing officials to approve VACISS as a software-defined guard.

Semantic Architecture Conversion Framework. Establish semantically enhanced, multitier architecture to compose open systems from legacy capabilities.

- Apply NSA's open source Ghidra to capture legacy software features and specify integrations between layers.
- Use xtUML to select templates for desired architectural patterns.

- Modify existing UI to intake parameters for optimizing templates.
- Use Open DDS to combine software.
- Build a GitHub Continuous Integration/Continuous Deployment environment.

BIO: Capt. Chris Gunderson, USN (Ret.), joined FTI in July of 2017 where he invented open system applications of traditional FTI decision support tools to streamline information system acquisition—including digital transformation via cloud migration. Prior to joining FTI, he served as a member of the faculty of the Air Force Institute of Technology and the Naval Postgraduate School, specializing in transformational innovation, adaptive acquisition and open system acquisition.

McAfee Cloud and Application Migration Solutions

Nick Graham, McAfee Cloud and Application Migration/Enterprise and Hybrid Cloud Strategy, McAfee • nick_graham@mcafee.com

ABSTRACT

McAfee CASB Solutions for Enterprise and Hybrid Clouds

McAfee discovers all cloud services in use by employees both on- and off-network, including thousands of cloud services uncategorized by firewalls and web proxies. The solution's usage analytics summarize cloud usage in aggregate and at the department and user level with traffic patterns, access count and trends over time, enabling IT to securely enable cloud services that drive productivity and growth.

McAfee MVISION Cloud protects data and stops threats in the cloud across SaaS, PaaS and IaaS from a single, cloud-native enforcement point. Gain visibility into all cloud usage and data through various analytics such as content, collaboration, access and cloud usage then apply unified policies to all cloud services across data at rest and in transit.

McAfee CASB Connect Inline for Custom Applications

Helps organizations securely accelerate their business by providing total control over data and user activity in custom-built cloud services

Key use cases would be:

- Enforce sensitive data policies for cloud services. Data loss prevention will prevent regulated data from being stored in the custom application. McAfee's MVISION Cloud content analytics engine will discover sensitive data created in or uploaded to a custom application.
- Perform forensic investigations with full context. Activity Monitoring will provide visibility into the custom application usage and accelerate post-incident forensic investigations by capturing a comprehensive audit trail of all activity within the application.
- Detect and correct user threats. User Behavior Analytics applies data science and machine learning to automatically build models of typical user behavior, which identifies behavior that may be indicative of a threat.

McAfee Container Security Related to the Devsecops Conversation Surrounding Application Migration

MVISION Cloud Container Security provides a unified cloud security platform with container-optimized strategies for securing dynamic and ever-changing container workloads and the infrastructure on which they depend.

The MVISION Cloud for Containers provides:

- Vulnerability assessment for container components—Evaluate the code embedded in containers at build time and periodically to ensure that known risks are exposed or mitigated to reduce the opportunities malicious actors have to land in a container workload.
- Cloud security posture management for container infrastructure and orchestration systems such as Kubernetes—Ensure that the environment's configuration is not a source of risk—Ensure that the configuration of the environment does not drift over time, exposing unintentional risk.
- NanoSegmentation for inter-container communication—Zero Trust: Always Verify Never Trust. Discover and monitor the behavior of network communications between container processes in a way that can deal with the ephemeral nature of containers and not rely on external factors such as an IP address—Detect abnormal communications and notify or block based on user preference—Detect changes in communication patterns between versions of containers as the application evolves over time—Leverage known good configurations as a way to secure workloads, as opposed to keeping up with known bad.

BIO: Nick Graham is a cloud security architect for McAfee where he is focused on enabling McAfee's government customers to use technology that helps them provide secure, efficient and effective national security missions. Graham has been working in government and commercial technology for more than 20 years. After his service as a SSGT in the U.S. Air Force during Operation Desert Storm and Operation Desert Shield, Graham has helped many government agencies address their security requirements.

Building the Army's Modern Information Architecture To Drive Innovation

Sherry Bennett, Chief Data Scientist, DLT - A Tech Data Company • sherry.bennett@dlt.com

ABSTRACT

Building the Army's modern information architecture requires critical attention to the confluence of operational processes and the information and applications used to advise and drive mission activities. Cloudera's Data Platform (CDP) and DLT's Data Innovation Software Factory (DISF) framework are singularly focused on providing the Army a solution and for re-engineering key business processes and application modernization, and building the foundation for a new information architecture to support any workloads, including ML and AI. From pilot to production, in an iterative, progressive manner the proposed solution and framework empowers the Army to prioritize innovation while balancing critical, tactical, day-to-day mission operations.

