From Digitally Disrupted to Digital Disrupter

Accenture Technology Vision 2014
Every Business Is a Digital Business

High performance. Delivered.

accenture

consulting | technology | outsourcing
TREND 3: DATA SUPPLY CHAIN
Yes, data technologies are evolving rapidly, but most have been adopted in piecemeal fashion. As a result, enterprise data is vastly underutilized. Data ecosystems are complex and littered with data silos, limiting the value that organizations can get out of their own data by making it difficult to access. To truly unlock that value, companies must start treating data more as a supply chain, enabling it to flow easily and usefully through the entire organization—and eventually throughout each company’s ecosystem of partners too.
Why now?

**Corporate data silos:** Data is the lifeblood of every digital organization, but businesses are struggling to access, share, and analyze much of the data they already have. Through 2015, 85 percent of Fortune 500 organizations will be unable to exploit big data for competitive advantage.

**Rising data volumes:** In addition to the data that organizations already collect, new external data sources are available, providing new opportunities for data insights. The digital universe is doubling every two years and is expected to grow to 40 trillion gigabytes (more than 5,200 gigabytes for every man, woman, and child in 2020).

**Maturing data technology:** The tools and technology required to build a data platform, ensuring data access and velocity, are available and in use. For example, a reported 20 percent of enterprises are already using NoSQL. With the foundation of these technologies, the integrated, end-to-end data supply chain is possible.

For years now, data has been talked about as the premier strategic IT asset. Today, it's so much more than that.

Business leaders now view data as among their most valuable assets too—some even call it the lifeblood of their organization. That's why they're implementing the newest big data tools, investing in advanced analytics applications, and purchasing the latest data visualization software.

Yet the reality is that these easily become one-off data fixes that contribute to data silos rather than provide an end-to-end data solution. Few companies have mastered the concepts at the foundation of modern data management—ideas such as the mobility and portability of data, its structure and velocity, data as a "saleable" product, and its valuation in open data exchanges.

Fewer still are comfortable with these concepts at scale. Most efforts to properly manage data are ad hoc at best. No wonder half of all companies are concerned about the accuracy of their data and a majority of executives are unclear about their analytics programs' business outcomes.
According to research by Gartner, 85 percent of *Fortune* 500 organizations will be unable to exploit big data for competitive advantage through 2015.\(^5\)

The high performers, however, will embark on a journey to ROI; they will liberate their data, generate value from it, and operationalize insights to drive strategic decisions through the organization. Key to their success will be managing their data like any core product—in the context of a supply chain. The supply chain begins when data is created, imported, or combined with other data. The data then moves, flows, and transforms through the supply chain, incrementally acquiring value. And while there may be diversions along the way, such as when a data “product” is removed for repairs (in other words, data cleansing), the supply chain ends with a valuable insight as its output. Guiding this movement is a data services platform. Analogous to the blueprint of a factory floor, this platform provides the structure for the intelligent transportation of data throughout the organization. It enables the effective supply chain—fit to strategy and designed to drive outcomes.

This means that business leaders now need to develop an end-to-end view of data in order to achieve their business goals. Accenture’s earlier Technology Vision reports have dived deeply into the keystone data topics, highlighting the value of data platforms, the importance of industrializing data services, and the need to think in terms of the velocity of data as well as its variety and volume. Now, all of those themes must come together in an end-to-end “supply chain” perspective that can help break down the data silos—usually built and “owned” by a single department—and enable data to flow freely for the benefit of the whole organization.

Working in this way, however, will change how the enterprise works with data at each stage. New sources of data can be leveraged. How data is managed and analyzed will evolve. And companies will begin to explore new ways to monetize their data. Understanding the implications at each of these stages will be critical for business leaders if they are to capitalize on the value of their data now and in the future and, importantly, at scale.
Enabling the data supply chain with a data services platform

The data supply chain must enable data movement. And in order for data to move, it must be made visible and accessible to those who need it when they need it. As such, the first step is to create a data services platform or federated data access layer, which provides a standard method of access to an organization’s curated and trusted (albeit varied and siloed) data in a time-relevant manner. Currently, only one out of five organizations integrates data across the enterprise.6 But those few are realizing great benefits. For instance, a major bank has been able to use Palantir’s Capital Markets solution to integrate 15 data sources into a single point of access. This has allowed hundreds of analysts and less-technical business users to do a variety of tasks, including performing regression analysis, developing hedges, and assessing investment risks.7

