DataInformed

# The Unrealized Promise of Analytics

## (and How Data Governance Can Help)

Sponsored By:

collibra™

# Table of Contents:

# Expert Q&A: How Data Governance Helps You Succeed with Modern Analytics

Stan Christiaens, co-founder and CTO, Collibra

Data is the currency of the digital age, and most organizations understand the imperative of leveraging their data to drive business growth. But several considerations impact the success of data management and business analytics efforts. In this Q&A, Stan Christiaens, co-founder and CTO at Collibra, shares his expert insights on how strategic data governance strategies can ensure data analytics and digital business success.

### 1. What are the requirements for modern businesses to succeed with analytics?

Many organizations have discovered that succeeding with analytics isn't easy. But in talking with our customers and prospects, I've found that the most successful analytics projects have four things in common.

1. **Look beyond the analytics tool:** Investing in an analytics or visualization tool is important, but the tool alone isn't enough to make you successful. For the tool to do its job, you need to ensure that the people using it can access all the data they need from across your organization. And the reality is that data lives – and will continue to live – in many systems and departments. By bringing together data from the far-reaching corners of your organization, users can leverage analytics to uncover insights and spark new ideas that wouldn't reveal themselves if the data remained siloed. Simply put, it prevents the scope of your analytics project from remaining locked into a specific tool or system. By bringing your data together and making it accessible to the people who need it, you gain greater value from your analytics initiative.

2. **Embrace automated analysis:** A processing capability that performs automated analysis of the data takes analytics beyond the classic report-writing scenario. It augments human analysis by using advanced technology to spot trends, highlight outliers, predict outcomes, and detect patterns across the various data sets that exist within your organization. This type of analysis combines data in ways that reveal hidden insights. And the output from this automation feeds reports so that everyone across the business can use this comprehensive analysis to drive business decisions.

3. **Evaluate your human capital:** You probably have people on staff who can bring together data from various systems into a central repository. But do you have the right people onboard to run the automated analysis I mentioned above? Maybe not. And as data continues to grow, the scale changes, and so do the data engineering skills required to operate the data machinery. You may need to invest in people with different skillsets to make analytics truly successful.

4. **Make a cultural shift:** Many organizations that say they are data-driven are not. And the number one reason why they are not data-driven is because they are not treating data as a strategic business asset. When you have a data-driven culture, data – and good governance of your data – underpins everything you do. The people in your organization look at data as something of value. They own it and protect it. Being data driven also means you need to accept that the answers you're looking for in your data may not exist. And that your first several attempts at unlocking hidden insight in your data may fail. When this happens (and it most likely will), you may need to take a step back, look at a different set of algorithms, or find the missing data. Finally, to embrace a truly data-driven culture, your organization needs to have a process in place to act on the results that your analysis uncovers. Don't let the results of the analysis sit on a chart in a PowerPoint presentation. Use them to drive business transformation, optimize processes, attack a new market, or create a new product.

### 2. How can organizations leverage data governance to improve their analytics?

Data governance is an important element in any successful analytics initiative. Without data governance, analytics users struggle to find, understand, and trust the data. And when users don't know where the data lives, what it means, or if it's right, they lose confidence in their analysis, or worse, present conflicting analyses that spark a data brawl. For your organization to realize the value that self-service analytics provides, your users must first know where they can find the data they need and how to access it. They also need to understand the data, including what each data source means and how the various groups and departments across the business are using it. They must understand the business context of data that's generated within the various systems and devices across your organization. And they need to understand the rules and policies surrounding the data, such as who can use it and for what purposes. Finally, they need to trust the data. They need to have confidence that the data they are using in their analysis is right. And this element of trust becomes even more critical when we look at automated analysis. As the number of machine learning devices that we use increases, so does our need for data confidence. Users need to know they can trust the outcomes of the analysis, even if the results are unexpected.

### 3. How does data governance change in the age of self-service analytics? What new challenges does it pose?

In modern business, it's no longer sufficient for a group of data scientists to handle analysis for the organization. Instead, businesses must empower all users – the data citizens – to find data and use it to move the business forward. And as the number of data sources – and the

number of people using the data – grows, so do the challenges associated with the data. Clearly, traditional data governance that focused on a specific system or department is no longer enough. In the age of self-service analytics, organizations must ensure that their data governance initiative spans the enterprise. It must cross departmental silos so that every data citizen can access the data that's available, know that the data they are using is accurate, and trust that the results they are generating are right. And when data governance is truly enterprise-wide, it fosters a collaborative environment that enables users to proactively tend to the data, rather than react when an analysis returns a negative outcome.

Implementing true data governance to support your self-service analytics initiative can be a challenge. But at its core, data governance is about enablement: enabling the business with data and making sure the right controls are in place. This is especially true in the case of self-service analytics. You need to listen to the needs of the business. For example:

- What data challenges are they struggling with?
- What data element or data sets are they not getting access to?
- Where are they losing trust or time?

Once you've answered these questions, then think about how data governance can ease these pains and enable smooth data use. And when you've determined which pains you can ease, then the best advice I can offer is to just get started. Focus your energy on making the initial areas of focus successful. Then, use this foundation to expand your efforts to other areas of the business. Before you know it, data governance will become business as usual for all your data citizens.

### 4. What data governance structures, organizations, platforms, and processes are required to realize the value of your investment in analytics?

To realize the value of your analytics investment, your organization must first look at data through the eyes of the business. It's no longer about storing and locking down the data. Today, it's about opening up the data and making it accessible to all data citizens. To do this, you should elevate data to its own department, with a dedicated leader charged with controlling the data chaos that likely exists across the business. For many organizations, this leader is the Chief Data Officer (CDO).

Second, you should invest in a governance structure and process supported by a data system of record. A system of record doesn't simply store all the data in one place. Rather, it connects all the data throughout your organization in a way that is meaningful for all data citizens. And it drives greater value for the business by integrating initiatives and projects across the organization. When it comes to analytics, the real value of a system of record is that it provides a holistic view of the data landscape across the business, not one that focuses specifically on a specific department or system. That way, data citizens can access all the data they need, resulting in greater insight that propels the business forward.

A system of record for your data also embodies all the processes that underlie the data. It gives data citizens a process for requesting access to data, for approving data, and for making changes to the data. And it empowers all data citizens to identify and fix bad data. It also keeps everyone who needs to know about the data in the loop, so a small change to one data element doesn't cause heated discussions about whose data is right.

### 5. What best practices should organizations follow to implement effective data governance policies in the digital age?

Getting started with data governance can be a challenge. But I've found that the most successful and effective data governance programs have a few things in common.