However, the process to drive innovation requires a framework that is flexible to support the development of a changing information architecture while simultaneously supporting existing environments. CDP and DISF deliver an agile and flexible environment that can adapt legacy software to meet changing operational needs. The following areas are the key determinants to drive the Army's modernization efforts.

Multifunction analytics: A robust Army information architecture, able to support the development and deployment of dynamic, evolving data solutions (IoT, AI, ML, etc.), needs to accommodate not only data at rest, but continuous IoT feeds and data produced by various systems/stakeholders across the network. As innovation and modernization are driving the Army's data plan, its information architecture must accommodate and manage a variety of sources, including, voice, video, imagery and telemetry, etc. (NOTE: CDP and DISF are equipped to handle many of the requirements articulated in the other solution areas pertaining to enablers, visibility, accessibility, understandable, trust, interoperability and security.)

Hybrid and multicloud: CDP and DISF are an elastic cloud experience with no silos and no lock-in. CDP and DISF provide an intelligent way to migrate applications and required workloads, rationalized on mission resources and priorities.

Unified security, governance and metadata: CDP supports strict enterprise data privacy, governance, migration and metadata management across all environments. Cloudera's data platform can process

workloads from multiple endpoints, including the tactical edge, while predicting key outcomes and utilizing machine learning, while leveraging a hybrid cloud environment to afford agility and elasticity as required. This can all be accomplished on an open platform where security and governance are applied at each stage of development in the data innovations framework. Moreover, the solution can deliver an improvement on readiness, threat detection and accommodate a "zero-trust" approach to cybersecurity.

Open platform. CDP and DISF align with Army's desire for a process and environment that ensures easy integration, which encompasses a variety of tools, apps and infrastructure, allowing users and operators alike to manage data in any cloud environment whatever the requirements, SLAs and regulations. CDP and DISF were built with an open source philosophy and enterprise-class, data-driven cloud architecture to promote seamless migrations without the risk of vendor lock-in. Cloudera's commitment to open source ensures easy integration on a platform that encompasses a variety of tools, apps and infrastructure, with open, backward-compatible APIs.

BIO: Sherry Bennett serves as chief data scientist for DLT Solutions. As the chief data scientist for DLT, she is responsible for the vision and oversight of DLT's Data Innovation technology stack, which accelerates the capabilities and development of analytical and AI-enabled organizations within the public sector.

Prior to her tenure at DLT, Bennett spent more than 15 years in higher education, where, in her roles as chief data officer and data scientist, she was responsible for establishing global data science teams and decision support services for universities. One of her most notable roles before joining DLT was serving as divisional vice president at Laureate International Universities where she led the creation and management of a global business intelligence and data science team, serving the information needs for a portfolio of universities across North and South America, Europe and Asia. Prior to joining DLT, Bennett was serving as the chief data officer at University of Maryland Global Campus, the largest online public university in the United States.

Bennett began her career in academia, serving as an assistant professor at Rice University, teaching statistics and international political economy courses to undergraduate and graduate students, as well as managing a quantitative research portfolio. Bennett has a Ph.D. and BA in political science from Michigan State University.

Secure Software Factory

Rick Stewart, Chief Software Technologist, DLT Solutions • rick.stewart@dlt.com

ABSTRACT

DLT offers a holistic approach to application modernization using innovative automation tools and platforms that will allow the Army to achieve elasticity, resiliency, broad access, efficiency, secure computing platforms, data standardization and compliance tools. DLT's Secure Software Factory (SSF) provides a framework of innovative, best of breed, platforms and tools that allow the Army to plan, build, test, deploy, release and monitor deployed services in a continuous manner that injects security controls early and often in the life cycle and monitors its health once in production.

The DLT SSF also provides the right traceability to measure effectiveness and conformance to requirements. In addition, the DLT SSF detects any defects or vulnerabilities in any stage of its life cycle so it can be remedied quickly. Activity artifacts also are captured continuously so that appropriate stakeholders can review past, existing, or future release enhancements, which can be assessed to ensure mission goals are on target or can be adjusted in an agile manner.

