applications for all the functional areas of business, and for government and scientific work. One popular use for data mining is to provide detailed analyses of patterns in customer data for one-to-one marketing campaigns or for identifying profitable customers.

For example, Harrah's Entertainment, the second-largest gambling company in its industry, uses data mining to identify its most profitable customers and generate more revenue from them. The company continually analyzes data about its customers gathered when people play its slot machines or use Harrah's casinos and hotels. Harrah's marketing department uses this information to build a detailed gambling profile, based on a particular customer's ongoing value to the company. For instance, data mining lets Harrah's know the favorite gaming experience of a regular customer at one of its Midwest riverboat casinos, along with that person's preferences for room accomodations, restaurants, and entertainment. This information guides management decisions about how to cultivate the most profitable customers, encourage those customers to spend more, and attract more customers with high revenue-generating potential. Business intelligence has improved Harrah's profits so much that it has become the centerpiece of the firm's business strategy.

Predictive analytics use data mining techniques, historical data, and assumptions about future conditions to predict outcomes of events, such as the probability a customer will respond to an offer or purchase a specific product. For example, the U.S. division of The Body Shop International plc used predictive analytics with its database of catalog, Web, and retail store customers to identify customers who were more likely to make catalog purchases. That information helped the company build a more precise and targeted mailing list for its catalogs, improving the response rate for catalog mailings and catalog revenues.

Text Mining and Web Mining

Business intelligence tools deal primarily with data that have been structured in databases and files. However, unstructured data, most in the form of text files, is believed to account for over 80 percent of an organization's useful information. E-mail, memos, call center transcripts, survey responses, legal cases, patent descriptions, and service reports are all valuable for finding patterns and trends that will help employees make better business decisions. **Text mining** tools are now available to help businesses analyze these data. These tools are able to extract key elements from large unstructured data sets, discover patterns and relationships, and summarize the information. Businesses might turn to text mining to analyze transcripts of calls to customer service centers to identify major service and repair issues.

Text mining is a relatively new technology, but what's really new are the myriad ways in which unstructured data are being generated by consumers and the business uses for these data. The Interactive Session on Technology explores some of these business applications of text mining.

The Web is another rich source of valuable information, some of which can now be mined for patterns, trends, and insights into customer behavior. The discovery and analysis of useful patterns and information from the World Wide Web is called **Web mining**. Businesses might turn to Web mining to help them understand customer behavior, evaluate the effectiveness of a particular Web site, or quantify the success of a marketing campaign. For instance, marketers use Google Trends and Google Insights for Search services, which track the popularity of various words and phrases used in Google search queries, to learn what people are interested in and what they are interested in buying.

INTERACTIVE SESSION: TECHNOLOGY

WHAT CAN BUSINESSES LEARN FROM TEXT MINING?

Text mining is the discovery of patterns and relationships from large sets of unstructured data—the kind of data we generate in e-mails, phone conversations, blog postings, online customer surveys, and tweets. The mobile digital platform has amplified the explosion in digital information, with hundreds of millions of people calling, texting, searching, "apping" (using applications), buying goods, and writing billions of e-mails on the go.

Consumers today are more than just consumers: they have more ways to collaborate, share information, and influence the opinions of their friends and peers, and the data they create in doing so have significant value to businesses. Unlike structured data, which are generated from events such as completing a purchase transaction, unstructured data have no distinct form. Nevertheless, managers believe such data may offer unique insights into customer behavior and attitudes that were much more difficult to determine years ago.

For example, in 2007, JetBlue experienced unprecedented levels of customer discontent in the wake of a February ice storm that resulted in widespread flight cancellations and planes stranded on Kennedy Airport runways. The airline received 15,000 e-mails per day from customers during the storm and immediately afterwards, up from its usual daily volume of 400. The volume was so much larger than usual that JetBlue had no simple way to read everything its customers were saying.

Fortunately, the company had recently contracted with Attensity, a leading vendor of text analytics software, and was able to use the software to analyze all of the e-mail it had received within two days. According to JetBlue research analyst Bryan Jeppsen, Attensity Analyze for Voice of the Customer (VoC) enabled JetBlue to rapidly extract customer sentiments, preferences, and requests it couldn't find any other way. This tool uses a proprietary technology to automatically identify facts, opinions, requests, trends, and trouble spots from the unstructured text of survey responses, service notes, e-mail messages, Web forums, blog entries, news articles, and other customer communications. The technology is able to accurately and automatically identify the many different "voices" customers use to express their feedback (such as a negative voice, positive voice, or conditional voice), which helps organizations pinpoint key events and relationships, such as intent to buy, intent to leave, or customer "wish" events. It can reveal specific product and service issues, reactions to marketing and public relations efforts, and even buying signals.