1. **Start with one step:** We've all heard the phrase "don't boil the ocean," and that's certainly true when it comes to data governance. The best data governance initiatives start with a small scope and expand once they've proven value. My advice is to pick the area or use case with the greatest need. Analytics is a common place to start. Work with the business owners for that area, and make them successful. Then use your experience to make the case for expanding to other use cases within the business.
2. **Let the business lead:** The most common scenario for a successful governance initiative is one where the business leads the effort. That way, the business is invested in the project, and is more likely to take ownership and drive adoption throughout the organization. And just to be clear, the most successful data governance initiatives do not just involve the business, but rather put them in the driver's seat to lead the project.
3. **Drive the project from the top down:** By establishing what's important at the top, your organization can create a governance model that aligns with organizational priorities and expectations.
4. **Invest in the right people and platforms:** When your organization is ready to make the investment in data, you must also make the investment in people and platforms. A Chief Data Officer, combined with an enterprise-wide data governance platform that serves as the system of record for your data, is a proven recipe for success.

### 6. What future developments will change the way we use data?

The future is already here, but it's going to continue to evolve over the years to come. The expectation that organizations will use data will continue to grow. And the need to become data-driven will become more urgent than ever before. The Internet of Things, for example, is moving from the halls of academia to become a part of our everyday lives. Think about it. We have thermostats that can connect to the internet. They collect information about you and apply it to do things like automatically enable and disable the thermostat based on your behavior or the weather. Simple examples like this one show the vast amounts of data that will become available for organizations to analyze. It makes the "big data" we've been talking about "even bigger data." And as IoT continues to move into the mainstream, the data it generates will become part of the pool of data users access for self-service analytics.

Mobile is also changing the way we use data. If we think back to five years ago, we had mobile devices and even smart devices. But not with the level of sophistication that devices have today. And that sophistication will continue to grow. Data is literally in people's hands all the time. And it's more than just housed in a cool app. It's readily available – and usable – right from their mobile device. That means that analytics can happen anywhere, not just in front of your computer in the office. And it's this access-everywhere mentality that will continue to drive the transformation to data-driven business.

### 7. How should organizations plan to adapt their data governance processes for the future?

When you mention data governance, many people still associate it with control: data policing, lack of agility, slow change, and more. But an often overlooked, but critically important, part of data governance is enablement. See, enablement is the part of the data governance process that creates peace of mind when it comes to data. And organizations should adapt and ensure that their data governance processes embrace enablement. One way to do this is through crowdsourcing data governance.

Now, do I believe that all organizations should go full crowdsource? No, just as I do not believe that the opposite is the solution. Every organization (and in more detail, every department, line of business, project team, etc.) has its own context: culture, maturity, priorities, even political factors or tight regulations. There is no single way of doing data governance. In some contexts, crowdsourcing will produce fantastic results, and in others it will be a failure. Crowdsourcing is, simply put, the most unlimited side of a control spectrum. And you need to figure out your ideal place on that spectrum.

This is why I am absolutely convinced that a data governance platform that aims to be successful needs a capability for operating model configuration: your roles, responsibilities, workflows, dashboards, views, use cases, and more. And because organizations are not carved from a single slab of stone, but rather cling together like parts of a living organism, the configuration needs to support federation; some parts of the business (e.g., analytics) will require crowdsourcing, while others live in a stricter world and require tight controls (e.g., risk and compliance).

As you grow, mature, and adapt your data governance processes, I urge you to determine where you fall on the crowdsourcing spectrum. ■

Stan Christiaens is the co-founder and CTO at Collibra and leads the global product organization. He's responsible for product management and UX, the Collibra Center of Excellence, and Collibra University, the Collibra online learning platform. Prior to founding the company, he was a senior researcher at the Vrije Universiteit of Brussels, a leading semantic research center in Europe, performing application-oriented research in semantics.

# Demystifying Data Governance: What it Is and What it's Not

*by Christophe Marcant*

Christophe Marcant, Vice President of Product Strategy, Stibo Systems

High-quality product and customer data are the lifeblood of any business. Enterprise systems, including ERP, supply chain management, CRM, and e-commerce all rely on complete, accurate, and consistent data to drive operational performance, sales conversion, and customer service. But in most organizations today, data quality remains a major challenge. Gartner pegged the average cost of poor data quality to an organization at $8.2 million a year, with 22 percent of respondents calculating their annual costs at $20 million or more. In a similar study, Aberdeen Research Group reported that "best-in-class" organizations were "three times more likely than other organizations to adopt data quality tools" and that those tools led to significantly more accurate and usable information. One thing is clear: Integrating product and customer data from a wide variety of different formats and data structures without considering data quality not only is costly but also can have a disastrous impact on the enterprise systems that rely on that data.

Poor quality data can affect nearly every area of a business. For example, goods can be delivered to the wrong address or customers can receive goods that don't match the description on the Web. However, there are other, more subtle effects of poor data quality that are far more damaging to the business: missed opportunities to upsell to a customer, not being able to negotiate purchasing discounts, and losing web sales because of inaccurate sizing data are just a few examples. In fact, Gartner reported that, "Information governance and master data management programs are central to an organization's success in assuring business outcomes and increasing business value from reusing enterprise information assets."

Companies are addressing the need for improved data quality by enabling line-of-business professionals to manage their business information proactively using a combination of Master Data Management (MDM) solutions with sound data governance processes. This creates a positive ripple effect and a distinct competitive advantage in all downstream systems that rely on complete, high quality, and timely master data.

Data governance continues to be a hot topic these days but, despite all of the articles and whitepapers dedicated to the subject, it seems that business leaders still are not clear on what data governance actually entails. Some of the confusion may revolve around the word "governance" itself. Instead of getting hung up on the word "governance," one way to consider data governance is to think in terms of quality, which is the fundamental aim of any data governance initiative and exactly where business leaders should be focusing their attention.

Formalizing your governance is a process that's often overlooked, as business users often perceive that other members of the organization are responsible for assuring data quality. Much of an organization's operational data already are part of an active management process but, to a large extent, the focus is on quantities and values. Areas that tend to be overlooked the most involve the reference or master data that drive many of the organization's actual business processes. Data governance aims to correct this by establishing formal management responsibilities for the quality of this data.

The key to establishing a solid data governance foundation is to shift from a reactive approach to a proactive approach. It's common to adopt data governance after poor data quality results in a bad business outcome or when no one takes responsibility for an error. Having a formalized, proactive data governance approach ensures that somebody is clearly responsible not only for fixing the disasters but also for reducing the likelihood of one occurring.

Before embarking on a data governance initiative, it is important to understand the options and some common areas of confusion:

**Are off-the-shelf data governance tools available?** Many tool vendors offer data governance solutions, and there are certainly tools that can help you govern data. These include tools that can enable you to store and communicate defined business rules as well as tools to measure data quality, identify compliance issues, etc. However, governance is really about the organization and the processes and responsibilities within which such tools can be deployed. For instance, a solution may support full lifecycle control of data, metadata management, data quality rules, and monitoring, but without the correct organizational support, the benefits of these governance tools will not be realized.

