

# Big data:

## What's your plan?

Many companies don't have one. Here's how to get started.

**Stefan Biesdorf, David Court,  
and Paul Willmott**

### The problem

Technical and organizational challenges associated with “big data” and advanced analytics make it easy to pursue costly or ineffective solutions or to become paralyzed into inaction.

### Why it matters

Exploiting data is an increasingly important source of advantage. Companies whose early efforts struggle risk getting lapped by competitors.

### What to do about it

Craft a big-data plan. A successful one promotes strategic dialogue at the top of the company and helps to shape investment priorities and to establish trade-offs. Such plans rest on three elements:

- assembling and integrating extraordinary volumes of new data to mine fresh insights
- selecting advanced analytic models that optimize operations and predict the outcomes of business decisions
- creating intuitive tools that translate the output of models into tangible business actions and training key employees in the models' use

**The payoff** from joining the big-data and advanced-analytics management revolution is no longer in doubt. The tally of successful case studies continues to build, reinforcing broader research suggesting that when companies inject data and analytics deep into their operations, they can deliver productivity and profit gains that are 5 to 6 percent higher than those of the competition.<sup>1</sup> The promised land of new data-driven businesses, greater transparency into how operations actually work, better predictions, and faster testing is alluring indeed.

But that doesn't make it any easier to get from here to there. The required investment, measured both in money and management commitment, can be large. CIOs stress the need to remake data architectures and applications totally. Outside vendors hawk the power of black-box models to crunch through unstructured data in search of cause-and-effect relationships. Business managers scratch their heads—while insisting that they must know, upfront, the payoff from the spending and from the potentially disruptive organizational changes.

The answer, simply put, is to develop a plan. Literally. It may sound obvious, but in our experience, the missing step for most companies is spending the time required to create a simple plan for how data, analytics, frontline tools, and people come together to create business value. The power of a plan is that it provides a common language allowing senior executives, technology professionals, data scientists, and managers to discuss where the greatest returns will come from and, more important, to select the two or three places to get started.

There's a compelling parallel here with the management history around strategic planning. Forty years ago, only a few companies developed well-thought-out strategic plans. Some of those pioneers achieved impressive results, and before long a wide range of organizations had harnessed the new planning tools and frameworks emerging at that time. Today, hardly any company sets off without some kind of strategic plan. We believe that most executives will soon see developing a data-and-analytics plan as the essential first step on their journey to harnessing big data.

<sup>1</sup>See Dominic Barton and David Court, "Making advanced analytics work for you," *Harvard Business Review*, October 2012, Volume 90, Number 10, pp. 78–83.

The essence of a good strategic plan is that it highlights the critical decisions, or trade-offs, a company must make and defines the initiatives it must prioritize: for example, which businesses will get the most capital, whether to emphasize higher margins or faster growth, and which capabilities are needed to ensure strong performance. In these early days of big-data and analytics planning, companies should address analogous issues: choosing the internal and external data they will integrate; selecting, from a long list of potential analytic models and tools, the ones that will best support their business goals; and building the organizational capabilities needed to exploit this potential.

Successfully grappling with these planning trade-offs requires a cross-cutting strategic dialogue at the top of a company to establish investment priorities; to balance speed, cost, and acceptance; and to create the conditions for frontline engagement. A plan that addresses these critical issues is more likely to deliver tangible business results and can be a source of confidence for senior executives.

## **What's in a plan?**

Any successful plan will focus on three core elements.

### **Data**

A game plan for assembling and integrating data is essential. Companies are buried in information that's frequently siloed horizontally across business units or vertically by function. Critical data may reside in legacy IT systems that have taken hold in areas such as customer service, pricing, and supply chains. Complicating matters is a new twist: critical information often resides outside companies, in unstructured forms such as social-network conversations.

Making this information a useful and long-lived asset will often require a large investment in new data capabilities. Plans may highlight a need for the massive reorganization of data architectures over time: sifting through tangled repositories (separating transactions from analytical reports), creating unambiguous golden-source

data,<sup>2</sup> and implementing data-governance standards that systematically maintain accuracy. In the short term, a lighter solution may be possible for some companies: outsourcing the problem to data specialists who use cloud-based software to unify enough data to attack initial analytics opportunities.

## Analytic models

Integrating data alone does not generate value. Advanced analytic models are needed to enable data-driven optimization (for example, of employee schedules or shipping networks) or predictions (for instance, about flight delays or what customers will want or do given their buying histories or Web-site behavior). A plan must identify where models will create additional business value, who will need to use them, and how to avoid inconsistencies and unnecessary proliferation as models are scaled up across the enterprise.

