

April 15, 2015

Big Data Looking Beyond the Hype

By Robert Balon

Seventy-seven percent of information technology (IT) leaders plan to take advantage of the potential offered by Big Data. Ninety percent believe investments in Big Data initiatives are worthwhile¹, yet more than half report:

- 🕒 Data is stored in disparate systems that don't "talk" to each other.
- 🕒 Data owners are not held accountable for ensuring data quality.
- 🕒 Data ownership is not clearly defined.
- 🕒 Quality of data is questionable.

Big Data offers many promises, but what is it? What steps can an organization take to ensure that investments in Big Data deliver tangible results to the business?

What's the Big Deal?

Big Data is often regarded as a reference to huge data sets that surpass the capabilities of traditional database tools. Several common frameworks recognize Big Data associated with large data volumes, rapid data accessibility, and a wide variety of data sources and types. The term "Big Data" is also frequently associated with Business Intelligence and Data Analytics efforts, with visions of sophisticated data mining tools and algorithms that harvest magical information.

All of these impressions are true, but harnessing the power of extreme data sets and computing power requires the data to be structured and organized in a way that can be referenced and used consistently. This is a significant challenge for most organizations.

Name, Address & Phone Number: The Realities of Dirty Data

Ninety percent of the world's data has been generated in the last two years. By the time you finish reading this article, people will have initiated 20 million Google searches, sent two billion emails and posted more than one million tweets.²

These are impressive statistics, but the reality is that most corporate data suffers from basic quality and consistency issues.^{3,4,5} Some common examples include data associated with fundamental client or customer data such as names, addresses, phone numbers, email addresses, and contact preferences. Data associated with these entities and attributes are frequently housed in disparate, fragmented data sources.

An entity as simple as "product", or attribute such as "part number", can have different meanings and inconsistencies in how the underlying values are stored. This can wreak havoc on any data



integration project unless a meaningful data cleansing effort is first completed. Numerous data projects have failed or faced significant challenges by ignoring data cleanup as a critical first step.

What Do We Do With All This Data?

On a cold Chicago February evening, Professor Florian Zettelmeyer from Northwestern University's Kellogg School of Management led a group of expert panelists in a heated discussion on Big Data and common misperceptions in front of a packed conference hall.⁶ What was the number one revelation? Big Data is not a technology challenge, it is a leadership challenge. In fact, the theme for the entire executive networking event was "Big Data Doesn't Make Decisions, Leaders Do."

Professor Zettelmeyer and his expert group of panelists spent more than an hour discussing concepts such as change management, business priorities, and well-defined processes and roles. What's going on? These concepts are traditionally associated with business initiatives, not data and technology projects. Yet, the group consensus was clear - the business priorities and leadership challenges associated with Big Data projects far outweigh the technical considerations.⁶

Once data has been cleansed, aggregated, compiled, and synthesized, many organizations then ask themselves "Now what?" Understanding how any data effort relates to business objectives and priorities is critical before significant costs are spent on technology and integration. Answering the question of "How will this data improve our business?" is essential in the first stage of any data project. It is quite likely that the data integration approach will change as priorities change. Recognizing data projects as an extension of business leadership (not merely a technology exercise) is vital.

Big Data & Cybersecurity

It is beyond the scope of this article to delve deeply into cybersecurity. However, many organizations regard Big Data and Cybersecurity efforts as similar endeavors. More data exposes greater cybersecurity risks.⁷ For example, more data proliferation is typically accompanied by more data sources and more individuals with greater access. Increased flexibility and connectivity heightens data security considerations.

Alternatively, more data can also be captured with respect to intrusions, access logs, and audit trails. Harnessing the power of data can certainly help mitigate cybersecurity risks even as more data poses greater risks.

Either way, the leadership and governance challenges associated with cybersecurity are quickly put into perspective. According to Kellogg's expert panelists, 90% of cybersecurity threats are caused by human failure to follow simple policies and procedures.⁶ Studies by IBM and Verizon recognize that more than 50% of cybersecurity attacks are initiated by outsiders that exploit weaknesses afforded by unsuspecting insiders that inadvertently left a door open for malicious code or unauthorized access.^{8,9} In this case, the leadership and governance aspects of Big Data once again overshadow the technical considerations.

In summary, Big Data initiatives require business leadership, not merely technology solutions. Successful data integration efforts recognize the importance of data quality and consistency. When business priorities and objectives are clearly defined, Big Data efforts can deliver compelling results.



References

1. Big Data: The Next Frontier, Teksystems, September 2013.
2. Data on Big Data, Marcia Conner, July 2014.
3. Data Cleaning: Problems and Current Approaches, Erhard Rahm & Hong Hao Do, University of Leipzig, Germany, 2000.
4. Data Cleansing Big Data: Scrubbing the Elephant, xplenty, May 2014.
5. Data Quality: The ugly duckling of big data?, TechRepublic, February 2014.
6. Kellogg Insight Live: Big Data Doesn't Make Decisions, Leaders Do, Executive/Alumni Panel With Professor Florian Zettelmeyer, February 2015, Chicago, Illinois.
7. Garner Debunks Five of the Biggest Data Myths, Business Intelligence & Analytics Summit 2014, Gartner, October 2014, Munich, Germany.
8. Cyber Security Intelligence Index, IBM Security Services, December 2014.
9. Data Breach Investigations Report, Verizon, 2014.

Kral Ussery LLC serves US public and private companies to protect and grow shareholder value, as well as non-profits and governments with internal controls and in combating fraud. We assist entities with governance and in all matters relating to financial reporting, including SEC compliance, internal controls testing and remediation, IT general controls, IPO readiness, M&A transactions, US GAAP compliance and implementation of new accounting standards. Visit us at www.KralUssery.com.

This is an article from the Governance Issues™ Newsletter, Volume 2015, Number 2, by Candela Solutions, which has since merged to form Kral Ussery LLC.

© Kral Ussery LLC. Copyright: The Governance Issues™ Newsletter is meant to be distributed freely to interested parties. However, any use of this article must credit the respective author and Kral Ussery LLC as the publisher. All rights reserved. Use of the newsletter article constitutes acceptance of our [Disclaimer](#) and [Privacy Policy](#).

To automatically receive the newsletter, go to www.KralUssery.com and register. Or, send a request to newsletter@KralUssery.com and we will register you.