An ISACA White Paper January 2014

GENERATING VALUE FROM BIG DATA ANALYTICS

ABSTRACT

Many enterprises are moving quickly to adopt “big data analytics”—specifically, the application of advances in analytics techniques to the rapidly-expanding pool of information that enterprises have at their disposal to enable better decision making. As this trend of adoption continues, information security, risk and audit professionals are likely to become increasingly aware of the possible technical and operational risk that may arise as a result of adoption in their enterprises. However, non-adoption can also carry its own risk—particularly in the arena of business competitiveness. To analyze risk holistically, practitioners need to evaluate both technical risk and the business risk, in equal measure. Understanding the “use case”—the reasons why big data analytics is appealing from a business perspective—can help ensure that both angles are considered and practitioners guide their enterprises to be optimally competitive.

ISACA®
Trust in, and value from, information systems

ISACA®

With more than 110,000 constituents in 180 countries, ISACA *(www.isaca.org)* helps business and IT leaders maximize value and manage risk related to information and technology. Founded in 1969, the nonprofit, independent ISACA is an advocate for professionals involved in information security, assurance, risk management and governance. These professionals rely on ISACA as the trusted source for information and technology knowledge, community, standards and certification. The association, which has 200 chapters worldwide, advances and validates business-critical skills and knowledge through the globally respected Certified Information Systems Auditor® (CISA®), Certified Information Security Manager® (CISM®), Certified in the Governance of Enterprise IT® (CGEIT®) and Certified in Risk and Information Systems Control™ (CRISC™) credentials. ISACA also developed and continually updates COBIT®, a business framework that helps enterprises in all industries and geographies govern and manage their information and technology.

Disclaimer

ISACA has designed and created *Generating Value From Big Data Analytics* (the "Work") primarily as an educational resource for governance and assurance professionals. ISACA makes no claim that use of any of the Work will assure a successful outcome. The Work should not be considered inclusive of all proper information, procedures and tests or exclusive of other information, procedures and tests that are reasonably directed to obtaining the same results. In determining the propriety of any specific information, procedure or test, governance and assurance professionals should apply their own professional judgment to the specific circumstances presented by the particular systems or information technology environment.

Trust in, and value from, information systems

3701 Algonquin Road, Suite 1010
Rolling Meadows, IL 60008 USA

Phone: +1.847.253.1545

Fax: +1.847.253.1443

Email: *info@isaca.org*

Web site: *www.isaca.org*

Provide feedback:
www.isaca.org/big-data-analytics

Participate in the ISACA Knowledge Center:
www.isaca.org/knowledge-center

Follow ISACA on Twitter:
https://twitter.com/ISACANews

Join ISACA on LinkedIn: ISACA (Official)
http://linkd.in/ISACAOfficial

Like ISACA on Facebook:
www.facebook.com/ISACAHQ

Reservation of Rights

ACKNOWLEDGMENTS

ISACA Wishes to Recognize:

Project Development Team

Bhavesh Bhagat
CGEIT, CISM, EnCrisp Chairman and ConfidentGovernance.com Co-founder

Zubin Chagpar
CISA, CISM

Victor Chapela
CRISC, Sm4rt Corp, CEO & Founder

William Gee
CISA, CRISC, FCA, FHKICPA, PwC, Partner

Norman Marks
CPA, CRMA, Independent Evangelist for Better Run Business

Jamie Pasfield
CGEIT, ITIL V3, MSP, PRINCE2, Pfizer, UK

Expert Reviewers

Dave Newell
CISSP, Loptr LLC, Founder

Norman Marks
CPA, CRMA, Independent Evangelist for Better Run Business

Peter O'Kelly

Rammiya Perumal
CISA, CISM, CRISC, Sumitomo Mitsui Bank, USA

Mike Saurbaugh
CISM, CRISC, CISSP, Corning Credit Union, Manager Information Security

ISACA Board of Directors

Tony Hayes
CGEIT, AFCHSE, CHE, FACS, FCPA, FIIA, Queensland Government, Australia, International President

