



SIGNALS FOR STRATEGISTS

## **Data ecosystems**

How third-party information can enhance data analytics

David Schatsky, Jonathan Camhi, and Craig Muraskin

Companies are increasingly seeking better insights by tapping into third-party data. Outside data can bring lots of opportunity, but using it effectively can be challenging.

**M**ASTERING DATA ANALYTICS is how companies avoid flying blind. Increasingly, this requires tapping into data from outside an organization's four walls. Growing numbers of companies are doing so in pursuit of an analytical edge. But using third-party data sources effectively can be very challenging. To boost the business value of their companies' analytics efforts, leaders can adopt key practices to navigate the complexity of third-party data.

## Signals

- In a recent survey, nearly half of companies reported using external data in their analytics activities.<sup>1</sup>
- Organizations in industries including financial services, logistics, technology, health care, retail, and the public sector are using external data to gain new insights that can help increase efficiency and revenue.<sup>2</sup>
- In another survey, 92 percent of data analytics professionals said their firms needed to increase use of external data sources; 54 percent said their companies plan to increase spending on it.<sup>3</sup>
- Half of analytics professionals surveyed recently reported that their companies are selling their own data.<sup>4</sup>
- To gain external data or services based on that data, companies in recent years have been executing mergers, acquisitions, and commercial partnerships.<sup>5</sup>

## Mining external data for insights is increasingly important

Companies know they can gain valuable insights by analyzing the data they generate from their operations. But internally generated information can leave gaps, and companies are increasingly moving to incorporate new, nontraditional, and external sources of data into their analyses. This data can include almost anything, from historical demographic and weather data to satellite imagery and private company information.

Companies increasingly operate as part of networks consisting of business partners such as suppliers, resellers, channel partners, regulators, and other stakeholders. These networks are often globally distributed and potentially affected by economic, political, and/or environmental factors. Analyzing external data can help companies see risks and opportunities that they would miss with inputs limited to data generated from internal operations, customers, and first-tier suppliers. Analyzing external data can illuminate how factors such as shifting consumer behaviors, competitor initiatives, or geopolitical events can affect a business.

As most business and technology professionals know, the volume of data being created, shared, and stored is increasing at an exponential pace. According to one study, the data stored in data centers will nearly quintuple by 2021 to reach 1.3 zettabytes globally by 2021.<sup>6</sup> (One zettabyte is equivalent to one trillion gigabytes.) Along with the volume of data available, the potential value of analyzing this data grows bigger by the day.

It's not surprising, then, that companies on the leading edge of data and analytics are more likely to make use of external data. An *MIT Sloan*

*Management Review* report published last year found that the companies making the most innovative use of data and analytics were more likely than others to leverage more external data sources, including social, mobile, and publicly available data.<sup>7</sup> A different study found that faster-growing companies were more likely to be planning to expand their ability to source external data than companies with lower growth rates.<sup>8</sup>

## How to get value from external data

External data sources are helping businesses personalize marketing offers, improve HR decisions, gain new revenue streams by launching new products or services, enhance risk visibility and mitigation, and better anticipate shifts in demand for their products and services. For instance, a major semiconductor manufacturer used third-party data to build models that could predict the best types of customers to target in marketing campaigns. This

external data helped train the models to identify potential targets that fit similar profiles to the company's most engaged customers. These "lookalike" models helped the organization optimize marketing spend and reduced a major campaign's cost-per-engagement by 75 percent.<sup>9</sup>

There are numerous other examples of analytics programs generating value with external data. Several startups monitor social networking data to try to predict patterns of external job-seeking behavior and retention risk; they claim their data is more predictive of an employee's likelihood of leaving than any internal data available. Others use geolocation and weather data to predict crop yields, helping farmers optimize their use of fertilizer. Retailers are using economic data and forecasts, data from suppliers, and geolocation data to better predict demand and reduce stockouts. Some companies are using satellite imagery to estimate mall traffic and predict retail sales; others are using aerial imagery to estimate oil inventories to better underwrite loans to refiners. (For select examples, see figure 1.)