In order to realize their data platforms, organizations must first make their disparate data sources accessible through data services—ranging from Lightweight Directory Access Protocol (LDAP) to Web services or application programming interfaces (APIs). Regardless of the specific type, the power of these data services is in their ability to mask back-end complexity and expose data in defined ways. Once all of the data sources have been opened up, one common solution is to make them accessible through a virtualized data layer, which unifies everything into a single view. Then, users can interact with this abstracted data platform in a standardized way. (Behind the scenes, they are actually being rerouted to the data sources.)

Many newer data platform strategies now depend on opening up each data source separately, but through a common standard access protocol. Increasingly, companies are turning to APIs to achieve this, and they’re often aided by API management platforms. Walgreens, for example, has opened up its prescription API to enable third-party developers to incorporate the technology to scan barcodes on prescription bottles, in order to make it easier for customers to refill their prescriptions.8
These are just two of many data platform solutions in a market of vendors offering a wide variety of methods to create this federated access layer. Traditional middleware providers such as Tibco Software and Red Hat have evolved their messaging solutions to provide that abstraction layer. Companies such as Apigee are touting their API management software as the solution. Additionally, Platform-as-a-Service (PaaS) providers such as Amazon and Windows Azure are offering readymade solutions already built into their clouds. And large database vendors—such as Oracle and Microsoft—are creating access layers to allow connections to data silos in order to better manage the flow of information throughout an organization.

In the end, there’s no one-size-fits-all solution; most enterprises will end up with a hybrid set combining many of these tools. But no matter what the solution, it’s important to understand that data access and data acceleration make the data services platform both possible and necessary—and thus help to realize the data supply chain at scale.

Accelerating data through the supply chain

Data access on its own isn’t enough—velocity is needed. Importantly, this does not mean that all data needs to move at top speeds at all times but rather that data needs to be prioritized on the data services platform—such that important, time-critical data is accelerated through the supply chain, while stale, less relevant data moves more slowly but still meets the demands of the business. Quick access to valuable data means that analyses can be performed, insights can be gained, and actions can be taken in the sometimes very small window of opportunity available to businesses. Historically, IT professionals have addressed this problem by giving precedence to “hot” data—data that is accessed frequently and saved onto high-performance systems that can store and retrieve it very quickly. For its counterpart “cold” data—tax records, say—they have used slower disk hardware or even tape backups in legacy systems.
While these strategies have worked, newer prioritization practices improve data acceleration by adding many more gradations of “data temperature”—or data tiers. This makes sense when businesses consider the wide range of how, when, and how fast users need to consume data. The next step is to enable dynamic movement of data through these tiers, meaning they can be “heated up” or “cooled down” at any time. The ability to seamlessly change priority over time based on business need improves data velocity—but it’s also an efficient and cost-effective capability.

Facebook shows how. Not long ago, the social network discovered that 8 percent of all Facebook photos accounted for 82 percent of its network traffic. It turns out that there is a marked drop-off in accessing photos as they age, meaning that the photos’ data may be top priority at some point but not always. This led to Facebook’s development of its own three-tiered data storage solution. Its software categorizes photos and stores them in the appropriate tier, which has dedicated hardware to increase savings—the lower tiers (for colder data) can store more photos and use less energy. While the hot, tier-one data can be accessed almost instantaneously, one of the unique aspects of Facebook’s solution is that even cold data can be retrieved quickly (most solutions require several hours) to better meet users’ expectations. In this way, Facebook is both increasing its photos’ velocity and saving on storage costs.9

Tiered data solutions allow for time-critical and commonly accessed data to be stored in data-centric caching structures, optimized for quick transport through the supply chain. But in all likelihood, most data will eventually end up in the “data lake”—used to store the vast quantities of an organization’s data that are less time-sensitive or used less frequently. Although historically the “data lake” was tape, technologies such as Apache Hadoop and Amazon Redshift have been able to keep costs low while maintaining data accessibility. With a greater number of data tiers, more data can be stored longer and at a lower cost, without interfering with access to important, time-relevant data.
More uses for more data sources

The supply chain process starts with ingesting data. These days, companies can use a wide variety of new data sources—including, notably, data that they do not control or own. For example, if a regional grocery chain wants to analyze its daily transactions over the past month, it should look beyond the data in its database to supplement its findings. Many companies are already tracking sentiment on social media sites; data can also be analyzed in the context of weather, characteristics of shoppers, events in the news, or virtually any new data dimension imaginable—if the relevant data can be located. Whether this "external" data is to be obtained from partners, Data-as-a-Service providers, or open data sources (free for anyone to use in any way they want), companies should capitalize on the business value that these new sources provide.