Allan Boardman
CISA, CISM, CGEIT, CRISC, ACA, CA (SA), CISSP, Morgan Stanley, UK, Vice President

Juan Luis Carselle
CISA, CGEIT, CRISC, Wal-Mart, Mexico, Vice President

Ramses Gallego
CISM, CGEIT, CCSK, CISSP, SCPM, Six Sigma Black Belt, Dell, Spain, Vice President

Theresa Grafenstine
CISA, CGEIT, CRISC, CGAP, CGMA, CIA, CPA, US House of Representatives, USA, Vice President

Vittal Raj
CISA, CISM, CGEIT, CFE. CIA, CISSP, FCA, Kumar & Raj, India, Vice President

Jeff Spivey
CRISC, CPP, PSP, Security Risk Management Inc., USA, Vice President

Marc Vael, Ph.D.
CISA, CISM, CGEIT, CRISC, CISSP, Valuendo, Belgium, Vice President

Gregory T. Grocholski
CISA, The Dow Chemical Co., USA, Past International President

Kenneth L. Vander Wal
CISA, CPA, Ernst & Young LLP (retired), USA, Past International President

Christos K. Dimitriadis, Ph.D.
CISA, CISM, CRISC, INTRALOT S.A., Greece, Director

Krysten McCabe
CISA, The Home Depot, USA, Director

Jo Stewart-Rattray
CISA, CISM, CGEIT, CRISC, CSEPS, BRM Holdich, Australia, Director

Knowledge Board

Christos K. Dimitriadis, Ph.D.
CISA, CISM, CRISC, INTRALOT S.A., Greece, Chairman

Rosemary M. Amato
CISA, CMA, CPA, Deloitte Touche Tohmatsu Ltd., The Netherlands

Steven A. Babb
CGEIT, CRISC, Betfair, UK

Thomas E. Borton
CISA, CISM, CRISC, CISSP, Cost Plus, USA

Phil J. Lageschulte
CGEIT, CPA, KPMG LLP, USA

Anthony P. Noble
CISA, Viacom, USA

Jamie Pasfield
CGEIT, ITIL V3, MSP, PRINCE2, Pfizer, UK

Guidance and Practices Committee

Phil J. Lageschulte
CGEIT, CPA, KPMG LLP, USA, Chairman

John Jasinski
CISA, CGEIT, ISO20K, ITIL Exp, SSBB, ITSMBP, USA

Yves Marcel Le Roux
CISM, CISSP, CA Technologies, France

Aureo Monteiro Tavares Da Silva
CISM, CGEIT, Brazil

Jotham Nyamari
CISA, CISSP, Deloitte, USA

James Seaman
CISM, RandomStorm, UK

Gurvinder Singh
CISA, CISM, CRISC, Australia

Siang Jun Julia Yeo
CISA, CPA (Australia), MasterCard Asia/Pacific Pte. Ltd., Singapore

Nikolaos Zacharopoulos
CISA, CISSP, DeutschePost–DHL, Germany

INTRODUCTION

New analytics tools and methods are expanding the possibilities for how enterprises can derive value from existing data within their organizations and from freely available external information sources, such as software as a service (SaaS), social media and commercial data sources. While traditional business intelligence has generally targeted "structured data" that can be easily parsed and analyzed, advances in analytics methods now allow examination of more varied data types.

These advances allow enterprises to make better business decisions and increase competitive advantage. This renaissance of analytics capability—"big data analytics"—can also introduce additional technical and operational risk. Enterprises should understand that risk can be incurred either through adoption or non-adoption of big data analytics; specifically, enterprises must weigh the technical and operational risk against the business risk that is associated with failure to adopt.

Information security, audit and governance professionals should take a holistic approach and understand the business case of big data analytics and the potential technical risk when evaluating the use and deployment of big data analytics in their organizations.

