Data Integration for the U.S. Intelligence Community

> *"Data is created constantly, often unknowingly and without permission. The number of data points that can be collected is virtually limitless—presenting, of course, both enormous intelligence opportunities and equally large counterintelligence challenges."*
>
> *– General David Petraeus, In-Q-Tel CEO Summit, March 2012*

Transform Data Into Actionable Intelligence

Since U.S. intelligence agencies are responsible for safeguarding against acts of terrorism and cyber warfare, they are always looking for new ways to gain new insights and detect new threats.

Analysts in the intelligence community need to identify relevant information and recognize potential relationships of that information to other data points. The intelligence community needs to collect, analyze, and disseminate key transactions, events, and communications and share information within and among organizations.

Detecting and Preventing Threats Amidst Massive Data Volumes

But gathering, categorizing, handling, storing, and securing data is a big challenge. New data comes to light every minute—even every second. It comes from many sources and in many formats. The sheer volume of data makes detecting and preventing threats more difficult.

This executive brief outlines how an enterprise data integration platform can help intelligence analysts gather and share enormous volumes and varieties of data and interpret the complexities within that data with unparalleled precision and sophistication. An enterprise data integration platform should:

- Detect suspicious activity by proactively identifying risks, anomalies, and unusual activities
- Deliver unprecedented insight through proactive monitoring and alerting
- Establish a data governance plan to use data efficiently across the organization

Supporting the Mission of Intelligence Gathering

The Informatica® Platform is the answer. This platform directly supports the gathering of intelligence by discovering, distilling, and delivering the information that the intelligence community needs to detect and prevent new threats.

This open, unified, economical, and comprehensive data integration platform handles the most complex data demands and meets the most stringent security and privacy requirements.
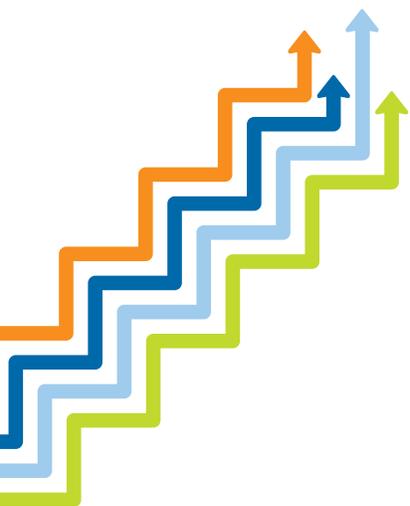
Unique Data Integration Needs of the U.S. Intelligence Community

The explosive growth of information, specifically unstructured data, is making the detection and prevention of threats more difficult. To transform large volumes of data into actionable intelligence, the intelligence community needs sophisticated data integration software that offers key capabilities. Let's explore them now.
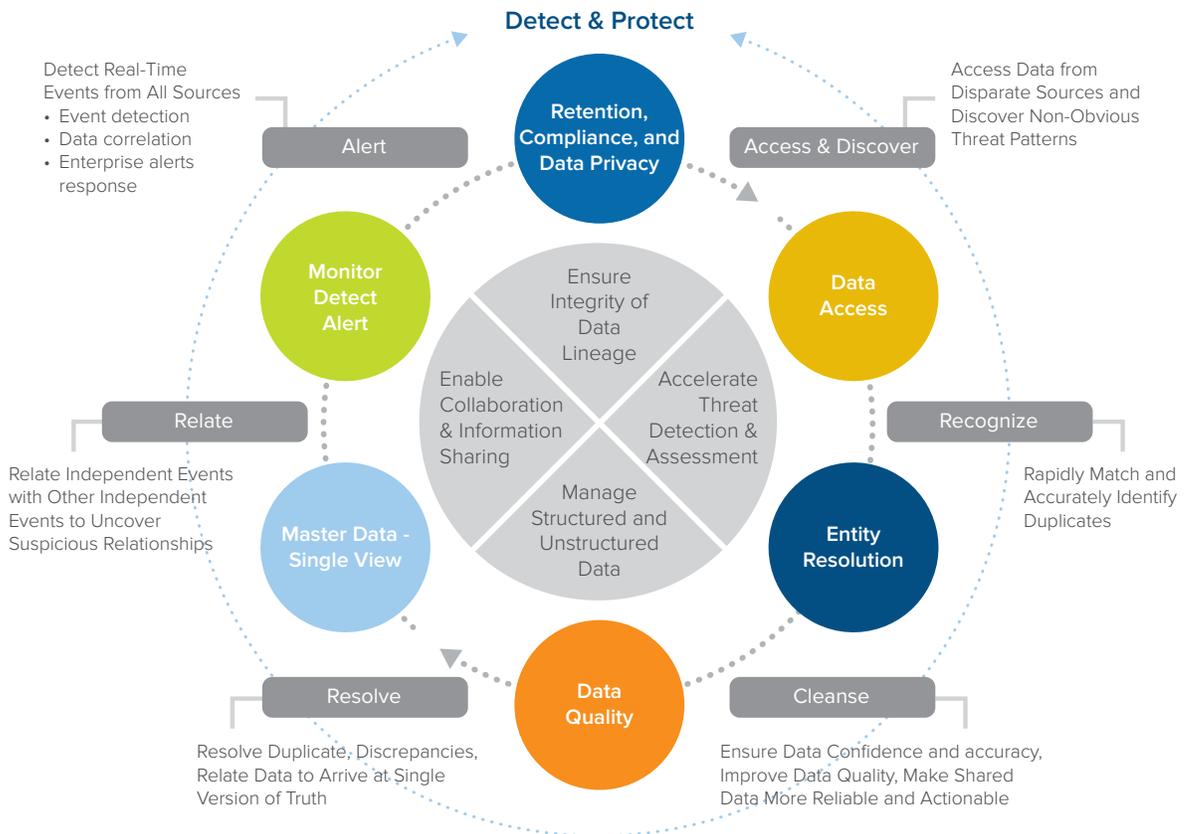
Access and Discover

