

TDWI RESEARCH

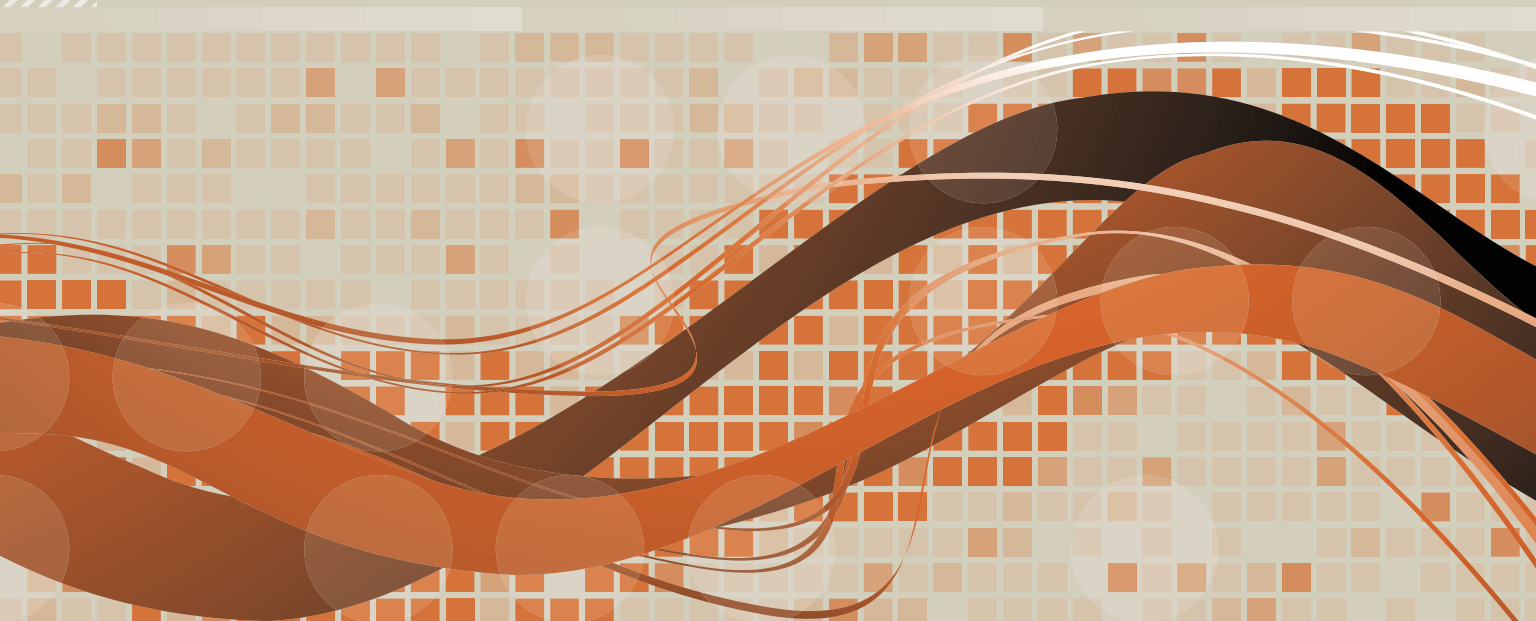
TDWI BEST PRACTICES REPORT

FOURTH QUARTER 2011

EXECUTIVE SUMMARY

BIG DATA ANALYTICS

By Philip Russom



Executive Summary

Big data used to be a technical problem. Now it's a business opportunity.

Oddly enough, big data was a serious problem just a few years ago. When data volumes started skyrocketing in the early 2000s, storage and CPU technologies were overwhelmed by the numerous terabytes of big data—to the point that IT faced a data scalability crisis. Then we were once again snatched from the jaws of defeat by Moore's law. Storage and CPUs not only developed greater capacity, speed, and intelligence; they also fell in price. Enterprises went from being unable to afford or manage big data to lavishing budgets on its collection and analysis.

Today, enterprises are exploring big data to discover facts they didn't know before. This is an important task right now because the recent economic recession forced deep changes into most businesses, especially those that depend on mass consumers. Using advanced analytics, businesses can study big data to understand the current state of the business and track still-evolving aspects such as customer behavior.

Big data is not just big. It's also diverse data types and streaming data.

If you really want the lowdown on what's happening in your business, you need large volumes of highly detailed data. If you truly want to see something you've never seen before, it helps to tap into data that's never been tapped for business intelligence (BI) or analytics. Some of the untapped data will be foreign to you, coming from sensors, devices, third parties, Web applications, and social media. Some big data sources feed data unceasingly in real time. Put all that together, and you see that big data is not just about giant data volumes; it's also about an extraordinary diversity of data types, delivered at various speeds and frequencies.

