

The Metadata Management Revolution: Integrated Metadata and what it means for BI

Introduction

We are living in a generation where nearly everything is automated, simplified and made to enable us to do what we do better. However, when it comes to BI metadata management, we have not yet seen technology that significantly transform the way BI groups work on their daily tasks. What we have seen are proprietary tools that support only a single BI platform, that usually have a narrow view of metadata and are distributed by a proprietary software provider. Moreover, in today's business setting we still scramble to find the right information in conflicting reports, where limited visibility and control of metadata make it extremely difficult for businesses to make quick and accurate decisions.

Today's progression in data management revolves around data governance and effective access to metadata—the frame of reference giving data its context and meaning. Metadata governance provides a comprehensive view into the flow of data, the ability to perform impact analysis, a common business terminology and accountability for the terms and definitions with accurate audit trails for compliance. In this context, metadata management becomes an important capability, enabling IT to oversee changes while delivering trusted data in a complex data integration environment.

This white paper discusses the ins and outs of the BI landscape, the challenges of using silo tools, and the revolutionary impact of centralized metadata management on BI groups. Today, the ability to navigate through multiple systems and tools quickly, accurately and effectively to locate relevant metadata is an absolute must for all BI groups.

The Problem

The BI landscape

The BI landscape is flooded and metadata is scattered all over the place. Whether we're talking about ETL, databases or reporting systems, there are many data management tools, created by many different vendors, and it's important not only to not get lost in this crowded space, but to also find the business value in each technology.

The challenge becomes even greater when BI groups must work with tools that are incapable of communicating with one another.

Silo tools don't cut it anymore

With so many silo tools out there, discovering and understanding metadata is extremely cumbersome for BI groups who have to manually trace the data journey. These tools are incompatible with each other and complicate the way BI professionals work.

Not only is this process extremely time consuming, but it is also error prone and costly to businesses. With a centralized repository for all metadata across the BI landscape, BI groups would be able to see the data journey clearly. With metadata centralized in one place, they would have full visibility, access to and control of metadata whenever they need it.

A TDWI report looking at the cost justification of metadata management shows that metadata management can help not only reduce payroll costs, but also to boost productivity by assigning more projects to current staff without raising the overall costs. Metadata management tools can save companies a lot of money and can help the business move faster.

The Technology

Integrated Metadata Management - a game changer

The creation of an automated metadata management capability improves the productivity of both users and implementers of BI solutions.

Metadata allows for the linking and integration of structured and unstructured data types. With the volume of data growing exponentially and so many different types of data out there, it is quickly becoming impossible to manually search for metadata or to know in advance the full impact of potential changes or development. Therefore, tools that can automatically discover metadata – beyond merely capturing the metadata created during the development and operation processes - are becoming an absolute must.

An automated, cross-platform metadata management system revolutionizes the way BI professionals work with metadata. With metadata centralized and accessible to everyone in the group, BI professionals will no longer lack visibility and control of metadata.

The most advanced cross-platform data discovery products can identify where data is located using the names of fields and tags, as well as the actual content of the data structures. This technology enhances the view of data flow, and the ability to perform impact analysis, with an automatic common business glossary. The management of metadata becomes an important capability enabling IT to oversee changes while delivering trusted data in a complex data integration environment.

One technology investment that should be considered in a data integration strategy is a central metadata repository that pulls all the metadata together. In this context, it has been demonstrated that such a repository provides an excellent way to govern data across the organization, provide both a business and technical view of the organization's data assets, and show an audit trail of data lineage suitable for regulatory reporting requirements.

Our Solution

The Octopai Technology

Octopai transforms the tedious BI manual maintenance process into a powerful experience, which creates a new shared metadata platform with a single enterprise view, enabling the technical group to search and trace all BI objects within seconds.

Octopai combines enterprise metadata technologies by using various plugins into a single, unified IDA [Integrated Data Architecture] platform. It automates reporting and auditing, turning days or weeks of manual work into just a few seconds and increasing technical staff productivity exponentially. Octopai's goal is to empower BI groups as well as other departments [e.g. Finance] in their daily tasks and business challenges by automating manual tasks, dramatically increasing throughput and shortening time to market while reducing risks of error.