**Is data governance the same as data maintenance?** The two are very closely linked through data quality, but they are actually independent functions. Maintenance organizations tend to be aligned with specific IT systems or with specific business units within the organization, whereas data governance is about a common set of rules to which business members should adhere. The key to understanding the difference is to understand the two parties' relationship to "standards."

As part of a data governance effort, an organization should create standards by defining a set of best-practices or principles that will ensure the organization creates and maintains good quality data. It is the role of the data maintenance teams to comply with these standards, but it is the role of data governance to define the standards and to ensure that they are being met.

**What exactly is data ownership?** Data ownership can be a very confusing term. For example, it is common for businesses to split data responsibility according to geography. For instance, members of the UK sales force may manage all customer data based out of UK, whereas the U.S. team takes responsibility for those data in the States. However, the optimal approach is to create a single group within the organization to be responsible for all customer data instead of having siloed, single "data owners."

The term "data owner" is actually a misnomer because, in practice, what is owned is not the data but the standards that guide users in how to achieve good quality. So while many departments may lay claim to the contents of the data, it is the data governance group itself that owns the structures and the quality rules.

**What constitutes a data-governance–oriented organization?** When viewed at a high level, data governance professionals perform two activities. But, in practice, these two activities can be very complex and can require a network of resources to achieve them. Specifically, the data governance team is responsible for the following:

- **Change Management.** Once the organization has defined a set of standards and aligned its data to it, it is important that any changes to these standards be controlled. For example, if the company defined that all dates are stored in the UK format of Day/Month/Year, then it will be problematic if somebody wanted to change to the American format of Month/Day/Year. It is the job of the data governance team to assess the impact of any such change, confer with any relevant stakeholders, measure the costs and benefits of such a proposal, and then – if the change is deemed appropriate – to manage those changes across all affected areas of the business.
- **Compliance**. Wherever there are rules, there is a requirement for policing. It is the role of data governance to be that police force – to measure the organization's compliance to any standards that it governs and to act to improve the level of that compliance.

Data governance is a methodology to exercise data control processes. It gets your team on the same page regarding data-quality issues and limits confusion and wasted time. Ensure that your data assets and MDM efforts don't go to waste by taking the time to institute the proper standards and processes. ■

Christophe Marcant is Vice President of Product Strategy at Stibo Systems, a global leader in multidomain Master Data Management (MDM) solutions. Contact him at chrm@stibosystems.com.

# How to Make Your BI and Analytics Dreams Come True

*by Chris Cooper and Simon Hankinson*

Chris Cooper, Director, Healthcare, Collibra

Simon Hankinson, Market Manager, Financial Services, Collibra

You've sold the vision. By becoming a self-service, data-driven organization, you'll increase revenue, beat the competition, and drive innovation across the business. You've purchased and deployed the latest data warehousing appliances and data management suites. You've spent months integrating data from the far corners of your enterprise. You've selected the hottest self-service business intelligence tools on the market. You've trained hundreds – or maybe thousands - of users, stood up centers of excellence, presented your pilot program results to executive management and the board. Your analytics program is live, and innovative ideas are waiting to reveal themselves in your data.

Or maybe not.

This story is all too familiar in today's business environment. Organizations worldwide are buying into the hype that business intelligence (BI) tools – in particular, self-service BI and analytics – will solve the many challenges they're facing in the market. And while the tools themselves are simple to deploy and easy to use, several challenges remain.

# The Unrealized Promise of BI & Analytics

## Data Integration is Difficult

Bringing together data from the far corners of the organization is no easy task. Not only is it time-consuming, but it's also resource-intensive. Think about where your data lives. The reality is that data lives – and will continue to live – in systems across your organization. And even though data integration is easier today than in the past, it's still no small task. To effectively integrate your data, you need a complete understanding of the organization's "data map" and the data's journey and relationship to other similar – and contradictory – data throughout the organization.

## Data Definitions are Inconsistent

If your organization is like most, there are multiple data definitions in use across the business. The most common example is the term "customer." In Finance, the definition of "customer" could mean an organization that pays for goods and/or services from your organization. In Sales, "customer" could mean any business that may purchase goods and/or services from your company, either today or in the future. These simple differences can wreak havoc on data analysis, especially when the results vary widely. And in extreme cases, these inconsistent definitions cause data brawls, where departments fight with one another about whose data is right. By rationalizing definitions across the business, you can prevent the data brawls from erupting and focus instead on the analysis at hand.

## The Data is Messy

Data quality issues are commonplace in today's business environment. But high quality data is key to making your analytics and BI initiative a success. If your users question the quality and trustworthiness of the data, then they are unlikely to use the data – or the shiny new BI tools. Your data will never be perfect, but ensuring that it's (relatively) clean and trustworthy is job #1 when it comes to driving successful BI initiatives. Because when users know they can trust the data, they are more likely to use it.

## The Data is Lost in a Data Swamp

As the number of systems in your organization proliferates, so does the amount of data available. And as the volume of data continues to grow, the question shifts from "how do I access the data?" to, more simply, "where do I even find the data?" Many organizations try to solve this issue by building more reports and dashboards. But this is not the solution to this problem. In fact, it makes the problem even worse by adding more data to the pile. Instead, organizations must focus on helping their users find the right data so they can use the BI tools more effectively.

## Data is More Important than the Tool

BI and analytics tools are only as good as the data they analyze. If your users can't find the data or know it's trustworthy, then the potential of the BI tool remains unrealized.

While each of these challenges is significant, they all lead back to one fundamental issue: the lack of data governance. By implementing enterprise-wide data governance, data citizens across your organization will be able to:

- Find the data they need
- Understand what it means
- Trust that it is right

And that's when the BI magic begins. See, when users can find, understand, and trust their data, then they become data citizens. They use data to drive better decisions, to spark innovation, and to gain competitive edge. Certified reports and watermarking, all made possible through data governance, give users confidence in their analysis. Now, their BI and analytics tools become the vessels that lead them to insights that can truly transform the business.

# Data Governance Industry Perspectives

## Data Governance in Healthcare

We all know healthcare data is complex. And many times, it's fraught with quality issues. The complexities of different clinical terminologies (ICD-10, SNOMED, LOINC, CPT) make simple things, like defining all the patients with a neurologic-related diagnosis, labor intensive and error prone. Even when tools are used to profile data in order to assess data quality, results are often abstract and disconnected from how that data is used in a report or analysis. In fact, many self-service analytics users are left without any real gauge of the quality of the data they might want to incorporate into their analysis.

Users know that not all data is of the same quality, completeness, and accuracy. But when you properly govern your data, you can increase users' confidence and trust that they are using the best data for their needs. And when the data is trusted, the use of analytics increases – and patient outcomes improve.