Big data analytics is the application of advanced analytic techniques to very big data sets.

Note that two technical entities have come together. First, there's big data for massive amounts of detailed information. Second, there's advanced analytics, which is actually a collection of different tool types, including those based on predictive analytics, data mining, statistics, artificial intelligence, natural language processing, and so on. Put them together and you get big data analytics, the hottest new practice in BI today.

Of course, businesspeople can learn a lot about the business and their customers from BI programs and data warehouses. But big data analytics explores granular details of business operations and customer interactions that seldom find their way into a data warehouse or standard report. Some organizations are already managing big data in their enterprise data warehouses (EDWs), while others have designed their DWs for the well-understood, auditable, and squeaky clean data that the average business report demands. The former tend to manage big data in the EDW and execute most analytic processing there, whereas the latter tend to distribute their efforts onto secondary analytic platforms. There are also hybrid approaches.

There are many types of vendor products to consider for big data analytics. This report discusses the types.

Regardless of approach, user organizations are currently reevaluating their analytic portfolios. In response to the demand for platforms suited to big data analytics, vendors have released a slew of new product types including analytic databases, data warehouse appliances, columnar databases, no-SQL databases, distributed file systems, and so on. There is also a new slew of analytic tools.

This report drills into all the aspects of big data analytics mentioned here to give users and their business sponsors a solid background for big data analytics, including business and technology drivers, successful business use cases, and common technology enablers. The report also uses survey data to project the future of the most common tool types, features, and functions associated with big data analytics, so users can apply this information to planning their own programs and technology stacks for big data analytics.



DOWNLOAD THE FULL REPORT

>> tdwi.org

EXECUTIVE SUMMARY

BIG DATA ANALYTICS

By Philip Russom

 **DOWNLOAD THE FULL REPORT**

>> tdwi.org

Table of Contents

Research Methodology and Demographics	3
Executive Summary	4
Introduction to Big Data Analytics	5
Defining Advanced Analytics as a Discovery Mission	5
Defining Big Data Via the Three Vs.	6
Defining Big Data Analytics	8
Why Put Big Data and Analytics Together Now?	9
The State of Big Data Analytics	10
Big Data Analytics Adoption	10
Benefits of Big Data Analytics	10
Barriers to Big Data Analytics	11
Big Data: Problem or Opportunity?	12
Organizational Issues.	13
Ownership and Control of Big Data Analytics	13
Big Data Analytics Can Have a Departmental Focus	14
Job Titles for Big Data Analytics	14
Best Practices in Big Data Analytics	15
Volume Growth of Analytic Big Data	15
Managing Analytic Big Data	16
Data Types for Big Data	17
Refresh Rates for Analytic Data	19
Replacing Analytics Platforms	20
Tools, Techniques, and Trends for Big Data Analytics	22
Potential Growth versus Commitment for Big Data Analytics Options	24
Trends for Big Data Analytics Options	26
Vendor Products for Big Data Analytics.	31
Recommendations	34

Research Methodology and Demographics

Report Scope. According to TDWI survey data, a new flood of user organizations is currently commencing or expanding solutions for analytics with big data. To supply the demand, vendors have recently released numerous new products and functions, specifically for advanced forms of analytics (beyond OLAP and reporting) and analytic databases that can manage big data. While it's good to have options, it's hard to track them and determine in which situations they are ready for use. The purpose of this report is to accelerate users' understanding of the many new tools and techniques that have emerged for analytics with big data in recent years. It will also help readers map newly available options to real-world use cases.

Survey Methodology. In May 2011, TDWI sent an invitation via e-mail to the data management professionals in its database, asking them to complete an Internet-based survey. The invitation was also distributed via Web sites, newsletters, and publications from TDWI and other firms. The survey drew responses from almost 360 survey respondents. From these, we excluded incomplete responses and respondents who identified themselves as academics or vendor employees. The resulting completed responses of 325 respondents form the core data sample for this report.

Survey Demographics. The majority of survey respondents are corporate IT professionals (58%), whereas the others are business sponsors or users (22%) and consultants (20%). We asked consultants to fill out the survey with a recent client in mind.