Using DLT's SSF as a platform for their digital transformation, the Army can realize its strategy to modernize and migrate thousands of applications and data to the cloud by providing innovative and integrated tools that protect data through the use of solutions that generate efficiency through automation.

DLT's SSF is built on a platform that allows for installation on-premises, a private or public cloud of their choice, and most importantly, the flexibility to have a hybrid cloud environment to run their workloads where they are most appropriate to run without changing implementation scripts per cloud provider.

The DLT SSF provides the Army an enterprise cloud and data ecosystems that are AI and ML ready, hybrid, protects Army data, increases its lethality at each echelon, and provides a flexible set of solutions for modernizing its workloads. By supporting agile methodology planning and a flexible cloud framework, the DLT SSF accelerates legacy software transformation to quickly meet changing operational environments, increase readiness, and improve cybersecurity.

From idea to production, DLT's SSF allows software development, IT operations, and security teams to utilize a single platform to collaborate, communicate, develop and deliver their workloads quickly at the speed their stakeholders desire, and perhaps expect; but with guard rails that assist them with rectifying issues early in the development process before they become tragedies in production. Utilizing technologies that employ artificial intelligence and machine learning, DLT's SSF can shorten cycle times by focusing on the most important quality controls without trying to gain speed by skipping important testing or scanning activities.

Most importantly, DLT's SSF provides observability to metrics related to all aspects of an application's life cycle so that continuous improvement activities can be implemented by the Army's IT teams. As more applications of various languages, or micro-service architectures, are implemented; the Army will have visibility into an application's life cycle as they are each on their own release cycle and should be deployed on need, not dependent on one another.

BIO: With more than 30 years of diverse experience in the IT industry, progressing through technical and leadership roles in the telecommunications, mobile entertainment, federal government and manufacturing industries, Rick Stewart has a unique perspective of providing DevOps-based solutions. His duties at DLT are in support of DLT to promote the sharing and adoption of best of breed technologies with within the federal government and state, local and higher education (SLED) enterprise accounts.

Data Protection Across Hybrid Environments

Rick Bueno, CEO and Founder, Cyber Reliant Corp. • rabueno@cyberreliant.com

ABSTRACT

The Army is challenged to protect its data in highly distributed environments. Army data owners are justifiably concerned to migrate their data to cloud due to real security risks. Cyber Reliant can address and mitigate those risks by applying the data protection directly to the data itself regardless of the CSP. Cyber Reliant Data can directly enable the Army to establish an enterprise cloud and data ecosystem that is hybrid and protects the data at each echelon and generates reinvestment opportunities for modernization. Built and designed to protect the Army's most sensitive data in the most austere cyber environment, Cyber Reliant applies data protection directly to the data itself. Traditional security relies on protecting the perimeter and on data encryption. Perimeter-based data protection strategies are important but are not enough as the perimeter has been and will continue to be breached. Encryption is also an important data protection strategy, but it is not enough, as encryption can also be broken with relative ease. The key to successful data protection is to apply the data protection to the data itself in a manner so difficult and complex that not even the most sophisticated attackers would know how to break through. Cyber Reliant products were designed and built by offensive information operations engineers who designed a data protection product to counter the most sophisticated state-level sponsored attacks. Cyber Reliant incorporates not just one technique but a series of innovative and specialized techniques to counter any offensive information operations attempt to data exploitation and exfiltration. Cyber Reliant has implemented advanced key management, disassociation, encryption, shredding, embedding and dispersion techniques that create a framework of data protection that has been adjudicated as Information-Theoretic Secure by the NSA. Information-Theoretic Secure implies a quantum resistance data protection methodology.

BIO: Rick Bueno, president and CEO of Cyber Reliant, is a U.S. veteran, entrepreneur and experienced visionary with a deep 35-year history of developing and executing cybersecurity strategic initiatives and solutions in the commercial, defense, intelligence, and special operations community. Prior to founding Cyber Reliant in 2010, Bueno served as the National Security Agency Information Assurance Directorate AD Afghan Mission Manager/NATO Special Operations Forces, for the ISR Task Force. In this role, he worked closely with NSA, USDI, CENTCOM, and others to develop secure communications strategies in support of NATO and Special Forces missions. Prior to his role at ISR TF, Bueno served as the chief strategic architect for the Director of National Intelligence (DNI). As a member of the director's action group, he was responsible for early establishment of the Intelligence Community's policies and processes for the information sharing environment and other activities within the DNI CIO scope.