As with any potential investment that is intended to support improvement in efficacy or efficiency of business activities, several key elements must be well understood to enable systematic strategic planning:

- Anticipated returns and potential impacts to competiveness through adoption
- Potential impact to the current operational ecosystem
- Opportunity cost for the investment (i.e., what else the enterprise might have invested in instead)
- Loss of value for investments already made

Objective and systematic analysis of these factors becomes increasingly challenging as the industry "hype" that surrounds a new technology or business trend increases: hype can, in some cases, create unfounded pressure to adopt, or it can create barriers to adoption in others.

For technologists, the potential technical, operational and compliance risk that is associated with maintaining and operating on a large volume of potentially sensitive data is very apparent; however, the business-relevant factors that provided the initial impetus for adoption of these analytics tools and methods may be less apparent. Addressing this risk is important—ISACA has already started to develop guidance about the impact of big data, by outlining governance and risk considerations that enterprises should address as they adopt (e.g., ISACA's white paper "Big Data: Impacts and Benefits"[1]).

However, understanding the business case is equally important—understanding the rationale for adoption, the anticipated return that the business hopes to achieve and the competitiveness impact to the business if the enterprise chooses not to adopt while its competitors do adopt. For information security, audit and governance professionals, lack of clarity about the business case may stifle organizational success and lead to role and responsibility confusion.

The intent of this white paper is to provide supplemental guidance about the business factors that drive the adoption of big data analytics, by explaining the reasons why adoption is compelling to the enterprise and supporting these reasons with real-world examples. Understanding the business case and the specific challenges of investing in big data analytics allows information security and audit professionals to balance controlling technical and operational risk with enabling the mission of the enterprise they support.

[1] ISACA.org, "Big Data: Impacts and Benefits," March 2013, *www.isaca.org/Knowledge-Center/Research/ResearchDeliverables/Pages/Big-Data-Impacts-and-Benefits.aspx*

UNDERSTANDING THE BUSINESS CASE

"Big data" refers to large, quickly growing or varied types of information ("high volume, high velocity, and/or high variety information assets that require new forms of processing to enable enhanced decision making, insight discovery and process optimization"[2]). Big data analytics is the application of emerging statistical, processing and analytics techniques to big data for the purpose of advancing the business—applying statistical models and techniques to business information to derive conclusions that are beneficial to that business.

Big data analytics is particularly appealing to many enterprises because, in many cases, they have already made some investment in both business analytics and the collection of large data sets, on which analytics can be applied.

This means that the foundation may already be in place from which to draw new conclusions, explore new ways of doing business and open up new avenues of competitive advantage.

What is this competitive advantage specifically? Some data suggest a direct correlation between the use of big data analytics and profitability. For example, one study cites an increase in overall profitability of six percent as a direct result of using big data effectively.[3] That metric, while appealing in the abstract, lacks sufficient underlying context and level of detail to be able to understand precisely how that correlation is made. Specific case studies, by contrast, provide a clearer picture of how increases in competitiveness are achieved and why these analytics techniques provide value. These studies show the imaginative uses to which preexisting data are leveraged as a result of better analytics techniques and the transformative impacts that are achieved.

[2] Gartner.org, "Big Data," *www.gartner.com/it-glossary/big-data/*

[3] Libert, Barry; "Why Boards Must Embrace Big Data," Directorship.com, 11 September 2013, *www.directorship.com/why-boards-must-embrace-big-data/*

UNDERSTANDING THE BUSINESS CASE

A study published by IBM® gives an example of an increase in competitive advantage as a result of the application of big data analytics in the healthcare sector. This study cites advances made by State University of New York (SUNY) at Buffalo that increased the efficiency of the genetic research analysis process. SUNY Buffalo was able to decrease the amount of time required to perform analysis of gene-environmental interactions analysis in support of Multiple Sclerosis research by 99 percent—reducing analysis time from 27.2 hours to 11.7 minutes.[4] This represents a performance improvement in this mission-critical process of two orders of magnitude—most likely a huge competitive leap for any enterprise, where it should prove feasible.

Performance improvement of a given process is obviously desirable, but competitive advantages also can extend into increases in quality of process output. Another published study reveals advances made by Columbia University Medical Center in locating complications in brain-injured patients. Through the application of advanced analytics techniques, they were able to shorten the time to identify possible complications by up to 48 hours.[5] This time savings translates directly to better patient care and potentially lifesaving results.