Through guided workflows, data stewards can reference discovered objects and connect them to data points in the IT landscape, ETL, databases, applications, reports, and other objects. This approach delivers a traceable and manageable path to enterprise data diversity that is growing in complexity, and often times creating sudden business risks. Octopai returns control of data to business users, data stewards, and IT architects responsible for decisions, data governance, and policies.

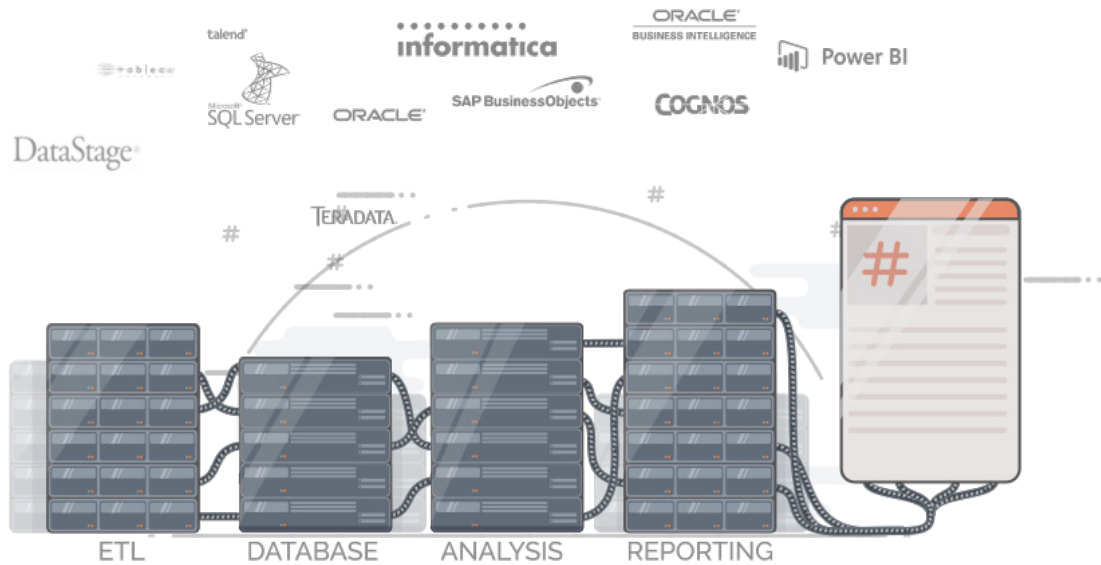
The platform enables the technical staff to understand, cleanse, transform, and deliver reliable and context-rich information to complex BI and reporting problems that cannot be tracked easily by traditional manual processes and thus eliminating the time delays associated with long and tailor-made manual solutions.

Octopai supports the automated search and tracking of the following objects:

ETL: Informatica, DataStage (IBM), SSIS (Microsoft), OWB.

Database: Oracle, SQL Server, Teradata, Vertica, Exadata.

Reporting & Analytical systems: COGNOS, Business Objects (SAP), Tableau, OBIEE (Oracle), QlikView, QlikSense, Power BI (Microsoft).



Understanding how Octopai works

1. Octopai developed "connectors"/ "extractors" which connect to all the different systems (ETL, DB and Reporting) and extract metadata items from each.
2. Octopai performs read-only operations on databases by using APIs (application program interface) and other communication methods to communicate within the applications.
3. Octopai developed connectors for each tool using a service bus for moving the metadata to a centralized repository.
4. In a subsequent step, Octopai establishes dependencies among the tools. (Analyzer) with an algorithm that knows how to create dependencies and cross connections of metadata items.
5. Octopai stores metadata in a DB repository in a unique way for quick presentation with detailed and lineage on the visualization layer

Future Outlook

In the next two years, Octopai is determined to link most of the BI platforms into a single integrated metadata environment, thus becoming the sole provider of cross-platforms enterprise metadata search engine.