Providing Enterprise Hybrid Cloud Management and Cloud Data Expertise

Bill Kodzis, Senior Vice President, Applied Insight • bkodzis@applied-insight.com

ABSTRACT

Within Altitude, users natively access either the Amazon Web Services (AWS) console or the Microsoft Azure portal via managed identity federation. Altitude utilizes role-based access controls (RBAC) to prescribe the associated privileges to define the resources users may access within the cloud. A unique benefit of this approach is that users receive native access to their CSP resources. Altitude is not a cloud broker. This is done without compromising security by defining policies associated with a user's federated identity, ensuring that users can only provision approved resources configured in a manner that aligns with security and compliance mandates.

To enable transitions between vendors, as well as collaboration between resources hosted in different CSPs, Altitude employs a vendor-agnostic network mesh that spans CSPs and on-premises environments, facilitating a true hybrid, multicloud environment. The networking implementation within Altitude utilizes an enterprise-scale IP management strategy, allowing virtual networks to be designed in a manner that achieves network-layer isolation, without precluding the ability for inter-account connectivity. By designing an IP schema that ensures unique address blocks across multiple CSP accounts/subscriptions, CSP vendors, and on-premises facilities, routing can be strategically implemented to provide granular network-layer control. This ensures efficient elasticity and resiliency without sacrificing security or imposing vendor lock-in.

Customers of Altitude require a framework that is able to quickly adapt to changing requirements and foster agility and operational readiness. To effectively manage cloud and data ecosystems in a hybrid and multicloud fashion at an enterprise scale, the company believes that the use of automation is a cornerstone to the strategy. Altitude provides an environment that facilitates rapid scalability and elasticity, allowing a hybrid or multicloud environment to dynamically scale in response to operational needs. By leveraging infrastructure-as-code principles, Altitude supports the automated provisioning of AWS accounts and Azure subscriptions with a consistent, scalable networking, authentication, security and governance approach.

Applied Insight (AI) has also implemented customized Altitude solutions that perform the collection, aggregation, enrichment, storage and transportation of data throughout classified, unclassified, cloud-based, hybrid, on-premises, virtualized and multicloud environments. Leveraging best-of-breed

technologies and cloud-native tools, the company can perform flexible and automated data gathering and storage for IC and DOD customers.

For an IC customer, AI technical teams have been the first organization to implement DataWave within both AWS and Azure CSPs. DataWave is a Java-based ingest and query framework sponsored by the NSA that leverages Apache Accumulo to provide fast, secure access to data sets. It enables a wide array of features including PKI integration, multitenant architectures, and providing data fusion across both structured and unstructured datasets. The company's engineers were the first to implement this data aggregation software within a multicloud environment and were also crucial to the development of the DataWave software itself, contributing to its open source codebase.

AI has previously provided data storage, application migration and enterprise hybrid cloud solutions for customers within the IC, DOD and federal/civilian spaces and would be honored to have the opportunity to do the same for the U.S. Army.

BIO: Applied Insight has worked closely with customers across the defense and Intelligence Community to overcome such hybrid, CSP-agnostic challenges. What the company has learned through those engagements has been integrated into Altitude, an automated cloud management platform that is currently deployed for thousands of users in government and industry. Altitude facilitates end-user software development, data analytics and machine learning at scale within the cloud. This is accomplished by providing end-users with federated access to the Altitude platform, exposing native cloud services and providing an uncompromised end-user experience.

U.S. Army Data Plan: Cohesity Modernized Cloud

Steve Grewal, Federal CTO, Cohesity • jtogher@cohesity.com

ABSTRACT

Cohesity is a software-defined data management company built by the Google File System chief architect and Nutanix founder Dr. Mohit Aron. Cohesity utilizes the principles of cluster computing combined with an infinite (Web-Scale) file system to bring unmatched data visibility, manageability, understandability and transportability to any mission or requirement.

