ORACLE
BIG DATA

Big Data Discovery: Five Easy Steps to Value

"Big data" could really be called "big frustration."

For all the hoopla about big data being poised to reshape industries—from healthcare to retail to financial services—people in the data trenches are telling a different story. Sure, there is plenty of data. But it's "messy," and it's not clear how to extract the nuggets of insight hidden in the swirls of information. Even when companies do figure it out, the process is cumbersome, time-consuming and overly dependent on highly skilled experts.

So is big data fact or fiction? Can it really be turned into a competitive advantage, or is it just a hyped-up creation of overly zealous reporters and marketers? The answer is that big data is the real deal, but it also has been over-hyped. The big data landscape has been confusing, and organizations have struggled to collect, store, prepare, query and analyze the data, let alone get insights out of it. As Forrester Research says in a recent report, "Big data rhetoric is at an all-time high." Clearly, it's time for a new approach.

ORACLE

> “Big data rhetoric is at an all-time high.”
>
> --Forrester Research

Oracle Big Data Discovery, the visual face of Hadoop, is the game-changing platform that eliminates all this confusion by creating a clear and quick path to value. Using visual analytic capabilities, Oracle Big Data Discovery works natively with Hadoop to transform raw data rapidly into business insight in five easy phases (see Figure 1):

Five Easy Steps from Data to Insight

Find relevant data

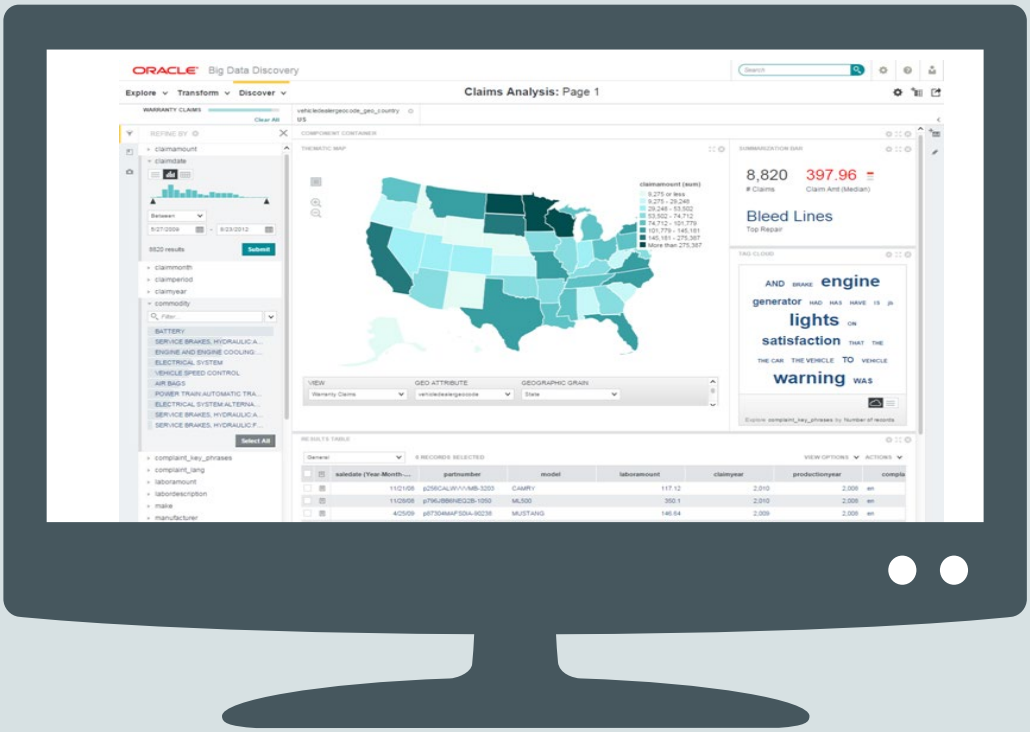
Explore the data to understand its potential

Transform and enrich the data to make it ready for analysis

Discover powerful new insights

Share those insights for enterprise leverage

Figure 1



Hadoop Hoopla

In the eyes of many businesses, big data and Hadoop go hand-in-hand. The Apache open-source software framework can store and analyze a staggering amount of information in many types of formats, such as videos, maps, documents and tweets. With its open-source origins and inexpensive hardware, Hadoop has been quickly adopted as a savior for handling the mass of unstructured data that is critical to achieving value from big data.

