

How Analytics Has Changed in the Last 10 Years (and How It's Stayed the Same)

by Thomas H. Davenport

JUNE 22, 2017



Photo by Ferdinand Stöhr

Ten years ago, Jeanne Harris and I published the book *Competing on Analytics*, and we've just finished updating it for publication in September. One major reason for the update is that analytical technology has changed dramatically over the last decade; the sections we wrote on those topics

have become woefully out of date. So revising our book offered us a chance to take stock of 10 years of change in analytics.

Of course, not everything is different. Some technologies from a decade ago are still in broad use, and I'll describe them here too. There has been even more stability in analytical leadership, change management, and culture, and in many cases those remain the toughest problems to address. But we're here to talk about technology. Here's a brief summary of what's changed in the past decade.

The last decade, of course, was the era of big data. New data sources such as online clickstreams required a variety of new hardware offerings on premise and in the cloud, primarily involving distributed computing – spreading analytical calculations across multiple commodity servers – or specialized data appliances. Such machines often analyze data “in memory,” which can dramatically accelerate times-to-answer. Cloud-based analytics made it possible for organizations to acquire massive amounts of computing power for short periods at low cost. Even small businesses could get in on the act, and big companies began using these tools not just for big data but also for traditional small, structured data.

INSIGHT CENTER

Putting Data to Work

SPONSORED BY ACCENTURE

Analytics are critical to companies' performance.

Along with the hardware advances, the need to store and process big data in new ways led to a whole constellation of open source software, such as Hadoop and scripting languages. Hadoop is used to store and do basic processing on big data, and it's typically more than an order of magnitude cheaper than a data warehouse for similar

volumes of data. Today many organizations are employing Hadoop-based data lakes to store different types of data in their original formats until they need to be structured and analyzed.

Since much of big data is relatively unstructured, data scientists created ways to make it structured and ready for statistical analysis, with new (and old) scripting languages like Pig, Hive, and Python. More-specialized open source tools, such as Spark for streaming data and R for statistics, have also gained substantial popularity. The process of acquiring and using open source software is a major change in itself for established businesses.

The technologies I've mentioned for analytics thus far are primarily separate from other types of systems, but many organizations today want and need to integrate analytics with their production applications. They might draw from CRM systems to evaluate the lifetime value of a customer, for example, or optimize pricing based on supply chain systems about available inventory. In order to integrate with these systems, a component-based or "microservices" approach to analytical technology can be very helpful. This involves small bits of code or an API call being embedded into a system to deliver a small, contained analytical result; open source software has abetted this trend.

RELATED VIDEO



▶ PLAY 2:44

The Explainer: Big Data and Analytics

What the two terms really mean -- and how to effectively use each.

SAVE SHARE SEE MORE VIDEOS >

This embedded approach is now used to facilitate "analytics at the edge" or "streaming analytics." Small analytical programs running on a local microprocessor, for example, might be able to analyze data coming from drill bit sensors in an oil well drill and tell the bit whether to speed up or slow down. With internet of things data becoming popular in many industries, analyzing data near the source will become increasingly important, particularly in remote geographies where telecommunications constraints might limit centralization of data.

Another key change in the analytics technology landscape involves autonomous analytics – a form of artificial intelligence or cognitive technology. Analytics in the past were created for human decision makers, who considered the output and made the final decision. But machine learning technologies can take the next step and actually make the decision or adopt the recommended action. Most cognitive technologies are statistics-based at their core, and they can dramatically improve the productivity and effectiveness of data analysis.

Of course, as is often the case with information technology, the previous analytical technologies haven't gone away – after all, mainframes are still humming away in many companies. Firms still use statistics packages, spreadsheets, data warehouses and marts, visual analytics, and business

intelligence tools. Most large organizations are beginning to explore open source software, but they still use substantial numbers of proprietary analytics tools as well.

It's often the case, for example, that it's easier to acquire specialized analytics solutions – say, for anti-money laundering analysis in a bank – than to build your own with open source. In data storage there are similar open/proprietary combinations. Structured data in rows and columns requiring security and access controls can remain in data warehouses, while unstructured/prestructured data resides in a data lake. Of course, the open source software is free, but the people who can work with open source tools may be more expensive than those who are capable with proprietary technologies.

The change in analytics technologies has been rapid and broad. There's no doubt that the current array of analytical technologies is more powerful and less expensive than the previous generation. It enables companies to store and analyze both far more data and many different types of it. Analyses and recommendations come much faster, approaching real time in many cases. In short, all analytical boats have risen.

However, these new tools are also more complex and in many cases require higher levels of expertise to work with. As analytics has grown in importance over the last decade, the commitments that organizations must make to excel with it have also grown. Because so many companies have realized that analytics are critical to their business success, new technologies haven't necessarily made it easier to become – and remain – an analytical competitor. Using state-of-the-art analytical technologies is a prerequisite for success, but their widespread availability puts an increasing premium on nontechnical factors like analytical leadership, culture, and strategy.



Thomas H. Davenport is the President's Distinguished Professor in Management and Information Technology at Babson College, a research fellow at the MIT Initiative on the Digital Economy, and a senior adviser at Deloitte Analytics. Author of over a dozen management books, his latest is *Only Humans Need Apply: Winners and Losers in the Age of Smart Machines*.

This article is about ANALYTICS

 FOLLOW THIS TOPIC

Related Topics: BUSINESS HISTORY

 Loading...

 Loading...

Comments

Leave a Comment

POST

6 COMMENTS

subha shunmugavel 2 months ago

analytics is very important and the blog is informative...

REPLY



 **JOIN THE CONVERSATION**

POSTING GUIDELINES

We hope the conversations that take place on HBR.org will be energetic, constructive, and thought-provoking. To comment, readers must sign in or register. And to ensure the quality of the discussion, our moderating team will review all comments and may edit them for clarity, length, and relevance. Comments that are overly promotional, mean-spirited, or off-topic may be deleted per the moderators' judgment. All postings become the property of Harvard Business Publishing.