TDWI BEST PRACTICES REPORT   THIRD QUARTER 2012

CUSTOMER ANALYTICS IN THE AGE OF SOCIAL MEDIA

By David Stodder

tdwi.org

Research Sponsors

Greenplum, a division of EMC

IBM

Informatica

SAP

SAS

Tableau Software

Teradata (including Teradata Aster and Aprimo)

Vertica, an HP Company

CUSTOMER ANALYTICS IN THE AGE OF SOCIAL MEDIA

By David Stodder

Table of Contents

About the Author

DAVID STODDER is director of TDWI Research for business intelligence. He focuses on providing research-based insight and best practices for organizations implementing BI, analytics, performance management, data discovery, data visualization, and related technologies and methods. He is the author of a TDWI Best Practices Report on mobile BI and analytics and TDWI Checklist Reports on data discovery and information management. He has chaired TDWI conferences focused on BI agility and big data analytics. Stodder has provided thought leadership on BI, analytics, information management, and IT management for over two decades. Previously, he served as vice president and research director with Ventana Research. He was the founding chief editor of *Intelligent Enterprise* and served as editorial director for nine years. You can reach him at dstodder@tdwi.org.

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, analytics, and data warehousing 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 BI professionals.

To support the program, TDWI seeks vendors that collectively wish to evangelize a new approach to solving BI, analytics, and data warehousing 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 David Stodder (dstodder@tdwi.org) to suggest a topic that meets these requirements.

Acknowledgments

Sponsors

Research Methodology and Demographics

Report Scope. Customer analytics and intelligence initiatives are undergoing major changes with the explosion in detailed customer behavior data, including that generated by social media. Today, organizations implement a variety of technologies to analyze customer data, including business intelligence (BI) tools, data warehouses, analytic databases, predictive analytics, data mining, data discovery, text analytics, social media analytics, customer data integration and master data management, and more. This report will examine organizations' experiences with customer analytics technologies in the age of social media and recommend best practices for improving customer intelligence and reaching customer-centric goals.

Survey Methodology. In February 2012, TDWI sent an invitation via e-mail to business and IT executives, business and data analysts, data scientists, BI, and data management professionals in its databases, asking them to complete an Internet-based survey. The invitation was also delivered via Web sites, newsletters, and publications from TDWI and other firms. The survey analysis drew from a total of 459 responses. A total of 377 completed every question. Answers from respondents who did not complete every question but provided enough answers for their input to be valuable are included in the results. Respondents who left too many questions unanswered or did not provide adequate responses were excluded. Thus, some questions have different numbers of responses, which is noted in the figures. Respondents who identified themselves as academics or students were excluded.

Survey Demographics. The largest percentage of survey respondents are business executives, sponsors, business or data analysts, and users (47%). The percentage of data and IT professionals (43%) was not far behind. Independent consultants accounted for 10%. Consulting and professional services respondents (22%) made up the largest industry segment, followed by financial services (12%), software and Internet services (11%), insurance (7%), and retail, wholesale, and distribution (7%). Most respondents reside in the U.S. (50%) or Europe (18%), but other regions account for 32%.

Other Research Methods. TDWI also conducted telephone interviews with business and IT executives, data analysts and scientists, BI directors, customer marketing specialists, and recognized BI, analytics, and data warehousing experts. TDWI also received briefings from vendors that offer related products and services.

Position

Business executives, sponsors/users	47%
Data and IT professionals	43%
Consultants	10%

Industry

Consulting/professional services	22%
Financial services	12%
Software/Internet services	11%
Insurance	7%
Retail/wholesale/distribution	7%
Telecommunications	7%
Education	5%
Healthcare	4%
Manufacturing (noncomputer)	4%
Computer or network manufacturing	3%
Government (federal)	2%
Advertising/marketing/PR	2%
Other*	14%

* "Other" consists of multiple industries, each represented by 2% or less of respondents.

Geography

United States	50%
Europe	18%
Asia/ Pacific Islands	10%
Canada	7%
South Asia	5%
Australia/New Zealand	5%
Africa	3%
Middle East	2%

Number of Employees

More than 100,000	12%
10,000–100,000	18%
1,000–10,000	32%
100–1,000	17%
Less than 100	21%

Based on 459 survey respondents.

Executive Summary

Becoming "customer centric" is a top priority today, and for good reason: as if it weren't important enough that customers buy products and contract for services, they now do much more than simply buy. Customers participate in social media networks and chat rooms; they write blogs and contribute to comment sites; and they share information through sites such as YouTube and Flickr. Their activities and expressions not only reveal personal buying behavior and interests, but they also bring into focus their influence on purchasing by others in their social networks.

By creating data trails in social media networks, participants generate new types of data that hold great potential for customer insight.

Social media networks have given organizations exciting new channels for marketing and customer engagement. Just as important, however, is that by creating data trails in social media networks, participants generate new types of data that hold great potential for customer insight. Advanced analytics technologies, modern data management and integration systems, and growing "big data" sources are enabling organizations to gain far greater depth and breadth of knowledge about customers, influencers, prospects, and the competitive marketplace. Social media data analysis can expand customer analytics by providing an unfettered, outside-looking-in view of an organization's brands, products, services, and competitors.

This TDWI Best Practices Report examines organizations' current practices and future plans for customer analytics technology implementations, with a special focus on how organizations are adapting to the knowledge potential as well as challenges of social media networks. The report offers recommendations for achieving greater return on investment (ROI) from customer analytics processes. This higher return is important because marketing functions are being held more accountable for the effectiveness of campaigns in delivering on financial objectives and their measurable impact on overall business growth. Business functions need customer insights not just for marketing campaigns, but also for informing the organization's sales, service, support, product development, and other key functions about customer feedback and trends.

Customer analytics, seasoned with insight from social media data, can enable organizations to make faster strides in predicting retention, attrition, and return rates, with the goal of reducing customer churn. Analytics can improve how organizations decide on characteristics for customer segmentation; social media can provide clues to emerging characteristics for definition of new segments. Firms can employ predictive modeling to test and learn from campaigns so that they are able to select the most persuasive offers to put in front of the right customers at the right time.

This report explores how organizations can use analytics to discover much sooner which message, interaction, or campaign actually had the most influence.

Speed is a competitive advantage in marketing. If an organization can analyze data faster and feed insights more frequently to its customer relationship management (CRM), marketing optimization, and campaign management applications, it will realize advantages over firms that are locked into seasonal campaigns, do not analyze customer behavior data, or are too slow in analyzing it. This report explores how organizations can use analytics to discover much sooner which message, interaction, or campaign actually had the most influence on triggering a customer purchase rather than attributing it blindly to the last thing that a customer did or saw.

Organizations are thus focused on evaluating and deploying a new breed of analytics and data management technologies to increase the speed of analysis and reduce latency in applying knowledge to marketing actions. These technologies include analytic databases, columnar databases, Hadoop and MapReduce, customer master data management (MDM), and predictive analytics tools. Nontechnical users in marketing and other functions are implementing social media analytics, business intelligence, data discovery, and visual analysis products to allow them to consume insights more easily and explore data on their own.

Customer Analytics and the Social Media Frontier

John Wanamaker, the nineteenth-century U.S. department store merchandiser often called the father of advertising, once famously said, "Half the money I spend on advertising is wasted; the trouble is, I don't know which half." Today, organizations are not dealing in just "halves"; with multiple channels to choose from for advertising, marketing, and customer engagement, they have complicated decisions to make about where to devote fractions of their resources to achieve the greatest impact. Organizations that are blind to how customers prefer to be engaged in each channel will be less efficient and effective with their marketing and fall short of goals for ROI. It is critical for organizations to build customer knowledge so that they know which actions in which channels are truly the most influential and relevant to a customer's purchase. They need to integrate information views effectively to support multichannel strategies that require coordinated actions across channels.

With multiple channels to choose from, customers today can make choices and dictate how they want to consume information about products and services. Customers are empowered; gone are the days when a market maker and business leader such as automaker Henry Ford could declare that "any customer can have a car painted any color that he wants so long as it is black." With customers exhibiting less loyalty and more selectivity about products, organizations need to be smarter. They need timely data and analytics to avoid losing customers to competitors who may be just a click away with products and services that are more in tune with customer preferences. Better customer intelligence is thus vital to more than just the marketing function; product development, services, and other functions in the organization need it.

Customer analytics, the focus of this TDWI Best Practices Report, is about implementing technologies and methods for knowing more about customers' behavior, their paths to purchasing goods and services, and what actions will engender greater loyalty among those who are most valuable. The goal is to derive accurate information and insights from traditional transaction and service data as well as different types of behavioral data sources so that organizations can better identify, attract, interact with, and retain customers.

Single Views of All Information. Many organizations seek the most comprehensive, 360-degree view of a customer rather than limit their understanding to one source or a diffuse and incomplete picture drawn from disconnected silos of information. They wish to integrate insights from behavioral data sources with those drawn from more traditional data sources such as transactions and service records. They also need to integrate different sources of behavioral data. Some of this data is generated by customers responding and reacting to an organization's activity, including product introductions and services, sales, and fulfillment processes through its own channels. Increasingly important, however, are other data sources generated by customers' independent activity in online social media networks, or offline as they visit physical stores and view messages on media such as billboards.

For years, leading organizations in retail, telecommunications, hospitality, gaming, financial services, and other industries have been implementing data mining tools and methods to discover buying patterns, affinities, and other indicators so they can be predictive about—and proactive with—customers. Along with data mining tools, additional technologies that have played a big role include enterprise data warehousing, business intelligence, and customer data integration (often called "customer MDM"). Organizations are now adding newer technologies, such as Hadoop and MapReduce, text analytics, and specialized tools for social media listening and activity analysis to discover insights from unstructured and semi-structured data.

Organizations that are blind to how customers prefer to be engaged in each channel will be less efficient and effective with their marketing and fall short of goals for ROI.

Many organizations seek the most comprehensive, 360-degree view of a customer rather than limit their understanding to one source.

Although organizations may not be able to replicate the mom-and-pop store, they can restore some of that intimacy if they use information effectively.

Customer analytics has risen in importance as businesses have grown in size and diversity and struggle to understand and anticipate customers' desires and concerns. Bigness has come at the cost of intimacy; although organizations may not be able to replicate the mom-and-pop store where customers are neighbors and known by name and reputation, they can restore some of that intimacy if they use information effectively. Strategic initiatives for one-to-one marketing, micro-marketing, finer segmentation, mass customization, and more depend on data insights from customer analytics.

Stepping into the Age of Social Media

The new frontier for customer analytics is social media. The advent of social media networks and related commenting and information sharing zones is a revolutionary change, full of both potential and challenges. In the social media sphere, customers are influencers, not just generators of sales transactions as seen though point-of-sale and e-commerce systems. Using social media networks, customers can influence each other by commenting on brands, reviewing products, reacting to marketing campaigns and product or service introductions, and revealing shared buying interests. Unlike casual conversations, the commentary and social network connections are recorded and can therefore be analyzed and measured. The result is a data tsunami: the actions and content generated by participants in social media create "big data" sources that are full of potential for tracking and understanding behavior, trends, and sentiments.

Discovery, including what-if analysis, is an important part of customer analytics because users in marketing and other functions do not always know what they are looking for in the data.

The biggest difficulty is filtering out the noise, but not so much that the trends, patterns, and other insights hidden in the raw data are lost through aggregation and filtering. The need to analyze raw, detailed data is a major driver behind the implementation of Hadoop. Organizations need an unstructured place such as Hadoop files to put all kinds of big data in its pure form, rather than in a more structured data warehousing environment. The reason is that depending on the intent of the analysis, what might be considered just "noise" in the raw data from one perspective could be full of important "signals" from a different perspective. Discovery, including what-if analysis, is an important part of customer analytics because users in marketing and other functions do not always know what they are looking for in the data and must try different types of analysis to produce the insight needed. They need to filter out noise, yet not be limited to standard, expected types of information such as what they might receive in a BI report.

Of course, to know how to filter out the noise, organizations need to define their purposes for analyzing social media data. Among the frequent targets for analysis are the following:

- Understanding sentiment drivers

- Identifying characteristics for better segmentation

- Measuring the organization's share of voice and brand reputation compared with the competition

- Determining the effectiveness of marketing touches and messages in buying behavior

- Using predictive analytics on social media to discover patterns and anticipate customers' problems with products or services

Social media data garners attention because it is new and exciting; however, it isn't necessarily the primary customer analytics focus in all organizations. Sources such as transactions, service records, call detail records, and Web clickstreams are often far more important for customer analytics. Data quality can also be a concern with social media. However, few organizations can ignore social media, because it is a living and adaptive ecosystem where consumers are actively participating around the clock and around the world. It is the public square of customer empowerment.

The Importance of Customer Analytics

Before delving further into specific areas of customer analytics and social media's impact, we can consider some context-setting results from TDWI's research about the general importance of customer analytics technologies and methods (Figure 1). Not surprisingly, we found that the business functions or operations for which respondents considered customer analytics most important were marketing (81%, with 52% indicating "very important"), sales and sales reporting (79%, with 45% "very important"), and campaign management (74%, with 47% "very important"). Market research (43% "very important") and customer service and order management (also 43% "very important") were also high among business functions regarded as critical to developers and consumers of customer analytics.

The business functions or operations for which respondents considered customer analytics most important were marketing, sales and sales reporting, and campaign management.