Collaborate.org is one example of an open data source, containing 5 petabytes-plus of data. It is a global collaboration platform (meaning that users are encouraged to contribute data back to the platform) that allows users to view data geospatially—such as satellite imagery and air quality. The U.S. state of Hawaii is using Collaborate.org to share data across organizations, and it is the platform for the Exemplary State Initiative, which monitors environmental efforts and enables early identification of natural disasters.

In similar ways, Beiersdorf, a global provider of skin care products, is using their own internal data along with syndicated data from research companies such as Nielsen to provide board members with market share development information across a variety of products, brands, and countries. Using SAP Demand Signal Management powered by HANA, Beiersdorf plans to automate its data integration process, leading to more accurate and accelerated market share insights. These examples demonstrate the new opportunity that businesses have to look beyond their four walls for data that will grant more informed data insights and, ultimately, more value.
Advancing data discovery

The process of discovering new insights to answer business questions is changing fundamentally as users get faster access to more data. Now, when data is manipulated as it moves through the supply chain, value can be added to and obtained from it in ways that were previously impossible. This is because data discovery allows businesses to discover answers to questions that they might never have known to ask in the first place. Previously, traditional business intelligence (BI) methods were the only way to answer prescribed business questions; they require multiple lengthy steps before a solution is possible. Now, however, data discovery helps discern the very questions that companies should be asking by uncovering insights in a visually interactive and rapidly iterative manner. Effectively, data discovery empowers users to “communicate” with data at close to the speed of thought—accelerating businesses' time to insight. Companies can and should be investing in this practice today.

So that businesses can better “communicate” with and analyze data, analytics are being embedded in data discovery tools (as they are in applications)—effectively enabling data scientists and less-technical business users alike to do data discovery more easily and intuitively. As an example, the Teradata Aster Discovery Platform enables data scientists to do data discovery and advanced analytics themselves, but it also allows them to extend those capabilities to business users by leveraging analytic functions exposed through BI or other tools. When business users, and even business leaders, are able to discover and answer their own business questions in a matter of minutes, intelligent insights quickly lead to intelligent actions.

For instance, Texas Medical Center (TMC) understands the importance of health care data to advance patient care. However, this data can also be extremely complex and difficult for anyone—from medical experts to data scientists—to analyze. Therefore, TMC is using Ayasdi’s technology—one that specializes in the analysis of high-volume, high-dimensional datasets—to empower users to find insights in their data. Using a visual representation of the data and built-in statistical tools, users can more
easily perform "semi-supervised analysis" of the data. The ability to interact with data in this quick and iterative way allows users to identify emerging patterns and accelerate their time to insight. This approach first proved its power when analyzing a breast cancer dataset (one that had been analyzed many times already); within just a few minutes, a new subset of survivors was identified.13

TMC plans to use Ayasdi for a variety of applications, from analyzing clinical and genomic data to drug repurposing.14 The truth is, every organization has unexpected insights waiting to be unearthed by data discovery methods. Now, it’s a matter of determining those insights and accelerating their time to insight.