## Data Governance in Financial Services

As banks continue to work diligently to build sustainable data governance to meet regulatory requirements such as BCBS 239, CCAR, and others, many of them are finding that a mind shift is required. It's no longer sufficient to think in terms of data management. For banks to ensure that they can efficiently sustain regulatory compliance now and in the future, they need to shift their thinking to focus on data governance.

Unfortunately, the reality for many financial institutions is they often lack two critical components for sustainable data governance: automated operating models and workflows, and reporting. First, data governance is often a largely manual process relying on Excel, SharePoint, and email. Processes use disparate technical tools to document data management, but lack the ability to connect these tools together and put the information in the hands of the groups that need them: report owners and data stewards. Second, financial institutions need reporting dashboards and analytics to demonstrate progress and, ultimately, compliance. Simply put, they need to prove that data governance is in place and under control. And their inability to report on data governance and prove data quality makes demonstrating compliance challenging at best.

When financial services institutions implement data governance practices, they do so in order that users, both internal and external, can have confidence in their data. And, they can use their data to prove the trustworthiness of reports. Now, financial services organizations can create a truly sustainable process to meet regulatory requirements today – and tomorrow. And with a foundation in data governance and trustworthy data, they are many steps closer to sustainable regulatory compliance.

## Getting Started with Data Governance

While many organizations want to get started with data governance, they often struggle with how best to do it. The answer is simple: just get started. First, pick one or two stewardship applications. Focus your resources on making them successful. That's how you'll prove value. Then, use this foundation to expand your efforts. Look at other data projects across the organization, and identify their pain. Give them the governance processes they need to be successful. And continue these steps until data governance is business as usual.

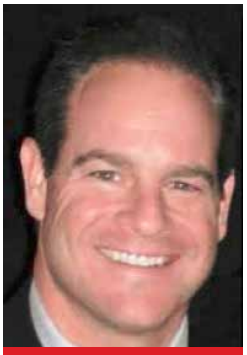
Realizing the promise of BI and analytics is no simple task. But data governance is clearly the secret to making the vision of BI and analytics a reality. So go ahead. Take the first step on your journey today. ■

**Chris Cooper** | After almost 15 years in healthcare data management with IBM, Oracle, and Informatica, Chris Cooper joined the Collibra team to help healthcare organizations get more out of their data-driven initiatives by focusing on data governance as a tool for engaging the business and clinical organizations. Chris is responsible for the Collibra Healthcare Sales and Strategy in the United States — a mission that aligns with his core belief that data and transparency are critical to providing the best healthcare.

**Simon Hankinson** | During his 20+ years of management consulting services experience at leading consultancies EY and PwC, Simon Hankinson helped financial services institutions design and implement information and data governance processes to support internal, external, and regulatory reporting including BCBS 239 and CCAR. Simon joined Collibra as the Global Financial Services Market Lead.

# Making Sense from Nonsense with Data Governance

*by Scott H. Schlesinger*



Scott H. Schlesinger, Principal, IT Advisory, Ernst & Young LLP

More and more organizations are seeking guidance to help them understand how to harness the power of big data. These organizations want to know how to efficiently manage massive volumes of complex (structured, semi-structured, and unstructured) data and how to leverage analytics (predictive, social, mobile) to gain better insights to improve operational efficiency, increase profits, and gain a better competitive position in the marketplace. Organizations across nearly every sector are looking for ways to tap into data that was previously trapped in unstructured sources, such as text documents, email, and social media sites, and leverage this previously untapped data to turn poor business decisions made using haphazard guesswork into well-considered and successful business decisions that improve overall performance.

Yet, while forward-looking insight is the ultimate goal, organizations must understand the data they collect and store. This data has to be identified, acquired, organized, filtered, and cleansed, then integrated and stored before it offers real value to the end business consumer. This basic "blocking and tackling," which I have preached for years as a business intelligence professional, is a critical underpinning for any future analytics initiative. This begins with a clearly defined and agreed-upon data governance strategy.

## What is Data Governance?

Data governance is the process of creating and agreeing to standards and requirements for the collection, identification, storage, and use of data. This should not be viewed as optional with any data-driven project. A data-governance initiative helps the organization set rules, policies, standards, and procedures, and define roles and responsibilities with respect to the overall management of data. Effective data governance allows for the efficient integration of new data sources and helps the organization realize value from this data – such as the ability to aggressively pursue new market opportunities and identify and capitalize on emerging business opportunities. Other key benefits of an effective data governance program include – but are not limited to – reduced cost for data storage, reduced cost for

rework often associated with poor data quality, increased confidence in data quality (often a result of improved data consistency), improved performance of technology solutions, and enhanced data security.

Data governance has not changed dramatically over the years, but the types and volume of data being collected, stored, and used for analytics has changed significantly. Big data governance requires governance over many different types of data (including metadata, or data about the data), not just what's in the legacy systems or relational databases. This requires a new understanding of the methods, processes, and tools that must to be deployed to deal with this big data. Organizations spend a great deal of time, money, and human capital developing big data programs that involve implementing data management solutions and ironing out the integration processes needed to tie it all together.

It is estimated that 80 percent of the data being created today is unstructured. This only exacerbates the issues as organizations begin to tackle their big data challenges around weaving together structured, semi-structured, and unstructured information. While data can provide game-changing insights to run the business, key leaders are left wondering what data they have, what data is needed to answer key organizational challenges, how the data should be integrated, where the data should be stored, what processes should be in place to assure data accuracy and security, and who is allowed to view the data. This is where a robust data governance plan can help and create order from chaos.

A data governance program must be in place at the outset of any large-scale technology project so that the resulting insights can be trusted to help the organization achieve value from the investment being made. This will ensure the organization gets the right people involved, can define and adjust the relevant processes for data management, and deploys the right technology solutions to address the needs of the business and the complexity and volume of data being managed. Not having a data governance program in place can result in misalignment between the data and overall business strategy and reduce trust between business and IT due to a fundamental lack of trust in the information being provided.

One of the most important aspects of data governance is alignment between IT and business. Having clearly defined roles and responsibilities and objectives understood from the outset is paramount to the success of any governance initiative. Agreement on the use case for this type of initiative and business value that is being sought will help gain support and funding for the initiative within the organization.

As for initial tactical steps that an organization should take, that is largely dependent on the organization's data maturity, ability to manage data, and understanding of the objective(s) of the key consumers of the data. Most data-governance programs begin with agreement between IT and business on the business needs/goals/expectations, the approach to data management, the scope of the work, the business and technology requirements, the design of the technology landscape, and the deployment of assets to realize/implement the solution.

Data-driven organizations of all sizes, across all industry sectors, are seeking better ways to manage the growing volume of data and create and disseminate relevant, accurate, trusted, and timely information. In this era of big data, an effective data governance program makes this possible. When governance is layered over the framework of a data analytics platform, the result is a holistic understanding that improves the company's ability to manage and measure ROI. The more that companies invest in strategic analytics, the more they will need a robust data-governance plan to extract results from their data and ensure immediate and long-term success. ■

Scott Schlesinger is a Principal within Ernst & Young LLP's National IT Advisory Practice, and serves as EY's Americas Leader for Business Intelligence and Information Management.