How important is it for your organization to implement customer analytics technologies and methods to benefit the following business functions or operations?
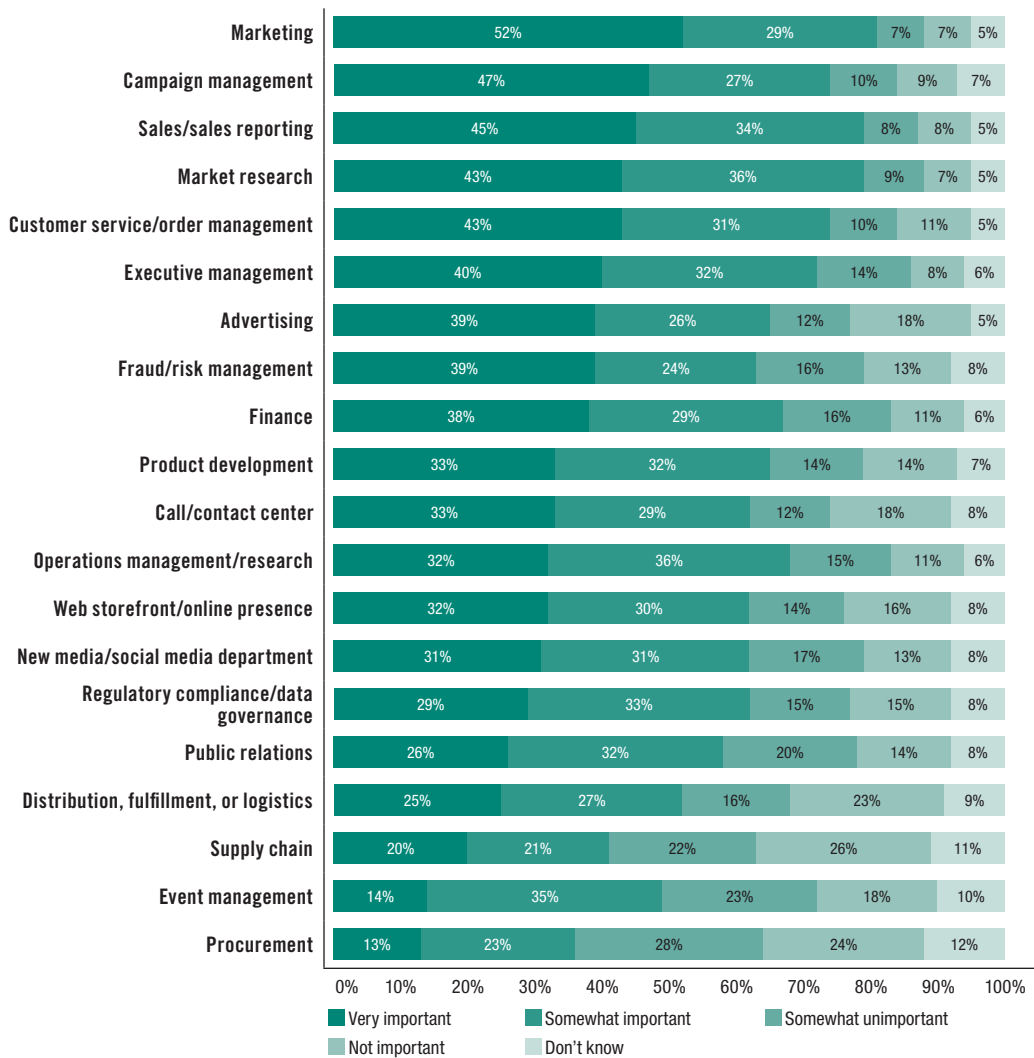

Figure 1. *Based on one answer per business function from 452 respondents.*

Marketing functions are key supporters of "data science," which is the use of scientific methods on data to develop hypotheses and models.

The marketing function, which in most organizations is empowered with the responsibility for identifying, attracting, satisfying, and keeping customers, is clearly the main stage for customer analytics. Marketing functions are becoming increasingly quantitative; they are replacing "gut feel" with data-driven decision making. Data drives the pursuit of efficiency and the achievement of measurable results. Marketing functions are key supporters of "data science," which is the use of scientific methods on data to develop hypotheses and models and apply iterative, test-and-learn strategies to marketing campaigns and related initiatives.

The research shows that sales functions can be important beneficiaries of customer analytics. Sales reports typically focus on providing visibility into the pipeline. Managers can use data insights to improve sales forecasting of potential revenues based on deeper knowledge of priority opportunities, most valued customer segments, and more. The office of finance can also benefit from more sophisticated analytics in sales reports. C-level and finance executives typically consume sales reports; 67% of respondents said that customer analytics was important for the finance function (with 38% indicating "very important"). (See Figure 1, previous page.)

Analytics can help service and order management functions move away from one-size-fits-all approaches.

Customer service and order management can use customer analytics to get a more subtle and substantial view of what actions most impact customer experiences and satisfaction. Contact centers can use customer analytics to help tune performance metrics closer to real time, so that each day's agents are guided, if not incentivized, to interact with customers in beneficial ways. In addition, analytics can help service and order management functions move away from one-size-fits-all approaches to customers and instead tune and tailor interactions more personally based on knowledge of particular types or segments, such as regions or nationalities. Finally, through integrated views of customer data and analytics, service and order management functions are able to work in better synchronicity with the organization's marketing, sales, and other business functions.

Customer Analytics Budgets and IT Control

TDWI Research found that IT and data management (59%) are the most common owners of the budget for customer analytics technologies and services (see Figure 2; respondents could select multiple answers). This reflects the central role played by BI and data warehousing systems, which are often managed by IT. (We will examine research about BI and data warehousing for customer analytics more extensively later in this report.) In interviews, TDWI did discover a growing budget role played by marketing and advertising functions. Nearly two out of five (38%) respondents said that this function has responsibility for the customer analytics budget in their organizations. Executive management (39%) is also a significant player in budgetary decisions.

In your organization, which of the following functions is responsible for the customer analytics technologies and services budget? (Please select all that apply.)

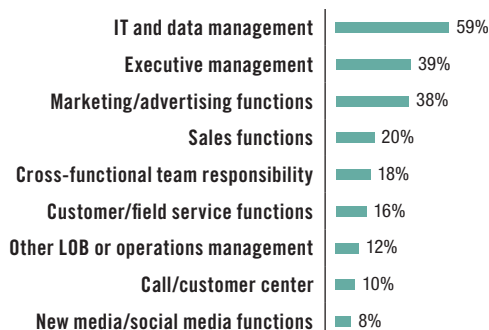| Function	Percent
IT and data management	59%
Executive management	39%
Marketing/advertising functions	38%
Sales functions	20%
Cross-functional team responsibility	18%
Customer/field service functions	16%
Other LOB or operations management	12%
Call/customer center	10%
New media/social media functions	8%

Figure 2. Based on 992 responses from 448 respondents; a little more than two responses per respondent, on average.

Whether located in IT or under the aegis of the corporate marketing function, specialists in customer analytics must often consult with globally distributed, departmental marketing teams as well as other business units to understand key business challenges and opportunities that should be considered in the development of models, algorithms, queries, and data files for analysis. In other words, customer analytics professionals must be able to live in both technology and business worlds and work with diverse teams from not only marketing, but also finance and operations, to develop accurate, consistent, and common metrics for evaluating results. The ability to move across functions is important for delivering holistic, or enterprise, benefits from customer analytics that go beyond marketing.

Customer analytics and the budget for analytic processes are often in the middle of tensions between IT and marketing. In interviews for this report, TDWI found that the growth in analytics implementation by marketing functions is putting stress on relations with IT over control of the data and who develops and runs analytic routines. The iterative, discovery-oriented qualities of predictive modeling and variable development don't fit well with IT's standard approach to gathering all user requirements at once and owning the development of a solution. "IT would ask us to identify the fields we wanted," a marketing data analyst said, "but we had to say, 'Gee, we won't know until we can look at what's available and start playing with it.'"

Data strategies established over many years for delivering BI reporting applications likely do not fit the needs of analytics. Organizations need to focus on resolving tensions between business functions and IT over analytics and data access before internal problems become obstacles to meeting strategic objectives.

Receptivity to Data-Driven Insights

Along with being able to think scientifically about data, models, and variables, customer analytics professionals must have courage: they must test common wisdom and hunches that are often held by powerful people in their organizations, including in executive management. In our research, about one-third (32%) of respondents said there was "strong acceptance" of data-driven customer analytics and 43% indicated "moderate acceptance" over "gut feel" or personal experience (see Figure 3). Only about 5% indicated moderate or strong denial of data-driven customer analytics in their organizations.

How would you rate the level of acceptance by business executives and managers of the use of data-driven customer analytics in place of "gut feel" or personal experience for decision making and process change?

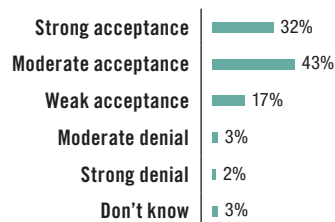|	
Strong acceptance	32%
Moderate acceptance	43%
Weak acceptance	17%
Moderate denial	3%
Strong denial	2%
Don't know	3%

Figure 3. *Based on answers from 457 respondents.*

Organizations are growing more receptive because they are interested in discovering how they can be more effective, not just efficient, in marketing. "Whereas other types of applications for e-commerce, fulfillment, or marketing automation help organizations determine *how* to get things

done (e.g., getting goods delivered at the right time, executing a marketing campaign), customer analytics helps organizations answer *who*, *what*, *when*, *where*, and *why* questions," said Scott Groenendal, program director of customer analytics market strategy for IBM Business Analytics. "They can find answers to questions such as: What channel should I communicate through? When is the best time to target this person, and why would they be receptive to this message?"

The role of customer analytics is not necessarily to replace creativity and personal experience, but to help decision makers come to fact-based conclusions.

Without a doubt, individual creativity and personal experiences will always be critical components of marketing decisions. The role of customer analytics is not necessarily to replace these, but to help decision makers come to fact-based conclusions through better knowledge of the organization's customers and markets. In addition, analytics are critical for scale. Just as automation is necessary to run hundreds or thousands of marketing campaigns, customer analytics processes are important for supplying intelligence and guidance to those automated routines. Customer analytics can provide the brains to match the marketing systems' brawn.

Benefits and Barriers to Implementation

Business Benefits Sought from Customer Analytics

In the words of business management guru Thomas Davenport, "Organizations are competing on analytics not just because they can—business today is awash in data crunchers—but also because they should." Davenport adds, "Business processes are among the last remaining points of differentiation. And analytics competitors wring every last drop of value from those processes."[1] Customer analytics helps organizations determine what steps will give them competitive advantages, increase profitability, and identify waste in business processes.

With products and services becoming commoditized at a rapid rate, customer loyalty can be elusive; innovation must be constant and must immediately address why an organization is losing customers. Information insights from analytics can help organizations align product and service development with strategic business objectives for customer loyalty. In addition, these insights can help organizations be selective in how they deploy marketing campaigns and customer-touch processes so that they emphasize features in new products and services that are important to customers.

Overwhelmingly, respondents cited giving executive management customer and market insight as the most important business benefit.

TDWI Research examined the business benefits sought from customer analytics (see Figure 4). The results show the importance of customer analytics in the boardroom; overwhelmingly, respondents cited giving executive management customer and market insight (71%) as the most important business benefit that their organization seeks to achieve from implementing customer analytics. This percentage rises to 81% when survey results are filtered to see only the responses from those who indicated "strong acceptance" of data-driven customer analytics over gut feel in Figure 3. The second highest benefit cited was the ability to react more quickly to changing market conditions (62%), which speaks to the need for customer data insights to help decision makers address competitive pressures from rapid product or service commoditization.

[1] Davenport, Thomas [2006]. "Competing on Analytics," *Harvard Business Review*, January.

What are the most important business benefits that your organization seeks to achieve from implementing customer analytics technologies and methods? (Please select all that apply.)

Give executive management customer/market insight	71%
React more quickly to changing market conditions	62%
Improve customer satisfaction with experiences and engagement	60%
Gain complete view of customer activity across channels	56%
Identify potential competitive advantages	50%
Discover what influences buying behavior	49%
Apply insights to product/service development	44%
Manage and target marketing mix	43%
Identify financial impact of marketing actions	42%
Develop more effective loyalty programs	32%
Manage brands effectively in social media	31%
Gain accurate attribution of conversions to marketing touches	26%

Figure 4. Based on 2,573 responses from 454 respondents; almost six responses per respondent, on average.

Business Benefits of Integrated Data for CEM. Improving customer satisfaction with experiences and engagement (60%) was the third-highest benefit sought. This result lines up with the importance of customer analytics for service and order management functions seen in Figure 1 (page 7). Customer experience management (CEM) processes, which involve monitoring and measuring customer interactions to understand perceptions about brands and products, are critical to identifying what steps an organization must take to build customer loyalty. CEM and similar applications are increasingly used in contact centers. An integrated view of different types of customer data—including transactions, service history, satisfaction survey results, and in some cases, social media data—is essential to CEM processes. CEM feedback and customer experience analysis are more valuable if they tap multiple integrated data sources.

Closely related to improving customer satisfaction is gaining a complete view of customer activity across channels, identified by more than half (56%) of respondents as an important business benefit sought. As mentioned earlier, the number of channels is increasing with mobile applications and social media now joining physical stores, Web sites, and other channels. Too often in organizations, each channel, division, or customer application has its own data silo, which makes it hard to consolidate information about individual customers and segments. Mergers and acquisitions can further complicate finding customer data relationships across sources. Organizations that can integrate customer data to achieve something closer to a total 360-degree view of customers can tap this information for analytics.

In many firms, the technology to achieve unified views is customer MDM. These systems enable organizations to establish a single reference or registry that unifies instances of customer data across sources. Organizations can use customer MDM tools to improve the completeness and quality of the data that is integrated for 360-degree customer views.

Immaturity of Social Media Channel Integration. Interestingly, less than one-third of respondents said that managing brands effectively in social media (31%) is a benefit they seek. This result suggests the immaturity of organizations in integrating their social media channel activity with customer analytics processes dedicated to other marketing channels. In the user story about comScore (next page), we can see how a leading-edge organization is relying on customer analytics, including social media feedback analysis, to stay ahead of commoditization by introducing new products and services specifically for that channel.

An integrated view of different types of customer data is essential to CEM processes.

Organizations can use customer MDM tools to improve the completeness and quality of the data that is integrated for 360-degree customer views.

COMSCORE CREATES CUSTOMER ANALYTICS PRODUCTS FROM BIG DATA

Riding atop one of the biggest of the big data engines in the world—and drawing data collected from around the world—comScore is in the business of converting information into data insights that it can offer as services. The company is a major source of digital marketing data. With nearly all top media properties directly tagging their digital content and providing census-based counts of activity, comScore can create highly accurate and timely views of what people are doing on Web sites, mobile devices, TVs, or other digital tools. The company's flagship Media Metrix service provides audience measurement and media planning services that are used by organizations to understand online media use, visitor demographics, and online buying power in the U.S. and globally so that they can make prudent marketing and advertising spending decisions.

The company's data resources are staggering, with billions of records captured every day from monitoring and tracking both human behavior and device readings and measurement. It employs a range of database technologies and tools to run its Unified Digital Measurement™ platform, enterprise data warehouse, and other resources.

