CLOUDERA AND ATTIVIO

Unify Business Users – and Unstructured Content – with Big Data

Organizations maintaining Big Data repositories today realize there is substantial new business value to be gained from the analysis of Big Data. However, a number of challenges exist which impede efforts to achieve that full value:

Business User Gap. Existing Big Data products offer cost-effective ad hoc analytical access to Big Data; however, these products leave a large gap for business users who must rely on highly-skilled technical resources to gain insight from their Big Data. Additionally, developing new Big Data applications or integrating with existing packaged applications is challenging.

Limited to Single Data Source. Big Data analytics are typically oriented toward a single, standalone source of data. This does not address the wider enterprise data challenge which Gartner Group calls “Extreme Information” – extremes in data volume, data velocity, and an extreme variety of types of enterprise data: structured, semi-structured and unstructured information.

Big Data Answers “What,” but not “Why.” Big Data is primarily made up of structured data and/or semi-structured data, such as clickstreams or logs, which can be analyzed to determine what business trends are taking place. However, explaining why those trends are taking place requires insight from the business signals contained within unstructured data, including customer sentiment, explanatory notes, and other vital contextual information. Big Data solution vendors do not leverage unstructured data, such as documents, email and free-flow text, to provide this critical context.

Solution: Attivio AIE/XT

Attivio has solved these Big Data challenges, including meeting the wider challenge of gaining the maximum business value from extreme information, with AIE/XT.

AIE/XT is an enterprise class unified information access (UIA) platform that enables rapidly-deployed, analytic-style applications and workflows that easily incorporate and converge existing Big Data infrastructures and enterprise unstructured information. In addition, AIE/XT seamlessly connects to Big Data stored within Cloudera clusters.

AIE/XT offers compelling new value for Big Data environments, including:
- Augmenting Big Data analytics with contextual insight and relevancy drawn from unstructured data
- Utilizing Big Data to enhance business processes, turning Big Data into Active Information that drives applications via search and SQL information access paradigms.

Attivio, Inc.

Website: http://www.attivio.com/

Company Overview
- Award-winning enterprise software company headquartered in Newton, Massachusetts USA
- Experts in unified information access (UIA), text analytics, enterprise search, business intelligence and data warehousing
- Global customer relationships in financial services, technology, government, telecom, manufacturing
- Our mission: Improving critical business results by delivering all the information that matters – both structured and unstructured data – in one combined information platform

Product Overview

Active Intelligence Engine™ (AIE), the only UIA platform that delivers complete analytical insight from fully integrated and connected structured data (databases) and unstructured data (free-flow text, documents, email, web content, etc.) accessible via search or SQL queries.

AIE/XT™ provides the power of AIE to Big Data environments, including Cloudera/Hadoop processed data
By eliminating the technological divide between structured and unstructured data, AIE/XT adds significant new analytical and discovery capabilities to drive new insight and business value. This is achieved through a powerful set of features, including:

- **Universal information repository.** AIE/XT integrates all relevant structured, semi-structured and unstructured data into a single repository for unified information access and analysis, with no need for complex data schemas. Structured databases retain their relationships such as key fields, alongside semi-structured and unstructured data for unified information access and analysis.

- **Business signals automatically drawn from unstructured data.** Far beyond merely storing documents, email, web content and other unstructured information, AIE/XT enriches these sources for actionable BI. Capable of extracting text from over 500 file formats, AIE/XT automatically mines and extracts key phrases and entities, categorizes documents, and assesses sentiment expressed in a document and about specific entities in documents, providing key “why” dimensions for analytic insight.

- **UIA application framework.** AIE/XT enables the building of UIA applications that provide relevant structured data and unstructured content together in a single, user-friendly view, bridging the gap between Big Data and business users. AIE/XT supports both SQL and search APIs for maximum flexibility in how business applications access unified information.

- **In-engine data integration and transformation.** AIE/XT’s embedded workflow layer automates data collection, integration and transformation including data cleansing, normalization, text analytics and data mining. In addition, as data comes into the system, triggers within workflows can send relevant data to user interfaces, other applications or directly to users via alerts. Big Data becomes Active Information through AIE/XT’s ability to initiate real-time business processes.

- **Real-time connectivity to unstructured and structured sources.** AIE/XT is designed for real-time streaming of new information into the Universal Information Repository. New structured and unstructured information sources can be added rapidly and be loaded in batch or streamed via AIE/XT’s ingestion API to AIE/XT’s workflow layer.

### Cloudera Connectivity with AIE/XT

AIE/XT enables the loading of Hadoop/Cloudera processed data using Hadoop connectors, including Hbase, HDFS and Hive. Additionally, AIE/XT enables the pushing of MapReduce output to AIE via client APIs. AIE/XT provides further data connectors to load data from existing data warehouses, data cubes or DBMS stores.

### From Big Data to 360-Degree Business Insight with AIE/XT

AIE/XT draws vital business signals and context from unstructured information, easily integrating those signals and context with Big Data for access and analysis by business users. The end result is an unprecedented 360-degree view of customers, products, business performance and more, enabling the discovery of business issues and opportunities that may have otherwise gone undiscovered. AIE/XT fulfills the demand of gaining new unparalleled insights from Big Data, as well as Extreme Information that exists throughout the organization, regardless of the content type.

### About Cloudera’s Distribution Including Apache Hadoop

Available for free download at www.cloudera.com/downloads, CDH delivers a streamlined path for putting Apache Hadoop to work solving business problems in production. Ideal for enterprises seeking a stable, tested Hadoop solution without proprietary vendor lock-in, CDH is the bridge between the insights of organizations using Hadoop in production and the continuous stream of innovations from the Apache community.

### About Cloudera

Cloudera is the leading provider of Apache Hadoop-based software and services and works with customers in financial services, web, telecommunications, government and other industries. The company’s products, Cloudera Enterprise and Cloudera’s Distribution including Apache Hadoop, help organizations profit from all of their information. Cloudera’s Distribution including Apache Hadoop is the most comprehensive Apache Hadoop-based platform in the industry. Cloudera Enterprise is the most cost-effective way to perform large-scale data storage and analysis and includes the tools, platform and support necessary to use Apache Hadoop in a production environment. For more on Cloudera, please visit www.cloudera.com.

### About Attivio

Attivio’s dynamic approach to UIA features rapid development, true incremental scalability and a small footprint that can be easily embedded in other products and applications. Business and IT leaders from a variety of industries – including financial services, e-business, media, technology, government and more – as well as system integrators and independent software vendors partner with Attivio to uncover new insights that drive innovation and reveal new opportunities, mitigate risk and contain costs. Please visit www.attivio.com for more information.