How to use, exploit, and contextualize big data, and how to avoid its misuse, have become societal issues. These issues matter to everyone because big data will play a key role in overcoming the current economic inertia and achieving the objective of inclusive growth—the involvement of the broadest possible spectrum of people in wealth creation.

Many doubt the wisdom of direct government intervention and increased spending to create jobs. Huge sovereign debt makes this approach problematic in any case. Traditional policy levers to address structural unemployment—such as retraining, increased labor mobility, deregulation, and investment in research—are having a diminished impact, while fiscal consolidation and austerity measures appear to be slowing the return to economic growth.

BIG DATA AS A DRIVER OF BUSINESS OPPORTUNITIES

Huge untapped opportunities exist in big data, but most commercial organizations in most sectors just do not know how to handle, identify, and exploit these opportunities. The management mindset must change.

This is also true of government, which can and must play a central role at the head of a broad coalition embracing business, academia, workers, and students to unlock the potential of big data.

If we can recognize big data as the new asset class that it is, the economic upswing could well match that of the second industrial revolution brought about by the mass production methods of Henry Ford and the scientific management techniques of Frederick Winslow Taylor.1

What is big data?

Big data is a popular term used to describe the exponential growth in the volume, variety, and velocity of data. At the same time that volumes of data are growing, the data used by organizations large and small are becoming increasingly variable, complex, and difficult to manage using established data management tools. An example is the highs and lows in data volumes created by web traffic originating in multiple sources, both external and internal to an organization. In 2011 alone, 1.8 zettabytes (or 1.8 trillion gigabytes) of data were created2—the equivalent of every person on the planet writing three tweets per minute for 1,210 years.

The term big data is therefore relative. It applies—per the assessment of leading information technology (IT) analyst the Gartner Group—when extreme information management and processing issues “exceed the capability of traditional information technology along one or multiple dimensions to support the use of the information assets.”3 This problem presents a huge opportunity: Gartner estimates that, by 2015, big data will directly create 4.4 million IT jobs globally, of which
Box 1: What does big data mean for the retail industry?

In its 2011 report, Big Data: The Next Frontier for Innovation, Competition and Productivity, McKinsey estimates that retailers who successfully harness big data could increase their operating margins by more than 60 percent. Retail is one of the most time-sensitive industries. Scott Zucker is Vice President of Business Services at Family Dollar, a grocery mega-chain with 7,100 stores in 45 states. Family Dollar relies on high-performance analytics to shrink data-processing windows from days to less than an hour.

"Big data allows us to look at product, time and location—our critical analytical levers—at a much lower level than we ever did before," Zucker says. "We might have looked at class or subclass, at total company, and then at month and sometimes at week. Now we’re looking at SKU, store and day. As we start going down to that level, the amount of information that we need to manage and analyze goes up exponentially."

"High-performance analytics lets you bring to market ideas, services, products and marketing plans much faster than you would ever think possible. No one ever does just one iteration of an analysis, right? There’s always the first iteration that goes to management, and then they want to look at it another way. We go back and forth for multiple iterations."

"Before high-performance analytics, that could take weeks or even a month. Now you can get data back in front of management the next day."[2]

Notes


1.9 million will be in the United States. With the multiplier effect, each of these additional IT jobs will create employment for three more people outside the tech industry in the United States, adding 6 million jobs to the economy.5

Likewise, a recent Centre for Economics and Business Research (CEBR) study has identified £216 billion worth of potential benefits to the United Kingdom alone through gains in efficiency, innovation, and creation driven by insights unlocked from big data (see Table 1).6

Moreover, according to research by Andrew McAfee and Erik Brynjolfsson of MIT, companies that inject big data and analytics into their operations show productivity rates and profitability that are 5 percent to 6 percent higher than those of their peers.6

This, however, could be just the tip of the iceberg. The overall impact may be far more difficult to quantify because, as was the case with Fordism and Taylorism, big data could be a “game-changer” with long-term effects that go way beyond improving the efficiency or creativity of how we do things today—in other words, big data could change the very nature of economic activity itself. Our work in the field with hundreds of SAS high-performance analytics clients indicates that big data will stimulate entirely new ways of doing things.