To keep its own customers satisfied and continue to realize return on its massive data resources, comScore is focused on creating new products. One of the most innovative and insightful of its recent products is comScore Device Essentials, which looks at online traffic across devices, such as smartphones, tablets, PCs, and e-readers. "We were able to take a large store of data—the total tonnage of Web consumption—and turn it around to answer a whole new set of questions," said Michael Brown, chief technology officer of comScore and one of its founding members. The product can draw rich detail from a stream of more than 25 billion events per day from over one million domains.

Products such as Device Essentials and comScore Social Essentials, used to derive insight from unstructured and semi-structured data generated in the social media arena, demand technology that can provide speed, scale, and flexibility without crushing comScore's staff with data management duties. "You don't want to tie your data scientists down with production and maintenance," said Brown. "Their time is a precious resource. If it takes five hours for every query, they can't get much done. They need to be able to do iterative development." Key to supporting these requirements is comScore's implementation of EMC Greenplum DB and Greenplum HR. comScore currently has a 100-node cluster (2,400 cores, 6.4 TB RAM, 540 TB disk) running Greenplum DB and a 75-node cluster (1,700 cores, 4.7 TB RAM, 1.1 petabytes) running Greenplum MR.

The Social Essentials product enables comScore to address marketers' questions that are pertinent in this age of social media, such as: What kinds of people are being exposed to my brands on Facebook and other media? Am I getting ROI from my Facebook presence, and how does this compare to what I receive from search engines, display advertising, and other channels? Through data analysis, comScore is able to identify influencers and show their "halo effect," or spreading influence over others. The product can show how the "lift" of someone becoming a Facebook fan impacts the bottom line. "You can see if the effort to convert someone into a fan is really worth it," Brown said.

Barriers to Adoption of Customer Analytics

What factors are giving organizations the most difficulty in adopting and implementing customer analytics technologies and methods? TDWI Research found that difficulty in identifying and quantifying hard return (46%) is the obstacle confronted by the largest share of respondents (see Figure 5; respondents could select multiple answers). This is followed closely by not having enough budget or resources (45%) and lack of skilled personnel or training (43%). This third factor is

particularly troublesome when organizations need enough data scientists to develop models and algorithms for customer analytics. "Data scientists are a scarcity," said one senior marketing director.

Which of the following barriers to adoption of customer analytics apply to your organization? (Please select all that apply.)
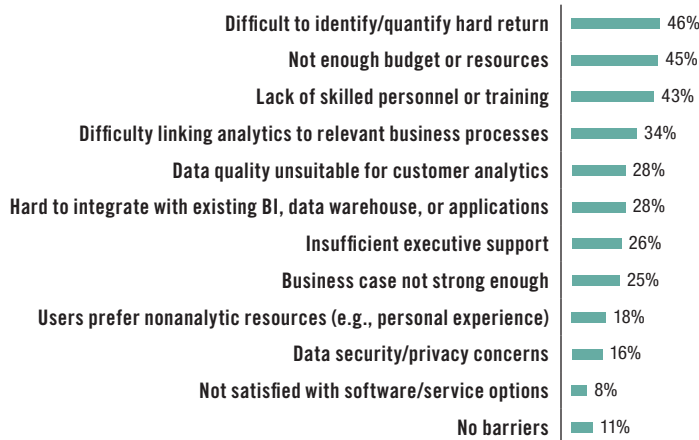
Barrier	Percentage
Difficult to identify/quantify hard return	46%
Not enough budget or resources	45%
Lack of skilled personnel or training	43%
Difficulty linking analytics to relevant business processes	34%
Data quality unsuitable for customer analytics	28%
Hard to integrate with existing BI, data warehouse, or applications	28%
Insufficient executive support	26%
Business case not strong enough	25%
Users prefer nonanalytic resources (e.g., personal experience)	18%
Data security/privacy concerns	16%
Not satisfied with software/service options	8%
No barriers	11%

Figure 5. Based on 1,477 responses from 451 respondents; a little over three responses per respondent, on average.

Difficulty in quantifying hard return and getting needed budget and resources are standard problems for any new technology investment. To overcome this obstacle, professionals in both business and IT can be persuasive with budget holders by identifying the cost to the business of inadequate information in each phase of a typical customer life cycle. Without customer intelligence at the beginning of the life cycle, organizations will face lost opportunity costs because they cannot identify the right segments for prospects. Later, when organizations try to pursue strategic goals such as "higher customer loyalty," the lack of knowledge will prevent them from realizing when in a customer's life cycle their loyalty is truly most valuable.

Without the insights from customer analytics, marketing managers will be unable to answer questions such as: When in the life cycle are customers most likely to churn? What types of products or services would prevent them from churning, and when should they be offered? When is it too costly to try to keep certain customers? Analytics can provide answers to these types of questions and help organizations improve marketing performance by quantifying customer value at different stages in the life cycle. Then, managers can align their deployment of resources to achieve the highest value and avoid the costs and inefficiencies of marketing to them at the wrong time.

Role of Analytics in Increasing Marketing ROI

Organizations are able to realize ROI from investing in customer analytics by improving the marketing function's efficiency and effectiveness. With better customer knowledge, they can optimize marketing spending by focusing resources on those campaigns with the highest predicted chances of success for particular segments. By using analytics to eliminate mismatches of campaigns targeting the wrong customers or using the wrong messages and offers, marketing functions can reduce wasteful spending and increase gains relative to costs.

TDWI Research examined the importance of accomplishing various marketing objectives for gaining ROI from customer analytics (see Figure 6). Using customer analytics to target cross-sell and

Professionals in both business and IT can be persuasive with budget holders by spelling out impact-of-knowledge gaps in a typical customer life cycle.

With better customer knowledge, firms can optimize marketing spending by focusing resources on campaigns with the highest predicted chances of success for particular segments.

To grow the business and achieve higher levels of profitability with existing customers, it is essential to expand the number and size of their purchases.

up-sell opportunities was the objective cited by the biggest percentage of respondents (54%). This objective is essentially about gaining more value from existing customers by knowing what additional (and bigger-ticket) products or services will interest them. Most organizations find it easier to keep existing customers than find new ones. However, to grow the business and achieve higher levels of profitability with existing customers, it is essential to expand the number and size of their purchases.

In your organization, which of the following marketing objectives are most important to achieve for customer analytics to deliver a return on investment? (Please select all that apply.)

Target cross-sell and up-sell opportunities	54%
Improve customer segmentation	49%
Predict retention, attrition, and churn rates	47%
Determine lifetime customer value	42%
Increase portfolio penetration per customer	39%
Optimize marketing across multiple channels	37%
Impact other business functions (sales, service, support)	35%
Forecast buying habits and lifestyle preferences	32%
Measure types of loyalty for campaign targeting	26%
Prioritize marketing e-mail messages	23%
Implement uplift, incremental, or true-lift modeling	18%
Increase speed of multivariate testing analysis	14%

Figure 6. *Based on 1,625 responses from 432 respondents; almost four responses per respondent, on average.*

Predictive analytics plays a major role in enabling organizations to investigate how customers in particular segments are spending over time and model their propensity to buy additional or more expensive products. Organizations can also analyze customer life cycles to determine when to attempt certain cross-sell and up-sell offers and what kind of products to offer at particular phases in the life cycle.

Some organizations (18% in Figure 6) are implementing an advanced technique called "uplift modeling" (also called incremental or true-lift modeling), which enables marketers to use data mining to measure the impact and influence of marketing actions on customers. With such insights, marketers can develop new kinds of predictive models for determining the best prospects for up-sell and cross-sell. Organizations can discover both the effectiveness of their campaigns and how to reduce marketing spending by focusing on only their best prospects. As firms scale up to execute large numbers of campaigns across multiple channels, the efficiency gained from predictive modeling can be critical to marketing spending optimization.

Understanding influence is the goal of using social media behavioral analysis to improve cross-sell and up-sell opportunities.

Understanding influence is the goal of using social media behavioral analysis to improve cross-sell and up-sell opportunities. This type of analysis focuses on the influence customers have on each other through social networks, chat rooms, review boards, and similar venues. It offers a second benefit in helping organizations tune their own cross-sell and up-sell efforts based on what they've learned from social media behavior.

Improving ROI with Smarter Segmentation

The objective cited by the second-highest percentage of survey respondents was improving customer segmentation (49%). Enabling smarter segmentation is a key objective of many customer analytics

implementations. Segmentation allows organizations to move further along the spectrum, away from one-size-fits-all, brand-level-only marketing and toward the "market of one": that is, personalized, one-to-one marketing.

Organizations have long used characteristics such as gender, household size, education, occupation, and income to segment customers. Data mining techniques enable organizations to segment much larger customer populations and, perhaps more important, determine whether to apply new characteristics that refine segmentation to fit the specific attributes of the organization's products and services. Customer analytics using data mining tools improves the speed of segmentation analysis over manual and spreadsheet efforts that are often used in less mature organizations. Speed is vital for marketing initiatives that are time sensitive—especially for those that need to provide cross-sell and up-sell offers to customers clicking through Web pages in real time.

The next most common objectives in the research were predicting retention, attrition, and churn rates (47%) and determining lifetime customer value (42%). Customer lifetime value analysis enables organizations to look at the entire life cycle of a customer relationship, calculate its worth at different stages, and determine the value of marketing to customers at each stage in the customer's "lifetime" of engagement with the firm. Organizations can employ analytics to understand how retention, attrition, and return rates align with the different stages.

Looking at two industries in a bit more depth, our research shows that firms in financial services and telecommunications take particular interest in customer lifetime value analysis. With the traditionally high rate of customer churn in those industries, organizations need to predict when customers are likely to leave, what types of actions or sentiments indicate that they are about to leave, and what offers the organization could make to prevent their departure. Along with customer transactions, our research shows that the data sources in which financial services and telecommunications firms are most interested are customer satisfaction surveys and call or contact center interaction records. Because of the volume and lack of structure of these sources, they are often ignored or analyzed manually, well after their generation. Text analytics and natural language processing tools and techniques can be used to drill down more rapidly into the enormous volume of records to uncover patterns, develop predictive models, and classify expressions of sentiment.

Churn can cost organizations heavily in the loss of profits from existing customers and the increase in expenditures to attract new ones. Attrition or churn analysis methods are aimed at discovering which variables have the most influence on customers' decisions to leave or stay. With experience, organizations can learn what attrition rates are acceptable or expected for certain customer segments and what rates are unacceptable and represent a true threat to the business. Predictive customer analytics can play a major role in enabling organizations to discover and model which customers are most likely to leave, and from which segments. The user story on the next page about XO Communications discusses the use of predictive analytics to reduce customer churn.

Social Media Analysis as Part of Mainstream Analytics. In our research, we asked organizations whether they are currently analyzing social media data or planning to do so as part of their efforts to achieve any of the marketing objectives identified in Figure 6. Just over one-quarter (26%; figure not shown) of respondents said that their organizations were doing so already; 22% said they were planning to do so within one year; and 16% planned to analyze social media data for their marketing objectives within two years. The top three objectives most cited by respondents who are already analyzing social media data were the same as those of all respondents: targeting cross-sell and up-sell opportunities; improving customer segmentation; and predicting retention, attrition, and churn rates. This suggests that as organizations become more experienced with social media data analysis, they will incorporate it into their mainstream customer analytics processes.

Data mining techniques enable organizations to segment much larger customer populations and, perhaps more important, determine whether to apply new characteristics.

Predictive customer analytics can play a major role in enabling organizations to discover and model which customers are most likely to leave.

USER STORY

XO COMMUNICATIONS APPLIES PREDICTIVE ANALYTICS TO REDUCE CUSTOMER CHURN

Customer churn is the existential battle that telecommunications companies must win to stay profitable and keep their footing in a competitive marketplace. Firms know that it is more costly to acquire customers than to retain them, and that lost customers go to competitors. If a telecommunications company can improve retention rates even 5%, it can increase profits significantly and create a strong customer base to buy more services and support market share growth.

XO Communications, based in Herndon, Virginia, is a leader in the business-to-business telecommunications services market. To attack customer churn, the company deployed IBM Business Analytics with a focus on implementing IBM's SPSS predictive analytics software and the industry-standard CRISP-DM methodology for all its data mining process steps. The company's main objective was to reduce churn and improve retention of its best customers in the midsize sector of the market, which is too large for its services managers alone to manage in the hands-on fashion that they normally apply to large, at-risk customers.

Why is using predictive analytics the right approach to addressing customer churn, as opposed to implementing BI and online analytical processing (OLAP)? "We couldn't just look at the data in dashboards and analyze it in standard BI ways," said Trent Taylor, who served as XO's strategic marketing director and director of customer intelligence and marketing programs during the project. "We probably spun our wheels for a year doing that. The criticism we received was, 'Oh, you're talking about a trend that's already happened; it's in the rear-view mirror. Tell me what's going to happen.'" Gaining a predictive understanding became "a very direct business imperative," according to Taylor. XO's sales leaders and regional VPs wanted to know what the company could do to prevent churn.

Taylor observed that you can't take a "find-the-root-cause" approach to churn. "It's like an accident investigation; it's not one thing but a series of failures. If you don't break the chain, it will keep on happening." Rather than try to isolate the answer or even the chain, XO used analytics to look more deeply at customer segments to determine who was most likely to churn. This is where the analytical requirements go beyond BI and OLAP. "With OLAP, you can only look at a small number of variables," said Cris Payne, XO's senior manager of customer intelligence for strategic marketing, who developed, ran, and built the predictive algorithms used by XO. "You can switch those variables up, but you really can't get the sense of the interplay of variables and correlations unless you start using predictive analytics."

XO began its quest with more than 500 variables, far more than could be handled with BI and OLAP tools but appropriate for the SPSS implementation. Payne, Taylor, and the XO team whittled it down 350; for any one data mining process, they may use only 15 or 20. "You don't know which 15 or 20 you need until you go through all 500 data sets," said Payne. Data preparation for the predictive analytics took the lion's share of their time.

Taylor noted a caveat about the variable reduction process. "Even when you do arrive at the 15 or 17 you need, those variables may only be valid for 12 to 18 months, depending on what you are trying to predict," said Taylor. "Then you have to go back to the drawing board and refresh your model." The beta weights attached to each variable in the logistic regression analysis Taylor and Payne used also influenced algorithm performance, adding more complexity.