The next step: cognitive computing

As the volume and variety of data grow, so too do the scale and complexity of the data supply chain, making it increasingly difficult to add to and get value from data as it is manipulated. Imagine it this way, on a standard supply chain: everyday, more and more raw goods (some of them new) are being delivered. Initially, this might sound great—more supplies mean more products—but it’s not what the machines were designed to handle, and workers can’t keep up with the maintenance required. The supply chain becomes clogged, hindering the creation of valuable products. But what if, instead, machines could be taught to leverage data, learn from it, and, with a little guidance, figure out what to do with it? That’s the power of machine learning—which is a major building block of the ultimate long-term solution: cognitive computing. Rather than being programmed for specific tasks, machine learning systems gain knowledge from data as "experience" and then generalize what they’ve learned in upcoming situations. Cognitive computing technology builds on that by incorporating components of artificial intelligence to convey insights in seamless, natural ways to help humans or machines accomplish what they could not on their own. At its most advanced, cognitive computing will be the truly intelligent data supply chain—one that masks complexity by harnessing the power of data to help business users ask and answer strategic questions in a data-driven way.
Although complex, large-scale cognitive computing may be beyond the reach of most companies, there are some cognitive computing capabilities that can be put to work in practical and affordable ways. Companies should focus on tackling well-defined problems on a smaller scale—where machine learning techniques can be leveraged to accomplish practical cognitive computing goals. For example, Tempo, the calendar application found on iPhones, uses data contained on the phone—from social media to email, location, and more—to "learn" about events and display relevant information to the user when requested. This smart personal assistant application masks the complexity behind the data supply chain as raw data is aggregated, analyzed, and turned into an actual event with value. Only then does the user naturally interact with the calendar through the phone's interface.

One interesting cognitive computing example comes from U.S. food company McCormick. Machines are now starting to use data to "sense" the world as humans do, and this extends to taste—with obvious benefits for the food industry. Using Enterra Solution's Cognitive Reasoning Platform, McCormick's FlavorPrint site asks customers to rate a variety of flavors in order to learn taste and, from that, creates unique taste preference profiles—or what it calls FlavorPrints. If customers provide additional information, such as cooking preferences, equipment, and typical pantry items, they can receive better personal product and recipe recommendations. As far as these customers can tell, they're providing just a few raw facts in return for a great deal of personalized value about taste—something almost everyone feels strongly about yet finds hard to quantify or specify. From McCormick's point of view, learning customers' taste preferences leads to better insights, product decisions, and, ultimately, ability to serve its customers.15

Cognitive computing can, and will, bring benefits to many industries, and it will fundamentally change the ways in which many businesses operate. It flips the problem of data volume and variety on its head and instead leverages it to enable the smart, interactive data supply chain. The ultimate goal is for any business user—from a CEO to a field worker—to be able to ask any business question and immediately get a data-driven answer from the masked data supply chain.
Although this technology may seem far off, there are already cases that prove its relevance. And by its very definition, with more data over time, cognitive computing technology will only learn more, adapt quicker, and improve. It’s important for business leaders to familiarize themselves with this technology now.

Realizing data value

By the final stages of the supply chain, a significant investment has been made in the ingestion, transformation, and analysis of data—and now that data is both accessible and sharable, companies have new opportunities to capitalize on its value. Of course, data is important internally, but companies must now realize that the value of data extends outside the organization as well. From forging new partnerships to creating new revenue streams, or even entering new markets, businesses now have more potential than ever to realize the true value latent in their data.

When companies open up their data for external access, they can profit by leveraging external developer talent to extend their platforms. It’s a win for developers, too; they can use these platforms to add value to their products much more easily and at a higher quality than if they had attempted to do so on their own. For example, one of the reasons Google Maps is so widely known and successful is that more than 800,000 websites use its data, accessed through its API.16

Now, companies can take advantage of the opportunities for data monetization—to sell data insights directly, share them through partnerships, or develop entire ecosystems around them (see the chapter “From Workforce to Crowdsource”). There are risks and rewards for each, of course, and some methods will be better suited for certain companies and industries than others, but they are all important strategies to consider. Kabbage, for instance, was able to find a unique use case for data from UPS. As a partner, UPS provides transactional shipping data to inform Kabbage’s health assessment of the companies it finances—determining those companies’ access to capital.17
Just as companies are finding new reasons for ingesting external data sources, there have to be companies out there to provide them. It’s not just about selling data; it’s also about strengthening partnerships and developing ecosystems around data to monetize it. Companies now have the opportunity to think outside of the box for new ways to realize and take advantage of the true value in their data.

**Adding one data supply chain, and then another**

Every day, as the amount of data grows, so too does the seemingly impossible task of realizing the data services platform and, ultimately, the data supply chain. Unfortunately, there are no shortcuts. The path is incredibly difficult and exceptionally long—in fact, it never quite ends. But it is also one of the most rewarding journeys that companies can make in their transformation to become truly data driven.