# The Secret to (Finally) Making Your Analytics Project a Success

*by Dan Sholler*

Dan Sholler, Director of Product Marketing, Collibra

A recent Gartner survey of CDOs showed that their primary responsibility is analytics. Yet users report that most of these analytics initiatives have not lived up to their promise. This "analytics gap" has several causes, but it is exacerbated by the inability of data scientists and other data professionals to find the right data. Giving all users – the data citizens – the ability to find the right data for analytics is top priority and should be one of the primary functions of any data governance initiative.

Many organizations are drowning in data, which makes finding the right data difficult. And even if you know you have the data, pulling the bits of data you need out of the huge collections that live across your organization is a daunting task. Unfortunately, there is no magic bullet, or magic machine learning algorithm, that can do this for you. While algorithmic approaches hold a lot of promise, they are just in their infancy today and should be thought of as a means of assisting subject matter experts. There is good news, though. With the proper infrastructure, you can harness the expertise of data citizens around the organization to organize your data, and then use that infrastructure to enable them to share their work.

## Stop Searching for Data

Creating a data catalog that organizes useful collections of data across existing boundaries is first step to making data discovery easier. Whether those boundaries are systems, organizations, or geographies, it is the cross-boundary visibility that drives many of the more significant insights from that data. The data catalog should help experts organize data in three ways:

1. **Provide an assisted mechanism to link data to meaning:** The business terms, rules, processes, KPIs, etc. that are related to data are the real pieces of information that are used to determine whether that data is a good fit for the analysis in question. And different views of the data provide different aspects that may feed that analysis. For example, understanding the subtle differences between the backend financial transaction view and the website interaction view can give vital clues to buying behavior. Each of these things reflect the same activity, but they have different meanings. And it is those differences that help create accurate predictions.

2. **Suggest using the work of your peers who may be working with similar analyses:** It is often the case that several people are working on different aspects of the same problem and may be able to leverage each other's datasets. The data catalog should recognize when the work you are doing is similar to your colleague's work and point you to datasets that they have already created. This suggestion will simplify the process of finding relevant and comparable data.

3. **Suggest other data that might be related to data you have organized into a data set:** This recommendation will help speed the process of determining what data to include. This is challenging, as there are many ways of determining this kind of similarity. But at the very least, it should determine when you have either structural or documented semantic relationships.

## Foster Trust in the Data

Discovering and organizing data cannot occur in a vacuum. You – and other data citizens - need to know that a formal data governance is behind it. Otherwise, you will not trust the data that you find. You need to know and understand its lineage, its quality, and the organizational responsibilities for that data so you can use it with confidence.

Unfortunately, many of the data catalogs that exist today do not have these capabilities. They present lists of information but lack ties with the governance of the information. Without these ties, it's difficult for you to distinguish between all the various copies and states of the data. And you have no idea which collections of data adhere to which policies, which have what quality, and other factors. You also cannot determine both how the data is used today and how it is formally intended to be used, which are often different things.

Furthermore, most people do not trust the quality of their data. Having the quantitative results in an understandable format helps you evaluate quality so you can determine the data's suitability for use in the data set. Also, many people naturally have a very hard time trusting something that is entirely controlled elsewhere. Having a direct connection to a data helpdesk is crucial so that all data users understand that they can fix any data problems they find.

These links to governance are not optional features of a catalog, but rather are essential to its proper functioning. Without them, the catalog might help you group, organize, or share. But it will not assist the development of new and useful analytics because it will lack the trust of the consumers of the data.

As your organization builds out its analytical capabilities, creating a data catalog is a powerful step. To gain value from this catalog, it needs to support the three capabilities that will truly help users and developers of analytics find the data. It also must be fully integrated into the data governance process and exhibit all of the governance capabilities (policies, quality, lineage, usage traceability, and repair) that create trust. Only in this way will you be able to get out of the analytics gap and deliver true insight to your organization. ■

Dan Sholler is an experienced software industry expert, with broad and deep experience in the data and software markets. Dan began his career as a developer and product manager for BI and reporting software, moved into integration and middleware. He spent several years at Gartner as an industry analyst in the software space, covering data management, middleware, application architecture, and SAP. He has spent time at various software companies, and joined Collibra as Director of Product Marketing.

# How Business and IT Can Find Middle Ground for a Data Governance Framework

*by Vijay Anand*

Vijay Anand, Senior Director of Product Marketing, MicroStrategy

Every day, I'm tasked with helping some of the largest organizations in the world unlock the value of their data. Their goal is to interpret large amounts of information and improve decision making in order to cut costs and identify new opportunities. According to the new Worldwide Semiannual Big Data and Analytics Spending Guide from research firm International Data Corporation, worldwide revenues for business analytics will grow to nearly $187 billion by 2019. Impressive, right? But in delving into that prediction, I wonder if organizations are as prepared as they need to be to begin collecting, analyzing, and acting on their data.

As businesses eagerly increase investment into initiatives centered around business intelligence, one task that is often overlooked is identifying someone who is responsible for keeping millions or billions of data points honest, organized, and insightful. After all, what good is that information if it can't be trusted and verified?

Different teams within an organization—despite sharing a common awareness around just how valuable company data can be—have different priorities, commitments, and agendas when it comes to that data. This is especially true for IT and the lines of business, the two key players in this space. Traditionally, the main business intelligence prerogative for IT was to focus on managing the data behind the reports. Business teams simply called on that information when relevant. However, as access to data became accessible to more employees both within the IT and business divisions and outside of them, it became increasingly possible to manipulate data of any kind.

## The Risk of Contaminated Data

Contaminated data is a dangerous realization for an IT or business department. Complications can originate due to information ownership, data collection processes, or technology standardization (or the lack of it). These inconsistencies often rapidly multiply and result in contaminated data where users unknowingly introduce unverified information and, worse, proceed to share it with others.

An employee probably doesn't have the malicious intent of contaminating your company's data accuracy, but a lack of technical training could lead to error without even knowing it happened. The ripple effects of such an occurrence can be devastating. Contaminated data can lead to excessive consumption of company resources, increased maintenance costs from a technical standpoint, and distorted results that end with bad and painful decisions. Reverse engineering a problem to sort through irrelevant, out-of-date, or erroneous industry data is tedious, takes up valuable time, and lets the competition get ahead.

## The Solution? Implement a Governance Framework

These situations can be avoided with a dedicated and unanimous nod towards data governance. A governance framework sets the parameters for data management and usage, creates guided processes for resolving data issues, and enables businesses to make decisions based on high-quality data and well-managed information. It's essential and, more often than not, a must-have for any organization that looks to pull precise insight and non-dubious business value from their data assets.