Predictive analysis processes put different demands on IT than most traditional, BI-oriented business-side requests, for which IT usually asks for all business requirements and then goes off to develop and deliver solutions. "The problem is that we don't know what's needed until we start playing with the data and applying our modeling software during the exploratory phase of data analysis," said Payne. "You can't crystallize the data needs into a single requirements document and say, 'Here, deliver this.'"

Finding ways to work productively with IT despite the differences in approach, plus gaining funding, are nontechnical challenges for business units trying to do predictive analytics and implement more robust software than BI tools and spreadsheet add-ons. XO's success with reducing customer churn offers good reason to believe that giving the business side the freedom and autonomy to explore data according to CRISP-DM or similar processes is the best way to arrive at predictive insights that improve strategies and change business outcomes.

Analytics Tools, Data Sources, and Techniques

As the number of channels for customer communication and interaction increases, so does the volume and variety of data. In large organizations, enterprise data warehouses dedicated to customer data already contain tens, if not hundreds, of terabytes of data from sales transactions and other events coming from online and physical stores, kiosks, and other kinds of data-gathering systems and applications. Now, with social media, clickstreams, and other behavioral data, the volume and variety are exploding. It is indeed "big" data; social networking sites such as Facebook, Twitter, LinkedIn, and MySpace have files containing petabytes of data, often in vast Hadoop clusters. Advertising concerns are recording tens of millions of events daily that organizations want to mine in near real time to identify prospects. Organizations want to use predictive models and score event and transaction records as they come in to gain insight into shopping behavior.

TDWI Research found that the data sources most commonly monitored for customer analytics are customer satisfaction surveys (57%) and customer transactions and online purchases (55%; see Figure 7). Just under half (44%) are monitoring Web site logs and clickstream sources. In addition to monitoring customer satisfaction surveys, about half (48%) of organizations surveyed are studying call and contact center interactions.

Enterprise data warehouses dedicated to customer data already contain tens, if not hundreds, of terabytes of data from sales transactions and other events.

In your organization, which data sources are monitored for customer analytics? (Please select all that apply.)

Customer satisfaction surveys	57%
Customer transactions/online purchases	55%
Call/contact center interactions	48%
Web site logs/clickstream	44%
E-mail response	44%
Social media	31%
Sales force automation data	28%
In-store/POS purchases	27%
Traditional advertising response	24%
Digital advertising response	19%
Chat records/text messages	11%

Figure 7. Based on 1,668 responses from 426 respondents; about four responses per respondent, on average.

Customer satisfaction surveys are traditionally conducted in person, on the phone, on Web sites, and through traditional mail and e-mail. The variety of modes leads to many different data types, including semi-structured data in forms and unstructured comments. Standard questions inquire about a customer's satisfaction with purchases, the services they received, and the company's brands overall. Other questions address the customer's likelihood of buying from the company again and whether they would recommend the firm to others. Some organizations add questions that are more

specific to their offerings or that can give them data about the customer's path to purchase—that is, to what marketing actions the purchase can be attributed.

Organizations can employ text analytics (or "voice of the customer" applications embedded with text analysis and natural language processing capabilities) to increase the speed, depth, and consistency of unstructured content analysis far beyond what can be done manually. Analysts can then integrate these findings with their customer satisfaction ratings and agent performance metrics to learn whether their current processes and practices for customer interactions are effective. More advanced analytics can look for correlations between satisfaction ratings, commented sentiments, and other records, such as first-call-resolution metrics.

Customer Satisfaction Input from Social Media Networks. Through social networks such as Twitter, organizations can pick up customer satisfaction feedback in real time. Social media is enabling companies such as Coca-Cola, Starbucks, and Ford to go beyond standard customer satisfaction data gathering to innovate by setting up and participating in communities to gain feedback from customers. A good example is MyStarbucksIdea.com, through which Starbucks customers can relate their experiences and offer ideas about how to improve the Starbucks experience, from drinks to foods to ambiance. Although Figure 7 shows that just 31% of respondent organizations are monitoring social media for customer analytics purposes, it is likely that this percentage will increase as organizations participate more in social media communities as Starbucks has done, and in this way collect more feedback data for analysis.

> **Social media is enabling companies to go beyond standard customer satisfaction data gathering to innovate by setting up and participating in communities.**

BI, OLAP, and Data Discovery for Customer Analytics

TDWI Research found that basic reporting (76%) is the dominant means by which organizations analyze data for customer analytics, followed by online analytical processing (OLAP) and BI (65%; see Figure 8). These are the most prevalent technologies even for organizations in our research that are advanced in analyzing social media data. The dominance is likely due to the ubiquity of these tools in organizations surveyed compared with more advanced data mining and predictive analytics tools, as well as the interest in exploiting existing tools and licensed software as fully as possible. BI and OLAP tools excel at querying, reporting, and analysis of historical data to understand and compare results to date or for specific time periods. Many users interact with BI and OLAP through dashboard interfaces that can integrate data access and visualizations and provide performance metrics such as key performance indicators (KPIs) focused on customer satisfaction.

> **"Data discovery" tools give users more self-service functionality for analytics and visualization.**

Easy-to-understand dashboards and data visualizations are important for giving nontechnical users in marketing, sales, and service functions self-service capabilities for working with data. "Data discovery" tools, which push beyond standard, IT-driven BI systems to give users more self-service functionality for analytics and visualization, help users get deeper into the data to perform what-if analysis and unearth more uncommon insights than what standard reports can provide. The user story on the next page describes how E.B. Lane implemented data discovery to monitor and analyze data coming from multiple channels, including social media.

What analytic tools or techniques are currently being implemented in your organization for customer analytics? (Please select all that apply.)
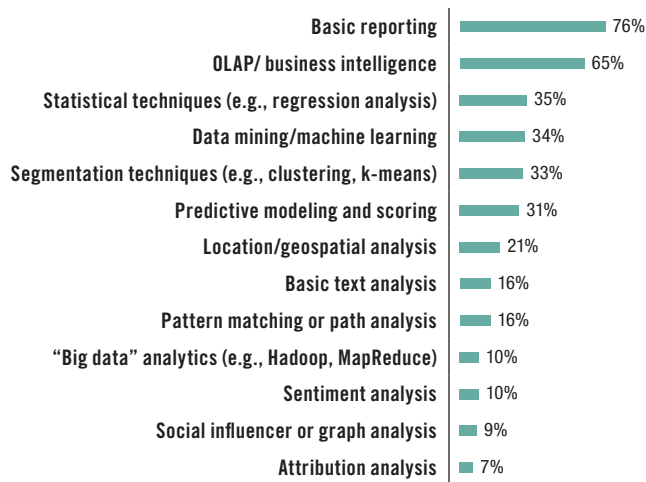
Basic reporting	76%
OLAP/ business intelligence	65%
Statistical techniques (e.g., regression analysis)	35%
Data mining/machine learning	34%
Segmentation techniques (e.g., clustering, k-means)	33%
Predictive modeling and scoring	31%
Location/geospatial analysis	21%
Basic text analysis	16%
Pattern matching or path analysis	16%
"Big data" analytics (e.g., Hadoop, MapReduce)	10%
Sentiment analysis	10%
Social influencer or graph analysis	9%
Attribution analysis	7%

Figure 8. Based on 1,546 responses from 426 respondents; nearly four responses per respondent, on average.

USER STORY

E.B. LANE USES DASHBOARDS TO INTEGRATE VIEWS OF ARIZONA LOTTERY'S CUSTOMER DATA

"Frenzy" is the best word to describe what happens when Mega Millions lottery jackpots soar, as happened early in 2012 when one reached a record $656 million. For the winners, the once-in-a-lifetime jackpot is a windfall; for state revenues, however, lotteries, Powerball, quick picks, and other games are critical, ongoing sources of revenue. To keep up a healthy revenue stream from lotteries and games, states focus on improving their marketing to attract younger players in the 21 to 35 age group. This means ramping up marketing on social media to go along with television and other more traditional channels.

E.B. Lane, the advertising, public relations, and interactive firm of record for the Arizona Lottery, has been implementing business intelligence dashboards and data discovery analytics to develop strategies and manage efforts aimed at revitalizing marketing campaigns to attract new customers. "We receive quite a sizeable nightly data feed with all the sales numbers across all games throughout the state of Arizona," said Shailesh Ghimire, director of digital strategy at E.B. Lane. Using Tableau, E.B. Lane posts the daily sales on a live dashboard that can show a geographical breakdown (based on ZIP codes) by store. Once the firm's media team uploads the media allocation information, Ghimire is able to combine it with nightly Web data from Google Analytics for Web performance as well as social media data, including "likes" and followers on Facebook and Twitter. E.B. Lane also produces video content for the Arizona Lottery that is accessible via YouTube, and the firm can pull information about the reception of that content as well.

The dashboard provides a timely, consolidated, and visual view with reach into social media data to give the E.B. Lane and the Arizona Lottery vital information about the performance of campaigns both in social media itself and in other channels. "We are able to see a simple but well-rounded view of how all of our marketing pieces are doing, and can drill into the data to understand better what's working and not working, such as what the engagement numbers from the Lottery's Facebook presence look like," said Ghimire. "When Mega Millions went on its run, we could monitor sales and look at how social media comments were increasing. The 'find a retailer' page was going crazy; people needed to know where they could buy a ticket. This surge was interesting because it told us that new

players were coming into the market. Regular players usually already know where to buy tickets." E.B. Lane could then tune its marketing to attract those types of players.

Ghimire said the consolidated view is important to raising the value of social media to the services that E.B. Lane can offer the Arizona Lottery. "The impact of seeing data across channels is huge," he said. "Social media doesn't become just an isolated discussion. It becomes: 'Wait a minute, now we can post numbers, drive people to the Web site, and see the impact of marketing changes on our sales.'" Using Tableau's visualization, E.B. Lane and the Arizona Lottery can see the rate at which people are engaging with the brand, especially in social media, through monitoring the rise in comments, retweets, and more.

In-Memory Computing for More Rapid Discovery Analysis

As organizations pursue faster "speed-of-thought" analysis for marketing decisions that can impact customer decisions in real time, in-memory computing will play a major role.

Some data discovery tools as well as more advanced BI and OLAP systems can exploit in-memory computing for analytics; users of these systems are able to tap even larger volumes of data and run algorithms and other analytic processes without going through the I/O bottleneck to access data stored on disk. As organizations pursue faster "speed-of-thought" analysis for marketing decisions that can impact customer decisions in real time, in-memory computing will play a major role. Preprocessing steps that are usually required for reducing the number of records moving through the I/O bottleneck—steps that have effectively slowed analysis and limited the scope of discovery—are not as necessary when all the data is loaded into memory.

In-memory computing is significant for rapid analysis of large volumes of customer behavior data from social media. To analyze data generated by social media networking services such as Twitter, Facebook, and LinkedIn, many organizations are implementing Hadoop and NoSQL technologies, which do not force a schema on the source data prior to storage, as traditional BI and data warehousing systems do. This lets the discovery analytics processes run against raw data. Customer analytics tools need to be able to consume data from sources such as Hadoop clusters and then integrate the insights into overall customer profiles.

Predictive Analytics, Data Mining, and Advanced Statistics Implementation

At the cutting edge of customer analytics are technologies and methods for data mining, advanced statistical analysis, and predictive analytics. These allow organizations to examine large data sets more rapidly to test hypotheses and forecast buying behavior, customer churn risk, household customer value, and other considerations. Data sources are varied for these technologies and methods; they include transaction data, clickstreams, satisfaction surveys, loyalty card membership data, credit card purchases, voter registration, location data, and a host of demographic data types.

Referring again to Figure 8, our research found that 34% of respondent organizations are implementing data mining and machine learning for customer analytics. Slightly more (35%) are using statistical techniques, including regression analysis. Similar percentages are working with advanced segmentation techniques such as k-means clustering (33%) and predictive modeling and scoring (31%). Taken together, these results reveal the penetration rate thus far for predictive analytics, data mining, and advanced statistical analysis of customer information.

Nearly 80% said "reacting more quickly to changing market conditions" is a business benefit that they seek.

Among respondents who indicated they are implementing tools and technologies for advanced statistical analysis, data mining, and predictive analytics, nearly 80% said "reacting more quickly to changing market conditions" is a business benefit that they seek. This indicates the importance of increasing the speed of analysis to achieving firms' objectives in implementing predictive analytics and data mining. The overall business goals are often to use predictive insights to anticipate, rather than

react to, customer behavior, as well as enable organizations to be proactive about addressing risk factors that increase customer churn.

Advanced analytics generally involves statistical, quantitative, or mathematical analysis and centers on developing, testing, training, scoring, and monitoring predictive models. Organizations implement predictive analytics and data mining to accomplish many objectives, including discovering which variables (often out of hundreds) are most influential in determining customer loyalty within certain segments. Organizations will build models to uncover patterns, affinities, anomalies, and other useful insights for marketing campaigns and for determining cross-sell and up-sell recommendations. The tools and techniques are also used for developing and deploying behavioral scoring models for marketing, deciding whether to adjust customers' credit limits for purchases, and a variety of highly time-sensitive analytic processes. The tools and techniques are generally applied against very large data sets and run in high-performance, massively parallel processing computing environments.

Moving Beyond BI and OLAP. When do organizations make the leap from BI and OLAP to predictive analytics and data mining? Actually, the two types of technologies are not mutually exclusive, although they are frequently implemented by different sets of users in different parts of organizations. They can come together when users are able to consume advanced analytics through visualizations in BI dashboards and use visual analysis to explore data more deeply to find out what is driving trends and patterns. Some organizations tie performance metrics and KPIs to the results of analytic processes; the analytic insights are then displayed to executives within the context of known business performance objectives. Nor are data mining and advanced statistics the only tools and methods used for decisions affecting customer marketing strategies; for example, managers will use OLAP and financial planning tools, often supported by in-memory computing capabilities, to increase analytical power and run queries that help them understand how best to align financial and marketing goals.

A marketing data analyst at a large retail company interviewed for this report summed up the difference between BI and OLAP and statistical and predictive analytics. "Whereas BI and OLAP systems have to keep referring back to the data for historical analysis or to do slicing and dicing for multidimensional analysis, once we've built our predictive model, we're actually done with the data," he said. "With the model done, we then look for new cases where we can apply the insight we have brought out of the data and can test our hypothesis." He noted that predictive analytics is part of the organization's test-and-learn feedback loop; the firm implements a cycle of continuous testing, learning, adapting, and then revamping models based on results from production deployments. They are able to use the tools and techniques to keep pace with the dynamic nature of customer behavior, and adjust marketing strategies to match that behavior.