Attensity's software integrated with JetBlue's other customer analysis tools, such as Satmetrix's Net Promoter metrics, which classifies customers into groups that are generating positive, negative, or no feedback about the company. Using Attensity's text analytics in tandem with these tools, JetBlue developed a customer bill of rights that addressed the major issues customers had with the company.

Hotel chains like Gaylord Hotels and Choice Hotels are using text mining software to glean insights from thousands of customer satisfaction surveys provided by their guests. Gaylord Hotels is using Clarabridge's text analytics solution delivered via the Internet as a hosted software service to gather and analyze customer feedback from surveys, e-mail, chat messaging, staffed call centers, and online forums associated with guests' and meeting planners' experiences at the company's convention resorts. The Clarabridge software sorts through the hotel chain's customer surveys and gathers positive and negative comments, organizing them into a variety of categories to reveal less obvious insights. For example, guests complained about many things more frequently than noisy rooms, but complaints of noisy rooms were most frequently correlated with surveys indicating an unwillingness to return to the hotel for another stay.

Analyzing customer surveys used to take weeks, but now takes only days, thanks to the Clarabridge software. Location managers and corporate executives have also used findings from text mining to influence decisions on building improvements.

Wendy's International adopted Clarabridge software to analyze nearly 500,000 messages it collects each year from its Web-based feedback forum, call center notes, e-mail messages, receipt-based surveys, and social media. The chain's customer satisfaction team had previously used spreadsheets and keyword searches to review customer comments, a very slow manual approach. Wendy's management was looking for a better tool to speed analysis, detect emerging issues, and pinpoint troubled areas of the business at the store, regional, or corporate level.

The Clarabridge technology enables Wendy's to track customer experiences down to the store level within minutes. This timely information helps store, regional, and corporate managers spot and address problems related to meal quality, cleanliness, and speed of service.

Text analytics software caught on first with government agencies and larger companies with information systems departments that had the means to properly use the complicated software, but Clarabridge is now offering a version of its product geared towards small businesses. The technology has already caught on with law enforcement, search tool interfaces, and "listening platforms" like Nielsen Online. Listening platforms are text mining tools that focus on brand management, allowing companies to determine how consumers feel about their brand and take steps to respond to negative sentiment.

Structured data analysis won't be rendered obsolete by text analytics, but companies that are able to use both methods to develop a clearer picture of their customers' attitudes will have an easier time establishing and building their brand and gleaning insights that will enhance profitability.

Sources: Doug Henschen, "Wendy's Taps Text Analytics to Mine Customer Feedback," *Information Week*, March 23, 2010; David Stodder," How Text Analytics Drive Customer Insight" *Information Week*, February 1, 2010; Nancy David Kho, "Customer Experience and Sentiment Analysis," *KMWorld*, February 1, 2010; Siobhan Gorman, "Details of Einstein Cyber-Shield Disclosed by White House," *The Wall Street Journal*, March 2, 2010; www.attensity.com, accessed June 16, 2010; and www.clarabridge.com, accessed June 17, 2010.

CASE STUDY QUESTIONS

1. What challenges does the increase in unstructured data present for businesses?
2. How does text-mining improve decision-making?
3. What kinds of companies are most likely to benefit from text mining software? Explain your answer.
4. In what ways could text mining potentially lead to the erosion of personal information privacy? Explain.

MIS IN ACTION

Visit a Web site such as QVC.com or TripAdvisor.com detailing products or services that have customer reviews. Pick a product, hotel, or other service with at least 15 customer reviews and read those reviews, both positive and negative. How could Web content mining help the offering company improve or better market this product or service? What pieces of information should highlighted?

Web mining looks for patterns in data through content mining, structure mining, and usage mining. Web content mining is the process of extracting knowledge from the content of Web pages, which may include text, image, audio, and video data. Web structure mining extracts useful information from the links embedded in Web documents. For example, links pointing to a document indicate the popularity of the document, while links coming out of a document indicate the richness or perhaps the variety of topics covered in the document. Web usage mining examines user interaction data recorded by a Web server whenever requests for a Web site's resources are received. The usage data records the user's behavior when the user browses or makes transactions on the Web site and collects the data in a server log. Analyzing such data can help companies determine the value of particular customers, cross marketing strategies across products, and the effectiveness of promotional campaigns.

DATABASES AND THE WEB

Have you ever tried to use the Web to place an order or view a product catalog? If so, you were probably using a Web site linked to an internal corporate database. Many companies now use the Web to make some of the information in their internal databases available to customers and business partners.