

# Information, Data, Transparency

Technological advances have resulted in an exponential growth in the amount of data generated by humankind. In today's connected society people constantly upload videos and photos to social media websites, share messages on blogs and Twitter, allow their cell phone's GPS to track their movements, and websites to follow their browsing patterns. To put the amount of data generated by modern society into context, more data crosses the Internet today every second than was stored in the entire Internet 20 years ago. This growth is driven by the strengthening of certain digital megatrends: the Internet of Things, smart devices, cloud computing and the rise of online communities. The fact that these megatrends are expected to continue growing, points to the fact that the volume of data generated will continue to increase and by 2020 humanity will generate 5 times more information than today.

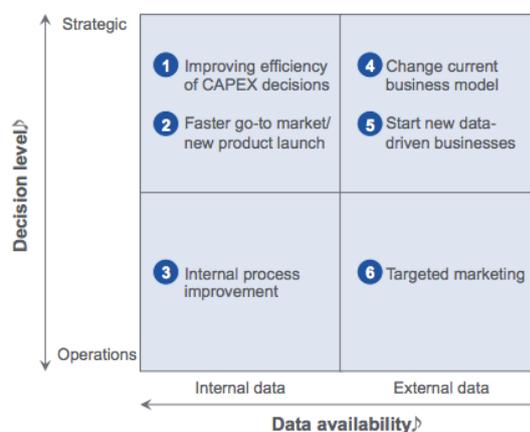
The drastic increase in the volume of data being collected, stored and analyzed by governments, universities and companies around the globe led to the creation of the term Big Data. Big Data analysis differs from traditional data analysis due to three factors: Volume, Velocity and Variety, enabling companies to undertake more relevant and timely decisions than was possible with traditional business-intelligence methods.

- **Volume:** The sheer volume of data available nowadays presents a significant challenge to traditional IT structures.
- **Velocity:** Along with the rise in data volume there has also been an increase in the velocity at which data is collected and analyzed. Velocity is required not only to store the information but also to ensure that the speed of the feedback loop between data gathering to actionable insight to decision making is as short as possible.
- **Variety:** Nowadays the source of data is extremely diverse, coming in as text from social networks, photos, videos, directly from sensor sources, and GPS signals from cell phones. The challenge of Big Data processing is taking this

diverse, unstructured data and extracting ordered meaning.

Due to these three unique characteristics most businesses are currently having trouble extracting a fourth 'V' from Big Data: Value. It is estimated that only about 0.5% of all collected data is properly analyzed. Concurrently, only 4% of companies are able to extract value from their data.

The way companies can generate value from Big Data initiatives may be classified in two axes: whether the initiative is operational or strategic, and whether it relies primarily on internal or external data. From there, six different ways companies can profit from Big Data may be derived:



Regardless of how businesses intend on reaping the rewards of Big Data, they should focus on three areas to immediately build the necessary capabilities and avoid the most common pitfalls when implementing such an initiative:

- **Adjust Corporate Culture:** Big Data analytics requires a corporate culture that allows experimentation and encourages non-traditional ideas. One of Big Data's greatest strengths is providing business with deeper insights that, at first glance, might even appear to be counter intuitive. Unless the company has a corporate culture open to new ideas, these valuable insights might not be spotted, or acted upon.
- **Build a secure technology platform:** Companies should implement a scalable and secure IT platform through the entire organization that allows for growth, both in size and complexity, of the Big Data initiative.

- **Enable the organization:** Putting in place operational processes that reduce the time of the feedback loop between data input and actionable decision is vital to fully capturing the value of Big Data. Furthermore, securing senior-level sponsorship is very important. Fortune Knowledge Group found that "62% of business leaders said they tend to trust their gut, and 61% said real-world insight tops hard analytics when making decisions." Unless that mentality is altered, Big Data initiatives are likely to fail. Finally, due to the lack of talented professionals in the field, attracting and retaining key skills/talents remains crucial to driving a Big Data revolution within companies.

However, even if you do all of the above, that might still not be enough to guarantee Big Data success due to the current state of the external environment.

### Current Big Data Environment

Three main players populate the current Big Data environment: Businesses/companies, governments/regulators, and individuals. The struggle between them will define the future of the industry. As it stands we are living in a time when there are no clear rules and the business landscape is chaotic. The main factors driving this instability are: *Regulation, Data Ownership and Privacy Concerns.*

### Regulation

This is the biggest challenge to the future development of Big Data, as the current regulatory framework is fragmented across countries and outdated and does not address some extremely important issues, such as:

- *Consent before data collection*
- *Definition of 'Personal Data'*
- *Data 'anonymization'*
- *The 'right to be forgotten'*
- *Relevant jurisdiction across the globe*
- *Liability issues between companies in case of data leakage.*

These issues need to be quickly resolved to create a stable environment for Big Data to flourish. Unfortunately, regulations are usually slow to catch up to technological developments and there are rarely unified initiatives. Furthermore, new regulations are normally unclear and do little to address the problem at

hand, sometimes even helping to create more uncertainty (e.g. the 'right to be forgotten').

