

MCKINSEY GLOBAL INSTITUTE

Competing through data: Three experts offer their game plans

Erik Brynjolfsson

Professor of management science at the Massachusetts Institute of Technology's Sloan School of Management

Jeff Hammerbacher

Cofounder and chief scientist of Cloudera

Brad Stevens

Butler University men's basketball coach

As big data creates new opportunities

and threats, it also demands new mind-sets from senior executives about the role of information in business and even the nature of competitive advantage. The perspectives that follow may help shake up your thinking and forge that new frame of mind.

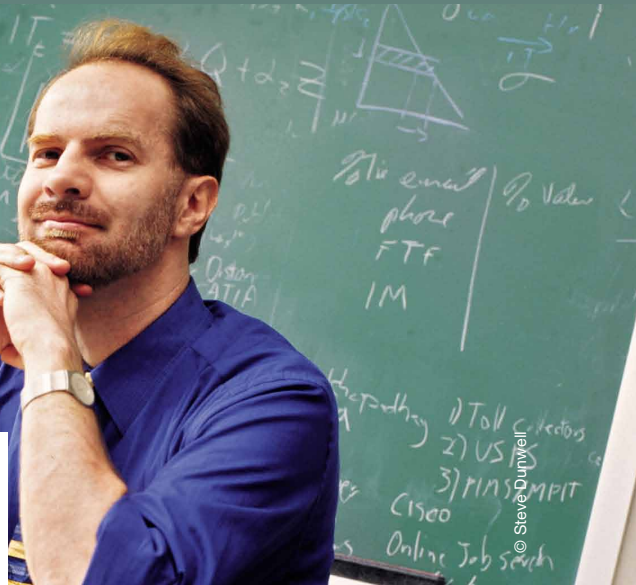
Massachusetts Institute of Technology (MIT) professor Erik Brynjolfsson explores the implications of intriguing new research about the relationship among data, analytics, productivity, and profitability. Jeff Hammerbacher, co-founder of the data-oriented start-up Cloudera, provides a view from the front lines about what it takes to harness the flood of data now at companies' collective fingertips. Finally, basketball coach Brad Stevens describes how, on a tight budget, he uses data that's powerful (even if not extraordinarily "big") to help his Butler University squad punch above its weight. Presented here are edited versions of interviews with each, conducted by McKinsey's Michael Chui and Frank Comes.

“Too many managers are not opening their eyes to this opportunity and understanding what big data can do to change the way they compete.”

The professor

Erik Brynjolfsson

Erik Brynjolfsson is the Schussel Family Professor of Management Science at the Massachusetts Institute of Technology's Sloan School of Management, director of the MIT Center for Digital Business, and one of the world's leading researchers on how IT affects productivity.



The data advantage

Most great revolutions in science are preceded by revolutions in measurement. We have had a revolution in measurement, over the past few years, that has allowed businesses to understand in much more detail what their customers are doing, what their processes are doing, what their employees are doing. That tremendous improvement in measurement is creating new opportunities to manage things differently.

Our research has found a shift from using intuition toward using data and analytics in making decisions. This change has been accompanied by measurable improvement in productivity and other performance measures. Specifically, a one-standard-deviation increase toward data and analytics was correlated with about a 5 to 6 percent improvement in productivity and a slightly larger increase in profitability in those same firms. The implication for companies is that by changing the way they make decisions, they're likely to be able to outperform competitors.

Becoming data driven

The prerequisite, of course, is the technological infrastructure: the ability to measure things in more detail than you could before. The harder thing is to get the set of skills. That includes not just some analytical skills but also a set of attitudes and an understanding of the business. Then the third thing, which is the subtlest but perhaps the most important, is cultural change about how to use data. A lot of companies think they're using data, and you often see bar charts and pie charts and numbers in management presentations. But, historically, that kind of data was used more to confirm and support decisions that had already been made, rather than to learn new things and to discover the right answer. The cultural change is for managers to be willing to say, "You know, that's an interesting problem, an interesting question. Let's set up an experiment to discover the answer."

“I think this revolution in measurement, starting with the switch from analog to digital data, is as profound as, say, the development of the microscope and what it did for biology and medicine.”