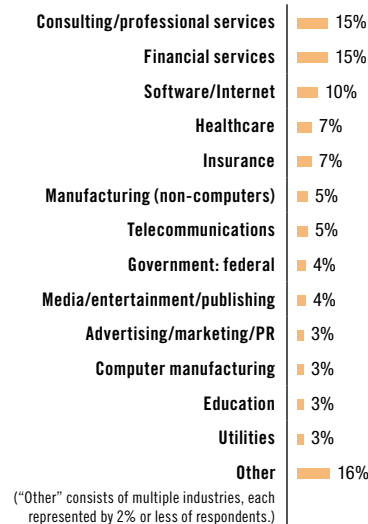
The consulting (15%) and financial services (15%) industries dominate the respondent population, followed by software (10%), healthcare (7%), insurance (7%), and other industries. Most survey respondents reside in the U.S. (56%) or Europe (17%). Respondents are fairly evenly distributed across all sizes of companies and other organizations.

Other Research Methods. In addition to the survey, TDWI Research conducted many telephone interviews with technical users, business sponsors, and recognized data management experts. TDWI also received product briefings from vendors that offer products and services related to the best practices under discussion.

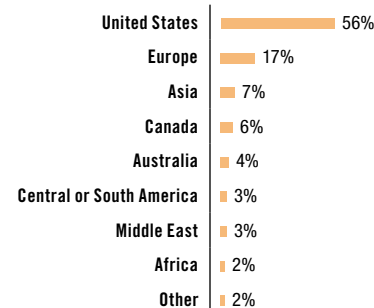
Position



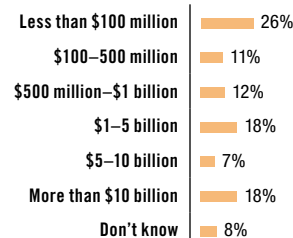
Industry



Geography



Company Size by Revenue



Based on 325 survey respondents.

About the Author



PHILIP RUSSOM is director of TDWI Research for data management and oversees many of TDWI's research-oriented publications, services, and events. He is a well-known figure in data warehousing and business intelligence, having published over five hundred research reports, magazine articles, opinion columns, speeches, Webinars, and more. Before joining TDWI in 2005, Russom was an industry analyst covering BI at Forrester Research and Giga Information Group. He also ran his own business as an independent industry analyst and BI consultant and was a contributing editor with leading IT magazines. Before that, Russom worked in technical and marketing positions for various database vendors. You can reach him at prussom@tdwi.org, @prussom on Twitter, and on LinkedIn at [linkedin.com/in/philiprussom](https://www.linkedin.com/in/philiprussom).

About TDWI

TDWI, a division of 1105 Media, Inc., is the premier provider of in-depth, high-quality education and research in the business intelligence and data warehousing industry. TDWI is dedicated to educating business and information technology professionals about the best practices, strategies, techniques, and tools required to successfully design, build, maintain, and enhance business intelligence and data warehousing solutions. TDWI also fosters the advancement of business intelligence and data warehousing research and contributes to knowledge transfer and the professional development of its members. TDWI offers a worldwide membership program, five major educational conferences, topical educational seminars, role-based training, onsite courses, certification, solution provider partnerships, an awards program for best practices, live Webinars, resourceful publications, an in-depth research program, and a comprehensive Web site: tdwi.org.

About the TDWI Best Practices Reports Series

This series is designed to educate technical and business professionals about new business intelligence technologies, concepts, or approaches that address a significant problem or issue. Research for the reports is conducted via interviews with industry experts and leading-edge user companies and is supplemented by surveys of business intelligence professionals.

To support the program, TDWI seeks vendors that collectively wish to evangelize a new approach to solving business intelligence problems or an emerging technology discipline. By banding together, sponsors can validate a new market niche and educate organizations about alternative solutions to critical business intelligence issues. Please contact TDWI Research Director Philip Russom (prussom@tdwi.org) to suggest a topic that meets these requirements.

Acknowledgments

TDWI would like to thank many people who contributed to this report. First, we appreciate the many users who responded to our survey, especially those who responded to our requests for phone interviews. Second, our report sponsors, who diligently reviewed outlines, survey questions, and report drafts. Finally, we would like to recognize TDWI's production team: Jennifer Agee, Bill Grimmer, and Denelle Hanlon.

Sponsors

Cloudera, EMC Greenplum, IBM, Impetus Technologies, Kognitio, ParAccel, SAND Technology, SAP, SAS, Tableau Software, and Teradata sponsored the research for this report.

TDWI RESEARCH

TDWI Research provides research and advice for business intelligence and data warehousing professionals worldwide. TDWI Research focuses exclusively on BI/DW issues and teams up with industry thought leaders and practitioners to deliver both broad and deep understanding of the business and technical challenges surrounding the deployment and use of business intelligence and data warehousing solutions. TDWI Research offers in-depth research reports, commentary, inquiry services, and topical conferences as well as strategic planning services to user and vendor organizations.



1201 Monster Road SW T 425.277.9126
Suite 250 F 425.687.2842
Renton, WA 98057-2996 E info@tdwi.org

tdwi.org