Advanced Analytics for Determining Marketing Attribution

According to TDWI Research (see Figure 8, page 19), only a small percentage of respondents (7%) is currently performing attribution analysis to determine which campaign or customer action actually triggered a purchase. Since the research shows that the vast majority of organizations surveyed are implementing primarily BI and OLAP tools for customer analytics, the low percentage could reflect the lack of adequate tools and techniques for answering one of the most vexing and complicated questions in customer marketing.

The frustration organizations encounter in attempting to determine attribution is further underlined by the research results shown in Figure 9. Only about 7% of respondents are "very satisfied" by their organization's ability to analyze customer data to determine marketing attribution.

Organizations implement predictive analytics and data mining to accomplish many objectives, including discovering which variables are most influential in determining customer loyalty.

"Whereas BI and OLAP systems have to keep referring back to the data … once we've built our predictive model, we're actually done with the data."

Interestingly, about three-quarters (74%) of those who indicated that they are very satisfied with their organizations' ability to do so are from organizations that currently employ data scientists (a topic discussed in the next section). Two out of five (40%) respondents overall are dissatisfied with their ability to determine attribution, while 37% are somewhat satisfied.

How satisfied is your organization with its ability to analyze customer data to determine marketing attribution? (That is, to determine how many of customers' interactions across marketing campaigns as well as organic, customer-driven interactions across multiple channels were included in the purchasing path that ultimately led to a transaction?)
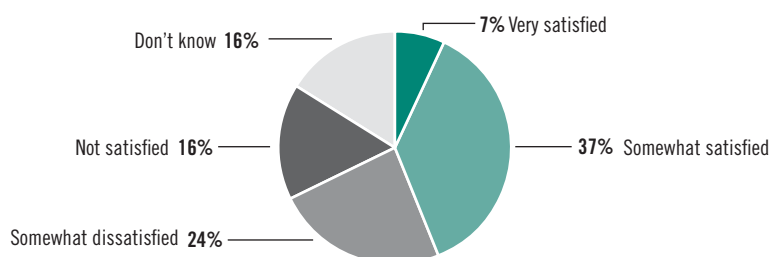

Figure 9. Based on answers from 434 respondents.

The proliferation of channels to include social media, search advertising, and more has only made the issue of attribution more complicated.

The proliferation of channels to include social media, search advertising, and more has only made the issue of attribution and determining the customer's path to a purchase more complicated. Rather than look more deeply, organizations will often reach for the easy answer. For example, a company might have spent a significant amount of money on e-mail promotion; since it was expensive, managers will be tempted to attribute to it all the credit for the purchase. Another typical approach is "last-touch" attribution; organizations will look at whatever the customer was exposed to last before making a purchase and put all the credit on that marketing action.

The reality could be quite different. With so much information available online, customers are likely to have done their own research using search engines, browsing Web sites, clicking on advertisements, and participating in social media networks. They may also have gone to physical stores, seen billboards, or heard radio or TV advertisements. Satisfaction surveys that include questions about how customers heard about a brand or product hardly give enough information to glean attribution. Disconnected information silos that are spread across marketing, sales, e-commerce, and other departments and functions make it difficult to get an integrated view of attribution.

Attribution analysis is a big-data problem, given the quantity and variety of data. Organizations that are performing attribution analysis will frequently employ Hadoop, MapReduce, and specialized analytic databases.

As more online customer behavior is recorded in Web logs and tracked through cookies and other observation devices, sizeable amounts of information are becoming available to organizations that seek a more accurate view of a customer's path to purchase. Attribution analysis is a big-data problem, given the quantity and variety of data. Organizations that are performing attribution analysis will frequently employ Hadoop, MapReduce, and specialized analytic databases so that they can run sophisticated algorithms against detailed data to find the correct path. The analysis can then be integrated with analysis of other data types and sources, such as those generated by offline customer activity.

The depth of analysis required usually involves many passes through the data. This can be difficult to do exclusively with SQL database systems because they require extensive recoding to explicitly tell the system each time which pages the query to needs to hit. In-memory computing can alleviate some of the pain of this slow and expensive process by making I/O to disk less of an issue, but

organizations will frequently look beyond SQL-only database systems and implement Hadoop and MapReduce alternatives.

Attribution analysis is important to understanding what kinds of campaigns most influence customer behavior. The analysis can help organizations determine where to allocate marketing resources to gain the highest level of success, as well as how to more accurately assign the percentage of credit due to specific marketing and advertising processes.

Filling the Role of the Data Scientist for Customer Analytics

Many organizations are trying to determine whether they need to hire "data scientists" to perform advanced customer analytics. Industry analysts and experts offer differing definitions of this person and role, but most organizations know one thing: there are not enough of them. A 2011 McKinsey Global Institute study warned of the shortage of data scientists, noting that "the U.S. alone could face a shortage of 140,000 to 190,000 people with deep analytical skills as well as 1.5 million managers and analysts with the know-how to use the analysis of big data to make effective decisions."

Our research found that 16% of organizations surveyed currently employ data scientists, and 6% plan to do so (see Figure 10). The results confirm what TDWI Research heard anecdotally through interviews for this report: that organizations are taking a variety of paths to expand their data science capabilities. About one-fifth are giving data scientist roles to current business, data, or statistical analysts (20%), and another 6% are training and shifting personnel to become data scientists. One-quarter of those surveyed have no plans for data scientists. A study of cross-tabulations in our research indicates that organizations with no plans for data scientists are less likely to be using Hadoop and MapReduce; they are also confronting a lack of budget or resources as the chief barrier to adopting customer analytics.

Our research found that 16% of organizations surveyed currently employ data scientists, and 6% plan to do so.

What is your organization's current experience with, and/or future plans for employing, "data scientists" to meet objectives for customer and social media data analytics?

We currently employ data scientists	16%
We plan to employ data scientists	6%
Current business, data, or statistical analysts perform data science roles	20%
We are training and/or shifting roles internally to create data scientists	6%
We engage services of contract data scientists on a per-project basis	5%
We have no plans for data scientists	25%
Don't know	22%

Figure 10. Based on answers from 423 respondents.

Rising interest in data science is not exclusively due to rising interest in analyzing social media data; there are other big data sources that are equally, if not more, important for data scientists to analyze. However, many organizations have little experience with, or understanding of, how to draw good information and insights from data generated by possibly hundreds of their own Facebook pages and Twitter identities, for example—much less the "fire hose" of all data coming from these social networks. Some data scientists consequently specialize in working with social media data and connecting insights from that realm into the organization's segmentation and buying behavior analysis being done against more traditional, internal sources of data.

Many organizations have little experience with, or understanding of, how to draw good information and insights from data generated by their own Facebook pages and Twitter identities.

Using Data Science to Gain the Holistic View. One quality that can make data scientists distinct from business or data analysts is if they take a more global or holistic view of how marketing campaigns are working across multiple channels. Business or data analysts are often dedicated to one department and focus on a specific set of campaigns or metrics. To identify which cluster of customers is the most influential in driving sales, for example, "holistic" data scientists can develop hypotheses that take in data generated by customer behavior from a number of sources, including Web sites, social media, service records, and sales transaction data. This broader level of insight can become critical to the branding, marketing, and CEM strategies of the entire organization rather than just one marketing campaign.

Companies that engage customers through their branded online services and social media sites have their own multi-terabyte, if not petabyte-sized fire hose of data, not even counting the feeds from Twitter or other social networks. Such organizations will dedicate data scientists to interpreting their own social media performance and seeking insights into customer preference trends. The user story below describes the role of Hadoop in support of data science for customer analytics at Ubisoft.

USER STORY
UBISOFT EMPLOYS HADOOP TO SUPPORT SCALE AND FLEXIBILITY FOR CUSTOMER ANALYTICS

Analytical priorities never stay the same, especially in industries where change is constant and the thing that could drive higher profit margins is a moving target. Ubisoft, one of the largest game publishers in the world, is increasingly looking for ways to create efficiencies and reduce development costs. The company employs Hadoop and MapReduce to capture data about customer behavior, as well as economic data sets, and make that data available for a variety of analytics aimed at learning about and measuring the different stages in each product's life cycle.

Hadoop plays a key role for Ubisoft in providing a data store for many uses. "Our Hadoop-based data warehouse acts as a federated data store where we can bring in all kinds of raw, detailed data," said Lenin Gali, senior director of game analytics services for digital publishing at Ubisoft. "Our ultimate goal is to empower our data analysts or engineers who require some data, but not in the aggregated form you would have in a data warehouse. They can interface with Hadoop, our Vertica data warehouse platform, or with another database or application. They can do an extract, transform, and load (ETL) or ELT process, or put the data back in the warehouse, or publish a report or analysis. You can mold it the way you want."

Gali says that the raw, detailed data is vital for the company's data scientists working on analysis of customer behavior. "With aggregated data, you can miss a lot of the signals that for science are very important." Hadoop data can be stored and archived, but often it is data with a brief shelf life. It does have structure, however. As Gali noted, "the term 'unstructured' is kind of a misnomer. Logs, for example, have a structure. All that most businesses have is structured or semi-structured data. We can put it in Hadoop files and determine what the delimiters are, the key values, whether the data is of fixed length, or whatever. Analysts can bring in data in its own form, mold it however they want, and then move it to wherever it is needed."

Applying Technologies for Social Media Data Analysis

TDWI Research found that "gaining deeper customer understanding" (56%) is the top objective that organizations seek to achieve from implementing customer analytics with social media data (see Figure 11). This is particularly the case for organizations that do not have good internal customer

transaction or service record data sources and have little history of collecting customer satisfaction or other behavioral information. Social media listening can provide an unprecedented window on customer sentiment and the reception of an organization's marketing, brands, and services. (As a reality check, 32% of organizations surveyed do not analyze social media data.)

Which of the following objectives does your organization seek to achieve by implementing customer analytics technologies and methods with social media data? (Please select all that apply.)

Gain deeper customer understanding	56%
Identify customer paths to buying decisions	31%
Monitor and measure sentiment drivers	30%
Determine value of social media engagement to marketing campaigns	29%
Discover new audience segments	27%
Gain insights for new product development	24%
Analyze social networks, links, and graphs	22%
Differentiate influencers from followers in social media	20%
Increase engagement beyond passive social media monitoring	19%
Analyze competition's "share of voice"	18%
Monitor and analyze social activity in real time	14%
Improve "long-tail" analysis of buying by small groups of customers	11%
We do not analyze social media data	32%

Figure 11. Based on 1,546 responses from 418 respondents; a bit more than three responses per respondent, on average.

Beneath the broad objective of gaining a deeper customer understanding, we can see in Figure 11 the relative importance of several tactical objectives. Nearly one-third (31%) seek to identify attribution, or paths to buying decisions, a topic that was discussed in the previous section. As they mature in their attribution inquiry and wish to examine much larger sets of detailed or unstructured data, organizations reach the limits of available services such as Google Analytics and Web site analysis applications. These typically provide only an aggregated view of data. Organizations need detailed data to support analysis of how marketing and engagement processes can improve personalized interactions with smaller customer segments or even individuals. Organizations may want to examine purchase paths and attribution across the multiple online channels as well as offline sources. Robust database and customer analytics systems are frequently necessary to handle the big data needed for this more complex analysis.

Discovering Customer Sentiment. About the same percentage (30%) of respondents seeks to monitor and measure sentiment drivers. Sentiment analysis enables organizations to discover positive and negative comments in social media, customer comment and review sites, and similar sources. Sentiment analysis often focuses on monitoring and measuring the "buzz" value, usually through volume and frequency of comments around a topic. However, many organizations want more analytical depth so that they can discover what the buzz is about, where it originated, and who is benefiting or not benefiting from it.

For more sophisticated sentiment analysis, text analytics tools play a big role. These tools employ lexicons, word extraction, natural language processing, pattern matching, and other approaches to examine social media users' expressions. Sentiment analysis can give organizations early notice in real time of factors that may be affecting customer churn; in Figure 11, the research shows that 14% are interested in monitoring and analyzing social activity in real time. Sentiment analysis is also

important to understanding competitors' relative strengths and weaknesses in the social sphere. Our research found that 18% of respondents are examining social media data to analyze the competition's "share of voice."

Applying Analytics to Find and Influence the Influencers

Data scientists and analysts can test variables to help identify social communities as "segments."

One of the biggest challenges can be simply deciding which social media sites' data to analyze. Organizations have to research where their customers are most likely to express themselves about brands and products. They need to spot influencers who have networks of contacts and take it upon themselves to play an advocacy role. About 20% of respondents are interested in differentiating influencers from followers in social media (see Figure 11). Link analytics tools and methods specialize in identifying relationships between users in social communities and enabling organizations to measure users' influence. With some tools, data scientists and analysts can test variables to help identify social communities as "segments." Then, as they implement segmentation models for other data sources, they can integrate these insights with social media network analysis to sharpen the models and test new variables.

Analytics are critical for enabling organizations to make the right decisions about when, where, and how to participate in social media. It isn't enough to just listen; organizations must insert themselves and become part of the conversation. Leading companies will start viral campaigns, for example, using Twitter hashtags for a topic; the campaign could be a component of a larger marketing strategy. They can then monitor social media to see what people say and analyze how the campaign is playing among influencers and across networks. "My belief is that the sweet spot for social media is not conversion, but nurturing," said Brian Ellefritz, vice president of global social media at SAP. "Whether it's in your community, through Twitter, or through Facebook pages, you want to build an increasing conviction that your company is the one to do business with. It's about establishing a belief system that becomes robust with the support of fans and followers. The question is how you measure that and create value out of that investment."

Social media conversations fall into a gray area from a regulatory perspective, which has made many firms keep their distance other than for advertising.

Social media data analysis can also be a key component of risk mitigation in the marketplace. Pharmaceutical companies, for example, need to follow social media to watch for early signs of negative consequences with drugs or other treatments. How participatory these organizations and healthcare providers can be is a sensitive issue because of regulatory reporting requirements. Social media conversations fall into a gray area from a regulatory perspective, which has made many firms keep their distance other than for advertising. Overcoming regulatory concerns in the healthcare industry could enable greater participation in social media by these firms; this issue is discussed in the following user story about the Digital Health Coalition.