To tap into this opportunity, business, government, and society as a whole all need to adjust the way they think and act. Without new thinking, the current excitement surrounding big data could easily lead to disillusionment. The hardware and software technology needed to solve the volume aspect of the problem is now in place. Today, you can buy a disk drive that can store all the music in the world for just US$600.7 The know-how exists as well. Companies that were “born digital”—such as Amazon and Google—have built their success on big data. We now need to extend their data-driven mindsets to more traditional businesses and the public sector. If this happens, big data can get the global economy back on track.

Boxes 1 through 4 provide some examples of the impact big data can have in the retail, utilities, healthcare, and public sectors.

Data-driven decisions

Dynamic pricing in the airlines industry is an excellent example of the potential impact of big data on economic activity. Dynamic pricing, based on the analysis of millions of transactions to calculate the best current price point, broadens the market and maximizes revenue. Online shopping is another good example. Online retailers not only track what customers buy, but also what they look at and do not buy, their navigation paths (clickstreams), their propensity to respond to promotions and reviews, their own reviews and recommendations, and so on. By capturing and analyzing these data, online retailers can build models and algorithms to predict what other products the individual customer will buy, as well as the next big consumer trends. Moreover, these algorithms constantly learn from every customer interaction.

---

Table 1: UK industry benefits of big data, £ million, 2011–17 (2011 prices)

<table>
<thead>
<tr>
<th>Industry</th>
<th>2011</th>
<th>2012–17</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manufacturing</td>
<td>5,965</td>
<td>45,252</td>
</tr>
<tr>
<td>Retail</td>
<td>3,406</td>
<td>32,478</td>
</tr>
<tr>
<td>Other activities</td>
<td>3,446</td>
<td>27,929</td>
</tr>
<tr>
<td>Professional services</td>
<td>3,039</td>
<td>27,849</td>
</tr>
<tr>
<td>Central government</td>
<td>2,517</td>
<td>20,405</td>
</tr>
<tr>
<td>Healthcare</td>
<td>1,460</td>
<td>14,384</td>
</tr>
<tr>
<td>Telecommunications</td>
<td>1,465</td>
<td>13,740</td>
</tr>
<tr>
<td>Transport and logistics</td>
<td>1,360</td>
<td>12,417</td>
</tr>
<tr>
<td>Retail banking</td>
<td>708</td>
<td>6,408</td>
</tr>
<tr>
<td>Energy and utilities</td>
<td>660</td>
<td>5,430</td>
</tr>
<tr>
<td>Investment banking</td>
<td>554</td>
<td>5,275</td>
</tr>
<tr>
<td>Insurance</td>
<td>517</td>
<td>4,595</td>
</tr>
<tr>
<td>UK economy (total)</td>
<td>25,087</td>
<td>215,964</td>
</tr>
</tbody>
</table>

Box 2: What does big data mean for utilities?

Most organizations never saw the era of big data coming. But U.S. Gas & Electric, a major energy retailer in 12 US states, has been watching closely.

“Our industry is on the cusp of smart meters,” says Greg Taffet, CIO of U.S. Gas & Electric. Taffet is referring to the digital devices that will deliver a steady stream of real-time demand and usage information from customer homes to utility providers. Electricity providers manually read meters once a month, feed the data into complex algorithms that take into account historical weather and demand patterns, and make purchasing and pricing decisions based on the results. “There is still a lot of interpretation of the data involved,” says Taffet.

Within the next five to ten years, smart meters will begin streaming usage data to both U.S. Gas & Electric and its customers, significantly affecting the company’s business model. Customers are likely to be more energy-conscious with usage data at their disposal. U.S. Gas & Electric will have an opportunity to offer new services and may even begin expanding into ancillary businesses, such as selling high-efficiency air conditioners or offering insulation services.

“We think this has the opportunity to benefit both our customers and our own business model,” says Taffet. He estimates that smart meters will result in 1,000 times the data coming through his systems. In preparation, Taffet is investing heavily in infrastructure, especially storage and processing capacity. “It is going to be a game changer,” he says.1

Note
1 Economist Intelligence Unit 2011, p. 22.

Box 3: What does big data mean for healthcare?