Progress becomes possible when the transformation process is viewed as a matter of small steps rather than one giant leap. So, we expect that leading organizations will start by establishing a data services platform, followed by implementing a single data supply chain for a specific outcome. Once that’s done, they will incorporate another—and another. It may not be easy, but it’s doable.

The implications of enabling the data supply chain are huge. In the first stage, organizations now have the opportunity to ingest new sources of data. In the second, the manipulation of data through new methods of data discovery adds significant value. The future of this phase is to mask its complexity and truly embed this value through cognitive computing technologies. And finally, businesses can now look externally to realize value from data in new ways.

It’s time for business leaders to start thinking about the entirety of the data supply chain—as an end-to-end process that is outcome driven and fit to strategy. Today, data should be more than just a premier IT asset; it should be a premier business asset. Let’s start treating it that way.
Your 100-day plan

In 100 days, begin to develop a comprehensive strategy around laying the foundation for your data supply chain.

- Start to build an inventory of your data, beginning with your most frequently accessed and time-relevant data—which will be given first access to your data platform and accelerated on it.
- Identify any manual, time-consuming data curation processes (e.g., tagging, cleansing) for potential replacement with machine learning algorithms.
- Identify data silos within your organization (e.g., HR, finance, engineering), along with corresponding data needs that are currently unmet across the business.
- Begin to simplify/federate access to trusted data. Create a strategy for standardizing data access via the data platform. Solutions may be hybrid, utilizing a combination of traditional middleware and API management, or even a PaaS offering.
- Prioritize your individual data supply chains to develop a road map for implementing the data supply chain at scale.
- While building your platform, start looking outside your company for external data sources that can be incorporated to complement existing data and help lead to more complete insights.

This time next year

In 365 days, begin the journey to ROI by building a supply chain that is designed to drive outcomes.

- Pilot an initial data supply chain targeting a single business function. Once proven, each subsequent data supply chain will build upon this initial framework.
- Accelerate data on your data platform to enable the practice of data discovery across your organization. Uncover value in your data by quickly asking questions, failing fast, and iterating to solutions.
- Empower your business users to be value creators in the data supply chain by giving them data discovery tools and training them to deliver insights.
- Reduce the manual effort of curating data by using machine learning algorithms to automate the repetitive processes in your supply chain.
- Find a specific use case that can be addressed by cognitive-computing techniques. Use this focused issue as a chance to experiment with and learn the ways in which cognitive computing can be applied to your organization.
- Investigate opportunities to monetize your data. These should include developing new partnerships or ecosystems around data insights, in addition to directly selling them.
Data supply chain

1 "Big Data Business Benefits Are Hampered by 'Culture Clash',' Gartner, September 12, 2013.
5 "Big Data Business Benefits Are Hampered by 'Culture Clash',' Gartner, September 12, 2013.
10 Collaborate.org.
CONTACTS

For more information

Paul Daugherty
Chief Technology Officer
paul.r.daugherty@accenture.com

Michael J. Biltz
Director, Accenture Technology Vision
michael.j.biltz@accenture.com

Prith Banerjee
Managing Director, Accenture Technology R&D
prithviraj.banerjee@accenture.com

www.accenture.com/technologyvision
About the Accenture Technology Labs

The Technology Vision is published each year by the Accenture Technology Labs, the research and development (R&D) organization within Accenture. For more than 20 years, the Technology Labs have helped Accenture and its clients convert technology innovation into business results. Our R&D team explores new and emerging technologies to create a vision of how technology will shape the future and shape the next wave of cutting edge business solutions.

The Accenture Technology Labs offers seminars on the Technology Vision, which provide a forum to discuss the trends in greater depth and explore the implications for your organization’s business.

To learn more about the Accenture Technology Labs, or our seminars, contact a member of the Technology Vision team.

About Accenture

Accenture is a global management consulting, technology services and outsourcing company, with approximately 281,000 people serving clients in more than 120 countries. Combining unparalleled experience, comprehensive capabilities across all industries and business functions, and extensive research on the world’s most successful companies, Accenture collaborates with clients to help them become high-performance businesses and governments. The company generated net revenues of US$28.6 billion for the fiscal year ended Aug. 31, 2013. Its home page is www.accenture.com.