No wonder the global Hadoop market is expected to expand from $1.5 billion in 2012 to $50.2 billion by 2020, an annual growth rate of 58.2 percent, according to Researchbeam.[2]

Exercise in Frustration

But despite the justified praise heaped on Hadoop, getting true analytics and insights from the technology remains a challenge. There are a few reasons for this:

+ Messy Data

Even though big data is enticing in concept, in reality much of the data is messy, with uneven quality and data sets that weren't designed to be put together. As a result, data scientists spend up to 80 percent of their time wrestling with the data rather than generating new business insights.[3]

+ Troublesome Traditional Tools

Traditional business intelligence tools are easily overwhelmed by the variety, volume and velocity of big data in Hadoop. Geared toward the tidy world of data warehouses, these tools are based on the notion that users will know what questions they want answered up-front. Traditional tools don't allow for free-form exploration, and they don't know how to answer questions no one has ever thought to ask—which is the very point of big data analytics.

In their efforts to resolve these issues, some companies are turning to emerging point solutions designed for Hadoop; however, these tools lack the end-to-end functionality necessary for turning raw data into analytics. Consequently, analysts constantly need to log in and out of different tools, disrupting their workflow and inhibiting their agility.

> 140,000 to 180,000 data scientist positions will remain unfilled by 2018
>
> --McKinsey Global Institute

+ Data Skills Gap

Specialized tools for working with data in Hadoop require the skills of highly trained data scientists. The trouble is, the experts who are capable of handling the full alchemy of big data—including statistics, programming, SQL, ETL and domain knowledge—are in short supply.

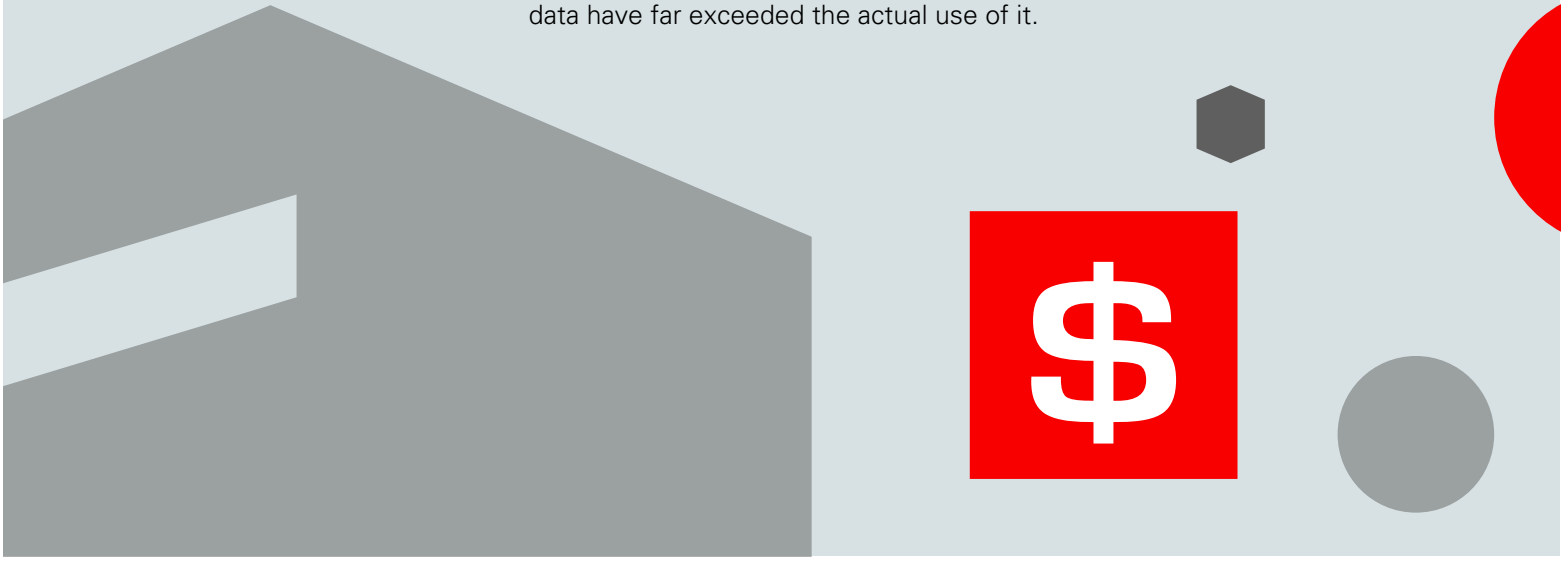
According to McKinsey Global Institute, 140,000 to 180,000 data scientist positions will remain unfilled by 2018.[4] Given this talent gap, there's little wonder that 37 percent of respondents in an Accenture study say they lack the talent to run big data and analytics on an ongoing basis.[5]