Healthcare examples provide proof points to illustrate the transformative impact that can be achieved in quality and process performance within a given sector. Healthcare is a natural early fit for application of advances in analytics because, in most cases, a wealth of information on which to apply analytics already exists. Healthcare providers retain information about patient treatments, outcomes and health history and numerous other factors that are potentially relevant to diagnostics (e.g., risk factors for disease and lifestyle). Therefore, healthcare providers are able to apply new analytics techniques directly to data that they already have available. For example, providers can analyze a given patient against a number of other similar cases to see which treatment methods are likely to be most effective; they can examine environmental, genetic or lifestyle factors to analyze how well particular treatment methods are likely to work for a given patient.

[4] IBM.com, "State University of New York at Buffalo: Substantial data analysis improves gene-environmental correlation identification to help develop new treatment for multiple sclerosis," 25 July 2013, *www-01.ibm.com/software/success/cssdb.nsf/CS/KPES-97WTG7?OpenDocument&Site=default&cty=en_us*

[5] IBM.com, "Data-driven healthcare organizations use big data analytics for big gains," February 2013, *www-03.ibm.com/industries/ca/en/healthcare/documents/Data_driven_healthcare_organizations_use_big_data_analytics_for_big_gains.pdf*

UNDERSTANDING THE BUSINESS CASE

When data are already present, the limiting factors to gaining these advantages are accessibility of those data for analysis and the availability of analytics techniques themselves.

Other sectors can realize benefits from big data analytics in the same way as the healthcare sector. Within the retail sector, for example, large-scale analytics techniques are being used to better understand customer purchasing behavior. For instance, men's clothier Brooks Brothers uses sales data from its 500-plus online and physical retail properties to dynamically adjust marketing efforts based on customer demand, to optimally arrange inventory items in marketing campaigns and to guide selection of clearance inventory.[6]

As with healthcare, retail applications can increase both speed and quality of processes through the employment of advanced analytics techniques. A study published by Alteryx, Inc. cited a 95-percent reduction in the assessment time for salon-enterprise Great Clips® to select new sites for their business[7] (illustrating an increase in process performance), and a study by Tiger Analytics cited increases of $90 million annually in margin improvements for a large, distributed retailer[8] (highlighting quality improvements to sales efforts).

What makes these improvements particularly compelling from a business perspective in many retail contexts is that, like healthcare, the data already exist. Data exist about customers, such as purchases they make and their receptiveness and responsiveness to marketing efforts. During the normal course of business, many enterprises collect large volumes of data about their customers—their habits, preferences, the specifics of individual transactions, fraud history, etc. When analyzed, these data allow enterprises to make changes and, subsequently, measure the performance of those changes. This allows those enterprises to dynamically shift inventory and/or pricing in response to consumer demand. In addition, data analytics allows enterprises to create better-targeted marketing campaigns and to better measure the efficacy of those campaigns, and to launch new products and service offerings in response to customer demand. From a business standpoint, investment in big data analytics is compelling because it leverages the otherwise latent or unused resource of already-collected data.

In addition to direct increases in the quality or throughput of processes, another aspect for enterprises to consider when analyzing the business case for big data analytics is the positive gains that might be realized by competitors if they choose to invest in getting more out of their data (thereby potentially realizing the benefits described previously), while the evaluating enterprise chooses to take no action.

This consideration is the inverse of the direct-value proposition outlined in the case studies discussed previously and reflects competitiveness as a two-fold equation. In this equation, the potential value directly realized as the result of an investment is weighed against the net-negative value if a competing organization chooses to invest when another enterprise does not.

If an enterprise chooses not to invest while its competitors do, this equates to an overall decrease in enterprise competitiveness.

Because most markets are "zero sum" (i.e., the number of customers is limited), the more competitive enterprise may be the one that invests in a response to technology gains—as publicized case studies highlight the value, the likelihood of investment among competitive peers and proliferated usage increases.