As with fresh data sources, companies eventually will want to link these models together to solve broader optimization problems across functions and business units. Indeed, the plan may require analytics “factories” to assemble a range of models from the growing list of variables and then to implement systems that keep track of both. And even though models can be dazzlingly robust, it’s important to resist the temptation of analytic perfection: too many variables will create complexity while making the models harder to apply and maintain.

## Tools

The output of modeling may be strikingly rich, but it’s valuable only if managers and, in many cases, frontline employees understand and use it. Output that’s too complex can be overwhelming or even mistrusted. What’s needed are intuitive tools that integrate data into day-to-day processes and translate modeling outputs into tangible business actions: for instance, a clear interface for scheduling employees, fine-grained cross-selling suggestions for call-center agents, or a way for marketing managers to make real-time decisions on discounts. Many companies fail to complete this step in their thinking and planning—only to find that managers and operational employees do not use the new models, whose effectiveness predictably falls.

<sup>2</sup>The practice of storing a unit of information only once across an enterprise to ensure accuracy.

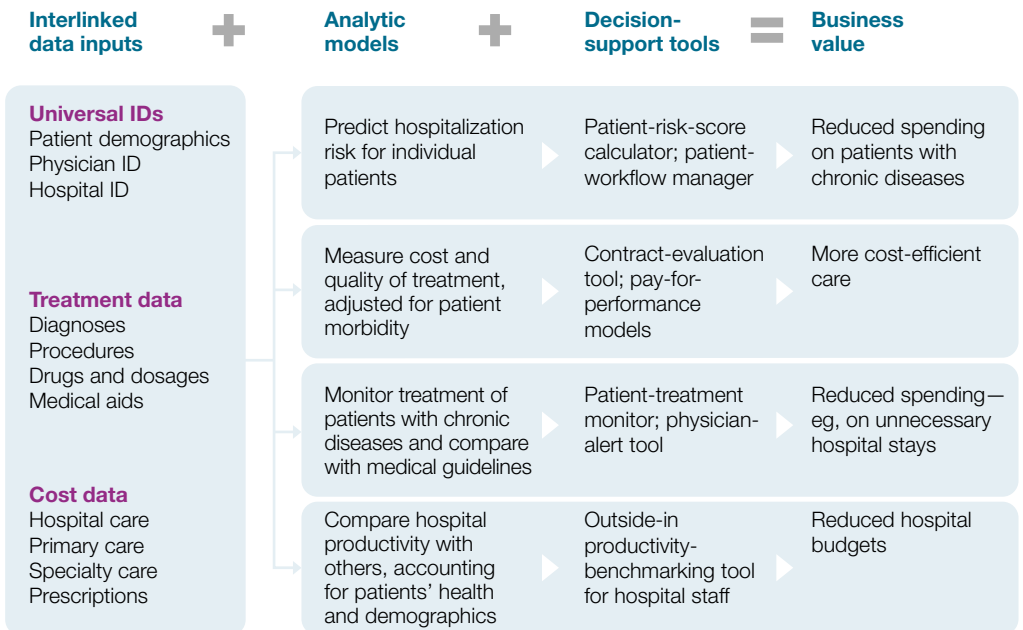
There’s also a critical enabler needed to animate the push toward data, models, and tools: organizational capabilities. Much as some strategic plans fail to deliver because organizations lack the skills to implement them, so too big-data plans can disappoint when organizations lack the right people and capabilities. Companies need a road map for assembling a talent pool of the right size and mix. And the best plans will go further, outlining how the organization can nurture data scientists, analytic modelers, and frontline staff who will thrive (and strive for better business outcomes) in the new data- and tool-rich environment.

By assembling these building blocks, companies can formulate an integrated big-data plan similar to what’s summarized in the exhibit. (For an expanded look at the exhibit, see this article on mckinsey.com.) Of course, the details of plans—analytic approaches, decision-support tools, and sources of business value—will vary by industry. However, it’s important to note an important structural similarity across industries: most companies will need to plan for major

Exhibit

**A successful big-data plan will focus on three core elements.**

Health-care industry, public-payer illustration



data-integration campaigns. The reason is that many of the highest-value models and tools (such as those shown on the right of the exhibit) increasingly will be built using an extraordinary range of data sources (such as all or most of those shown on the left). Typically, these sources will include internal data from customers (or patients), transactions, and operations, as well as external information from partners along the value chain and Web sites—plus, going forward, from sensors embedded in physical objects.

To build a model that optimizes treatment and hospitalization regimes, a company in the health-care industry might need to integrate a wide range of patient and demographic information, data on drug efficacy, input from medical devices, and cost data from hospitals. A transportation company might combine real-time pricing information, GPS and weather data, and measures of employee labor productivity to predict which shipping routes, vessels, and cargo mixes will yield the greatest returns.