But let's be clear; implementing a data governance framework isn't easy and there isn't a one-size-fits-all approach to how that framework should look. For business and IT departments to find common ground and influence insightful, data-fueled decisions, they must collaborate around a governance framework and lay a foundation for data that the entire organization can trust.

That's easier said than done, and why I like to approach this partnership through three main disciplines. I call them the "Three P's" of data governance: product, process, and people. Only when all three are working together can your IT and business teams establish a framework that the entire organization can adopt.

## Product: Is the right technology in place?

Putting the right technology in the hands of both business and IT users is possibly the easiest part of this process. Technology should enable business teams to control the who, what, where, when, and why of data entry so different functions within the organization aren't able to influence information that doesn't pertain to them.

Are teams able to collaborate within the technology? Does it provide the necessary workflows for IT to easily "promote" business user data mashups to a centralized/certified model? Are you able to quickly and accurately monitor and identify anomalies or determine business impact to help quickly provide teams with the information they need?

Technology plays a crucial role in an overall data governance strategy, and if IT and business teams understand up-front what the technology should enable their teams to do, they can find the solution or solutions that fit their organization best.

However, technology is just one part of the framework.

## Process: How will everyone's data-related needs be met?

According to the NewVantage Partners 2016 Big Data Executive Survey, business and IT partnership was cited as the number-one factor in ensuring successful adoption of data-driven initiatives. Successful organizations have developed a common process on how data is organized, managed, and processed that is built around a set of data governance principles and practices.

Don't wait to adopt a data governance program after poor data leads to a bad business outcome or after an error occurs that no one takes responsibility for. Have a proactive process in place with clear responsibilities drawn out and the technology to support those needs.

Data ownership is a shared responsibility for business and IT teams, and governance must be managed as a business function like finance or human resources. A collaboratively built process will begin with the priorities specific to each team and clearly defined roles for all involved.

Open and upfront communication is the best way to ensure everyone's needs are met. Business teams must be transparent about their needs and respect that accurate information takes time, while IT must prioritize turn-around time for business users and keep business aware of the technical commitments necessary to generate this information.

A clear process allows data-driven organizations in any industry to build a bridge between technologists and decision-makers.

## People: Are the right people identified?

Employees are the centerpiece of the governance puzzle, and there are new roles being created to help curate data and manage the processes we talked about. The chief data office (CDO) or data steward roles are critical new players whose goal is to curate data and foster communication between teams. In fact, according to Gartner, the CDO position will be filled at 90 percent of large companies within the next three years.

There's great value in establishing a liaison and mediator between business and IT team leads. This helps business teams work with IT to maintain information protection, governance, and data quality while also working with business representatives to create value from data assets faster.

The governance protocol then moves down the ladder to all aspects of the business where data is involved. Each business unit needs a representative to make sure that their team is up-to-speed on the process for inputting and drawing data and trained with the technology that enables them to do so.

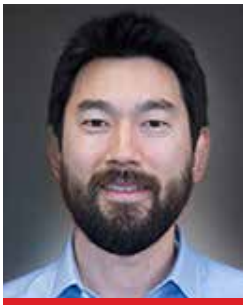
Data governance is not just about technology. It's about key stakeholders and employees creating processes and best practices to properly organize, validate, and derive business value from their own information. Without the appropriate framework in place to allow

the product, process, and people to work together, an organization's entire data strategy may rest on shaky ground and bear the risk of contaminated information. It's imperative that business and IT teams work closely together in developing a process, implementing technology to support the process, and identifying the people to manage it all. ■

Vijay Anand is Senior Director of Product Marketing at MicroStrategy. With over ten years of experience, he has served in several capacities including roles in consulting, technology, and marketing. His main areas of focus are self-service analytics and Software-as-a-Service BI. A graduate of Duke University, Mr. Anand has previously worked with General Electric and TATA Consulting Services and has developed his own start-up business.

# Balancing Governance and Empowerment in a World of Data Chaos

*by Dean Yao*



Dean Yao,
Director of Marketing,
Jinfonet Software

When it comes to business intelligence, the modern IT organization faces a dilemma: the demand from users to self-serve their analytics needs versus upper management's goal to ensure information is consistent with organizational goals and government regulations. Balancing both the goal of empowerment through self-serviceability, and governance poses a challenge.

If the scale is heavily tipped towards governance, organizations can lose agility and in turn miss out on opportunities that arise in a rapidly changing business landscape. However, governance is necessary in many organizations to ensure information integrity.

On the other hand, too much empowerment can lead to chaos and legal issues, which poses a threat to the bottom line. At the same time, empowerment enables users to find critical information that may benefit the organization.

Ultimately, the goal of implementing business intelligence is to strike a balance between both governance and empowerment. Business intelligence grants the potential for a strong competitive advantage, but proper implementation of any technology is critical to its success within the enterprise.

## The case for governance

In earlier days, BI vendors and data warehouse purists attempted to manage data as a "single version of truth." Many people now recognize that as a pure fairytale. Data can manifest and reside anywhere and enterprise data warehouses cannot contain all organizational information.

However, there is still a strong need for enterprises to govern all their data to achieve the utmost accuracy and consistency of reporting. Allowing business users to link multiple data sources into a report is seen, at the enterprise level, as running the risk of providing reports that may be inaccurate, inconsistent, and misleading. The challenge is how do enterprises govern these diverse sets of data?

The answer to this issue is for organizations to use corporate definitions of information, ensure templated reporting is managed, and metadata is enforced throughout the enterprise. Report templates allow users to easily compare information across business units and act as measures to ensure consistency. Metadata is critical to the goal of governance. This is because metadata is the essence of how data can be transformed to information thus making the case that the hallmark of governance is consistent usage of metadata.

## What does empowerment truly mean?

Empowerment is a critical enabler because the primary benefit of BI is to provide the ability for users to explore data, understand information and use it to facilitate decisions that benefit the organization.

Today, empowerment is often boxed into the concept of "self-service" BI, which has become nebulous and means something different to different people. Part of the confusion is the fact that people use information in different ways.

For example, developers have the goal of integrating customized reports into internal and external applications for corporate stakeholders and business customers, respectively. Business analysts bridge the gap of technical expertise and business requirements to simplify complex information that meets the business users' needs, which is to scan and drill into information for proactive decision making.

The overall goal of empowerment is to tailor the analytics platform to meet the specific needs of the different BI users within the organization. Understanding their needs, creating a strong collaboration between IT and business units will help craft a solution that truly empowers them to perform their analytics and get their job done.

## Balancing Governance and Empowerment

The CXO suite must clearly show the importance of information management governance. Governance is not only critical to IT but to all departments within the enterprise. The need to identify where governance is required, what challenges may arise, and how to create policies that address those challenges is paramount to the enterprise.