To identify nonobvious patterns or threats, intelligence analysts need to access all data—regardless of type, structure, or location (e.g., on premise, with a third party, or in the cloud)—from a wide variety of disparate sources.

Recognize

Data formats, quality, validity, trustworthiness, and completeness vary greatly across systems. The intelligence community needs to be able to rapidly identify duplicate data. Identity data needs to be matched and resolved to detect financial relevancy, unobvious affiliations, and events that relate to identities.

Cleanse

Intelligence analysts need to have confidence in the data they use. Data must be accurate, current, and reliable. Authoritative and trustworthy data must be available to all stakeholders and data domains, for all projects and applications—on premise or in the cloud.

Detect & Protect

Detect Real-Time Events from All Sources
• Event detection
• Data correlation
• Enterprise alerts response

Alert

Access Data from Disparate Sources and Discover Non-Obvious Threat Patterns

Access & Discover

Retention, Compliance, and Data Privacy

Monitor Detect Alert

Data Access

Ensure Integrity of Data Lineage

Enable Collaboration & Information Sharing

Accelerate Threat Detection & Assessment

Manage Structured and Unstructured Data

Relate

Relate Independent Events with Other Independent Events to Uncover Suspicious Relationships

Recognize

Rapidly Match and Accurately Identify Duplicates

Master Data - Single View

Entity Resolution

Resolve

Resolve Duplicate, Discrepancies, Relate Data to Arrive at Single Version of Truth

Data Quality

Cleanse

Ensure Data Confidence and accuracy, Improve Data Quality, Make Shared Data More Reliable and Actionable

To detect and protect against threats, U.S. intelligence agencies need actionable intelligence. This community needs data integration software that transforms data into actionable intelligence.

Resolve

Duplicate data needs to be merged to create a single, trusted version of the truth. Intelligence agencies need to access, evaluate, and seamlessly integrate data from multiple data sources to create a golden record—the most reliable, accurate, timely, and comprehensive master file across individuals, groups, events, and transactions.

Relate

Relationships, patterns, associations, and relevant hierarchies need to be revealed. Entity linking, relationship resolution, and identity matching enable agencies to sort and mine relevant data across disparate systems. The goal is to create a single, comprehensive view of people, events, organizations, and transactions.

Alert

Intelligence analysts need to correlate data that could indicate suspicious activities or threats. The ability to apply user-defined analytical rules is essential to exposing hidden, distributed, and meaningful event patterns in unstructured or transaction data.

Govern

Intelligence agencies must maintain complete and accurate records. Keeping these sensitive records secure and confidential is paramount. They must also carefully adhere to retention mandates and comply with privacy requirements.

