WHITE PAPER
08.13 | EB 6839

# BIG DATA ANALYTICS: TURNING INSIGHT INTO ACTION

## NEW TECHNOLOGIES ACCELERATE DATA DISCOVERY

TERADATA

TIBCO™ Spotfire®

## EXECUTIVE SUMMARY

The core of business intelligence (BI) is the transformation of data into insight and insight into action that can add value to the enterprise. But the rise of structured and unstructured data, which is known as big data, has radically transformed the BI function.

While big data has increased the opportunities available to businesses, it also creates more challenges to capturing, storing, and accessing information. In particular, the advent of big data underscores the time value of information. Since the utility of some data points declines very quickly, there's a premium on accelerating the process of turning data into information and information into action.

This paper describes how companies can combine a unified data architecture (UDA) with complex event processing software to thrive in the era of big data.

## RISE OF BIG DATA

Take, for example, the geospatial data generated by mobile phones. This data is routinely captured and stored by robust data warehouses and can be analyzed thoroughly. If the goal is to understand usage patterns to optimize a cellular network, a "real-time" understanding isn't necessary. But if the goal is to understand where Customer A is right now and to provide an offer based on his or her location, capture and analysis of geospatial data must occur in something like real time—a very different technological challenge.

Dealing with both of these types of scenarios simultaneously demands a unified data architecture, which includes a robust platform for the collection and storage of all kinds of structured and unstructured data and a platform for data discovery to generate insights, opportunities, and actions. To complement such an architecture and enhance its ability to deal with the data tsunami in real time, companies can also turn to complex event processing software that includes advanced analytics and data visualization capabilities.

BI and analytics platforms have been tremendously effective at using structured data, such as that found in customer relationship management (CRM) applications, to optimize functions throughout an enterprise. But in the big data era, the BI challenge has changed dramatically, in terms of both goals and execution:

- **Volume:** The volume of data has grown by orders of magnitude. A recent article in *The New York Times* stated that the size of the entire digital universe in 2005 was 130 billion gigabytes; in 2013, it was up to 40 trillion gigabytes. Today's enterprise environments routinely contain petabytes of data.
- **Variety:** The variety of data has evolved from the traditional structured datasets of enterprise resource planning (ERP) and CRM to data gathered from user interactions on the Web (clickstreams, search queries, and social media) and mobile user activity, including location-based information.
- **Velocity:** The velocity of data accumulation and change is accelerating, driven by an expanding universe of wired and wireless devices, including Web-enabled sensors and Web-based applications.

This revolutionary shift places significant new demands on data storage and analytical software, and poses new challenges for BI and database professionals. It also creates powerful opportunities for discovering and implementing new strategies that generate competitive advantage.

Realizing these opportunities requires two things: the technological capacity to gather and store big data, as well as new tools for turning data into insights and, ultimately, value.

## NEED FOR SPEED

One way to look at the evolution of BI is to think in terms of "data at rest" and "data in motion." Traditional BI primarily uses data at rest: "steady state" customer information, purchasing history, inventory, and other kinds of information. While such information does change constantly, it can be considered at rest for analytical points in time. Even extremely dynamic information (e.g., the geospatial coordinates for a population of cell phone users) can be used for "data at rest" analyses. The point is that if you're making a decision about what to do next month or quarter or even next week, data at rest should be able to give you the answers you need.

In contrast, data in motion can be used to make decisions about what to do now, or in the next five minutes, if the right data can be brought to bear on the right question. Using the cellular phone example, you could rely on collated, historical GPS data (data at rest) from the entire population of cell phone customers to make decisions on adding cell towers to a particular region. But if you want to identify the customers within one mile of a retail location and create an incentive for them to visit the store, you need to act on their GPS information while the data is still in motion.

Take another hypothetical example: Imagine a clothing retailer with two different chain stores in the same shopping mall. A customer visits Store A and uses an opt-in loyalty card to pay for her purchase. The loyalty card has already told the enterprise a lot about the shopper, such as purchasing history and demographic information. This is data at rest. But with the swipe of the customer's card at the point-of-sale terminal, the retailer gains a valuable piece of data in motion: The customer is in Store A. She's buying. And she's near Store B. Why not give her an offer connected with Store B right then, through a text on her phone? (Note that this new piece of data in motion ultimately becomes a piece of data at rest, adding to the company's intelligence about the shopper, that store, etc.)