In its 2011 report, Big Data: The Next Frontier for Innovation, Competition and Productivity, McKinsey estimates that the potential value from data in US healthcare could be more than US$300 billion per year.1

“In healthcare, it’s a tidal wave of data. And our ability to restructure and change our culture is almost entirely informed by these data,” says Dr Jim LaBelle, corporate vice president of quality, medical management, and physician co-management at Scripps Health, the San Diego–based company that includes five hospitals.

For several years, Dr LaBelle has been overseeing an effort to change the culture at Scripps, from measuring quality almost entirely by the performance of physicians to measuring quality by the performance of processes, systems, and teams. “We are looking at monitoring variation around processes and driving out waste and supporting better care by developing a management system and partnership with the medical staff,” LaBelle says.

To inform its approach to these changes, Scripps collects and analyzes variation data. For example, in anticipation of re-engineering its emergency room procedures, Scripps looked at masses of data on wait times (such as the door-to-doctor metric) and cross-referenced the information against the type of injury, tests that were ordered, and how long it took to discharge the patient. “Then we did extensive simulation of our processes using real-life data, modeling how new and different processes might work,” LaBelle says.

Scripps found that the triage process added an unnecessary and wasteful step in getting patients from the door to a doctor. It was adding time and cost to the system, and not adding significant value. The company eliminated it. “We were able to reduce door-to-doctor time, add capacity to our emergency rooms, and improve the quality of our service,” LaBelle says.

Other examples of how we have seen big data analytics boosting our clients’ businesses include the ability to:

- recalculate entire risk portfolios in minutes and understand future possibilities to mitigate risk; 8
- analyze millions of SKUs to determine optimal prices that maximize profit and clear inventory; 9
- better understand customers to optimize product assortments; 10
- send tailored recommendations to mobile devices at just the right time, while customers are in the right location to take advantage of offers; 11
- analyze data from social media to detect new market trends and changes in demand; 12 and
- use data mining to detect fraudulent behavior. 13

In each case, success is determined by how effectively the organization (1) harnesses data and uses them creatively, (2) builds models that enable it to predict better and to optimize outcomes, and (3) transforms itself so that it is more agile in acting on insight. It is this last requirement that poses the greatest challenge and it is here that government can play an active supporting role, as discussed below.

GOVERNMENT BIG DATA INITIATIVES TO TACKLE UNEMPLOYMENT

Government is one of the largest users of data. It must now take the lead both as an exemplar and as an enabler of big data best practices. McKinsey estimates that the governments of developed European Union countries could save more than €100 billion (US$149 billion) in operational efficiency improvements alone by using big data. 14 Our work with government agencies demonstrates that far more can be saved by using big data to reduce fraud and tax evasion.

Big data can also help government to make the leap from “fail and fix” to “predict to prevent.” A recent
study conducted by Global Pulse, in partnership with SAS Institute, using linguistic analytics, demonstrated how government agencies could harness big data from social media to help formulate policies to address unemployment. The primary goal of the research was to compare the qualitative information offered by social media with unemployment figures. We first selected related conversations from blogs, forums, and news from the United States and Ireland between June 2009 and June 2011.

Figure 1 illustrates the project workflow of the study. For all documents (blog posts, tweets, etc. in the public domain), we assigned a quantitative mood score based on the tone of the conversations—for example, happiness, depression, anxiety—it contained. We also quantified unemployment-related documents that dealt with other topics, such as housing and transportation, in order to gain insight into populations’ coping mechanisms.

We analyzed these data in two primary ways. First, we correlated mood scores with the unemployment rate to discover leading indicators that forecast rises and falls in the unemployment rate. For example, the social media conversations in Ireland categorized as showing a confused mood preceded variations in the unemployment rate with a lead time of three months. Second, the volume of documents related to coping mechanisms also showed a significant relationship with the unemployment rate, which may give insight into the reactions that can be expected from a population dealing with unemployment. For example, the conversations in the United States around the loss of housing increased two months after unemployment spikes.