Deliver

Trusted master data must be synchronized with and delivered to downstream applications and data warehouses for use by analysts and data scientists to detect and prevent threats.

INTELLIGENCE AGENCY ONBOARDS DATA SETS FASTER

- Onboarded data sets in < 1 day
- Eliminated need for hand coding
- Reduced costs by 60% (projected)

A U.S. intelligence agency set a goal to onboard a single data source in 90 days.

But data quality issues and a lack of reusable mappings and services kept the agency from meeting its goal.

As a result, the agency was having trouble complying with data retention regulations.

With the Informatica Platform, the agency onboarded new, complex raw data sets in less than a day. Data was delivered faster using a "create once, deploy everywhere" model for ETL, data transformation, and data quality mappings, services, rules, and logic. Data quality issues were identified and resolved, and future ones were prevented by cleansing data during ingestion—before problems spread into production data warehouses.
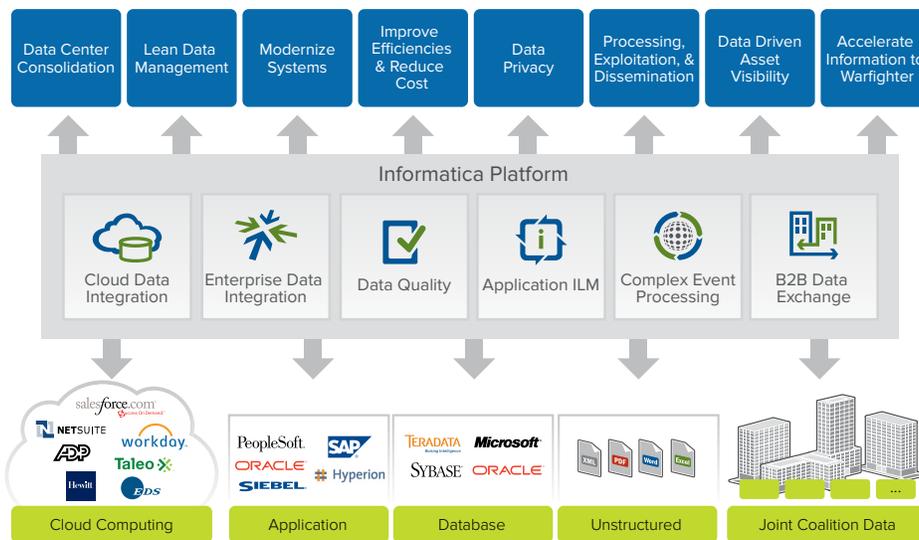

The Informatica Platform

The Informatica Platform is the ideal data integration software for the U.S. intelligence community. This open, unified, economical, and comprehensive data integration platform:

- Supports intense analysis of people, organizations, and watch lists by persistently searching, matching, deduplicating, and resolving data sets that contain information about high-risk individuals

- Empowers intelligence analysts to collect, correlate, and analyze multiple data sources and identify data anomalies without needing IT's help

- Ensures the responsibility, transparency, and accountability of data by capturing and tracking the pedigree and lineage of all data sources

- Improves productivity by onboarding new, large, and complex raw data sets quickly, eliminating hand coding, and by providing reusable data mappings and services

- Reduces annual data management costs by eliminating the need for error-prone hand coding and consolidating multiple data sources into a single and common environment

- Extends the reach of search and discovery capabilities by providing JWICS and SIPR data integration and data quality services for TIRs, IIRs, and other complex unstructured data sources

The Informatica Platform has helped intelligence agencies surmount the challenges of gathering, categorizing, handling, storing, and securing data. It has provided a cost-effective, secure, and flexible platform for enabling their IT organizations to handle the most complex data requirements and to meet stringent retention and privacy mandates.

The Informatica Platform empowers the intelligence community to detect and prevent threats by accessing and integrating huge volumes and types of data from any source and by sharing a single, trusted view of data in and among intelligence agencies.

INFORMATICA®

Worldwide Headquarters, 100 Cardinal Way, Redwood City, CA 94063, USA
phone: 650.385.5000   fax: 650.385.5500   toll-free in the US: 1.800.653.3871
informatica.com   linkedin.com/company/informatica   twitter.com/InformaticaCorp