The value of this new data point is significant—she's in a buying mood, she's near another branded store in the company's portfolio—but it has an extremely short half-life. A BI system capable of putting that data to work has a lot to do in a very short window:

- Recognize a valued customer when she enters Store A.
- Match her customer profile with an appropriate and targeted promotion for Store B.
- Deliver that promotion to her phone.

And it has to do it all within five or ten minutes; because if it takes much longer, she's more than likely to be too far from Store B to take advantage of the offer. In this example, the data's half-life is extremely short.

The diagrams in Figure 1 highlights the business value gained by accelerating decision-making process that leverages all data — data at rest and data in motion. Both plot, in hypothetical ways, the time value of action between each step of the decision-making chain that begins with a business event (e.g., transaction data coupled with locational information); the observation and collection of the relevant data (combining that real-time data in motion with existing customer segmentation data at rest); analysis of that data (understanding potential promotions based on the interplay between the data in motion and the data at rest); a decision on a course of action (perhaps a time-limited discount or alert about a nearby promotion); and action on that decision in something like real time (e.g., a text message or smartphone app notification or coupon delivered at the point of sale).

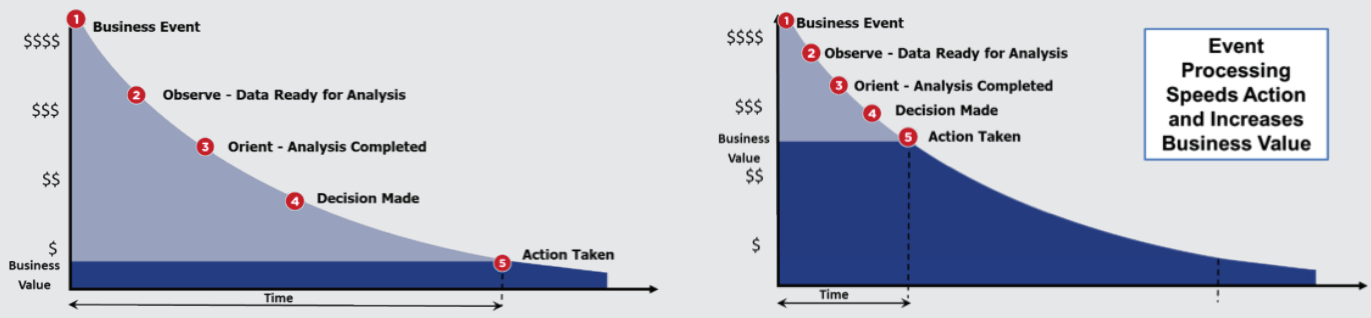


Figure 1: Businesses can drive a higher business value through an accelerated "insights to action" process.



TIBCO Spotfire® TERADATA®

Whether the enterprise is relying on data at rest or data in motion, the BI process (event>observation and collection>analysis>decision>action) does not change. What does change is the slope of the value-versus-time curve, which is dramatically steeper. If it takes 45 minutes from event to action, the shopper will be in her car and half way home. If event to action can be accomplished in two or three minutes, there's a good chance of reaching her in time for the offer to be acted upon.

## NEW DEMANDS ON BI TECHNOLOGY

Many BI systems are already capable of handling the efficient capture and analysis of highly structured data, and they are well-equipped to generate insights and possible actions with predetermined assumptions using a wide variety of tools and solutions. The most robust BI systems, those that make use of a UDA, are extremely good at the timely capture and analysis of both structured data and the huge realm of unstructured and multi-structured data, generated by embedded sensors, loyalty cards and programs, and Web interactions from Google searches, Facebook posts, Twitter feeds, and clickstreams. By weaving together an enterprise data warehouse, a discovery platform, and a data storage and refining platform, a UDA eliminates the expensive and time-consuming step of moving project-specific data into a middle tier for analysis and enables companies to act on some newly acquired data points within very short time windows—often as little as five or ten minutes.