FIGURE 1  
**How companies have used external data**

External data uses	Examples of relevant external data
A credit union personalized marketing offers for members based on customer profiles developed with the aid of external data <sup>10</sup>	<ul style="list-style-type: none"> <li>• Traditional consumer demographic data—age, gender, address, etc.</li> <li>• Social media data</li> <li>• Geolocation data</li> </ul>
A digital media company used external data to better predict and improve employee retention rates <sup>11</sup>	<ul style="list-style-type: none"> <li>• Postings from job websites</li> <li>• Government economic and labor sources</li> <li>• Social media data</li> </ul>
A logistics company uses external data to predict disruptions to clients' supply chains <sup>12</sup>	<ul style="list-style-type: none"> <li>• Social media data</li> <li>• Online news postings</li> <li>• Data from suppliers</li> </ul>
An agricultural giant launched a new service to help farmers predict and optimize crop yields <sup>13</sup>	<ul style="list-style-type: none"> <li>• Geolocation data</li> <li>• Weather data</li> <li>• Internet of Things (IoT) data</li> </ul>
A grocer uses external data to improve demand forecasting and reduce stockouts <sup>14</sup>	<ul style="list-style-type: none"> <li>• Weather data</li> <li>• Data from suppliers</li> <li>• Economic data and forecasts</li> </ul>

Source: Deloitte analysis.

## The challenges of using external data

Access to external data is getting easier in some ways, but it can still be daunting. Organizations report a wide variety of business and technical challenges in deriving insights from external data.<sup>15</sup> (Figure 2 summarizes some of these challenges.) Among the business challenges are the size and complexity of the data-provider market, which can make it hard to identify the right data sources and partners. Negotiating acquisition of data can be arduous, depending on factors such as whether ongoing access to data is needed for refreshing machine learning models, usage restrictions, whether the vendor wants a share of revenue gained from the data, and liability if the data proves to be inaccurate or tainted. This process can involve lengthy risk and legal reviews of vendor contracts and licensing agreements. The ongoing management of a growing roster of data-sharing relationships and partnerships can be taxing as well.

The technical challenges include fundamentals such as assessing data quality and accuracy: A variety of studies have demonstrated that third-party data can be riddled with inaccuracies.<sup>16</sup> There can also be inconsistencies between external and internal data to resolve before performing an analysis. Data preprocessing such as cleansing and formatting it for analysis is time-consuming. Some estimates suggest that this can account for 80 percent of the effort in data analysis projects. And securely storing

and cataloging data in an easily accessible manner can require updating information management processes and capabilities designed to handle only internal data. The longer it takes to work through these challenges, the less time available to react to market trends and external events with agility.

### TAP INTO A DATA ECOSYSTEM

Research suggests that most companies haven't yet developed the capabilities necessary to use external data effectively. To close this gap, companies may find it helpful to think of themselves as participants in a *data ecosystem*, which some have defined as a network of actors that directly or indirectly consume, produce, or provide data and other related resources.<sup>17</sup>

To be good at using external data means being competent in identifying, evaluating, procuring, and preparing external data in a consistent and timely manner. Companies will need a continuous process for identifying, engaging with, and evaluating new external data sources and partners and, when appropriate, integrating these data sources into analytics processes or product offerings. Maximizing the value of external data often requires integrating it with internal data for a more insightful analysis.

Companies may find it valuable to form a cross-functional group that acts as the organization's interface to the broader data ecosystem. This group could draw on competencies from multiple areas—including product management, business analysis, data science, legal, and procurement—to address

FIGURE 2  
**Challenges with using external data**

Business challenges	Technical challenges
Complex data-provider market	Data quality
Negotiating purchase and liability terms	Data preparation
Managing data-provider relationships	Storing, securing, and cataloging data

Source: Deloitte analysis.

both the organizational and technical challenges mentioned above. Some organizations have created special roles charged with scanning the third-party data market and matching business requests with relevant sources, a role Gartner has described as “data curator.”<sup>18</sup> Curators can help companies quickly identify and assess data sources matched to business needs, while reviewing external data sets for quality and accuracy using consistent evaluation processes. Companies that have an effective tech-scouting capability may look at that group for inspiration.