USER STORY

DIGITAL HEALTH COALITION SEEKS TO UNLEASH SOCIAL MEDIA INNOVATION IN HEALTHCARE

Social media innovations are changing how customers, product and service providers, and other parties communicate and collaborate in most industries—but healthcare and pharmaceutical firms have been lagging. "Regulatory concerns are holding innovation back," said Mark Bard, founder of the Digital Health Coalition (DHC), a nonprofit organization that is bringing together leaders and entrepreneurs from technology provides such as SAS, healthcare providers, pharmaceutical companies, policy-making agencies, academia, and more to help the industry overcome hurdles to fuller digital communication and collaboration. "At the same time," Bard added, "regulatory concerns can be a crutch; if we can get together [in the DHC] to address those problems, we should be able to make it easier for companies to experiment a little bit more."

Healthcare and pharmaceutical companies' difficulties with social media shine a spotlight on one of the biggest problems that organizations have in using this channel to improve how they communicate and interact with customers: content control. Who is responsible for posts on third-party platforms such as Facebook? "It should be that if you have the ability to alter, censor, or remove posts, you are responsible for that page now—not just from a legal perspective but from a control perspective," said Bard. However, many companies back off from such responsibility out of fear that they could be exposed to legal or regulatory issues if they take explicit control. As a result, many healthcare and pharmaceutical companies limit their "social strategy" to advertising or Facebook pages with closed comments.

Early in 2012, DHC issued its "Guiding Principles and Best Practices for Companies and Users," a document that provides a consensus view of the organization's membership about what "control over content" really means. DHC believes that eliminating the existing confusion over content control will enable all parties in the industry, including the U.S. Food and Drug Administration, to feel more confident about innovation in social media. A freer flow of information through social media would improve not only communication and collaboration, but also the quality of information that is useful for behavioral analytics about many matters of concern, including the effectiveness of patient care and treatment.

Selecting and Accessing Internal and External Social Media Data

The top social media data sources that respondents in our research are currently accessing are internal interaction records such as voice of the customer (VOC) logs (31%; see Figure 12). Another 15% of respondents plan to access these data sources within one year. This suggests both the understandable immaturity of organizations' pursuit of social media data sources for analysis as well as a desire to apply advanced analytics tools and methods to internal customer information sources that they may view as more important.

Which of the following social media data sources is your organization currently accessing or planning to access for customer analytics, brand, or marketing management purposes?
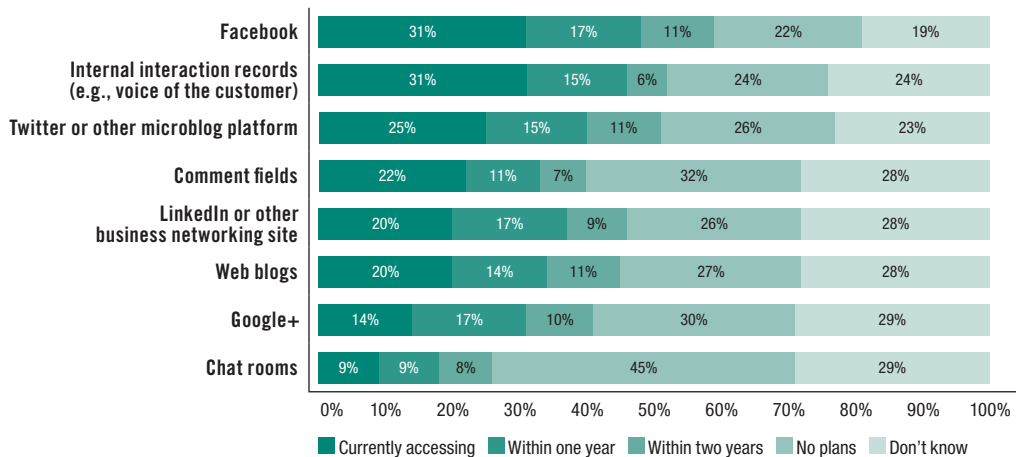

Figure 12. Based on 3,242 responses from 404 respondents; about eight responses per respondent, on average.

Text analytics tools, discussed earlier in the context of sentiment analysis, are established for analysis of internal content such as VOC records. These records are often poorly exploited because they are left to either agents or overtaxed data analysts to examine manually. Text analytics can automate discovery and provide greater analytical consistency and depth. Some organizations view social

media data analysis as a kind of extension of VOC analysis, except that customers' feedback and comments are less structured and are expressed outside the constraints of forms or questionnaires that require direct responses.

Overcoming Challenges to Accessing Social Media Network Data

TDWI Research found that Facebook (31%) and Twitter (25%) are the most common social media data sources that respondents currently access (referring again to Figure 12). LinkedIn (20%) and the relatively new Google+ (14%) are next. Another 17% plan to access Facebook within one year, and 15% plan to access Twitter data. Regarding LinkedIn and Google+ sources, an additional 17% of respondents plan to access each of these sources within one year.

Facebook and Twitter both present challenges for social media data analysis. The "big pipe" of data each site produces is no longer available to all analysts and developers. Facebook has many controls built into its application programming interface (API) to guard the privacy of its users and limit data access to its own proprietary use. Thus, external parties are essentially limited to what they can scrape from Facebook page descriptions. Except for fans' "likes" and comments recorded on a company's own Facebook pages, Facebook's data richness is mostly kept inside for its own advertising and behavior analysis.

Twitter's data fire hose consists of all public tweets, adding up by some estimates to more than 350 million tweets per day. This complete access is now available only through Twitter's partners. One such partner is Gnip, which for a price provides full Twitter streams through monitoring, BI, and discovery analytics tools such as Tableau. Gnip is strictly for performing analytics rather than for displaying tweets. Through a partnership with the Web development firm Automattic, Gnip can access content on Wordpress and Jetpack blogging and comment platforms, which enables Gnip to integrate these sources with Twitter data for social media analysis. DataSift is another Twitter partner with full access. Organizations can also to work with specialized analytics services providers such as Dataminr that are Twitter partners. Organizations that are not partners of Twitter and wish to develop their own analytic applications are limited to the Twitter streaming API, which provides near-real-time access to small subsets drawn from the tweet fire hose.

Finding the Right Role for Hadoop and MapReduce

As discussed earlier, Hadoop and MapReduce are critical technologies for big data customer analytics. Apache Hadoop's data storage and Hadoop Distributed File System (HDFS) can accommodate analysis of both structured and unstructured (or semi-structured) data. This allows organizations to use Hadoop to gain an integrated view of complex customer behavioral data that is usually separated into incompatible silos.

One-third of respondents find it important to integrate Hadoop, MapReduce, or related technologies with their BI and data warehouse systems.

Hadoop is also attractive because it is inherently scalable, running on commodity shared-nothing clusters, and can cost less than licensing systems from the big database companies. However, organizations have to factor in the technical talent needed to develop, implement, and maintain Hadoop programs, which is a reason many prefer to work with Hadoop as part of their existing data warehouse and database systems. TDWI Research found that 35% of respondents find it important to their organization's customer analytics objectives to integrate an implementation of Hadoop, MapReduce, and/or related technologies with their BI and data warehouse systems (Figure 13). One-fourth of respondents found such integration unimportant.

How important is it to your organization's customer analytics objectives to integrate implementation of Hadoop, MapReduce, and/or related technologies (e.g., Pig, NoSQL, or Hive) with BI and data warehouse systems?
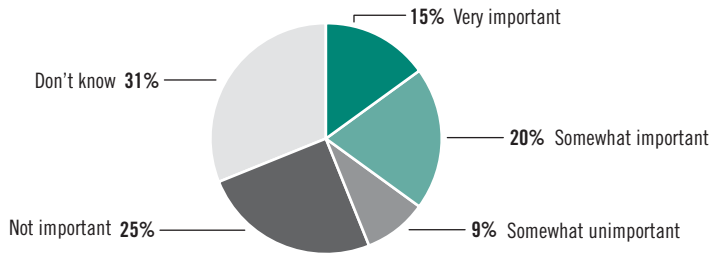

Figure 13. Based on answers from 393 respondents.

MapReduce, often mentioned in the same breath as Hadoop, is a framework for developers to write functions and process data on shared-nothing clusters, which has long been a difficult art. Consisting of "map" and "reduce" functions, it relieves developers of having to know the innards of distributed programming to produce code. MapReduce was developed by Google for its search engine and thus is particularly useful for analysis of unstructured data. MapReduce libraries and embedded, user-defined functions from database vendors can enable organizations to integrate SQL and MapReduce so that BI and analytics tools can work with a powerfully scalable parallel platform underneath. Database and application vendors can create built-in functions for specific kinds of analytics to run on the platform.

Small Role for Hadoop and MapReduce for Social Media Analytics. TDWI Research shows that the largest shares of respondents currently do not have plans to implement Hadoop, MapReduce, and/or related technologies such as Pig or Hive for social media data sources (41%) or for other customer data sources (40%; see Figure 14). Only 6% are currently implementing the technologies for these purposes, with somewhat higher percentages planning to implement them within one and two years. Among the small percentage of respondents who indicated that their organizations are currently implementing Hadoop, MapReduce, and related technologies, more than three-quarters of them (77%) are satisfied with their ability to determine marketing attribution. The social media data sources that most respondents overall are currently accessing are comment fields and internal interaction records such as VOC records.

Among the 6% whose organizations are currently implementing Hadoop, MapReduce, and related technologies, more than three-quarters are satisfied with their ability to determine marketing attribution.

Is your organization currently implementing or planning to implement Hadoop, MapReduce, and/or related technologies (e.g., Pig, NoSQL, or Hive) to support analysis of customer data, including the social media data sources identified in Figure 12?
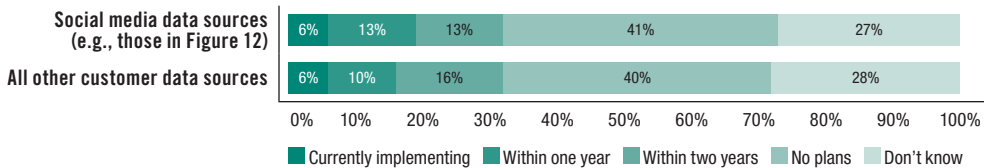

Figure 14. Based on 786 responses from 409 respondents; about two responses per respondent, on average.

The user story below illustrates how one organization used social media analysis to get a deeper look at customer experiences.

USER STORY

MZINGA CONQUERS SPEED AND SCALE CHALLENGES OF SOCIAL MEDIA ANALYSIS

"In the social media space, responding to things that happened two minutes ago is almost irrelevant." That's how Navdeep Alam, director of data architecture at Mzinga, Inc., explained the speed pressures his organization faces. Mzinga, a leading software-as-a-service (SaaS) producer of social software and analytics solutions, has 40 million users working with 15,000 different sites. For the company's OmniSocial business software suite, Mzinga relies on analytical speed to enable customers to create, design, and manage social and knowledge-sharing experiences.

Mzinga's analytics are critical to realizing business and marketing value from social media. "There is a difference between how you use social media to acquire customers, versus helping customers, versus creating a developer portal, and so on," said Alam. Mzinga's solutions address customer experience, support, and engagement; brand awareness; and employee experience through learning, development, recruiting, and related objectives. Mzinga can analyze customer behaviors across deployments of its solutions, whether they are implemented as part of closed communities on the Internet or embedded in other types of communities. "We need to be able to measure behaviors on both sides of the social graph, whether it is actions that you take or actions done unto you."

Mzinga's interest in deriving more knowledge from behavior data began about four years ago. "At that time," said Alam, "when you rated something on one of our sites, we always maintained just the final static state of the data. What became more interesting to us was how did that product on some customer site get a rating of 'four'? Well, it was because hundreds of people voted a two, a three, a five, and a six. But who were these people, and when did they rate the products? So, we started creating and collecting behavior data."

What Alam quickly learned was that the behavior data volume was big. "Our top client alone generated over a billion page events per month, and we had hundreds of customers. We hadn't seen this type of scale before. Plus, we wanted to do advanced analytics—to throw math at the questions. We didn't want to just sum, aggregate, or group the content. And we needed to have these queries perform with relatively low latency. We couldn't wait overnight for the query to give us a solution the next morning because often our customers need to know what's going on right now."

Hadoop was not the solution because it could not solve Mzinga's low-latency analytic requirements. Plus, the company needed data independence that reliance on the Hadoop Distributed File System (HDFS) could not provide. Mzinga chose to put its behavior data on the Aster Data Systems (now Teradata Aster) platform. "We are able to create optimized data partitioning schemes that support a recompute architecture against our analytics," said Alam. The Aster platform plays a key role in supporting the recompute architecture, which takes advantage of local data and previous computations to greatly lessen dependence on slower batch processing. "When users of our high-end, C-level dashboards ask who the 25 most influential users are on their systems, our architecture can support frequent recomputation of queries against all the data in our system, even as we are loading data. This enables us to do iterative analytical queries, including where our behavioral algorithms are tied to N-path queries to find patterns," Alam said.

Mzinga uses the Aster platform to perform many types of advanced analytics for identifying and forecasting connections, relationships, and influence among individuals and groups. By being able to keep up with the speed at which social media data is generated, Mzinga is able to provide business value to its customers and claim a competitive advantage.

Data Management and Integration Strategies

Customer analytic processes sink or swim with the quality of the underlying data. Data management runs the gamut from providing high-octane data performance for analytics to delivering high-quality, integrated data and protecting it through governance for privacy, security, and other regulatory requirements. Customer MDM systems and procedures are vital to giving executives, managers, and analysts a 360-degree view of a customer's information and interactions records. Customer data typically lives in many different systems and applications; with customer MDM, organizations seek to develop unified definitions, models, and representations of customers from data in different systems.

Full treatment of customer data management, integration, and MDM is beyond the scope of this report. These topics are covered in depth in other TDWI Best Practices Reports.[2] Here, we will take a brief look at some of the most significant data management and integration concerns within the context of supporting customer analytics.

TDWI Research found that the leading data management concern for customer analytics is the consolidation of all customer data for a comprehensive view (see Figure 15). More than one-third (36%) called it their "most challenging" issue and another 43% said it was "challenging." Large organizations in particular find that their customer data is distributed across multiple databases and applications held by marketing, sales, service, finance, and other departments. In addition, organizations are frequently in flux due to mergers, acquisitions, sales territory realignment, and other activities that can create chaos in customer data.