Meanwhile, companies' traditional analysts often are left to sit on the sidelines, unable to collaborate with the data scientists. All in all, it's a poor use of people power.

> 37 percent of companies feel they do not have the talent to run big data and analytics on an ongoing basis.
>
> --Accenture

+ Questionable Value

Without a clear path to value, even companies that believe in the transformative power of big data are left to question how to achieve it. In an ROI-focused world, it's difficult to justify large technology investments for big data analytics based on a "gut sense" of a payoff. When viewed in this light, the big data movement can seem like so much hype. And certainly, the reams of reports and articles about big data have far exceeded the actual use of it.



Five Phases of Value

Oracle Big Data Discovery addresses the issues that have stymied big data analytics. With it, companies finally are able to leverage their raw data in Hadoop to provide demonstrable value. Oracle Big Data Discovery delivers value through a five-phase process: find, explore, transform, discover and share.

+ Find: Pinpoint Relevant Data

A retail analyst who wants to improve the results of a marketing campaign has lots of potential data to sift through—customer tweets, loyalty program details, contact center complaints and more. However, determining which of that data is timely and trustworthy isn't easy.

Using the intuitive interface of Oracle Big Data Discovery, the analyst can navigate a rich catalog of all the raw data in a Hadoop cluster and quickly identify what's relevant. Searching the data is as easy as shopping online.

+ Explore: Understand Data Potential

Understanding the potential value of data consumes a lot of analysts' time. For instance, an analyst for an auto manufacturer seeking to streamline its manufacturing processes would likely endure many false starts when exploring the mass of information related to the engine-building process, from poorly scheduled lunch breaks to disconnects between suppliers.

Oracle Big Data Discovery dramatically speeds the data exploration process. Analysts can sort by information potential, with the most interesting attributes appearing first. In addition, analysts can easily experiment with different combinations of data to understand correlations, so they can rapidly determine whether the data set is worthy of more attention. The system also helps them quickly get a handle on data quality and other key elements, preventing time and money from being wasted on projects with limited potential.

+ Transform: Intuitive, User-Driven Data Wrangling

Typically, data in Hadoop needs to be manipulated and prepared before it can be used for analytics. With Oracle Big Data Discovery, analysts use an intuitive spreadsheet-like approach to transform big data for use in analytics. At the same time, the data can be enriched to infer location and language or detect topics, themes and sentiment buried in the raw text.

Rather than spending 80 percent of their time on data preparation, analysts can quickly transform even massive volumes of big data, making it available for the entire enterprise and freeing them to spend the bulk of their time on analytics.

+ Discover: Unleash Creativity

Discovering big data insights requires creativity, which can be difficult to hire for or develop in-house. With Oracle Big Data Discovery, enterprises can get more out of their analytics talent through tools that automatically blend data for deeper perspectives and help analysts see new patterns in rich, interactive data visualizations.