## A variety of ways to connect to the ecosystem

Organizations looking to connect to a data ecosystem can turn to a wide and growing variety of data and insights providers. Gartner Group categorizes data services, for instance, by the level of insight they provide:<sup>19</sup>

**Simple data services.** Data brokers collect data from multiple sources and offer it in collected and conditioned form. The data is used as additional input to a decision process by a person, an application system, or a device in an IoT ecosystem.

**Smart data services.** Data is enhanced by applying analytical rules and calculations. The results often take the form of scores or the tagging

of objects, as in services from marketing data providers and credit ratings agencies.

**Adaptive data services.** Customers submit data pertaining to specific analytical requests. Providers combine that data with data from other sources.

There are other ways to segment this dynamic marketplace as well. For instance, some providers specialize in serving industry sectors such as hedge funds or health care providers. In addition, consulting and systems integration services providers are reacting to client demand for new insights from external data by leveraging publicly available data or information from third-party data partners. These providers then integrate those sources with clients’ internal data and perform custom analysis.

## THE NEW DATA FRONTIER

Companies are making growing use of data from third parties. To get more value from their data analytics efforts, companies should consider enhancing their ability to identify, evaluate, and contract for new external data through a data ecosystem management program under a chief data office that links to business, IT, and legal teams. The pressure on companies to innovate and to improve the efficiency and effectiveness of their operations is unrelenting; they cannot afford to relent in their pursuit of insights to help them do so. For many companies, effective use of external data is a critical new frontier.

## Endnotes

1. Gartner, "Gartner survey shows organizations are slow to advance in data and analytics," press release, February 5, 2018.
2. See, for instance: DHL, "DHL supply watch: Machine learning to mitigate supplier risks," press release, May 22, 2017; Teradata, "Using advanced analytics cuts improper payments, insurance fraud," December 11, 2018; Genpact, "Sustaining a pharma company's good name with third-party risk management," 2018; Tonya Garcia, "The weatherman really can make it rain for these retailers," MarketWatch, October 23, 2015; Steve Banker, "Using weather data to improve supply chain resiliency," *Forbes*, June 29, 2016.
3. Forrester, "Digital is driving the next generation of data marketplaces," December 2017.
4. Jennifer Belissent, "The age of alt: Data commercialization brings alternative data to market," Forrester, June 22, 2017.
5. For merger and acquisition examples, see: Adam Satariano and Alan Bjerga, "The weather-predicting tech behind \$62 billion Monsanto bid," Bloomberg, June 9, 2016; Jonathan Varnan, "IBM bought the Weather Company because weather affects nearly everything," *Fortune*, October 28, 2015; Stephen Nellis, "Microsoft's Nadella banks on LinkedIn data to challenge Salesforce," Reuters, April 24, 2017; MarketWatch, "Moody's to acquire Reis, a leader in commercial real estate data," August 30, 2018; Patricio Robles, "How healthcare companies are getting creative to acquire data," Xeim, March 13, 2018. For partnership examples, see: Pfizer, "Pfizer Partnerships"; Mary Tyler March, "Komatsu, Trimble partner for 3-D construction data sharing," Construction Dive, July 26, 2017; Ed Cara, "Genetic data share: 23andMe partners with drugmaker GlaxoSmithKline," Genetic Literacy Project, August 6, 2018.
6. Cisco, "Cisco Global Cloud Index: Forecast and methodology, 2016–2021 white paper," November 19, 2018.
7. Sam Ransbotham and David Kiron, "Using analytics to improve customer engagement," *MIT Sloan Management Review*, January 30, 2018.
8. Jennifer Belissent, "Deep dive into data commercialization at Forrester's Data Strategy & Insights 2018 Forum," Forrester, November 28, 2018.
9. Oracle Data Cloud, "Understanding the value of combining first and third-party data," September 2017.
10. Ransbotham and Kiron, "Using analytics to improve customer engagement."
11. Josh Bersin and Michael Gretczko, *People data everywhere*, Deloitte University Press, February 27, 2015.
12. DHL, "DHL supply watch."
13. Monsanto, "The Climate Corporation transforms data into value with new Climate Fieldview integrated digital ag platform," September 2, 2015.
14. James Ovenden, "Are companies using their external data?," Innovation Enterprise, February 24, 2018.
15. Pitney Bowes, "Survey reveals global leaders struggle to access and effectively use data," press release, December 12, 2017.
16. Ross Benes, "Why is third-party data still often wrong?," Digiday, January 5, 2017.

