Analytic Technology™ Industry Roundtable

Understanding the Impact of the Analysis Exchange Model 1.0

Government agencies of all sizes rely upon analytic tools to make the best decisions about everything from evaluating travel expenses to identifying potential terrorists. Yet as essential as these tools are to the agencies' missions, the unique software from different analytic technology vendors can often present a time-consuming, expensive, and incomplete picture that can slow mission success. To address this, the Analytic Technology Industry Roundtable released the Analysis Exchange 1.0 to help government analysts share information across numerous analytic tools from different manufacturers—and even other agencies. This paper gives an overview of the Analysis Exchange, as well as a quick look at three key use cases: ***Fraud, Waste, and Abuse****,* ***Threat Assessment****, and* ***Cyber Forensics****. The Analysis Exchange 1.0 is an open source technology, which will be made available via GitHub in November 2017.*

Seeing Only Part of the Picture

Government information analysts like knowing "what the other hand is doing." The government invests billions of dollars each year in analytic and analysis technologies to help them gather relevant data to assess matters ranging from the commonplace to national security. Multiple analytic tools are often required to address the myriad facets of analytic requirements. No single tool acts as the ubiquitous solution that answers all analytic questions; so analytic organizations often use multiple tools to investigate different aspects of the same problem. Although each analytic tool may prove successful for its intended purpose, varying technologies from different vendors—and even purpose-built government systems—form silos. As a result, this may be

providing only a partial picture of what is actually available within the data, limited to insights accessible with only a single tool. As a result, bad actors of all kinds can slip past scrutiny.

For instance, if you were analyzing your agency's travel expenses, you would want to know if a contractor who claimed fifty miles of mileage reimbursement actually only needed to drive twenty miles to reach their destination. Such an anomaly would be easy to see if you could quickly link travel vouchers against mapping software and automate a non-compliance alert mechanism prior to reimbursement. But the reality is every government agency uses analysis and analytic technologies from a variety of companies using vendor software and that can make it very difficult to pull together relevant information in a comprehensive manner. To bridge these technology silos, government analysts must frequently send requests via email for information to be converted into a readable format and then wait—sometimes for weeks—for it to become available. As frustrating as that scenario is from an efficiency standpoint, consider the ramifications of the same kind of delay in identifying a terrorist threat.

It's a Lock: Government and Industry Perspectives about Vendor Technologies

While it may seem like the government and the analytics technology industry would have opposing ideas about vendor-unique technologies, the fact is they are both quite concerned about being locked in—or out. Government agencies, with their mandates to be efficient and cost effective, do not want to be locked into using specific vendors. They want the flexibility to take advantage of better pricing, features, and improved technologies. However, the government acquisition process requires substantial time and effort. For instance, if an agency purchases five new analytic technologies from different vendors, then it is unlikely that they will be willing to open up the bidding process for another vendor in the short term if it means dealing with vendor technical incompatibilities.

Similarly, analytics industry technology companies do not want to be locked out of opportunities—possibly for years—because of interoperability issues. This is particularly

relevant in the analytics industry, where technology improvements come along quickly, and the company that may have lagged behind six months ago releases a new product that is clearly the outstanding choice. In addition, they do not want to go to the expense and time of developing adapters to be compatible with every other possible vendor in their customer's environment.

Government agencies do have an alternative: create a bespoke solution based on a single data format / data store and build or customize tools to act natively against this data. This approach eliminates the data silo, ensures that all tools see all data and all insights, and that analysts have access to the essential capabilities. However, this approach requires significant development and customization, typically with a higher cost and longer delivery time. This approach can also be less adaptable: the introduction of new requirements or a new tool requires more customization. Organizations with large budgets and mission stability may find the benefit of this approach compelling, but this approach is not practical for many other organizations with limited budgets or rapidly changing missions.

When the Analytic Technology Industry Roundtable was formed in 2015, its members decided to address the challenge of loosely coupling vendor software, with all its negative implications. Roundtable founding membership—comprised of fifteen companies, advised by government and hosted by The MITRE Corporation—designed the architecture and technology for the Analysis Exchange (AE) as their first major project. Since its formation, the Roundtable has doubled in size and includes many more industry teams that are part of the consortium and the working groups.

The Components of the Analysis Exchange Model 1.0

The Analysis Exchange Model 1.0's two main components are its architecture and its ontology—defined as "a set of concepts and categories in a subject area or domain that shows their properties and the relations between them." The AE's architecture is a hub that brokers analytic artifacts, knowledge, and results between analytics and analyst customers. The architecture's principal components include the Collaboration Services, the Transfer Service, and the

Knowledge Store. The Collaboration Services includes the software that drives and manages the workflow of analytics interacting with the AE. The Transfer Service includes the software that moves results between the AE and external providers or consumers, and the key subcomponents of this are the adapters, which convert external formats and content into a form that matches the AE's ontology. The Knowledge Store is where shared results that have been adapted are held within the AE.

The AE's ontology defines the content and format of results held within the AE, which are drawn and adapted from various sources. This ontology was designed to have a wide coverage and be extendable, but also provide the specific support needed for the use cases exhibited in the Analysis Exchange Model 1.0. In the case of both the architectural components and the ontology, an overarching goal was to provide examples that could demonstrate repeatability for any number of use cases that would draw on multiple analytic sources from industry.