Overall, in this initial research, Global Pulse underlined the potential of online conversations to complement official statistics by providing a qualitative picture demonstrating how people are feeling and coping with respect to their employment status. The conversations that provided insight ranged from the banal, such as “my beer budget will obviously be cut” to the heartbreaking, “a few more months and we’ll have to seriously consider a bankruptcy” and “sorry water bill, this month I will have to pay the electric, next month the student loan.” Taken together, and tagged by mood score, the conversations revealed strong correlations with the unemployment rate, providing leading indicators that unemployment will rise or fall.

Thus, the study showed how linguistic analytics could provide government with the predict-to-prevent capabilities needed to take action before a problem manifests itself. At the level of the individual, this could mean that retraining is made available months before a job loss is experienced, thereby reducing dependence...
on benefits. The high market penetration of social media among young people makes this especially relevant for youth unemployment programs.

While this was a proof-of-concept project, it shows the potential of using social data to influence policy. Building larger databases over time and using richer geographical information related to the inputs would allow more detailed analysis and more nuanced approaches at the regional level.

PROPOSALS FOR LABOR MARKET TRANSFORMATION

In the first two sections of this chapter, we discussed how big data can impact the economy in the private and public sectors, both by spurring innovation and growth and by giving government deeper insight into the needs of citizens. For big data to influence the economy further, we suggest several measures that are needed to create the right labor market conditions for big data–driven growth.

**Needed now: A big data skills-for-growth program**

Talent shortage is the greatest obstacle to realizing value from big data. Based on current trends, by 2020 the world will generate 50 times the amount of information and 75 times the number of “information containers” it uses now, while IT staff to manage it will grow less than 1.5 times.\(^{16}\)

Today’s youth is digitally literate to a degree older generations could never have imagined, yet the world is not producing anywhere near enough data scientists. Investigating big data to answer a business question typically involves a “mashup” of several analytical efforts, and this requires a new breed of professional.\(^ {19}\) We need data scientists who are also domain specialists in all sectors, from chief digital officer down to entry-level workers. Our current educational institutions are behind the curve. There are few university programs that address big data analytics, let alone that provide degrees in data science, and there are virtually no schemes to retrain people in big data skills. We need more—far more—workers who are trained in using information to identify and execute business opportunities.

Putting the tools and methods of analytics into the hands of the workforce would industrialize the information-based service economy, much as Frederick Winslow Taylor’s and Henry Ford’s innovations industrialized factory management. In Singapore, the Infocomm Development Agency (IDA) has established a High-Performance Analytics Centre of Innovation, the first of its kind in Asia. Its role is to train professionals in data management and analytics, and to generate intellectual property through co-development with institutes of higher learning.

**Box 4: The public sector can use big data to match skills to jobs**

Singapore’s Ministry of Manpower has developed an SAS analytics solution that draws information from a variety of departmental sources to support its operations planning, case management, and the early detection of potential workplace and employment issues. This has enabled it to put resources in place in a timely manner to give employers greater visibility into skills availability, to identify and close skills gaps, and to offer a more targeted service to both employees and employers.\(^ {1}\)

Analytics can also be applied to anticipate employment needs effectively within a public-sector organization. Recruitment has traditionally been very slow in the US public sector (it currently takes an average of 105 days to fill a post at a federal agency).\(^ {2}\) However, the state of North Carolina is using an SAS analytics application called NC WORKS that enables the state government to proactively manage and forecast talent needs. It provides the workforce with the intelligence needed to respond to the changing workforce demographics, including an aging and retiring staff.\(^ {3}\)

**Notes**


**Big data to match people to jobs more effectively**

Despite high levels of unemployment, companies continue to experience significant skills shortages. In a recent survey of European decision makers, 43 percent reported that they are currently facing at least a moderate shortage of required skills.\(^ {20}\) Often, the skills and location of unemployed workers do not correspond to the skills and location of positions available. Big data can help predict these gaps and mismatches before they become critical, and can put plans and programs in place to address those gaps.

Government requires better analytics to profile its data about the unemployed to identify specific characteristics, plan appropriate interventions, and then track the impact of measures taken over time. Better analytics can simplify job searches, automatically provide jobseekers with options, identify the capabilities they lack to qualify for certain jobs, and direct them to the necessary programs for retraining.