Too many managers are not opening their eyes to this opportunity and understanding what big data can do to change the way they compete. They have to be ready to show some vulnerability and say, "Look, we're open to the data" and not go in there saying, "Hey, I'm gonna manage from the gut. I have years of experience and I know the answers to this going in." I think, historically, a lot of managers have been implicitly or explicitly rewarded for that kind of confidence. You have to have a different kind of confidence to be willing to let the data speak.

One CEO told me that when he pushed this attitude, he had to change over 50 percent of his senior-management team because they just didn't get it. Obviously, that was a painful thing to have to do. But the results have been very successful. And they require that level of aggressiveness by top management, if it really wants to end up in that group of leaders as opposed to the laggards.

Required skills

Having enough data to get a statistically significant result is not a problem. There's plenty of data. So the skills often have more to do with sampling methodologies, designing experiments, and working these very, very large data sets without becoming overwhelmed. If you look inside companies, you also see a transformation in the functions that are using data. CIOs are discovering that, more and more, it's the marketing people and the people working with customers—customer relationship management—who have the biggest data needs. These are the people CIOs are working with most closely. This is part of a broader revolution as we move from just financial numerical data toward all sorts of nonfinancial metrics.

Often, the nonfinancial metrics give a quicker and more accurate measure of what's happening in the business. I was talking to Gary Loveman—the CEO of Caesar's Entertainment, formerly Harrah's, and a PhD graduate of MIT. He's used some of these techniques to revolutionize what's happening in that industry. But, interestingly, increasingly what he measures is customer satisfaction and a lot of other intermediate metrics. He said that customer satisfaction metrics were much quicker and more precise metrics of what was happening in response to some of the policy changes that he put in place.

Think of it this way. If customers end up satisfied or dissatisfied, that will affect the probability of their coming back next year. Now, next year's financial results will be affected as a result. And you could, in principle, try to match up the experience the customer had this year with future years' return rates. But a much quicker way of getting feedback on which processes are working is to look at customer satisfaction when you put process changes in place.

The new landscape

I think this revolution in measurement, starting with the switch from analog to digital data, is as profound as, say, the development of the microscope and what it did for biology and medicine. It's not just big data in the sense that we have lots of data. You can also think of it as

“nano” data, in the sense that we have very, very fine-grained data—an ability to measure things much more precisely than in the past. You can learn about the preferences of an individual customer and personalize your offerings for that particular customer.

One of the biggest revolutions has involved enterprise information systems, like ERP, enterprise resource planning; CRM, customer relationship management; or SCM, supply chain management—those large enterprise systems that companies have spent hundreds of millions of dollars on. You can use the data from them not just to manage operations but to gain business intelligence and learn how they could be managed differently. A common pattern that we’re seeing is that three to five years after installing one of these big enterprise systems, companies start saying, “Hey, we need some business intelligence tools to take advantage of all this data.” It’s up to managers now to seize that opportunity and take advantage of this very fine-grained data that just didn’t exist previously.

The path ahead

There’s some good news and there’s some not-so-good news. The good news is that technology’s not slowing down, and the pie is getting bigger. Productivity is accelerating. And that *should* make us all better off. However, it’s not making us all better off. Over the past 20 years or so, median wages in the United States have stagnated because a lot of people don’t have the skills to take full advantage of this technology. And, unfortunately, I don’t see that changing any time soon unless we have a much bigger effort to change the kinds of skills that are available in the workforce and have a set of technologies that people can tap into more readily.

This flood of data and analytical opportunities creates more value for people who can be creative in seeing patterns and for people who can be entrepreneurial in creating new business opportunities that take advantage of these patterns. My hope is that the technology will create a platform that people can tap into to create new entrepreneurial ventures—some of them, perhaps, huge hits like Facebook or Zynga or Google. But also, perhaps equally important for the economy, hundreds of thousands or millions of small entrepreneurial ventures, eBay based or app based, would mean millions of ordinary people can be creative in using technology and their entrepreneurial energies to create value. That would be an economy where not only does the pie get bigger but each part of the pie—each of the individuals—benefits as well.

“If you can understand consumer behavior and get your hands around as much behavioral data as possible to better guide product decision making, then every penny you can eke out is increasing your margins and allowing you to invest more.”

The data entrepreneur

Jeff Hammerbacher

Before cofounding Silicon Valley software start-up Cloudera in 2009, at the age of 26, Jeff Hammerbacher was a quantitative analyst on Wall Street and one of Facebook's first employees.



