



Wrangling Actionable Insights from Organizational Data

Koverse Eases Big Data Analytics for Those with Strong Security Requirements

The amount of data created and stored by organizations around the world is doubling in size every two years, and is expected to grow from 800,000 Petabytes in 2010 to a predicted 35 Zetabytes (35 million Petabytes) according to research from IDC. ¹

Needless to say, flourishing data growth challenges organizations as they struggle to find ways to derive meaning from all of the information generated. The struggle is especially difficult for those in the public sector because the data is: 1) growing exponentially; 2) arriving from multiple channels; 3) gaining in complexity; 4) requiring strictly mandated security controls.

Yet, nearly every organization fully recognizes how the ability to derive value from data is crucial to gaining an edge and achieving mission-critical goals. Rising from the challenges is a fundamental question: How does an organization separate the signal from noise within their data, so that they can develop actionable insights?

2010 | 800,000 Petabytes

2020 | 35,000,000 Petabytes

1. "The Digital Universe in 2020: Big Data, Bigger Digital Shadows, and Biggest Growth in the Far East," IDC, John Gantz and David Reinsel December

THE CURRENT DATA LANDSCAPE

The need to turn data into action is not a new challenge. For 20 years software companies have designed data analytics tools to help organizations make better use of information.

Most systems were primarily single vendor solutions, including applications, databases and even hardware that placed limitations on interoperability, and often led to costly upgrade scenarios. Various components included:

Business Intelligence Systems

81 solutions were developed to aid analytic flexibility and move away from those traditionally rigid enterprise applications. BI systems offer powerful capability, but as with enterprise applications, the features that make BI tools powerful also limit their use. BI solutions require a significant amount of planning and knowledge of the underlying structure of data, and thus prove inflexible when adapting to rapidly changing or transient data. And BI tools also struggle to handle large volumes of data.

Knowledge Management (KM) Systems

Just as BI solutions address highly structured data, KM tools are designed to resolve unstructured data challenges. KM products are oriented around user-created data, including email messages and documents.

As opposed to making analytic decisions, the data in KM systems is primarily used for information-sharing and 'subjective' decision-making. In an ideal world, this user-created data would be used seamlessly alongside structured data to make business decisions, but there is still much work to do to accomplish this goal.

Enterprise Applications

This includes HR systems, accounting and billing systems, CRM, and supply chain management apps, among others. These systems contain and manage an organization's most critical organizational data.

There are many options in this category, and most products do their jobs quite well. But when it comes to data analytics, these products are rigid by design, to ensure strict enforcement of established business rules. As a result, they feed data to more analytically inclined systems, to operationalize the data.

Log Management and Analysis

In the last decade or so, applications often referred to as 'intelligence applications' have emerged. Similar to BI solutions, these tools can help organizations better understand their data. However, they do not use the core operational data that resides in BI systems.

Instead, log management applications operate with other data, including server logs, website activity and social media data. The importance of these systems is growing, but they are still limited in scope as these tools are designed to address specific use cases.

Unfortunately, all of these systems, tools and applications suffer from limitations that force them to store specific data, using specific repositories, built for specific use cases. In other words, they all create data silos.



THE BIG DATA EVOLUTION

In addition to silos and integration limitations, organizations of all kinds also face difficulties working to derive value from new and existing data sources, which have been made more complex as data has increased in scale, frequency, and complexity.

Luckily, Koverse offers a remedy to reduce organizational big data pain. The Koverse platform can help any organization to become more data-driven, by delivering actionable insights that help increase effectiveness in nearly every aspect of daily operations.

KEY FEATURES

Koverse consolidates three primary elements required to derive actionable insights:



Collect

Koverse incorporates a feature called Universal Ingest, which enables Koverse to quickly ingest data from any source. Ingest is schema free, requiring minimal to no pre-processing.

Unlike current data warehouse platforms, Koverse requires no advanced understanding of the logical structure of data to ingest and support discovery. In addition, no complex ETL processes are required to maintain data flows. Organizations can go from ingesting data, to deriving meaningful and actionable insights, in hours, rather than days or weeks using other analytic tools.



Analyze

Koverse features Adaptive Workflows to unify disparate data sets and allow organizations to use embedded analytics for immediate insight.

Embedded analytics are derived iteratively from raw data using Adaptive Workflows, and are 'pre-computed.' In other words, Koverse leverages the broadest possible range of permutations so insights such as pattern recognition, anomaly detection, categorization and recommendation are instantaneously available.