For example, if a telecom analyst wants to investigate the reasons for customer churn, he can use Oracle Big Data Discovery to mash up or join different data sets. This will reveal a whole new perspective; for example, it might show that customers in a certain geographic region using a certain handset are canceling their accounts because of a technical glitch that is disrupting service.

The data also can be filtered with keyword search and guided navigation, providing the consumer-like experience that users increasingly demand today from even the most complex enterprise technology.

+ Share: Drive Collaboration

Oracle Big Data Discovery fulfills the promise of democratizing big data analytics by enabling the results to be shared and published. Suddenly, the information can become a focal point of enterprise collaboration and collective discovery.

Entire teams can share projects, bookmarks and galleries of snapshots, enabling them to collaborate and iterate. Analysts, meanwhile, can publish their data transformation and enrichment results back to Hadoop, securing the work they've done to maximize the value of the data.

Excelling in the Data Ecosystem

Oracle Big Data Discovery not only unlocks the analytic power of big data, it also preserves existing infrastructure investments and plays well in the corporate technology ecosystem. It blends with other tools that access Hadoop, such as advanced statistical tools (Oracle R for Hadoop), business intelligence tools (Oracle Business Intelligence) or any other product designed to work with data in Hadoop (Pig, Hive or Oracle Big Data SQL).

Consequently, Oracle Big Data Discovery removes technical barriers, leveraging the power of Hadoop for maximum scale and performance while at the same time integrating with the wider big data ecosystem.

+ Big Data: Finally, the Value Is Now

Business leaders understandably have a love/hate relationship with big data. Although wary of the hype, they also are acutely aware that their future success relies on it. Some 79 percent of executives say that failing to embrace big data puts companies at risk of losing their competitive position, according to an Accenture survey. Those executives expect big data to impact their business in multiple strategic ways over the next five years (see Figure 2).[6]