## Three key planning challenges

Every plan will need to address some common challenges. In our experience, they require attention from the senior corporate leadership and are likely to sound familiar: establishing investment priorities, balancing speed and cost, and ensuring acceptance by the front line. All of these are part and parcel of many strategic plans, too. But there are important differences in plans for big data and advanced analytics.

### 1. Matching investment priorities with business strategy

As companies develop their big-data plans, a common dilemma is how to integrate their “stovepipes” of data across, say, transactions, operations, and customer interactions. Integrating all of this information can provide powerful insights, but the cost of a new data architecture and of developing the many possible models and tools can be immense—and that calls for choices. Planners at one low-cost, high-volume retailer opted for models using store-sales data to predict inventory and labor costs to keep prices low. By contrast, a high-end,

high-service retailer selected models requiring bigger investments and aggregated customer data to expand loyalty programs, nudge customers to higher-margin products, and tailor services to them.

That, in a microcosm, is the investment-prioritization challenge: both approaches sound smart and were, in fact, well-suited to the business needs of the companies in question. It's easy to imagine these alternatives catching the eye of other retailers. In a world of scarce resources, how to choose between these (or other) possibilities?

There's no substitute for serious engagement by the senior team in establishing such priorities. At one consumer-goods company, the CIO has created heat maps of potential sources of value creation across a range of investments throughout the company's full business system—in big data, modeling, training, and more. The map gives senior leaders a solid fact base that informs debate and supports smart trade-offs. The result of these discussions isn't a full plan but is certainly a promising start on one.

Or consider how a large bank formed a team consisting of the CIO, the CMO, and business-unit heads to solve a marketing problem. Bankers were dissatisfied with the results of direct-marketing campaigns—costs were running high, and the uptake of the new offerings was disappointing. The heart of the problem, the bankers discovered, was a siloed marketing approach. Individual business units were sending multiple offers across the bank's entire base of customers, regardless of their financial profile or preferences. Those more likely to need investment services were getting offers on a range of deposit products, and vice versa.

The senior team decided that solving the problem would require pooling data in a cross-enterprise warehouse with data on income levels, product histories, risk profiles, and more. This central database allows the bank to optimize its marketing campaigns by targeting individuals with products and services they are more likely to want, thus raising the hit rate and profitability of the campaigns. A robust planning process often is needed to highlight investment opportunities like these and to stimulate the top-management engagement they deserve given their magnitude.

## 2. Balancing speed, cost, and acceptance

A natural impulse for executives who “own” a company’s data and analytics strategy is to shift rapidly into action mode. Once some investment priorities are established, it’s not hard to find software and analytics vendors who have developed applications and algorithmic models to address them. These packages (covering pricing, inventory management, labor scheduling, and more) can be cost-effective and easier and faster to install than internally built, tailored models. But they often lack the qualities of a killer app—one that’s built on real business cases and can energize managers. Sector- and company-specific business factors are powerful enablers (or enemies) of successful data efforts. That’s why it’s crucial to give planning a second dimension, which seeks to balance the need for affordability and speed with business realities (including easy-to-miss risks and organizational sensitivities).

To understand the costs of omitting this step, consider the experience of one bank trying to improve the performance of its small-business underwriting. Hoping to move quickly, the analytics group built a model on the fly, without a planning process involving the key stakeholders who fully understood the business forces at play. This model tested well on paper but didn’t work well in practice, and the company ran up losses using it. The leadership decided to start over, enlisting business-unit heads to help with the second effort. A revamped model, built on a more complete data set and with an architecture reflecting differences among various customer segments, had better predictive abilities and ultimately reduced the losses. The lesson: big-data planning is at least as much a management challenge as a technical one, and there’s no shortcut in the hard work of getting business players and data scientists together to figure things out.

At a shipping company, the critical question was how to balance potential gains from new data and analytic models against business risks. Senior managers were comfortable with existing operations-oriented models, but there was pushback when data strategists proposed a range of new models related to customer behavior, pricing, and scheduling. A particular concern was whether costly new data approaches would interrupt well-oiled scheduling operations. Data managers met these concerns by pursuing a prototype (which used a smaller data set and rudimentary spreadsheet analysis) in one



region. Sometimes, “walk before you can run” tactics like these are necessary to achieve the right balance, and they can be an explicit part of the plan.

