ovum

Thriving in the Age of Big Data Analytics and Self-Service

The new shape of BI

Tom Pringle, Surya Mukherjee & Tony Baer

Table of contents

Executive Summary

Data in every form, at any scale, is going mainstream. No longer viewed as a liability and IT headache, data has become a focal point for executives. It is now viewed as capital, one of an organization's most valuable assets, due to its potential to provide business value.

Technology is also becoming more accessible – lower-cost processing and storage are only part of the story. Cloud delivery, consumer applications, and ubiquitous mobility have fundamentally altered how users engage with their data, creating new expectations about ease-of-use and access.

Traditional business intelligence (BI) generally depends on a request/supply model, with a small centralized team of experts serving specific internal audiences. This model has been successful for many organizations, but its ability to scale to a far broader audience is limited. Getting users the data they want when they want it is the ongoing challenge for IT, representing a fine balance between speed and the governance and quality control that is IT's charter. But users now expect even easier access to data that is vastly greater in volume and variety, and for the data to be current – hours or minutes old, rather than months or quarters – regardless of its size or rate of change.

Starting with big data analytics does not, however, mean throwing away the investments made in traditional enterprise data and tools. The insights that traditional tools offer cover many of the core functions of business and provide the basis for delivering a more complete picture, enriched with the new insights made possible by big data analytics. Achieving this advantage, however, demands both technology and cultural change, enabling the merging of traditional and new analytical technique to uncover new insight while embedding evidence-based decision-making into the core processes that drive the business.

The new age of analytics and Oracle

Oracle has been a key participant in the development of these trends, offering hardware and software for handling use cases encompassing big data and traditional analytics, information discovery, data wrangling, self-service analytics, and real-time operational transactions. Oracle has extended its portfolio to provide customers a unified big data management solution strategy that interweaves big data analytics with traditional and exploratory BI, while providing common and consistent methods to govern and support the discovery and consumption of big data.

The New Shape of Enterprise BI

Cloud and mobile deliver on consumer-style expectations

Cloud deployment and mobile access are helping IT make data and analysis tools more widely available. Cloud business applications have had two effects on analytics: data generated and stored in the cloud has grown to the point where analyzing it in place, rather than shipping it to on-premises systems, has become a necessity; and cloud delivery cuts the long deployment cycles typically associated with on-premises solutions. In both cases, the speed and efficacy with which data can be accessed and analyzed has grown, delivering greater access to a broader audience.

Mobility has been a growing feature of corporate IT for many years, however it is only recently with the advent of smartphones and tablet computers that the means to consume large quantities of data on the go has become a reality. Mobility is more than just consuming information while out of the office, it is full-scale interaction that allows in-the-moment data capture and response, opening a

range of use cases – from retail assistants using customer relationship management (CRM) insights to better serve customers while they're in the store, to factory managers adjusting production using real-time analysis. The range of mobile use cases continues to grow, opening new opportunities to make data-driven insight available to more decision-makers.

Self-service reduces risk and unleashes the benefits of access

The traditional BI supply/request model is based on a built-in throttle. Given current data volume and variety and associated demand for access, traditional BI models simply cannot serve this expanded audience. Self-service is therefore essential for success.

Self-service, as part of an exploratory analytics strategy that incorporates discovery, can lead to richer insights on existing and new data. However, like all new models, it needs to be supported by a change in culture as well as technology. For self-service to work effectively, IT departments must enable fast access to diverse data that drives discovery. A self-service environment will likely fail if it does not allow access to rich data and content. On many occasions, such data lies inside the enterprise but is poorly used because people do not know about it, or because of overly restrictive policies that prevent access.

Self-service in the form of shadow BI (exploring data and performing analytics with unsanctioned tools and infrastructure) is a potential free-for-all, increasing rather than addressing the complex challenges organizations face. The rise of shadow BI is an understandable response to the need for creativity and freedom to work with data, as business users look to manage and operate at their own pace, under their own control – understandable, but potentially risky.

Ovum view

Ovum advises IT departments to embrace a shift in their role from data controllers to data enablers for the business, providing governed access, by way of self-service tools, to managed data sets from enterprise applications and the enterprise data warehouse. In making organizational and big data sets available quickly, IT not only ensures that data is properly governed and protected but also dissuades users from acquiring data in unsecure, static flat file dumps – which most ad-hoc BI users do, and which leads to more data silos.