But to put this data to work in something approaching real time demands the emerging capability of complex event processing (CEP). CEP solutions can effectively capture raw data, quickly recognize and separate what is useful from what is not, executing analytics in-memory and transforming analysis almost immediately into action. In short, CEP helps the enterprises sift through the "noise" of digital footprints to find the "signal" that alerts them to opportunities—to separate the proverbial wheat from the chaff.

### FINDING AND OPERATIONALIZING OPPORTUNITIES

A major Nevada gaming company is using big data analytics and data discovery to find interesting new opportunities to build business today and tomorrow. The company is already adept at collecting and monitoring how much guests earn in its rewards program, and now it has begun taking data collection and data discovery to an entirely new level.

Consider this scenario: A young woman, already a member of the gaming company's loyalty program, plans a bachelorette party in connection with her best friend's wedding. She has an app on her phone that can alert her to offers before, during, and after the trip.

A couple of weeks before the party, the casino uses the app to deliver a "VIP card" to the woman's entire group. The card enables her friends to take advantage of any offer, easily enroll themselves in the loyalty program, and, if they use the app, receive special offers and notifications in advance of checking in.

The woman who initiated the trip gets a special incentive designed around her gaming preferences. At one point during her stay, she plays blackjack, and when she doesn't fare particularly well, the system automatically sends her two free tickets to a comedy show that evening. When she and her friends wake up the next morning, they're offered discounts to the buffet and spa. As the weekend winds to a close, the group is offered a discount on a future weekend getaway.

In this example, big data analytics combines data at rest (traditional structured data, e.g., what games at what stakes has the guest played in the past? what property services has she used on previous visits?) with data in motion (e.g., what is the guest doing right now? what offers are available at this moment that are appropriate given the guest's past and present actions?) to identify opportunities and take advantage of them in what is essentially real time.

## COMPLEX EVENT PROCESSING: IDENTIFY AND ACT

CEP requires several interconnected steps that revolve around the two poles, insights and action, and include the following:

- **Capture:** Data capture and collection is where the cycle begins. For example, if a guest takes a trip to a Las Vegas casino, it would include past experiences at the resort, recent events/activities, as well as what the guest is doing at this very moment. Typically, data is collected in a UDA and made accessible through the data warehouse and data discovery platforms.
- **Understand:** Data is only useful as the raw material of insight and understanding. Historic customer information is typically used for segmentation and, in turn, to create appropriate offers. But to move beyond segmentation to individualization requires a more complex, data-rich and event-driven analysis, where data discovery and interactive visualization capabilities become paramount.
- **Model:** Many customer insights are intuitively obvious. It hardly takes technology to recognize and reward the top 1 percent of a gaming company's customer base. But modeling the remaining 99 percent (who probably represent 80 percent of total revenue) requires more robust analytics capability.
- **Anticipate:** Newly found insight is operationalized in the CEP platform, allowing it to act on desired events. In the example of the Nevada gaming company (see sidebar), the casino's up-to-date BI system monitors a guest's individual data streams in real time, which enables the company to anticipate opportunities for making offers that will be attractive to a particular guest.
- **Decide:** The power of real-time data analysis is its ability to enable real-time decisions. In the casino example, if a guest is losing at blackjack and is about to walk away disappointed, the casino could provide free show tickets, increasing customer satisfaction in a cost-effective way.
- **Act:** Once the system and the BI expert have determined a particular response to this complex chain of events, it must be acted upon quickly. An automated offer must be sent within a time frame that allows the recipient to take advantage of it, whether it's a pair of tickets to a show or a free drink at a popular lounge.
- **Monitor:** The efficacy of offers and promotions must be measured. Automated, data-driven offers are easy to track and evaluate, which can then enable companies to refine the entire cycle of collection, understanding, and modeling.