17. For a review of literature about data ecosystems and an effort to define the term, see Marcelo lury S. Oliveira and Bernadette Farias Lóscio, "What is a data ecosystem?," dg.o, May 30–June 1, 2018.
18. Mike Rollins and Andrew White, "Build a data-driven enterprise," Gartner, August 8, 2018.
19. Douglas Laney, Alexander Linden, and Mario Faria, "Understand the data brokerage market before choosing a provider," Gartner, November 27, 2018.

## About the authors

**DAVID SCHATSKY** is a managing director at Deloitte LLP. He tracks and analyzes emerging technology and business trends, including the growing impact of cognitive technologies, for the firm's leaders and its clients. Schatsky is based in New York. Connect with him on Twitter [@dschatsky](https://twitter.com/dschatsky).

**JONATHAN CAMHI** is a senior consultant at Deloitte LLP. He analyzes trends related to emerging technologies such as artificial intelligence, blockchain, and the Internet of Things. Camhi has several years of experience as a journalist and researcher writing about emerging technologies and their impact on businesses. He is based in New York. Connect with him on LinkedIn at [www.linkedin.com/in/jonathan-camhi-48347531](https://www.linkedin.com/in/jonathan-camhi-48347531).

**CRAIG MURASKIN** is managing director of the Deloitte US Innovation Group. He works with firm leadership to set the group's agenda and overall innovation strategy. Muraskin is based in New York. Connect with him on LinkedIn at [www.linkedin.com/in/craig-muraskin-b7b2a629](https://www.linkedin.com/in/craig-muraskin-b7b2a629).

## Acknowledgments

The authors would like to thank **Aniket Dongre** of Deloitte Support Services India Pvt Ltd.

## Contact

**David Schatsky**

Managing director

US Innovation

Deloitte LLP

+1 646 582 5209


[dschatsky@deloitte.com](mailto:dschatsky@deloitte.com)



# Deloitte.

## Insights

Sign up for Deloitte Insights updates at [www.deloitte.com/insights](http://www.deloitte.com/insights).

 Follow @DeloitteInsight

### **Deloitte Insights contributors**

**Editorial:** Matthew Budman, Blythe Hurley, Abrar Khan, and Rupesh Bhat

**Creative:** Molly Woodworth

**Promotion:** Nikita Garia

**Cover artwork:** Molly Woodworth

### **About Deloitte Insights**

Deloitte Insights publishes original articles, reports and periodicals that provide insights for businesses, the public sector and NGOs. Our goal is to draw upon research and experience from throughout our professional services organization, and that of coauthors in academia and business, to advance the conversation on a broad spectrum of topics of interest to executives and government leaders.

Deloitte Insights is an imprint of Deloitte Development LLC.

### **About this publication**

This publication contains general information only, and none of Deloitte Touche Tohmatsu Limited, its member firms, or its and their affiliates are, by means of this publication, rendering accounting, business, financial, investment, legal, tax, or other professional advice or services. This publication is not a substitute for such professional advice or services, nor should it be used as a basis for any decision or action that may affect your finances or your business. Before making any decision or taking any action that may affect your finances or your business, you should consult a qualified professional adviser.

None of Deloitte Touche Tohmatsu Limited, its member firms, or its and their respective affiliates shall be responsible for any loss whatsoever sustained by any person who relies on this publication.

### **About Deloitte**

Deloitte refers to one or more of Deloitte Touche Tohmatsu Limited, a UK private company limited by guarantee ("DTTL"), its network of member firms, and their related entities. DTTL and each of its member firms are legally separate and independent entities. DTTL (also referred to as "Deloitte Global") does not provide services to clients. In the United States, Deloitte refers to one or more of the US member firms of DTTL, their related entities that operate using the "Deloitte" name in the United States and their respective affiliates. Certain services may not be available to attest clients under the rules and regulations of public accounting. Please see [www.deloitte.com/about](http://www.deloitte.com/about) to learn more about our global network of member firms.