These questions pave the way for how software evaluations and defining specifications are made within the organization to ensure the needs of every user are met. Business units and IT need to work closely on these discussions to demonstrate the importance of governance. Given that, governance should not be seen as limiting software's ability to business, but a method of ensuring software provides all of the access that fits the business model.

Once governance policies are created and business units' needs are understood, empowerment becomes a matter of what technology meets user needs while providing features to meet governance goals.

When people feel their access to information is limited, they become frustrated. If software access limitations are in line with known governance rules, users can discuss those rules with management and then adapt the system to meet new exceptions or rules. The understanding of governance lays the groundwork for a lasting process of balancing empowerment and governance.

By working together, business units can find the right balance between empowerment and governance to ensure users can use information to their benefit while governance allows for information accuracy, consistency, and compliance. ■

Dean Yao brings over 10 years of experience in software marketing and product management. Prior to working at Jinfonet Software, Dean was a senior product manager at cloud computing startup Nimbula (acquired by Oracle), where he focused on technical best practices, competitive marketing, and product strategy. Dean was also in technical marketing at VMware, specializing in virtualization clustering and resource management products.

# 3 Big Data Housekeeping Measures You Can No Longer Overlook

*by Dan Potter*

Dan Potter,
Chief Marketing Officer,
Datawatch

As the big data marketplace moves closer to a point of mass maturity, business leaders have begun to take new approaches to implementation and utilization. Advanced analytics solutions have made their way into a range of industries and regions, and companies that successfully align these investments with core goals and requirements will enjoy more progressive improvements to operational sustainability, intelligence, and general performance.

However, there is some housekeeping that must be addressed as organizations embark on big data and analytics initiatives.

Data preparation, information governance, and security are three fundamental elements of effective analytics strategies, and yet each has been largely ignored by many organizations in the rush to realize the promise of big data.

## Data Preparation

Business leaders across industries now use big data analytics technology for a wide range of processes, objectives, and management needs. But while the technology is there, studies have shown that return on investment has been elusive at best for the vast majority of adopters. In fact, business analysts claim that 80 percent of their time is spent preparing data for analysis, and they still never seem to have the information they need.

Self-service data preparation is a critical, yet often overlooked, factor in the analytics process. Anyone can easily connect to relational data, CSV format, and other standard, structured data. But the data that provides the most analytical value often is locked away in multi-structured or unstructured documents, and it seems impossible to use this information without rekeying the data or asking IT for help. And with the volumes of data being created each day in various locations and formats, business users and data analysts don't have time to wait for a specialist to create and run a report, or time to grapple with IT to gain only limited access to data repositories.

Business users (aka non-IT experts) must be able to quickly and easily access all types of data – including multi-structured and unstructured sources such as PDFs, text reports, and Web pages, as well as real-time streaming data – across a variety of internal and external sources. Self-service data-preparation technology can enable users to extract, cleanse, prepare, and blend this otherwise unworkable data, transforming it into high-value information for solving business problems. Data experts, meanwhile, are liberated to spend the majority of their time on analysis instead of data preparation.

## Information Governance

Many organizations today are also struggling to reconcile information governance with strong analytics performance. One of the biggest benefits of self-service analytics lies in the ability to rapidly combine and analyze data from a variety of sources. However, this can sometimes introduce serious governance challenges, given that half of this data typically comes from sources that aren't managed by IT.

While most organizations have well-defined strategies for governing data that lives in managed systems, such as enterprise applications or data warehouses, analysts often need to pull data from non-managed sources, like CSV or text extracts from transactional systems, personal spreadsheets, third-party reports, or semi-structured content. Without proper governance, this can create big headaches around version control, data breaches, reconciliation, auditing, etc.

As big data becomes a more central aspect of corporate strategy, organizations must take important steps toward optimal information governance and then tailor their initiatives, policies, and strategies to adapt to the world of advanced analytics. When governance comes off the rails due to an advanced analytics project or any other reason, the chances of maintaining tight control over information and privacy while simultaneously enjoying high returns on big data investments will be inherently lower.

## Security

For all of the benefits big data provides, many professionals – especially in IT – remain fearful about the security issues it poses. And you can't blame them. IT departments are on the front lines, tasked with a never-ending battle to mitigate risks introduced by big data's volume, velocity, and variety. Additionally, advanced analytics tools and responsibilities are still somewhat new in the context of modern business intelligence solutions, causing many firms to struggle when searching for the right balance between protection, privacy, transparency, and return on investment.

Introducing data-masking approaches as part of the data-preparation process is a great first step in protecting sensitive information. However, more work needs to be done to ensure that the deployment of wide-reaching big data programs is both profitable and positive, rather than representative of much greater risks to information integrity and security. The goal is to get more out of analytics investments without bolstering risk levels.

## A New Era in Big Data and Analytics

The importance of security and information governance on big data and analytics implementations is clear, so why have many organizations failed to give them the attention they deserve? The answer is simple: they have historically impeded business processes and prevented business users and data analysts from doing their jobs effectively. But self-service data preparation has changed this and made data discovery and advanced analytics winning propositions for both end-users and IT.

The age of self-service analytics dawned years ago, after data and sources became so locked down by IT that users lost their ability to access a wide breadth of data for visualization and analysis. To get immediate results, analysts began resorting to the sources that were generally available to them, namely Excel. And once they started showcasing the insights that could be developed from such a small amount of valuable data, the self-service analytics movement took off. Suddenly, the desire to leverage Excel data for immediate business value became far more important than taking the time to track which data sources were available, who was accessing them, and how information was being repurposed and changed to support analytics processes. It was the "Wild West" of the data world. No one knew where data was coming from or who was managing it. Information was floating around without auditing or classification. And security and governance were neglected because they slowed down analysts' ability to do their jobs.

Today, there's a new sheriff in town – self-service data preparation, which is now being recognized as the answer to the big data security and governance challenge and a necessary component of any data discovery or advanced analytics implementation. Self-service data preparation drastically reduces the time and effort that analysts spend on prep work and enables them to leverage the widest variety of sources while keeping these corporate assets protected.

In today's big data and analytics landscape, business users can now be autonomous without causing disorder, and companies can leverage their intelligence investments while proactively mitigating threats. New approaches and technologies deliver the ease of use and agility that business users want, as well as the scalability, automation and control that IT demands. It's time to address security, governance, and data preparation head-on. Companies can no longer afford to sweep these three housekeeping items under the rug. ■

Dan Potter is chief marketing officer at Datawatch Corporation. Previously, Dan led the product marketing and go-to-market strategy for IBM's personal and workgroup analytics products and the online community and social media strategy for IBM's AnalyticsZone.com initiative. Formerly, he held senior positions with Oracle, Progress, and Attunity, responsible for identifying and launching solutions across a variety of markets, including data analytics, cloud computing, real-time data streaming, federated data, and e-commerce.