### Data Ownership

The concept of data ownership is another issue that needs to be addressed. Usually individuals are unaware of the rights they are giving up and how businesses may use their data.

### Privacy Concerns

The last big driver of the current instability in the Big Data environment is related to privacy concerns emanating from individuals. Unless companies are able to regain users' trust, they might start to suffer the consequences of abusing users' privacy. Some companies have caught on to this and players like Microsoft are already calling for better defined and clearer Big Data privacy laws. Ultimately governments and regulators around the globe will have to decide whether to side more with individuals to protect privacy, or with companies to foster innovation and economic development. It remains to be seen if a healthy balance between the two will be achieved.

Each of these three drivers (regulation, data ownership, and privacy concerns) can move in various directions, thus shaping the future of Big Data and data/information privacy. From the possible outcome we foresee three possible scenarios for the future of the next 10 to 15 years:

### Future Scenarios

#### Digital Wild West

The Digital Wild West is a continuation of the present situation, a world where governments are constantly lagging behind in their regulations, passing moderate laws that are primarily technology based and that quickly become outdated. In this scenario, the lack of a comprehensive regulatory framework will lead to companies continuing to own data and individuals constantly battling companies to try to maintain their privacy. This vision is supported by the long history of most governments being extremely inefficient in creating proper regulatory frameworks. Additionally, the increase in lawsuits of individuals against companies like Google and Facebook over data ownership and privacy prove that these are real issues that will remain in place unless properly regulated. Finally, there has been an increase in demand for encrypted

messaging, signaling that people are ever more concerned with their data privacy and security. In this scenario companies can get an upper hand by focusing on gathering more data and maintaining data quality. While these strategies are applicable in any environment they become crucial in the Digital Wild West. The lack of proper regulation makes it less restrictive for companies to gather as much data as possible. Companies should therefore invest heavily in technologies that increase the volume of data gathered to better predict the social trends and behavioral patterns of customers in general, exploiting regulatory loopholes if needed. In the end a company that owns more data than its competitors will be able to better predict the behaviors and trends of larger populations, thus gaining a competitive advantage. As individuals 'battle' companies over their privacy, it is likely they will start encrypting their data, and purposely send wrong information to companies. Companies must therefore increase the amount of data sources (variety) they tap into to get better insights and detailed profiles about customers. Collecting and sorting through the 'real and relevant' data will become increasingly complex and vital for business.

### **Digital Eden**

In this scenario governments and regulators are able either to establish the necessary regulatory frameworks, or the necessary incentives for companies to self-regulate over issues of data privacy and ownership.

Individuals realize the value of data and have 'won the battle' with companies over its collection and usage. In this case individuals have complete control over exactly what data they share and when the data should be erased from company records. Examples of this might be in personalized medicine where patients have great incentives to share as much about themselves as possible and even allow for DNA collection. Companies are either forbidden by law to violate individuals' privacies or have no economic incentive to do so due to risk of retaliation by their own customers and even other industry players.

In this scenario companies may profit by clearly demonstrating the value of data sharing to customers and developing trend-spotting capabilities. Being close to your customer is always a necessity, but this is doubly true in a

world where data is harder and harder to come by. Companies need to demonstrate clearly to customers the benefits of sharing their data to be able to gain deeper insights and better serve them. In this world, targeted marketing is much more difficult to execute successfully and word of mouth will be even more important than today. Though access to private data will be restricted, companies will still have access to public data. Therefore companies should focus on building in-house capabilities for trend-spotting based on the publicly available information. Due to free public data access from every potential competitor, being able to extract value from the public data internally becomes crucial to gaining a competitive advantage.

### **Big Brother**

In the final scenario, society as a whole has realized that the benefit of increasing the amount of freely available data outweighs any privacy concerns. In fact, governments will put in place laws and regulations that encourage companies to gather as much data as possible and share it, while making it more difficult for individuals to hold on to their privacy. Privacy becomes an outdated concept and companies will own what today would be classified as personal data, and freely decide on how to use it.

Under this scenario companies should primarily invest in customizing services for target customers through predictive analytic capabilities. Freely accessible data for companies means the competitive advantage will not be in its collection or storage, but rather on analysis. Companies will have to focus on constantly being ahead of their competitors in terms of using predictive analytic capabilities to offer customized services, as product and service differentiation will be increasingly difficult to sustain.

### **Conclusion**

Regardless of which scenario plays out, there are certain actions that will be valuable across any of them. The first should be following the actions mentioned in section 4: *Adjusting the corporate culture, building a secure technology platform, and enabling the organization.* Secondly, companies must communicate clearly with customers by implementing strategy, governance and systems focused on

transparency. Eventually, trusting that the company is not mishandling private data or deceiving customers into sharing more than they would be willing to, will be the main driver to retaining said customers. Companies that exploit these guidelines will see their reputation rapidly decrease and in turn their customer base diminish. In turn, companies that gain and foster customers' trust will have access to even more data. In today's world where data is an increasingly valuable resource, this will give companies a distinct competitive advantage.