Data management runs the gamut from providing high-octane data performance for analytics to delivering high-quality, integrated data and protecting it through governance.

The leading data management concern for customer analytics is the consolidation of all customer data for a comprehensive view.

In your organization, which of the following data management concerns are the five most challenging when it comes to customer analytics, including social media data analysis?
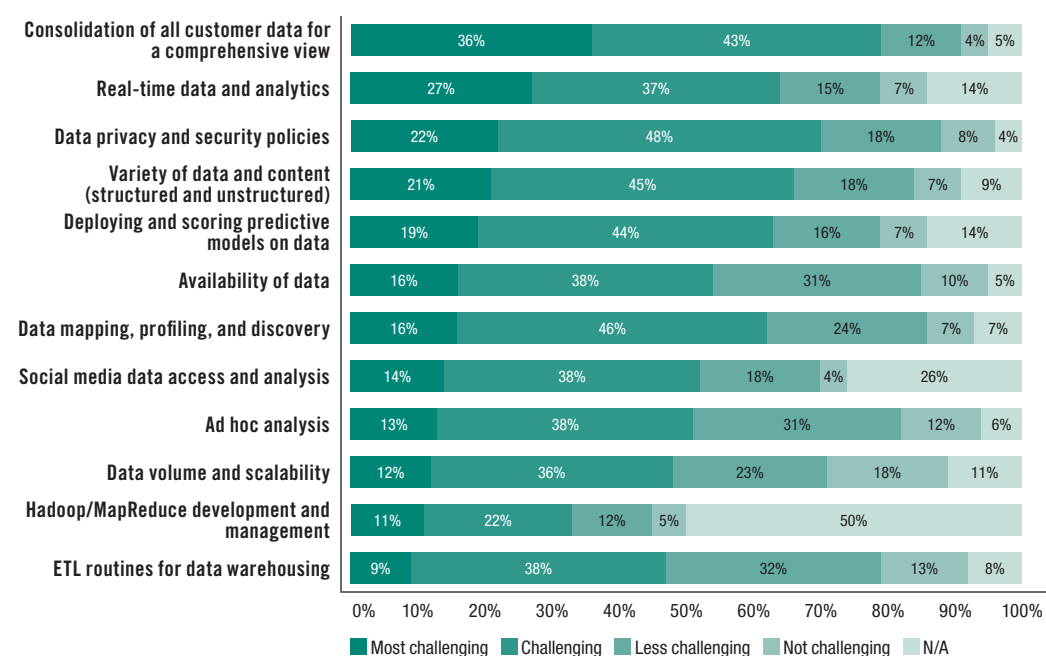

Figure 15. Based on 4,456 responses from 390 respondents; over 11 responses per respondent, on average.

To remedy problems, many organizations create an enterprise data warehouse to consolidate all customer data in one place. If this approach is impractical, organizations will build federated systems that use global metadata or customer MDM to track the customer data "in place" so that it may be

accessed by users of BI and analytics applications and managed more effectively for data governance and quality. Nearly two-thirds (62%) of respondents indicated that data mapping, profiling, and data discovery—three key tasks that are often part of data consolidation efforts—were challenging to their organizations, although just 16% described these tasks as "most challenging." These three tasks are often joined with data quality initiatives to enable marketing functions to eliminate wasted efforts and customer dissatisfaction that result from using incorrect customer data.

Addressing Real-Time Requirements for Customer Analytics

The need for speed is a key driver for organizations deploying analytic databases.

Moving toward real-time data and analytics is a significant challenge. Referring again to Figure 15, TDWI Research found that about one-quarter (27%) of respondents regard real-time data and analytics as one of their "most challenging" issues, and another 37% termed it challenging. Speed is a key reason organizations choose to deploy analytics software and marketing automation applications. The faster organizations can analyze customer data, including data generated in social media, the sooner they can put their marketing processes in motion to respond to changes in customer demand and adjust to newly uncovered behavioral clues. The need for speed is a key driver for organizations deploying analytic databases, including columnar databases, so that they can avoid delays encountered with transaction-oriented database systems and designs that are not well suited for analytic workloads.

Speed is also a reason some organizations will try to adjust analytics so that the point of decision happens sooner, at the moment events are detected in real time. Credit scoring for purchase and loan approvals processes, for example, can be implemented with predictive models that pick up event data as it streams into systems. With speed part of the basis of competition, operational intelligence systems such as these cannot function within enterprise data warehouse architectures that take too long to extract, load, and transform the data. Deploying and scoring predictive models is a challenge for 63% of respondents; 19% described it as "most challenging."

Meeting Data Governance, Quality, and Integration Needs

Data privacy and security policy concerns hover over customer marketing and analytics due to customers' sensitivity about how their behavioral and identification data elements are being shared, analyzed, and managed. Concerns are rising as well about monitoring and analysis of user activity in social media. Facebook, Google, and other social networking providers are under constant scrutiny for the privacy implications of their new features and services, particularly those that result from link analysis. In our research, 22% said data privacy and security policies were "most challenging" and 48% said they were challenging. Organizations need to establish and adhere to governance policies that cover privacy, security, and other regulatory issues for customer data that they are analyzing and sharing.

Only 10% of respondents considered their customer data of the highest quality, with no fragmentation, duplication, or inconsistency.

Reducing Chaos with Data Quality and Integration Processes. Data quality and consistency are a constant challenge with customer data. Internal sources can be conflicting, incomplete, and incorrect; common external data sources can be out of date; and social media data is inconsistent for even basic user information. Our research found that only 10% of respondents considered their customer data of the highest quality, with no fragmentation, duplication, or inconsistency (see Figure 16). Clearly, respondents feel that their organizations have work to do to address common data quality problems, which may be preventing them from delivering high-quality data for customer analytics.

What is the extent of data quality of your customer and related data within and across the information sources? (Please select the answer that best fits your organization.)
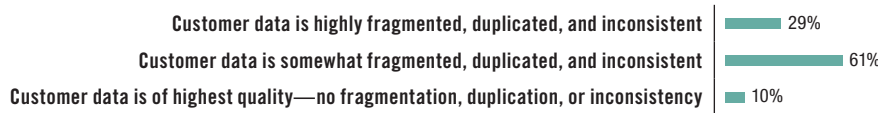
Customer data is highly fragmented, duplicated, and inconsistent	29%
Customer data is somewhat fragmented, duplicated, and inconsistent	61%
Customer data is of highest quality—no fragmentation, duplication, or inconsistency	10%

Figure 16. Based on answers from 385 respondents.

Data quality processes are often part of broader customer MDM processes for gaining a consistent and reliable view of all customer data. Fragmented data costs organizations by dragging down both customer-facing operations and analytical processes. In our survey, we asked respondents which customer data integration steps are undertaken by their organizations to prepare data for customer analytics (see Figure 17). The largest percentage is cleansing, updating, and completing data from multiple sources (61%). Just over half (52%) are taking steps to consolidate records, link records from disparate sources, and purge duplicates. Slightly more respondents (53%) are addressing data quality issues, including discovery of where data is located in incorrect fields.

Fragmented data drags down both customer-facing operations and analytical processes.

Which of the following customer data integration steps are undertaken by your organization to prepare data for customer analytics? (Please select all that apply.)
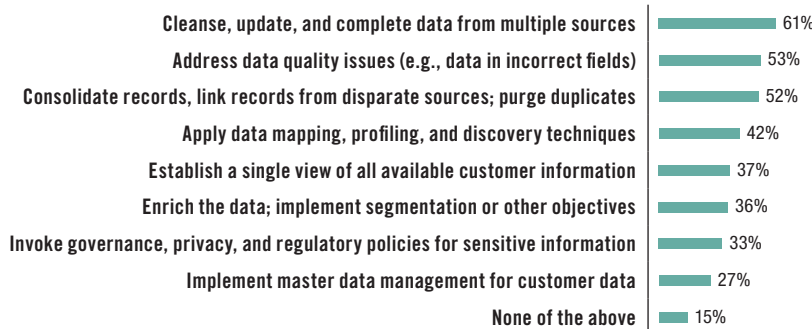
Cleanse, update, and complete data from multiple sources	61%
Address data quality issues (e.g., data in incorrect fields)	53%
Consolidate records, link records from disparate sources; purge duplicates	52%
Apply data mapping, profiling, and discovery techniques	42%
Establish a single view of all available customer information	37%
Enrich the data; implement segmentation or other objectives	36%
Invoke governance, privacy, and regulatory policies for sensitive information	33%
Implement master data management for customer data	27%
None of the above	15%

Figure 17. Based on 1,345 responses from 378 respondents; about 3.5 responses per respondent, on average.

Implementing Customer Master Data Management. Customer MDM can help organizations streamline data integration by establishing a reference source or registry that provides a common set of business-oriented customer definitions. Organizations that have established MDM are able to use this higher-level reference to connect business names or objects in multiple data sources, often by using the sources' metadata repositories or documentation about data schemas. Integration routines can employ an MDM resource to eliminate redundant steps in data integration, quality, and profiling routines.

TDWI Research found that 27% of respondent organizations are implementing MDM for customer data, suggesting that MDM still has a ways to evolve in most organizations. Although most mainstream MDM systems do not yet reach out to social media, some technology solutions are now making the extension possible. Advanced customer MDM can enable organizations to integrate corporate customer profiles in internal systems with social customer profiles and external sources, including Facebook, LinkedIn, and other social media.

Twenty-seven percent of respondent organizations are implementing MDM for customer data, suggesting that MDM still has a ways to evolve in most organizations.

Accessing Structured and Unstructured Data Sources

Customer data warehouses are the predominant data systems that users access (58%), according to our research, with departmental data marts the second highest (38%; see Figure 18). Not to be forgotten are spreadsheets; more than half of respondents (57%) are using them as sources for customer data analytics. As noted above, our research found that MDM is still maturing; only 17% of organizations are giving users access to MDM reference data sources.

Just over one-quarter (27%) of respondents are providing access to documents, XML data, or unstructured content. Textual content, whether semi- or unstructured, makes up the lion's share of big data in most organizations. XML is an important standard for organizations to begin to bridge structured and unstructured information and, along with HTML and related standards, to increase the semantic richness of content on Web pages and in other forms and documents. As standards are adopted for content, it will grow easier for customer analytics tools and analytic database systems to reach into content sources for customer information and blend it with structured data from BI and data warehouse systems.

Standards will make it easier for customer analytics tools to reach into content sources for customer information and blend it with structured data.

In your organization, which of the following data systems, applications, or services do users access to gain a comprehensive view of customer data for performing analytics? (Please select all that apply.)
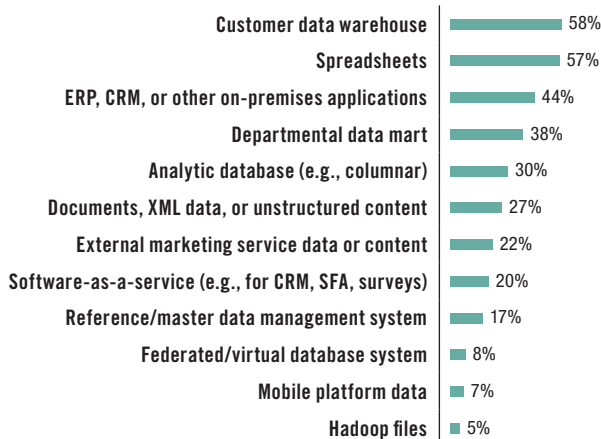

Customer data warehouse	58%
Spreadsheets	57%
ERP, CRM, or other on-premises applications	44%
Departmental data mart	38%
Analytic database (e.g., columnar)	30%
Documents, XML data, or unstructured content	27%
External marketing service data or content	22%
Software-as-a-service (e.g., for CRM, SFA, surveys)	20%
Reference/master data management system	17%
Federated/virtual database system	8%
Mobile platform data	7%
Hadoop files	5%

Figure 18. Based on 1,252 responses from 376 respondents; a little over three responses per respondent, on average.

Vendor Products

The firms that sponsored this report are among the leaders and innovators in providing tools, technology platforms, applications, and services for customer analytics, including social media analytics, as well as customer data integration and management. To get a sense of where the industry as a whole is headed, here is a brief look at the portfolios of these vendors. (Note: The vendors and products mentioned here are representative, and the list is not intended to be comprehensive.)

Greenplum, a division of EMC

EMC Greenplum, the centerpiece of EMC's Big Data Division since its 2010 acquisition, provides software and services for big data warehouses and analytics. The Greenplum Database, based on the open source PostgreSQL, offers a scalable and versatile system built on a shared-nothing, massively parallel processing (MPP) architecture. Offering an open source Hadoop distribution, EMC Greenplum is an important player in the Hadoop big data analytics arena. EMC Greenplum packages its Hadoop (based on MapReduce's distribution) with the Greenplum Database into a family of appliances: the Greenplum HD Data Computing Appliance Modules. These are designed to relieve data scientists of data configuration, integration, and design complexity so that they can focus on discovery and testing, particularly with unstructured and semi-structured customer behavior data and social media data. EMC Greenplum's Unified Analytic Platform brings these elements together in an integrated fashion as well for an extended platform architecture beyond a single appliance. The company also sports an enterprise data cloud platform, Chorus. The Chorus offering has social networking, workflow, agile development, and other tools and features to improve data-scientist collaboration on big data analytics.

IBM

The Business Analytics software portfolio from IBM, headquartered in Armonk, New York, includes SPSS predictive analytics products, Cognos BI products, and others. IBM's Business Analytics portfolio is therefore able to address customer analytics requirements from both predictive analytics and BI perspectives. Users can implement SPSS statistics, modeling, data collection, and deployment tools to determine, for example, which marketing campaigns and channels will be most effective. Predictive analytics results can be displayed in the BI environments; the PA tools can also be fed data that BI users are analyzing. Marketing directors and managers can employ Cognos TM1 to look at marketing campaigns from a financial planning perspective to determine the optimal mix for gaining the maximum return on investment. Software for the planning and execution phases can also be integrated with IBM Smarter Commerce solutions such as those for the market phase, including digital marketing optimization, social media marketing, marketing resource management, and more.