At a health insurer, a key challenge was assuaging concerns among internal stakeholders. A black-box model designed to identify chronic-disease patients with an above-average risk of hospitalization was highly accurate when tested on historical data. However, the company’s clinical directors questioned the ability of an opaque analytic model to select which patients should receive costly preventative-treatment regimes. In the end, the insurer opted for a simpler, more transparent data and analytic approach that improved on current practices but sacrificed some accuracy, with the likely result that a wider array of patients could qualify for treatment. Airing such tensions and trade-offs early in data planning can save time and avoid costly dead ends.

Finally, some planning efforts require balancing the desire to keep costs down (through uniformity) with the need for a mix of data and modeling approaches that reflect business realities. Consider retailing, where players have unique customer bases, ways of setting prices to optimize sales and margins, and daily sales patterns and inventory requirements. One retailer, for instance, has quickly and inexpensively put in place a standard next-product-to-buy model<sup>3</sup> for its Web site. But to develop a more sophisticated model to predict regional and seasonal buying patterns and optimize supply-chain operations, the retailer has had to gather unstructured consumer data from social media, to choose among internal-operations data, and to customize prediction algorithms by product and store concept. A balanced big-data plan embraces the need for such mixed approaches.

### 3. Ensuring a focus on frontline engagement and capabilities

Even after making a considerable investment in a new pricing tool, one airline found that the productivity of its revenue-management analysts was still below expectations. The problem? The tool was too complex to be useful. A different problem arose at a health insurer:

<sup>3</sup>A model, based on algorithms that analyze a customer’s purchase history, to predict the next product or service the customer is likely to buy. It then makes a specific recommendation.

doctors rejected a Web application designed to nudge them toward more cost-effective treatments. The doctors said they would use it only if it offered, for certain illnesses, treatment options they considered important for maintaining the trust of patients.

Problems like these arise when companies neglect a third element of big-data planning: engaging the organization. As we said when describing the basic elements of a big-data plan, the process starts with the creation of analytic models that frontline managers can understand. The models should be linked to easy-to-use decision-support tools—call them killer tools—and to processes that let managers apply their own experience and judgment to the outputs of models. While a few analytic approaches (such as basic sales forecasting) are automatic and require limited frontline engagement, the lion's share will fail without strong managerial support.

The aforementioned airline redesigned the software interface of its pricing tool to include only 10 to 15 rule-driven archetypes covering the competitive and capacity-utilization situations on major routes. Similarly, at a retailer, a red flag alerts merchandise buyers when a competitor's Internet site prices goods below the retailer's levels and allows the buyers to decide on a response. At another retailer, managers now have tablet displays predicting the number of store clerks needed each hour of the day given historical sales data, the weather outlook, and planned special promotions.

But planning for the creation of such worker-friendly tools is just the beginning. It's also important to focus on the new organizational skills needed for effective implementation. Far too many companies believe that 95 percent of their data and analytics investments should be in data and modeling. But unless they develop the skills and training of frontline managers, many of whom don't have strong analytics backgrounds, those investments won't deliver. A good rule of thumb for planning purposes is a 50–50 ratio of data and modeling to training.

Part of that investment may go toward installing “bimodal” managers who both understand the business well and have a sufficient knowledge of how to use data and tools to make better, more analytics-infused decisions. Where this skill set exists, managers will of course want to draw on it. Companies may also have to create incentives

that pull key business players with analytic strengths into data-leadership roles and then encourage the cross-pollination of ideas among departments. One parcel-freight company found pockets of analytical talent trapped in siloed units and united these employees in a centralized hub that contracts out its services across the organization.



When a plan is in place, execution becomes easier: integrating data, initiating pilot projects, and creating new tools and training efforts occur in the context of a clear vision for driving business value—a vision that’s unlikely to run into funding problems or organizational opposition. Over time, of course, the initial plan will get adjusted. Indeed, one key benefit of big data and analytics is that you can learn things about your business that you simply could not see before.

Here, too, there may be a parallel with strategic planning, which over time has morphed in many organizations from a formal, annual, “by the book” process into a more dynamic one that takes place continually and involves a broader set of constituents.<sup>4</sup> Data and analytics plans are also too important to be left on a shelf. But that’s tomorrow’s problem; right now, such plans aren’t even being created. The sooner executives change that, the more likely they are to make data a real source of competitive advantage for their organizations. ○

<sup>4</sup>See Sven Smit, Lowell Bryan, and Chris Bradley, “Managing the strategy journey,” [mckinseyquarterly.com](http://mckinseyquarterly.com), July 2012.

*The authors would like to acknowledge the contributions of Toos Daruvala, Amit Garg, and David Kang to the development of this article.*

**Stefan Biesdorf** is a principal in McKinsey’s Munich office, **David Court** is a director in the Dallas office, and **Paul Willmott** is a director in the London office.