---

[6] Ravindranath, Mohana; "Brooks Brothers, national retailers analyze 'big data' from sales to adjust marketing." WashingtonPost.com, 22 September 2013, *http://articles.washingtonpost.com/2013-09-22/business/42299416_1_brooks-brothers-analytics-sales-data*

[7] Alteryx.com, "Case Study: Great Clips®," *http://www.alteryx.com/sites/default/files/resources/files/case-study-great-clips.pdf*

[8] Tigeranalytics.com, "Markdown Optimization," *www.tigeranalytics.com/markdown-optimization-case-study.html*

CHALLENGES

Enterprises need to ask the following questions to address key potential challenges before they can, with confidence, realize the gains from big data analytics:

- Does the enterprise have the people, process and technology in place to build capabilities that will make productive use of data that the enterprise has collected?
- Has the enterprise established roles and responsibilities and identified stakeholders?
- Does the enterprise have (or can it get) data on which to apply advanced analytics?

Not every enterprise will be equipped to make use of big data analytic techniques. Some enterprises may be missing key skills in their existing personnel, or they may be missing critical portions of the technological ecosystem. Also, the technical ecosystem may not be laid out in a way that allows the techniques to operate; and enterprises may lack the processes to gain access to data and make use of the intelligence they collect as a result of the application of these methods. These questions are important to ask because investments made before readiness is fully achieved may be inefficient, suboptimal in terms of the results they produce or, as a worst case, represent needless expense.

With respect to the people portion of the capability triad, new skill sets may be required to support emerging analytics capabilities. While many vendors are actively seeking to "democratize" big data analytics, i.e., developing tools to render powerful broad-scope analytics capabilities that are accessible to enterprises without specialized expertise, the short term necessitates specialized skills in enterprises that are seeking to leverage these data rapidly. If an enterprise does not have these skills already in-house (the normative case), it may need to hire new personnel with the requisite skills. In the short-to-intermediate term, these resources may come at a premium cost due to demand for those skills. For example, The Harvard Business Review cites "Data Scientist" (a specialized role with a hybridized blend of technical and statistical skills) as the "Sexiest Job of the 21st Century"[9] with correspondingly higher compensation as a result. These already-scarce and in-demand skills are likely to remain challenging to acquire for the near future. A recent survey from The Data Warehousing Institute (TDWI) cites "inadequate staffing or skills for big data analytics" as the current top barrier for implementation of big data analytics in enterprises.[10] Demand will eventually slacken as personnel specialize and equilibrium is reached, but, until then, these personnel may be harder to find and more expensive to maintain.

[9] Davenport, Thomas H.; D. J. Patil; "Data Scientist: The Sexiest Job of the 21st Century," Harvard Business Review, October 2013, *http://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century/*

[10] Russom, Philip; "TDWI Best Practices Report: Big Data Analytics Fourth Quarter 2011," tdwi.org, 2011, *http://tdwi.org/research/list/tdwi-best-practices-reports.aspx*

From a process standpoint, existing silos should be evaluated to determine whether individual business units, departments and personnel are willing and able to share information and act on information received. This consensus needs to happen so that analysis can be performed (disparate sources of data may need to be consolidated to operate on them) and so that the derived conclusions can be put to productive use. Enterprises need to consider that these areas may not share information currently and may have a history (depending on the culture) of competitiveness, antagonism, or resistance to outside influence. These cultural barriers can impede open and collaborative exchange of important data elements and act as a barrier to adaptation in response to conclusions drawn. This consensus among silos can extend beyond the department level and down to the level of individual personnel. For example, key stakeholders may not know precisely where key data elements reside or how to access those data elements. Likewise, "Shadow IT" (technology adopted without the direct oversight or, in many cases, awareness of the IT organization) can complicate information sharing because technology adopted without centralized oversight may represent a significant repository of critical information, and lack of central awareness of the information may limit the ability to include it in the scope of analysis. Because of these factors, some degree of organizational self-awareness is required to think through supporting processes and identify potential problem areas before enterprises undertake significant investment in big data analytics.