Benefits of the Analysis Exchange Across Government and Industry

The AE was designed to help government and industry achieve common goals and efficiencies. Government analysts can be assured that any of the ever-growing number of vendors participating in the AE can share information across multiple platforms. And participating vendors know they can continue to build upon their competitive advantages through vendor-unique features, while being responsive to their customers by sharing information with other vendors' systems. Due to the "plug n' play" nature of the Analysis Exchange Model 1.0, vendors are better able to compete for acquisitions more efficiently using fewer resources. This allows industry to build better, more cost-effective solutions and provides the government with access to more responsive, cost-saving tools to achieve their missions.

The benefits of the Analysis Exchange include:

- Using a more interoperable, consistent, and repeatable model allow vendors to compete for acquisitions more efficiently, saving time and money.

- Products and systems can become more effective at meeting government needs and requirements with less investment of resources.
- The AE can break down the data silos that multiple analytic tools typically create.
- Member companies can adopt consistent standards more quickly, saving them and the government time and money.
- The AE can develop a set of common standards and protocols for the analysis and analytics community.
- The AE can create enriched products for government use based on the results of different companies' products.

Three Use Cases

While the AE can be designed to support a virtually unlimited array of analytic combinations, the Roundtable is initially focusing on three key use cases that they believe will be particularly relevant to government analysts. In each use case, the AE has developed an ontology that will help analysts mine their existing rich sources of data. From this data, the AE will help analysts gain a more complete picture—and possibly reveal anomalies and discrepancies—that will help them identify bad actors at work.

Use Case – Fraud, Waste, and Abuse

Many thousands of government employees, contractors, and citizens receive reimbursements for travel from federal agencies each year. While the majority of people record their travel expenses honestly, it is an area that is ripe for fraud, waste, and abuse. By reporting fraudulent mileage, hours worked, out-of-pocket expenses, and other scams, bad actors of all kinds steal millions from the federal government. Oftentimes, the information an analyst/data scientist would need to catch these potential offenders already exists within different analytic systems—but not in a coordinated form.

With the AE, analysts can ask their IT departments to pool the information from different analytic systems into a common ontology. For instance, the AE provides a consistent structure for information from different sources relevant to a travel voucher, which data scientist/analyst can use to identify outliers/anomalies that require further investigation. Was there a requested reimbursement for a patient cancelled appointment, or a no-show appointment? Or did a beneficiary travel out of town for a commodity that was readily available just miles from his office? The AE empowers analysts to use all the tools at their disposal to gain a clear picture and identify potential risks.

Use Case – Threat Assessment

Threat assessments are a proven approach to lower the risk of targeted violence by proactively identifying, assessing and mitigating threats, yet the resource intensive methods utilized can make the process prohibitive in both time and money. Analytic capabilities are needed that reduce the costs of doing complex data exploration and increase the effectiveness in identifying trends and patterns of targeted violent events over time and space. These capabilities can help government analysts more effectively identify and assess threats in immediate time to control and neutralize specific threat actors.

Conflict events do not occur in a vacuum, and many other contributing factors can happen simultaneously. Even after complex geopolitical events unfold, it may be difficult to gather and visualize a complete picture of the full context. For instance, when a terrorist incident occurs, it is important to gather not only the facts of the event, but also other contemporaneous related information. Did a terrorist group have other simultaneous activity? Is there a larger pattern that recurs throughout typical attacks that can be discovered? The AE allows an analyst to collect information from various data and analytic sources that provide enrichment and depict them all together.

Use Case – Cyber Forensics

Cyber security experts are challenged to stay ahead of threats posed by hackers who continuously outwit common security systems and the individuals responsible for security. Sophisticated developers who spent years honing their hacking techniques are now commoditizing their expertise and selling it in a kit on the dark web. Networks where millions of dollars are invested in security are being hacked by bad actors using a $500 dollar laptop and some innovative techniques. The massive amounts of data and array of security tools of modern networks may provide the indicators and solutions to a cyber-attack; however, the complexity of existing solutions makes it difficult to discover answers in real time.

Cyber attacks are numerous, sudden, and require a speedy reaction time. Relevant data for forensically analyzing an attack may have to be drawn from different sources. When an attack occurs, an analyst principally wants to know who perpetrated the attack and the extent of the damage. Answers to these questions often come from multiple analytic sources. Given the brisk nature of cyber attacks, an analyst also wants to proactively monitor a network for threats and be able to stop them. This also requires multiple analytic sources to build a graph of interrelationships of the cyber environment and the potential bad actors. Drawing from sources of prior attack evidence, analysts can discover likely new threats in the graph to minimize the unknowns and increase readiness. The AE can assist in pooling information from a variety of disparate sources to gather data for both reactive and proactive cyber analyses.

By addressing these use cases, the Roundtable will showcase a variety of capabilities within the Analysis Exchange that demonstrate the power of this approach. Using the underlying architecture and principles developed here, this approach can be extended to cover a multitude of new use cases. Moving forward, the AE is fostering a new culture of collaboration to the mutual benefit of government and industry.