Informatica

Informatica, based in Redwood City, California, is a leading data integration and management software provider. The company has deepened its portfolio in recent years, including MDM software. Organizations use Informatica MDM to provide reliable and enriched 360-degree customer views by creating corporate customer profiles from disparate internal sources and integrating them with social customer profiles in external sources such as Facebook, LinkedIn, and Twitter. Informatica 9.5, released in May 2012, addresses several important areas for customer data integration and management. These include greater capabilities for data quality, cleansing, and governance, including data privacy, enabling you to profile data sources to find sensitive data and apply unified policies across sources. The 9.5 release also helps tap social and Hadoop sources for customer data analysis. Regarding Hadoop, the 9.5 release enables organizations to embed routines and business logic in Hadoop for data transformations, cleansing, and more.

SAP   SAP, the global enterprise software provider headquartered in Weinheim, Germany, augments its application portfolio with social media intelligence technology from NetBase, of Mountain View, California, through a reseller partnership. NetBase's cloud-based solutions are centered on the company's natural language processing technology, which comprises text analytics and machine learning, and also includes Web crawling technology and a user interface. SAP integrates the NetBase with its applications to provide solutions to address a number of business and marketing needs. These include an on-demand social customer engagement that allows contact center agents to interact with, or follow up with, customers who are tweeting or putting up Facebook posts. Other integrated solutions add social media intelligence for service enhancement, marketing campaign management, corporate brand analysis, Balanced Scorecard (for tracking social media sentiment along with traditional dimensions), competitive analysis, and gaining a 360-degree customer view. In addition, in early 2012 the company introduced SAP BusinessObjects Predictive Analysis, which offers tools for designing predictive models as well as advanced visualization and visual discovery. Important for real-time customer analytics is that this solution can implement embedded predictive functions in SAP HANA to provide in-memory and in-database performance advantages.

SAS   SAS, based in Cary, North Carolina, offers a wide range of software and services for business and customer analytics. The company's Customer Intelligence suite has an Integrated Marketing Management (IMM) Framework that aligns and integrates the planning, operational, campaign management and execution, and analytic functions across the entire marketing process. The IMM framework includes information and analytics solutions for resource management; customer link analytics (including tools for profiling, segmentation, and targeting); customer value and profitability management; and event-triggered marketing. SAS's social media analytics offerings give organizations the ability to discover hidden social network links and spot influencers. The company's social media analytics include tools for discovering sentiment, attribution, social graph analysis, and more. SAS technologies aim at enabling organizations to apply their analytics to social operations. This would include giving customer service representatives insights so that in real time they can spot comments in social media and respond with offers, comments, or other feedback coordinated with overall strategies. SAS Enterprise Miner and High Performance Analytics offer technologies for gaining deeper customer insight through data mining, predictive analytics, and more.

Tableau Software   Seattle-based Tableau offers self-service business intelligence and data discovery tools. The company's products are known for their rich data visualization and data interaction capabilities in its mobile, desktop, and server editions. Tableau's self-service capabilities are important for customer and social media data analytics because the primarily nontechnical users working with the data need the flexibility to ask what-if questions, try new hypotheses against different sets of data, and see the impact of marketing projects in an ongoing fashion rather than in a monthly or quarterly report. Tableau supports broad sharing and collaboration on data visualizations and analysis through Tableau Public, a free service on the Web. Users, including media sources, post their visualizations on Tableau Public, where they are publicly accessible. Users can supply data from spreadsheets or other application data sources for visualizations loaded into Tableau Public. Anyone can then download the visualizations and use them in blogs or Web sites.

Teradata addresses customer analytics and intelligence from three perspectives: integrated data warehousing, big data analytics, and business applications, which are addressed by Teradata, the Teradata Aster platform, and Aprimo, a division of Teradata, respectively. Teradata, which for decades has been a leader in data warehousing, has expanded its portfolio into an integrated ecosystem that can address analytical processing needs with appliances, enterprise data warehouses, in-database analytics, specialized columnar systems, and more. The company also offers the Teradata Analytic Private Cloud, which can address on-demand performance requirements for customer intelligence. Teradata acquired the Aster MapReduce Platform in 2011. This system is suited to new types of customer analytics and new forms of big data, including social media data and other semi- or unstructured data. The Aster Platform offers speed and scale for testing hypotheses and learning from the results more quickly, which enables organizations to run more tests with more variables. Version 5.0 of Aster's Analytical Foundation supports multi-channel attribution analysis and customer behavioral analysis. The Aprimo integrated marketing management application can work with the database platforms for digital marketing optimization. Aprimo enables organizations to track marketing execution across channels from within one system.

Teradata (including Teradata Aster and Aprimo)

Vertica, based in Billerica, Massachusetts, focuses on real-time business analytics involving very large and complex data. The company's core technology is its columnar database, which it views as a performance improvement over standard, row-based relational databases for access needs that are common for many kinds of analytics. Vertica pairs its column orientation with data compression and encoding for more efficient storage and use of I/O bandwidth. The Vertica Analytics Platform functions within an MPP architecture. Together, these technology features can increase speed and reliability for organizations pursuing real-time analytics of customer behavior so that they can improve experiences while customers are engaged over the Web or on mobile devices. Vertica also supports in-database analytics to improve performance and reduce the need for extraction to separate systems. Vertica's 5.1 release, announced in February 2012, includes integration with HP's Autonomy IDOL 10, a platform for extracting meaning from audio, video, social media, and other types of content as well as structured data. This "HP Next Generation Information Platform" broadens the company's portfolio for analyzing different types of customer-generated information.

Vertica, an HP Company

Recommendations

Setting Strategies for Customer and Social Media Analytics

Predictive analytics can help organizations anticipate the results of active strategies.

Use social media data analysis to support an active, not passive, social media strategy. In competitive, fast-moving markets, organizations cannot just passively listen to and analyze social media data. The analytics should plug into strategies for engaging users and customers on social networks and comment sites. Predictive analytics can help organizations anticipate the results of active strategies. Special events such as tweet-ups can build on customer data analysis and create positive exchanges and engagement.

Take a holistic view of the potential contributions of social media data analysis. Understanding behavior in the social sphere can have a positive impact, not just on marketing and sales functions, but also on services and other processes in the organization. Marketing executives should use social media insights to improve brand awareness and reputation throughout the organization.

Chief marketing officers today have to be more than just brand stewards and marketing strategists. They are accountable for the ROI of marketing projects and campaigns.

Give CMOs and marketing executives the ability to understand financial impact. Chief marketing officers today have to be more than just brand stewards and marketing strategists. They are accountable for the ROI of marketing projects and campaigns and need to know the financial impact of marketing decisions. Business intelligence and OLAP capabilities can help CMOs align marketing with financial objectives.

Apply analytics to gain a more accurate understanding of marketing attribution. "Last-touch" attribution may be easy to affix, but it blinds the organization to the real impact and effectiveness of the whole series of marketing, engagement, and sales efforts along a customer's path to purchase. Powerful analytics, along with big data, can help organizations get a better understanding of what truly affected a customer's decision to make a purchase.

Deploying and Implementing the Right Tools and Technologies

Evaluate specialized analytic database technologies for customer analytics. Technologies such as columnar databases and Hadoop/MapReduce systems that are tailored to large-scale analytic workloads can help organizations increase the speed and depth of customer analytics. Organizations should evaluate these systems, including their packaging as appliances, which can reduce the IT function's configuration, design, and maintenance responsibilities.

Analytics need to be iterative, fluid, and self-creating rather than tied to more rigid and static paths that have worked for BI reporting and data warehousing.

Don't look at customer analytics through BI and data warehousing glasses. Analytics operate differently—that is, with less structure—from traditional BI and data warehousing. Analytics need to be iterative, fluid, and self-creating rather than tied to more rigid and static paths that have worked for BI reporting and data warehousing. Predictive analytics and modeling can help organizations fill in the gaps for BI users accustomed to simple dashboards and more limited OLAP functionality.

Give nontechnical users self-service data discovery and visual analysis capabilities. Users in marketing, sales, and service functions need actionable information and guided nontechnical means of exploring the data. Visual analysis is often the best way for these users to work with data; they can see immediately what is important and can more easily share insights with colleagues. Organizations should take advantage of new technology capabilities in tools for data visualization and self-directed exploration.

Supporting Analytics with Data Management, Governance, and Integration

Improve customer data integration with MDM and consolidate sources where possible. To give analysts and nontechnical users the 360-degree view of customer data they need, IT and data management functions develop plans and implement technology that can enable greater customer data integration. Where possible, silos and data marts that contain conflicting or duplicate customer data should be consolidated into enterprise data warehouses. Implement customer MDM to streamline and centralize integration, data quality, profiling, and related tasks to avoid inconsistency and reduce costs.

Ensure good integration between social media data analysis and other customer analytics processes. Social media data analysis, by opening a window on customers and prospect behavior that internal data and service records do not capture, can complement the organization's other customer intelligence and analytics processes. However, in many organizations the two areas are separate. Organizations should forge closer integration between these analytical processes.

Integrate social media data analysis with CRM and customer-oriented applications. Organizations should avoid letting social media data analytics, including monitoring applications, become new data silos. Nearly all applications that touch the customer, including those used in contact centers, fulfillment, and product development, should have a social media component. They could benefit from being able to consume social media data analysis.

Establish data governance to ensure adherence to privacy and security rules and regulations. Customer data is not only sensitive to the organization's strategies; it is also sensitive to the customers. The organization's underlying data management must monitor the use and sharing of sensitive customer data, including personally identifiable information. Data governance committees should exercise oversight to ensure adherence.

Recognize and address tensions between IT and marketing over analytics. Customer analytics and the budget for analytic processes are often in the middle of tensions between IT and marketing. IT is used to owning development and data access, and gathering user requirements all at once. Data scientists and marketing analysts need to test predictive hypotheses and explore the data before they know exactly what data they need. Organizations should resolve the tensions before they become obstacles to meeting strategic objectives.

Implement customer MDM to streamline and centralize integration, data quality, profiling, and related tasks to avoid inconsistency and reduce costs.

Organizations should avoid letting social media data analytics, including monitoring applications, become new data silos.

GREENPLUM **EMC²**

Greenplum, a division of EMC
www.greenplum.com

Greenplum, a division of EMC, is driving the future of big data analytics with breakthrough products that harness the skills of data science teams to help global organizations realize the full promise of business agility and become data-driven, predictive enterprises. Greenplum's products embody the power of open systems, cloud computing, virtualization, and social collaboration, enabling global organizations to gain greater insight and value from their data than ever before possible. EMC Corporation (NYSE: EMC) is a global leader in enabling businesses and service providers to transform their operations and deliver IT as a service.

INFORMATICA®
The Data Integration Company™

Informatica
www.informatica.com/mdm

Informatica is the world's number one independent provider of data integration software. Thousands of organizations rely on Informatica for maximizing return on data to drive their top business imperatives. To maximize return on data, Informatica increases the value of data by delivering relevant, trustworthy, timely, authoritative, actionable, accessible, holistic, and secure data. At the same time, Informatica lowers the cost of data by bringing down business costs, labor costs, software costs, hardware costs, and storage costs. Visit www.informatica.com/mdm for demos, videos, and white papers about how Informatica MDM helps improve and enrich your customer information with social network data.

IBM

IBM
www.ibm.com/analytics

IBM Business Analytics software delivers the actionable insights decision-makers need to achieve better business performance. IBM offers a comprehensive, unified portfolio of business intelligence, predictive and advanced analytics, financial performance and strategy management, governance, risk and compliance, and analytic applications.

With IBM software, companies can spot trends, patterns, and anomalies, compare "what-if" scenarios, predict potential threats and opportunities, identify and manage key business risks, and plan, budget, and forecast resources. With these deep analytic capabilities our customers around the world can better understand, anticipate, and shape business outcomes.

SAP

SAP
www.sap.com

Headquartered in Walldorf, Germany, SAP is the market leader in enterprise application software. Founded in 1972, SAP (which stands for "Systems, Applications, and Products in Data Processing") has a rich history as a true industry leader. SAP applications and services enable more than 183,000 customers worldwide to operate profitably, adapt continuously, and grow sustainably. With revenue (IFRS) of €12.5 billion for the year 2010, SAP has more than 55,000 employees and sales and development locations in more than 50 countries worldwide.

SAP is listed on several exchanges, including the Frankfurt stock exchange and NYSE, under the symbol "SAP."

SAS

www.sas.com/solutions/crm

Transitioning to a customer-focused business strategy provides the best opportunity to drive profitable revenue growth. SAS enables this through a full suite of customer intelligence solutions fueled by the world's best analytics, giving organizations the power to:

- Find the most profitable growth opportunities.

- Take the best marketing actions.

- Maximize cross-business impact.

Used at more than 55,000 sites in 129 countries, including 90 of the top 100 companies on the 2011 *Fortune* Global 500 list, SAS is proud to celebrate more than 35 years of giving customers THE POWER TO KNOW®.

Learn how top companies are using SAS at www.sas.com/solutions/crm.

Teradata (including Teradata Aster and Aprimo)

teradata.com

Teradata is the world's largest company focused on integrated data warehousing, big data analytics, and business applications. Our powerful solutions portfolio and database are the foundation on which we've built our leadership position in business intelligence and are designed to address any business or technology need for companies of all sizes.

Only Teradata gives you the ability to integrate your organization's data, optimize your business processes, and accelerate new insights like never before. The power unleashed from your data brings confidence to your organization and inspires leaders to think boldly and act decisively for the best decisions possible.

Tableau Software

www.tableausoftware.com/trial

Tableau Software helps people see and understand data. Tableau's award-winning software delivers fast analytics, visualization, and rapid-fire business intelligence on data of any size, format, or subject. The result? Anyone can get answers from data quickly, with no programming required. From executive dashboards to ad hoc reports, Tableau lets you share mobile and browser-based, interactive analytics in a few clicks. More than 7,000 organizations, including some of the world's largest enterprises, rely on Tableau Software. See how Tableau can help you by downloading the free trial at www.tableausoftware.com/trial.

Vertica

www.vertica.com

Vertica, an HP Company, is the leading provider of next-generation analytics platforms enabling customers to monetize all of their data. Vertica's elasticity, scale, performance, and simplicity are unparalleled in the industry, delivering 50x–1000x the performance of traditional solutions at 30% the total cost of ownership. Vertica powers some of the largest organizations and most innovative business models globally, including Zynga, Groupon, Twitter, Verizon, Guess Inc., Admeld, Capital IQ, Mozilla, AT&T, and Comcast.