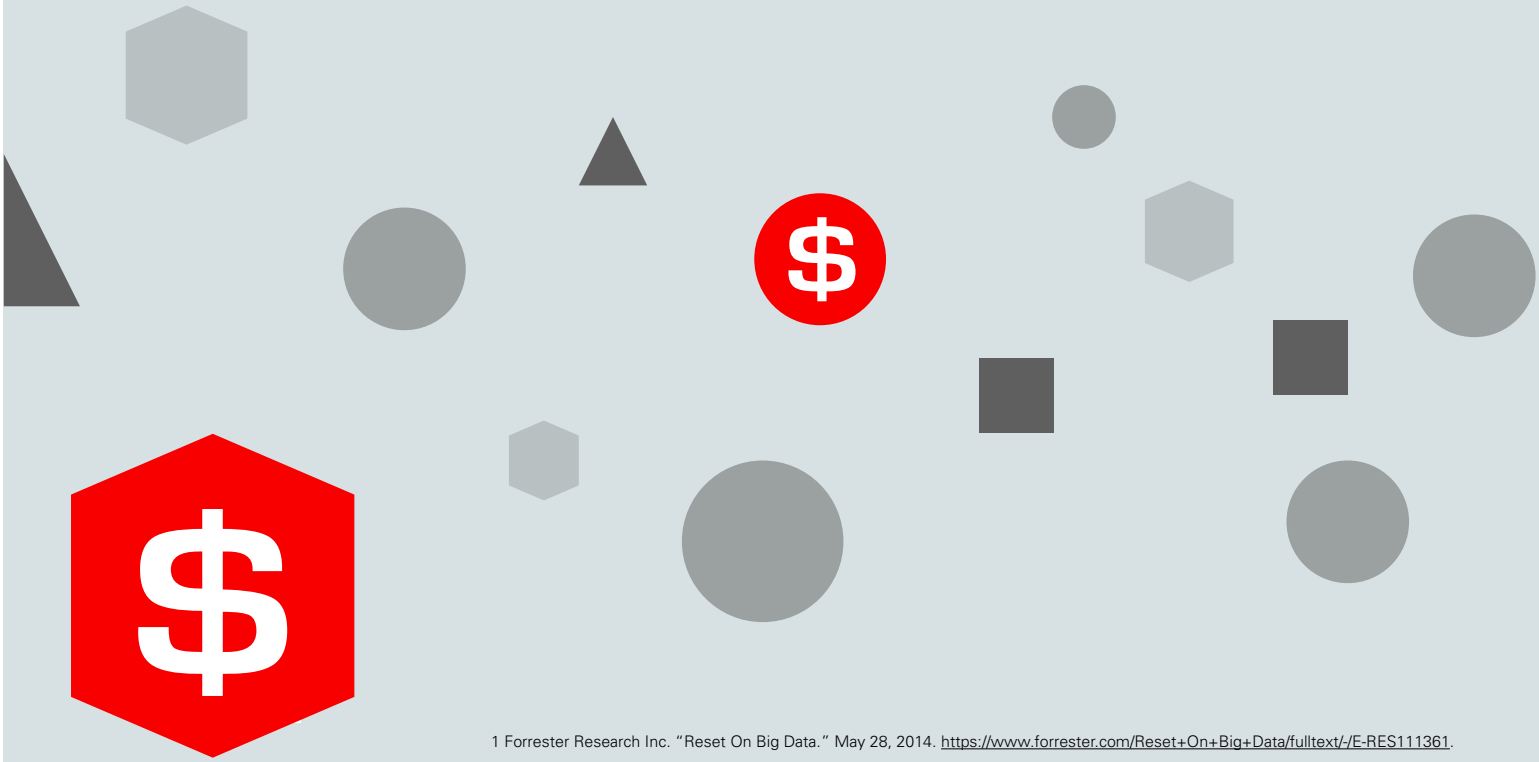
Potential for Transformation

	Top Impact	Top 3 Impact
Impacting customer relationships	37%	63%
Redefining product development	26%	58%
Changing the way we organize operations	15%	58%
Making the business more data-focused	8%	48%
Optimizing the supply chain	9%	47%
Fundamentally changing the way we do business	5%	27%

Figure 2 Source: Accenture

Fortunately, there is now a way to move quickly from big data hype to big data value. With Oracle Big Data Discovery, businesses can rapidly turn raw data into actionable insights without relying solely on specialized talent. They can extend the alchemy of big data to more people in the enterprise, creating entire teams that work collectively on insight discovery, improving the efficiency and extending the expertise of their existing analytics staff. In short, with Oracle Big Data Discovery, businesses now have a platform that can enable their own transformation; they can embrace analytics and use new insights—hidden in big data—to make strategic and game-changing decisions quickly, ensuring their success well into the future.

> 79 percent of executives believe that "companies that do not embrace big data will lose their competitive position and may even face extinction."
>
> --Accenture



1 Forrester Research Inc. "Reset On Big Data." May 28, 2014. https://www.forrester.com/Reset+On+Big+Data/fulltext/-/E-RES111361.

2 Researchbeam. "Global Hadoop Market." March 2014. http://www.researchbeam.com/hadoop-market.

3 Lohr, Steve. "For Big-Data Scientists, 'Janitor Work' Is Key Hurdle to Insights." The New York Times, Aug. 17, 2014. http://www.nytimes.com/2014/08/18/technology/for-big-data-scientists-hurdle-to-insights-is-janitor-work.html?_r=0.

4 Boulton, Clint. "Universities Can't Train Data Scientists Fast Enough for CIOs." The Wall Street Journal, Nov. 10, 2014. http://blogs.wsj.com/cio/2014/11/10/for-cios-universities-cant-train-data-scientists-fast-enough/.

5 Accenture. "Big Success With Big Data." April 2014. http://www.accenture.com/SiteCollectionDocuments/PDF/Accenture-Big-Data-POV.PDF.

6 Accenture. "Big Success With Big Data." April 2014. http://www.accenture.com/SiteCollectionDocuments/PDF/Accenture-Big-Data-POV.PDF.