Big data and visual discovery go hand in hand

Extracting insight from big data is a major challenge, as location, formats, velocity, and complex tools make big data analysis uniquely challenging. Deriving value from big data means creating insight that encompasses both traditional and nontraditional sources of data. In other words, getting the complete picture requires the big picture and the exact picture.

In addition to structured data from familiar enterprise sources, analysis can include text from messaging or social networks; weblogs, sensory data, and other machine data; and metadata derived from rich media (e.g., images, audio, and video). For instance, supplementing CRM data from a call center or website with social media sentiment provides a better view of buying behavior and the impact of customer satisfaction for both individuals and groups, offering critical insight about whom or what events they have influence over. This does not replace traditional enterprise data-driven analytics for building customer profiles, segmentation, credit scoring, or related analytics. Big picture analytics enriches traditional analytics by providing greater context, filling in the blanks regarding knowledge about the customer.

Big data analytics is the practice of getting value from these less familiar data sets. Teams may not always have clear answers as to which data sets they should include, or what questions to ask. Discovery is essential so teams can rapidly evaluate hypotheses before moving on, to "fail fast," because identifying correct hypotheses invariably requires testing false ones. In discovery, users probe further for answers about the root cause: why leading indicators for customer preference, security threats, evidence of fraud, or unexpected disruptions to operations are occurring. Ultimately, exploratory analytics enriches traditional analytics by validating that organizations are asking the right questions.

IT challenges in the new analytics landscape

The long-standing IT challenge to balance delivering requirements and maintaining control is significantly strained by the new hurdles - and opportunities - presented big data's volume, variety, and velocity.

Expectations have changed

In addition to working with a more demanding and self-sufficient set of stakeholders, IT is expected to reinvent itself to deliver new value to the business without sacrificing data stewardship: a process that requires collaboration across the business.

- **Users expect consumer-style service as well as technology –** It's not just expectations about technology that have changed. Users across the organization expect an increased level of service alongside friendlier technologies, which means that processes oriented around tight control need to evolve, so that enforcing governance is no longer the purview of a select few in a central hub, but a shared charter. In this new world, everyone engaged in analytics becomes a data steward.
- **Value-building analytic use cases demand a new approach –** Blending multiple data types, in near-real time, is at the heart of delivering business differentiation through data-driven insights.

Governing data is getting harder

Data is the fuel that powers the analytical engine; the expanding array of sources, data types/velocity, and new users, mean this long-standing issue is only getting more difficult.

Necessary data no longer just resides inside the firewall

Traditionally, the data required for operational and regulatory reporting was internal; this is not always the case anymore.

Multiple sources, types, user groups, requirements, devices, etc. are rapidly growing the challenge of effectively governing data. Frameworks of data governance must evolve to incorporate the new reality. Data governance processes must be extensible to support the requirements of big picture analytics, in addition to traditional exact picture BI.

Governance and security

Governance still remains a challenge even for structured data; enforcing a stringent governance protocol across both structured and semi-structured data sources is doubly challenging. Too many padlocks on data inhibit discovery and exploration, which is the primary aim of self-service tools;

however, opening up sensitive data sources to analysis puts organizations at a higher risk of misuse and loss.

Data governance that accommodates self-service tools is an area of the market that is still immature. Some self-service tools that offer a strong front end for data exploration are still developing the security features that enterprises expect. Enterprises must therefore ensure that self-service tools either include a strong governance framework or integrate with existing governance solutions to supplement their native capabilities. This approach works well when the governance solution, data-staging areas, and the exploration solution are easily or natively integrated, allowing metadata and access details to be shared across these applications.

For the SQL side of the world, there are mature and well-defined governance solutions with which self-service solutions can integrate. On the NoSQL side, there are a growing number of solutions.

Data integration is never-ending

Enterprises are faced with an ever-growing data integration need and must decide which sources warrant the investment. More than integration, this is truly a data engineering task that encompasses trying to understand how to integrate data from external sources with variable data quality and validity. Data tools must offer the ability to investigate data from multiple sources, making it easy to determine quality and improve validity before, during, and after incorporating it into analytics projects.