Act

Interactive data exploration and discovery make it easier for users to quickly find information and extract meaning. Discovery occurs in sub-second response times, using both schema and schema free query with access to all available data. This feature is also supported by an advanced, tightly integrated indexing and query capability.

Instead of search, where the user asks a specific question of the data, data discovery lets the data inform the user about what is known. Query results are served to thousands of users at a time, in seconds, making it easier to directly embed analytic results and raw data into other applications.

SECURITY AND ACCESS CONTROL

Koverse provides analytic capabilities in a scalable, secure multi-tenant environment, built using well-established technologies such as Hadoop.

Unlike alternative solutions that claim security and scalability as competitive advantages, Koverse believes these elements are fundamental to an 'enterprise class' platform, and critically important to government audiences, in particular.

Koverse provides **multi-tenant** and **self-service access** to information. In government, agencies typically utilize working groups of users and applications, and are accustomed to defining explicit access and sharing patterns for both users and applications.

The Koverse platform's self-service properties allow data owners to **define how data can be shared** within a working group, and across the greater organization.

The 'Collections' portion of Koverse is used to hold data from either an external data source, or the materialized result of an embedded analytic. Collections have two roles. They enforce multi-tenant, role-based access. And they allow applications and users to **quickly hone in on specific data** for analytics and search.

To store the terabytes, to petabytes of information needed, the Koverse platform **scales in both storage and processing capacity**, affordably. Further, it provides its capabilities with high levels of availability. Koverse understands how fault tolerance, data resiliency, elasticity and ease of use are absolutely crucial to gaining broad adoption across an organization.

"Only a tiny fraction of the digital universe has been explored for analytic value."

— IDC report, "The Digital Universe in 2020: Big Data, Bigger Digital Shadows," John Gantz and David Reinsel, December, 2012

KEY BENEFITS

The Koverse platform does not replace existing analytics applications. Instead, it complements BI tools and other enterprise applications. When Koverse is added to an existing environment, organizations can expect to realize the following significant benefits, almost immediately:

Regulatory Compliance Gains

Government organizations face a uniquely large number of regulatory requirements. To achieve compliance with FISMA, HIPAA and the full range of federal requirements, agencies need full visibility into their operational components. This visibility into operations is largely only possible when an organization can access and process a wide range of data, including everything generated by personnel, contractors, applications and constituent-facing transactions, and convert that into straightforward summarizations used to address specific compliance requirements.

The Koverse platform gives any organization the access, visibility and analysis capabilities desperately needed to generate such summaries, in a secure and controlled manner, resulting in clear and more complete answers to compliance questions, in less time.

Organizational Knowledge Gains

By now, everyone understands data is generated by multiple users, using multiple channels and various applications. To further complicate matters, not all users are employees or members of the organization. Take, for example, user forum comments, blog posts, constituent emails, support calls, or Tweets captured for future use.

Koverse allows users the flexibility to search across all agency data without being restricted by structured vocabularies. This means users are more likely to find the sum total of information needed to answer specific questions, and make more-informed decisions.

Information Sharing Gains

Within government agencies, working groups create their own data. To achieve strategic objectives, these groups are frequently required to share information, though this may not always be easy to do, because it tends to be difficult to integrate data from multiple sources using currently available technologies. In addition, questions of access control and security make information-sharing complicated.

Working groups are not allowed to enable confidential information to remain visible to those individuals not authorized to access to such information. Yet that information may be directly tied to other data that has significant value, when shared. Koverse addresses both the integration and security challenges by unifying data into a single repository. The Koverse 'Data Lake' includes both role-based and data-based access control and security measures, unlocking data for those who may not previously have been able to find it, while simultaneously protecting sensitive or classified information from those who aren't authorized to access it.

CONCLUSION

Data is separated into silos today because of the technical limitations of the tools available to meet analytic needs. Those same limitations now prevent organizations from addressing rapidly changing needs and data from multiple sources.

Koverse eliminates those limitations by providing a new data infrastructure that acts in concert with existing data repositories and applications, while enabling greater support for new and big data and rapidly changing data types.

About The Author

ThunderCat Technology is a Service-Disabled Veteran Owned Small Business that delivers technology services and platforms to the federal government, along with Fortune 500 companies.