If “industrialized,” such an approach can enable unemployment agencies to be more proactive in matching people to jobs and jobs to people. From a big data perspective, the process is not that different from what many large companies are already doing to identify trends and match future supply and demand.
Information and communication technologies to match jobs to people more effectively

The old Catch 22 is at work: workers cannot acquire skills and experience because employers only want workers with skills and experience. Deskilling has been a feature of employment booms and may provide part of the answer.21 The second industrial revolution was possible because companies such as Ford introduced production techniques that overcame the shortage of skilled engineers—opening up employment opportunities for unskilled and semiskilled workers. More recently, the media industry—once dependent on skilled typesetters, graphic artists, and other craft workers—has entered a new age largely as a result of deskilling, facilitated by digital technology.

The public sector can take the lead here. Policymakers should ask how they can redesign workplaces to reduce reliance on scarce managerial talent. Deskilling in public-sector organizations is likely to mean less top-heavy bureaucracy, flatter hierarchies, and greater workplace democracy, supported by big data to provide objective insight.

There is no reason why such an approach could not succeed in the private-sector corporate environment as well. In fact, this approach is highly likely to result in more agile and competitive enterprises as decisions are made more swiftly, but based on scientific analytics rather than executive fiat or company politics.

Restructure labor markets to optimize skills creation

In the era of big data, how do we optimize the labor value creation and delivery chain for a world where business must adapt and transform itself more and more rapidly? We must question our very perception of what constitutes a "job" and what constitutes a "profession."

In an information-led economy, knowledge process outsourcing organizations (KPOs) will assume a more prominent position in the employment landscape. KPOs provide sources of technical talent, with the knowledge workers often located remotely from the customer.

Although the KPO model has been most closely associated with information and communication technology companies, it can be extended to other areas such as legal processes and research, intellectual property and patent-related services, engineering services, web development applications, CAD/CAM applications, clinical research, publishing, and marketing services. The advantage of KPOs is their flexibility. They do away with the traditional recruitment process, overcome barriers to labor mobility, and are low risk for the employer while offering high rewards and variety for the employee. The fast-changing nature of the digital economy means there will be increasing demand for people who want new challenges rather than routine.

KPOs provide domain knowledge (such as expertise in IT, legal, marketing, or accountancy) to organizations that do not want to move such knowledge in-house, enabling these organizations to focus on core-competency areas that generate business growth. For example, KPOs might enable an electronics company to focus more on its core competence—developing innovative electronic circuitry—instead of employing people who file patents or run internal IT systems.

CONCLUSION

As we have shown here, forward-thinking governments in economies from Singapore to Ireland are already taking positive steps toward inclusive growth through the creative use of big data and analytics. Others must follow.

The opportunity is very easy to grasp. Big data can deliver insight. With the application of high-performance analytics to big data, public and private organizations can get the intelligence they need to support decisions in hours or even minutes instead of days and weeks. In simple terms, this will enable businesses to move away from the traditional intuitive management approach, which we would characterize as “fail and fix” or “fail fast” to one we would characterize as “predict to prevent” and “predict to perfect.”

Fail and fix, though always wasteful, can work in boom years when there is margin for error. It does not work in the “new normal” of economic inertia. The fear of failure is too great. The crisis of 2008 should have spelled the end of the fail-and-fix approach. SAS High-Performance Analytics, which uses parallel processing and advanced statistical techniques, can reveal previously unseen patterns and relationships in big data. It can enable governments and financial institutions and regulators to avoid the meltdowns that have characterized the financial landscape in recent years—and it can support business creation, business efficiency, and business innovation. For that to happen, mindsets must change to put more trust in analytics and the people who can interpret data.

NOTES

1. Henry Ford (1863–1947) was the American industrialist and sponsor of the development of the assembly line technique of mass production. Fordism is a concept used in various social theories and management studies about mass production and related socioeconomic phenomena. The term was introduced by Antonio Gramsci in 1934 in his essay “Americanism and Fordism,” in his Prison Notebooks. Frederick Winslow Taylor (1856–1915) was an American mechanical engineer who sought to improve industrial efficiency. Taylorism is a theory of management that analyzes and synthesizes workflows. Its main objective was improving economic efficiency, especially labor productivity.


5. CEBR 2012.

6. McAfee and Brynjolfsson 2012.
REFERENCES