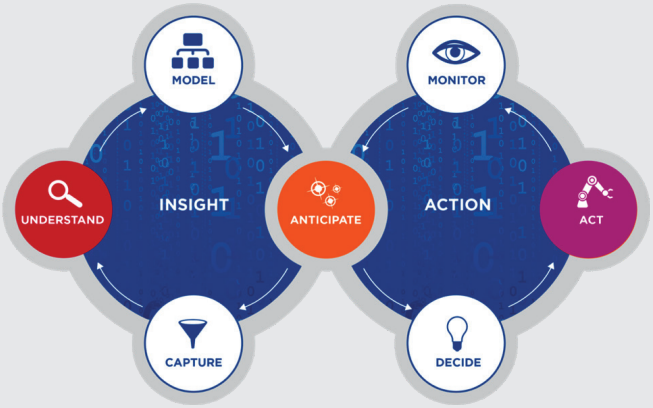


Figure 2: Complex event processing revolves around two poles: insight and action

This continuous, closed-loop process incorporates strategic and operational intelligence and in doing so helps users analyze what works and what doesn't, refine event-driven models, and generate new insights, actions, and opportunities for individualized offers.

## TERADATA AND TIBCO: PUTTING IT ALL TOGETHER

Technology leaders foresee, understand, and respond to changes in the business environment. As the era of big data has evolved, Teradata and TIBCO have quickly understood the increasing importance of data in motion and have partnered to create a technology platform capable of taking advantage of the rich opportunities.

The Teradata® Unified Data Architecture™ is a framework for organizations to address all types of data. It leverages the complementary value of best-in-class technologies from Teradata and open source Apache™ Hadoop®. This enables business users—not just IT experts—to ask any question, against any data, with any analytic, at any time. In short, the Teradata Unified Data Architecture™ enables users to consolidate siloed data streams into a single reusable strategic asset.



TIBCO™ Spotfire®

TERADATA®

The Teradata Unified Data Architecture™ handles the critical data discovery function with support from TIBCO Spotfire visual analytics, which provides Teradata users with a complete analytics workflow platform. Together, Teradata and TIBCO can handle extreme data access and data mash-ups, combining data from the Teradata integrated data warehouse, Teradata Aster® Discovery Platform, and Hadoop with transparent access methods and connectors. Moreover, Spotfire has an easy-to-use, graphically intuitive format that can be readily shared and used throughout the enterprise.

This Teradata and TIBCO Spotfire solution offers many benefits to users:

- **Single view of the business:** The Teradata Unified Data Architecture™ provides a data-rich, consistent, and comprehensive window into an enterprise, increasing decision-making effectiveness and enabling users to operate with the most accurate and timely information.
- **Fastest time to action:** The addition of TIBCO's high-performance big data analytics CEP platform increases decision-making productivity by enabling anyone in the enterprise to anticipate situations and opportunities, to ask relevant and timely questions, and to get answers that lead to decisive actions.
- **Visibility into the unknown:** The Teradata Aster Discovery Platform and Spotfire solution provides a robust big data analytics platform that allows users to discover unseen trends and patterns in large, complex datasets with intuitive visualizations that foster faster identification of strategic opportunities and threats.
- **Self-service data discovery:** The Teradata and Spotfire solution allows users across the enterprise to explore data and get answers without the need for specialized, in-depth data modeling. This reduces dependence on IT and dedicated BI resources and greatly accelerates decision making.
- **Encapsulated expertise:** Companies can master complex business challenges with deep, repeatable analytic solutions that leverage the knowledge of database administrators, analysts, and decision makers, along with the industry expertise of Teradata and Spotfire.
- **Universal adaptability:** Users can leverage a single enterprise-wide analytics data engine and data discovery platform to address analytical needs for any business or technical user, from front-line strategists to back-office managers.

## CONCLUSION

Big data is already a fact of life for enterprises, but the sheer volume and massive complexity of big data can feel overwhelming. Companies suddenly must struggle with making sense of and creating opportunities from both data at rest and data in motion, from structured, unstructured, and multi-structured data.

And yet big data also represents an expanding set of powerful new opportunities. And while it is unknown exactly how it will be used one, three, or five years down the road, the imperative of turning big data into competitive advantage means no one can afford the luxury of waiting to see how things shake out. The window of opportunity for action is getting shorter—the sense of urgency stronger.

The good news is that the tools to take advantage of big data are here today. A unified data architecture and substantial advances in analytical and data visualization software enable companies to realize the potential of the big data era and minimize its risks.

For more information, contact your Teradata representative today or visit **www.teradata.com**.

10000 Innovation, Drive Dayton, OH 45342 teradata.com



EB-6839 > 0813