A perennial challenge for data warehousing, data integration is magnified with big data. Not only are the data sets larger, but so are the number and variety of sources. Traditional data integration approaches will not scale because of the sheer number, variety, and increasingly dynamic nature of data sources.

Ovum view

To use big data for exploratory analytics, enterprises need integration technology that is flexible enough to handle different data sources, regardless of its origin. To that end, big data integration tools must work with diverse underlying architectures encompassing appliances, in-memory computing environments, cloud-based systems, and with varying sources encompassing databases, flat files, and machine data. Depending on analysis requirements, the systems should be able to analyze data at its source or move data between relational and non-relational stores. They also must be able to adapt to events or streams of data (that may or may not be persisted) – what Ovum calls fast data. Integration is not an end in itself; once integrated, the system must be able to harvest data from myriad systems and business applications to blend such data on the fly and perform analysis.

How to start now to meet the challenges

Fast, flexible technology

Oracle's analytics ecosystem provides technology that enables organizations to balance the equation: users can get the access they need to the data they want, and IT can continue to provide the governance and security the organization expects.

Changing the access equation: Oracle Analytics Cloud

The cloud has emerged as one of the most effective ways for enterprises to adopt new technology or update existing solution capabilities. Given that many enterprises are looking to adopt new and refreshed analytical capabilities, the cloud is a logical means to access them. Oracle Analytics Cloud (OAC) is Oracle's delivery of a growing range of key analytical capabilities through the cloud. The first two major components of this expanding cloud portfolio are Business Intelligence Cloud Service, which offers BI technology-as-a-service, and Transactional Business Intelligence, which provides embedded analytical capability for Oracle's software-as-a-service applications. . From this initial offering, near-term expansion includes availability of cloud-delivered Big Data Discovery and Big Data (Hadoop-as-a-service).

Engaging with the data: Visual Analyzer

A recent addition to BI Cloud Service, Visual Analyzer offers organizations a unique combination of highly visual self-service discovery together with industrial-strength operational analytics. Visual Analyzer is not just about eye-catching visuals, it includes the ability to blend different data types and perform complex analysis in a point-and-click interface, all in the context of a governed data environment. Freeing new users to visually explore corporate data (and other sources) is, perhaps, the primary driver to delivering greater value from that data – to do so in a governed way is the means by which enterprises can adopt these capabilities with confidence.

Making Big Data easier: Oracle Big Data Discovery

One of the biggest hurdles in the widespread adoption of Hadoop-based big data analytics is the level of complexity those solutions and the underlying data can present. While there is a rising demand for PhD-level data scientists, the real need is to make the power of big data available to the mainstream, Excel-wielding, business users.

Oracle Big Data Discovery provides the visual face of analytics in Hadoop. Through a web browser, users can explore different data sets and conduct analytics. Oracle Big Data Discovery uses machine learning algorithms to simplify the process of working with less familiar data sets on Hadoop, helping point users to the right data sets and signals to look for. For instance, in a traditional web shopping cart abandonment analysis, the tool can identify potential failure signals in weblog data. Through further exploration, a user can uncover the source of the problem. Oracle Big Data Discovery is available today as part of Oracle's Big Data Appliance and also runs on commodity hardware; in future it will be available as a service in the Oracle Public Cloud as well.

Extracting value from the enterprise: Oracle Business Intelligence Enterprise Edition

Oracle Business Intelligence Enterprise Edition (OBIEE) is relied on by enterprises day-to-day as one of the tools used to effectively manage their business. It delivers a wide range of BI capabilities that executives, power users, and a host of analysts and business users expect, and incorporates the enterprise-grade security, reliability, and scalability standards that organizations demand from a vendor like Oracle.

Expected standards are not, however, enough to satisfy the new needs of an expanded user base. As discussed previously, visualization and ease of use are the keys to satisfying users, who expect a high degree of interactivity coupled with simplicity and access.