Lastly, the technology implementation plays a role in determining organizational readiness. In many cases, new tools are required to support the analysis to be conducted, and capabilities for data storage and computation may need to be evaluated to ensure sufficiency. Moreover, sufficient data on which to operate need to exist and be accessible to analysts. Data sources must be identified, which involves locating structured data (e.g., data organized in a relational database) and unstructured data (e.g., data stored *ad hoc* on a file system or in a loose collection.) Identifying data sources can likewise involve data in a variety of different formats, including video, audio, images and text. Computational resources may need to be expanded to enable operation and analysis of these data.

COMPARING RISK
OF ADOPTING TO RISK OF NOT ADOPTING

Information security and audit practitioners who are evaluating a big data analytics initiative in their enterprise need to weigh the management and mitigation of the technology risk of adoption against the business risk to the enterprise if they choose to not adopt big data analytics.

Technical and operational risk should consider that certain data elements may be governed by regulatory or contractual requirements and that data elements may need to be centralized in one place (or at least be accessible centrally) so that the data can be analyzed. In some cases, this centralization can compound technical risk.

For example:

- **Amplified technical impact**—If an unauthorized user were to gain access to centralized repositories, it puts the entirety of those data in jeopardy rather than a subset of the data.
- **Privacy (data collection)**—Analytics techniques can impact privacy; for example, individuals whose data are being analyzed may feel that revealed information about them is overly intrusive.
- **Privacy (re-identification)**—Likewise, when data are aggregated, semi-anonymous information or information that is not individually identifiable information might become non-anonymous or identifiable in the process.

These risk areas can cause some practitioners to be understandably wary. However, analytics efforts can be used to offset risk by applying the tools and techniques to security-event information, transaction information for the purpose of detecting fraud or to other technical information for risk-reduction purposes. Stockpiles of security-relevant information, such as user and system activity, can be logged and examined the same way as more business-facing data can be logged and examined. The same analytics techniques and tools that streamline and increase the quality of business processes can likewise streamline and increase the quality of other risk-mitigation processes. Tools purchased and analytics techniques that are acquired to help enable business-facing efforts can, with planning, be adopted by information security and risk management areas to help advance their goals as well.

COMPARING RISK
OF ADOPTING TO RISK OF NOT ADOPTING

The security and audit practitioners' consideration of risk can and should be holistic.

Data analytics techniques can achieve a compelling competitive advantage to the business, which should not be discounted when evaluating the possibility of new security or privacy risk that could occur.

Said another way, if an enterprise elects not to employ these techniques, there is a risk to the business, because competitors will capitalize on the opportunity. This result could have ramifications just as serious to the enterprise as a security or privacy breach or dreaded business continuity implications. A holistic view of the risk in an enterprise should seek to account for all sides of the risk equation, including the following:

- Business value of adoption
- Business risk of non-adoption
- Technical/security/privacy risk that may increase depending on the implementation used to support the big data analytics approaches at the technical level
- Possible risk-offsetting benefits of the technology at the technical level

CONCLUSION

A number of business dynamics make the application of new and better analytics appealing to enterprises.

By looking at how these analytics techniques are transforming enterprises in real-world scenarios, the value becomes apparent as enterprises start to realize dramatic gains in the efficiency, efficacy and performance of mission-critical business processes.

The business case is made even more compelling by the fact that most enterprises already have in place the foundation for analysis in the form of more data than they can currently use productively. Most enterprises already retain a large amount of data, such as information about their customers, metrics about the performance of internal business processes, data about information systems and their technology ecosystem, transactional information about sales and marketing and numerous other data items about how they do business.

Understanding this business case can help security, audit and governance practitioners in two ways: It helps them to understand the motivation and rationale driving their business partners who want to apply big data analytics techniques within their enterprises, and it helps balance the risk equation so that technical risk and business risk are addressed. Specifically, while some new areas of technical risk may arise as a result of more voluminous and concentrated data, the business consequences of not adopting big data analytics may outweigh the technology risk.