The open-source advantage

I was Facebook's first research scientist. The initial goal for that position was to understand how changes to the site were impacting user behavior. We had built our own infrastructure to allow us to do some terabyte analytics, but we were going to have to scale it to up to petabytes.¹ We realized that instead of continuing to invest in infra-

¹ Under the International System of Units, a terabyte equals one trillion bytes, or 1,000 gigabytes. A petabyte is equal to 1,000 terabytes.

structure, we could build a more powerful shared resource to facilitate business analysis by working with the open-source community.

In founding Cloudera, I saw a path to a complete infrastructure for doing analytical data management. It would be made up of existing open-source projects as well as open-source versions of a lot of the technologies that we had built out internally at Facebook. Cloudera would be a corporate entity for pursuing those goals and ensuring that it wasn't just Facebook that would be able to use this technology but, really, any enterprise.

Data leaders

When we started Cloudera, we didn't have a core thesis around where the technology would be adopted or what the market was going to look like. Early adopters were clearly in the Web and digital-media spaces. But in terms of traditional industries, the federal government surprised me. They really are the leaders in multimedia data analysis—working with text, images, video. In the intelligence agencies, I've seen more sophistication than in commercial domains.

I was also surprised to see the retail space. Retailers had very large volumes of data, and because many were branching out into e-commerce, they had a lot of Web logs and Web data as well. There is an arms race going on right now in retail. If you can understand consumer behavior and get your hands around as much behavioral data as possible to better guide product decision making, then every penny you can eke out is increasing your margins and allowing you to invest more.

Financial services was one sector that I had hoped would be an early adopter, but these companies tend not to look at their businesses as a whole in the same way that retail does. Data management is thought of as project specific, even to the point where individual trading desks could have their own chief technology officers. Our technology tends to work best as a shared infrastructure for multiple lines of business.

Where this is headed is learning how to point this new infrastructure for storing and analyzing data at real business problems, as well as growing the imagination of businesspeople about what they can do when a variety of experts analyze the data. If you can digitize reality, then you can move your world faster than before.

Building a big data function

You need to make a commitment to conceiving of data as a competitive advantage. The next step is to build out a low-cost, reliable infrastructure for data collection and storage for whichever line of business you perceive to be most critical to your company. If you don't have that digital asset, then you're not even going to be able to play the game. And then you can start layering on the complex analytics. Most companies go wrong when they start with the complex analytics.

When deciding how to incorporate analytics expertise into an organization, you have to be honest about what your organization looks like—your capacity to hire and your long-term vision for what that organization is going to be. There isn't one right answer. Yahoo! built a centralized group called Strategic Data Solutions to run the entire gamut. Rather than just building a small group of people primarily focused on marketing analytics, the company took an end-to-end view, extending from data storage to the actual P&L. In our group at Facebook, because we were a very fast-moving organization, we were much more of a platform—a service organization for the rest of the company.

The rise of the 'data scientist'

I tried to articulate this title of data scientist in a book I put together with O'Reilly Media.² I now actually see people describing themselves as data scientists in their job titles on LinkedIn and scientists talking about themselves as data scientists. So it's evolving. People realize that there is a gap between the current role of statistician or data analyst or business analyst and what they actually want. They are grappling with the set of tools and the set of skills that they need. Across the whole research cycle, it's a combination of skills that social scientists understand, plus additional programming skills, plus the ability to do aggressive prioritization. And, of course, a good grounding in statistics and machine learning.³ That collection of skills is difficult to find.

²Jeff Hammerbacher and Toby Segaran, eds., *Beautiful Data: The Stories Behind Elegant Data Solutions*, Sebastopol, CA: O'Reilly, 2009.

³Machine learning is a form of artificial intelligence in which algorithms allow computers to make decisions based on data streams.

“I can have all the data I want to have—
but I still have to communicate it
to our players. It has to get into their
minds. And they have to utilize it.”



The coach

Brad Stevens

Brad Stevens is head coach of the Butler University men's basketball team.

Coach Stevens holds the National Collegiate Athletic Association (NCAA) record for most games won in the first four years as a Division I head basketball coach. Among those wins was a series of thrilling NCAA tournament games that brought his Butler University team to the championship final in 2010 and 2011.