Cohesity DataPlatform is a software-defined data management solution for sensitive data and apps. Native integration with public clouds helps users take advantage of their scalability and cost effectiveness for multiple-use cases from backup to disaster recovery, all while keeping control of data and preventing vendor lock-in. Cohesity natively integrates with FedRAMP-certified cloud environments including Amazon Web Services, Microsoft Azure, Google Cloud Platform and Cohesity-Powered CSP to extend an on-premises data protection solution to proven public cloud services.

Cohesity's vision of any-to-any data portability enables all echelons to potentially access mission-critical data and applications. This capability is multicloud, and extends into the traditional data center and to the edge for tactical data settings. The same software runs in all locations, which gives system engineers and operators a short ramp to acquire necessary design and operational knowledge and skills. Central management of all data is provided by ML/AI-based Helios "manager of managers" system. With network connectivity to the edge, core and cloud, all data repositories can be quickly inventoried and managed. Helios can manage what apps are running on each data cluster. Helios can monitor key Cohesity infrastructure telemetry to not only manage data systems' health but also assist in detecting and proactively addressing cybersecurity incursions.

Helios manages data flows from cluster to cluster through Cohesity's replication engine. The advanced engine uses Cohesity's deduplication and compression engines to ensure that only unique, changed bytes are sent over the wire. This efficiency is required in today's MDO.

Cohesity's approach to managing data en masse hinges on key technologies linked by a hybrid cloud data store. The underlying software defined storage system is SpanFS. Individual instances of SpanFS reside in virtual, physical or cloud-based data centers. Data is shared between data centers using Cohesity's replication, archiving and tiering algorithms. As SpanFS and the related data mobility processes share a common deduplication and compression capability, all data motion over the wire is super-efficient as only unique, changed data segments must be transported across distant network

links. This distinct combination of tools is the core construct for Cohesity's DataPlatform, which is a fully functional data fabric.

On top of this data fabric, Cohesity has developed a unique set of capabilities to keep servers, databases, file systems and applications in a state of readiness for quick deployments. The Cohesity architecture stages container-based, distributed applications side by side the data on the storage nodes. A combination of third-party and Cohesity-developed application give data managers a full suite of tools to manage the data as needed for mission requirements. Additionally, new apps for emerging requirements can easily be developed to help minimize capability gaps.

BIO: Steve Grewal is Cohesity's federal CTO. Grewal is the former DCIO/SES for General Services Administration (GSA), and former CIO/SES U.S. Department of Education. Harvard alumni | Tech/Cyber Executive | Board Member | Investor | Strategist.

Understanding Data for Cloud Migration via Data Trust and ML

Jim Evans, Federal Account Executive, Varonis Public Sector • jevans@varonis.com

ABSTRACT

Varonis has a large track record of aiding customers in the lead-up to, transition to, and end state environment of hybridized cloud environment. The Varonis Data Security Platform can assist organizations with migration projects by answering key questions around relevant data, such as: What is the sensitivity of the data (classification, PII, PHI, etc.)? Is it stale? Can it be removed before cloud migration? These data insights empower stakeholders in the lead-up to a cloud migration by ensuring that an organization is only moving the necessary data assets to the cloud environment.

In the transition to the cloud, the Varonis Data Security platform can help ensure that data security is paramount throughout the process. Data processes can be enforced to maintain that permissions on-prem are maintained in the cloud. Also, especially important in cloud environments is that all access is audited no matter where the data lies. If users move data to/from the cloud, all actions are seen in one pane of glass, giving a comprehensive understanding of what is occurring no matter where users are accessing their data. These actions are all added to a user's user behavior profile, which utilizes machine learning (ML) to ensure that users are accessing data appropriately not only to their past behavior but also to the behavior of their peers. In addition, all other metadata streams that Varonis would collect on premises are also available in the cloud. Entities can tag and identify sensitive content, assign and edit permissions, lock down data and ultimately make sure users are accessing it appropriately. Varonis is uniquely poised to not only prepare large enterprises for the move to the cloud, but also secure the environment once it gets there.

BIO: Jim Evans leads the Army team at Varonis Public Sector and has supported the American warfighter for 33 years.

WHAT IS AFCEA?

AFCEA is a member-based, non-profit association for professionals that provides highly sought-after thought leadership, engagement and networking opportunities. The association focuses on cyber, command, control, communications, computers and intelligence to address national and international security challenges.

The association has more than 30,000 individual members, over 130 chapters and more than 1,600 corporate members. For more information, visit www.afcea.org.