Sophisticated analytics across the entire analytics spectrum

An expanded range of approaches to handle a growing range of data challenges necessitates a toolbox of technically sophisticated, yet easy-to-use and -deploy analytical capabilities. Few software vendors are in a position to offer this breadth of capability, and it is only among the longest-established vendors of analytical technology that enterprises will find a suite capable of meeting the majority of these requirements. Oracle's capabilities and heritage in the delivery of solutions that address the management of data and its transformation to actionable business insight make it one of those vendors.

Data governance for the new analytics

Oracle takes a holistic approach to governance by combining its governance tools to present a single view of enterprise security to the customer. These include:

- **Data protection**: Oracle offers comprehensive data encryption capabilities for protecting data privacy and meeting regulatory requirements. It includes enterprise-grade authentication, authorization, and auditing that can greatly simplify the process of securing Hadoop deployments.
- **Extension of Oracle database security to Hadoop and NoSQL through Oracle Big Data SQL:** Oracle Big Data SQL enables users to run a single query across multiple data sources, extending the advanced security capabilities of Oracle Database to Hadoop and NoSQL. Oracle Database security solutions offered include data redaction, privilege analysis, and controls that manage user access to data per policies.

For a self-service environment to be secure, it is important that there is provision for role-based user accounting for access control, authorization, and authentication. As far as possible, this capability should be internal to the self-service solution. Using Oracle Big Data Appliance and/or Oracle Big Data SQL, administrators are given the same set of tools that are used to manage permission in the SQL database world. Using these tools, administrators can grant individual privileges, create roles and assign specific privileges to them, or create secure application roles.

Commodity hardware is not the only option: Oracle Engineered Systems

The depth and breadth of hardware solutions is expanding, from commodity to dedicated, process-tuned systems. A fit-for-purpose approach that delivers the right level of capability to meet specific data management and analytic needs has become necessary, drawing on the benefits each group of hardware technologies can offer.

Today's IT organizations have a range of options for deployment. Commodity hardware has lowered barriers to implementing systems delivering higher performance and scale, but they don't necessarily crack another issue: accelerating time to benefit. Engineered systems and cloud deployment can accelerate time to benefit by eliminating the requirements for customers to perform time-consuming system configuration and the time and capital expense of physical deployment, respectively. As a

result, both can offer attractive total cost of ownership (TCO) benefits over conventional approaches of buying and installing software and hardware. As such, both have become inviting options for deploying analytic systems.

Oracle provides enterprises with a range of hardware technology from its Engineered Systems portfolio that includes the Big Data Appliance, Exadata Database Machine, and Exalytics In-Memory Machine, among others, and is evolving the portfolio reflecting its commitment to this important capability. The Engineered Systems solutions are not, however, exclusively about addressing the need for high performance – as pre-integrated systems that combine hardware and software, they offer deployment economics and cost of ownership that recognize the value of performance in the context of budgetary reality.

Ovum view

Ovum believes the value-creating opportunities inherent in big data are compelling to the point that enterprises in every industry should be exploring them.

Ovum suggests that the thoughtful application of new technologies in combination with existing investments enables a broader approach that incorporates big data, offering organizations the means to transform potential value to tangible benefit.

These benefits are most easily realized by enabling a far-larger user group to work with data. Self-service discovery, increasingly the dominant approach to business user-friendly analytics, crosses the divide between small and big data, building a bridge that is the means by which enterprises will be successful in the age of big data analytics.

Appendix

Methodology

This white paper was compiled through the use of vendor interviews, extensive secondary research, and analyst insight across Information Management and its related markets.

Authors

Tom Pringle, Practice Leader, IT – Information Management

tom.pringle@ovum.com

Surya Mukherjee, Senior Analyst, IT – Information Management

surya.mukherjee@ovum.com

Tony Baer, Principal Analyst, IT – Information Management

tony.baer@ovum.com

Ovum Consulting

We hope that this analysis will help you make informed and imaginative business decisions. If you have further requirements, Ovum's consulting team may be able to help you. For more information about Ovum's consulting capabilities, please contact us directly at consulting@ovum.com.

Copyright notice and disclaimer

CONTACT US

www.ovum.com
askananalyst@ovum.com

INTERNATIONAL OFFICES

Beijing
Dubai
Hong Kong
Hyderabad
Johannesburg
London
Melbourne
New York
San Francisco
Sao Paulo
Tokyo