Before joining Butler, which is located in Indianapolis, Indiana, and has just 4,500 students, he was a marketing associate at the global pharmaceutical group Eli Lilly. In the following interview, Stevens explains how focusing on the numbers has helped improve his team's game.

The Quarterly: *How have things changed in basketball with regard to the use of data and analytics?*

Brad Stevens: You know, I'm a bad person to ask about that because I'm 34. The data's always been an important part of my job. I've always looked at it through that lens, even when I was a young assistant. This is how I work best. For me, it's incredibly interesting. There are complexities that you can really study using numbers. We don't have access to the highest end—we're not sitting here with NBA⁴ money to invest in a numbers-and-research department. But I think you can speak to your team with numbers and give your players pretty clear-cut and defined examples of what they need to do to get better.

The Quarterly: *If you had an infinite budget, what sorts of things would you do?*

Brad Stevens: The first thing is that I'd have one of the positions on our staff, or maybe a whole group on our staff, working on statistics. They would look at game planning and how players are most effective: what they're doing when they're most effective, where they are on the court—really show players the exact way that they are most effective in different areas of the game. That's an incredibly useful teaching tool.

The Quarterly: *In the absence of those resources, that staff, what do you do?*

Brad Stevens: I first break down all of the statistics that I can on opponents to try to get my mind wrapped around what their trends are. I'll look for how many three-point attempts per field goal attempt⁵—that tells you what kind of team they are right away. You can look at offensive-rebound percentages. Defensive- and offensive-turnover percentages. How teams shoot against them. What they defend well. What they try to defend well.

Then there's the ability to cut film on computers and to do so quickly. We can watch all of somebody's moves off of a ball screen. All of a person's moves going left. All of the post moves, going to the middle or going to the baseline. Whatever the case may be. And we can really

⁴The US National Basketball Association.

⁵For an explanation of basketball terminology, visit www.fiba.com/pages/eng/fc/baskBasi/glos.asp.

“As we get to the latter part of the season, I’ll spend a lot more time asking, ‘What’s happened in the past five games? What are they doing differently from a statistical standpoint? What have they improved on? What have they regressed in?’”

determine their effectiveness from that. We obviously hope that the film validates the statistics and we can figure out what’s unique about what players do.

One thing that you have to be careful of is not getting caught up in just season statistics. Teams change. And as we get to the latter part of the season, I’ll spend a lot more time asking, “What’s happened in the past five games? What are they doing differently from a statistical standpoint? What have they improved on? What have they regressed in?”

Of course, I can have all the data I want to have—but I still have to communicate it to our players. It has to get into their minds. And they have to utilize it. So you can’t inundate them. You can’t take three seconds to make a decision in basketball. It’s a game that moves too quickly for that. There’s no huddle in between plays; there’s not a moment in between every pitch. You’ve got to have thoughts in your mind about what the people that you’re playing against like to do, and what you do best, and at the same time you can’t be inundated with those thoughts or it’ll affect the way you play. That makes communicating data and simplifying it for the players incredibly important.

The Quarterly: *Can you say more about how you simplify data, how you engage your players?*

Brad Stevens: You’ve got to figure out how they react, how they best comprehend, how they best learn in a team setting, how they best learn in an individual setting, and go from there. Each team’s different, each player’s different. And, you know, it may mean bringing in a guy who has a mind for numbers and saying, “The bottom line is that, right now, you’re shooting 43 percent. You’re a better shooter than that. If

you make one more shot a game, you're probably at 48 or 49 percent. How can we make it so you're one more shot effective for a game?"

The Quarterly: *Was there one game or a couple of games where this really played out and made a difference?*

Brad Stevens: Every game we play in. There's not a game when this wouldn't have played a major role. We're not the most talented, so we have to be good in these little areas. Sometimes, you know, the numbers hurt you. You believe one thing, and then the other team has a night that's unique. But more times than not, the score takes care of itself, as Bill Walsh⁶ says. ○

⁶Bill Walsh coached the US National Football League's San Francisco 49ers to three Super Bowl titles (1982, 1985, and 1989). His book *The Score Takes Care of Itself: My Philosophy of Leadership* (Portfolio, August 2009), published two years after his death, was coauthored with his son, Craig, and with Steve Jamison.

Michael Chui is a senior fellow at the McKinsey Global Institute and is based in McKinsey's San Francisco office; **Frank Comes** is a member of McKinsey Publishing and is based in the New Jersey office.