# Overcome These 5 Challenges to Manage Data Overload

*by Rob Consoli*

Rob Consoli,
Chief Revenue Officer,
Liaison Technologies

As companies are increasingly recognizing, data is the new currency in business. Enterprises harness the power of their data and apply it to improve daily operations in many ways. But the more forward-thinking businesses are taking it a step further by using data to drive innovation and disrupt their industries. In this way, companies are moving toward a data-inspired future.

Companies that want to compete and win in a data-driven economy have to find a way to leverage the power of their data and extract maximum value. But that can be a challenge as the volume of data grows and new sources of information come online in a multitude of formats, threatening data overload. Data overload, if not carefully prepared for and mitigated, could prevent an organization from enjoying any of the benefits of a data-driven economy.

There are five potential challenges associated with data overload that an organization needs to be prepared to encounter:

**1) Analysis paralysis from too much information and too many sources:** There is already an enormous influx of data streaming into the enterprise from multiple sources, but it's set to increase exponentially. Analysts predict that the Internet of Things (IoT) will comprise 200 billion connected devices by 2020. When IoT information is combined with other data sources, including cloud applications and social media, the volume of information can quickly overwhelm businesses.

It is a well documented fact that too many options can prove paralyzing for consumers, and businesses aren't immune from that phenomenon. Without a solution that enables them to effectively handle the influx of information and harmonize data from multiple sources, businesses will face disrupted data workflows.

**2) Silos created by fragmented data solutions:** Big data works when companies can glean insights from a unified data pool. But too often, businesses face data fragmentation. They work with a range of big data tools that each address one part of the operation, including functions like data storage, cleansing, API management, data visualization, and more.

This piecemeal approach to data management results in multiple silos, which make governance and compliance incredibly challenging. Meanwhile data quality, security, and visibility decrease while expenses and inefficiency increase.

**3) Data generation and resource disadvantages for small and mid-sized businesses:** Big data is expensive; it requires an investment in resources to generate, process, and store all that information. Large companies like big box retailers have the cash and infrastructure to make big data work for them — they have assets like cameras, consumer apps, and point-of-sale software to generate and make sense of data so that they can continuously improve the customer experience.

But small and mid-sized businesses typically don't have the resources to monitor, influence, and predict customer behavior. And even those that do usually do not have a sufficiently large customer base to generate macro-level insights. Small and mid-sized businesses have to leverage all the innovation and creativity available to them in order to find ways to make the big data revolution work on their budget and with their customer base.

**4) Faulty decision management processes:** With big data analysis, more decisions are left to machines. That leaves the decision-making process vulnerable to the inclusion of faulty variables. In 2010, there was a stock market crash — the Flash Crash — that occurred due to a faulty decision management process that relied too extensively on algorithms. That incident demonstrated that it is unwise to leave machines completely in charge.

Businesses that use big data must guard against making similar mistakes that can wreak operational havoc downstream. To prevent a Flash Crash-like catastrophe, companies must balance the efficiency of machine algorithms with the superiority of human judgment and make adjustments as necessary.

**5) Backward-looking analytics that don't foster innovation:** It's important to acknowledge that big data algorithms are by nature backward-looking. Users propose a hypothesis, crunch historical data points, and review outcomes that fall into predetermined ranges.

This process can yield incredibly valuable insights, but its focus on historical data points doesn't readily foster innovative thinking. It merely provides a starting point. Companies that want to facilitate calculated risk-taking to drive disruptive change will usually need to look beyond math-based big data, leveraging it while also using creativity to innovate.

These five challenges present significant barriers for companies that seek to use big data to its full potential. But the good news is that big data practices and technology are maturing, making these barriers easier to overcome.

There are now solutions on the market that enable companies to leverage the cloud for the integration and data management support they need, regardless of business size or in-house expertise. One example of such a solution is a new approach to integration and data management called data Platform as a Service (dPaaS). dPaaS is a cloud integration and data management model named for its ability to provide PaaS functionality at the point of data analysis without encumbering users with the particulars of the underlying data capture, integration, or management mechanics.

Businesses that access such innovative solutions to overcome the five challenges of data overload with a comprehensive integration and management approach will be ready to embrace a data-inspired future. ■

Rob Consoli is the Chief Revenue Officer for Liaison Technologies. He brings over 25 years of technology industry experience and has a demonstrated track record of successfully building teams and helping growth-oriented companies navigate cultural and process transitions as they expand operations and global reach. In this pivotal role, Rob leads Liason's North American efforts to strategically position and sell the company's cloud-based integration and data management solutions, as well as increase its sales to meet the company's growth objectives. Consoli holds a Master of Science from Southern Methodist University and a Bachelor of Science from Auburn University.

# Bridge the CxO Gap with a Data-Driven Approach

*by Kelle O'Neal*

Kelle O'Neal, CEO, First San Francisco Partners

There has been a lot of noise both in the press and in social channels about how to identify the compelling business case to get senior leaders to "buy in" to the investment for improved information management and governance. I'm going to share a secret with you: They are already bought in!

All you have to do is look at the cover of any business journal to find examples of how data (or a lack thereof) has impacted a business. A recent infographic showed that bad data may cost businesses as much 10-25 percent of an organization's revenue each year, and Gartner surveyed a wide variety of companies and found that, on average, data quality issues cost them an estimated (and staggering) $14.2 million annually.

Any executive knows that, in order to properly run their business, they need to understand their expenses, their revenue, their risk, and their future opportunity – all categories of information – using aggregated data to provide a picture of the current state of the business. In addition, effective information management is crucial to compliance with regulations as well as country- and state-specific laws and regulations about data privacy. Whether it's for the purposes of regulatory reporting and compliance, decision making, or just ad hoc investigation and analysis, executives already know that it's important to have accurate, reliable, and auditable data to run their business.

But if senior leadership are already bought in, then why don't they behave like they are? And why don't they approve the budgets for projects that can improve their data quality?

Let me share another secret with you: They think you are already doing it. Because executives have to certify the validity of the data that is distributed externally to customers and investors, they assume that the data are already accurate, current, and controlled. What they may not be aware of is the amount of effort (often manual) that is needed to get the data to the state in which it is trusted. Addressing this is problematic because requesting the funding to automate the manual data sorting and cleaning and improve the process of data sourcing and reporting requires admitting that there are inefficient processes in the first place. Putting together this sort of proposal means honestly assessing the inefficiencies and identifying how improvements will increase productivity and reduce costs in the long term.

There's another reason that executives don't seem to behave like they are bought in to the requirement to better manage and govern data: Executives don't necessarily see the link between data management and governance and their corporate goals. This is partially our fault as an industry. We need to recognize that there is a business language that is different than a data language. In business-speak, the budgetary process consists of a series of trade-offs and decisions about what investments will enable the organization to meet their corporate goals in the most effective way possible. Therefore, in our data-speak, when we are creating our roadmaps and seeking funding, we need to understand explicitly how the data enables, or impedes, the organization's ability to meet its corporate objectives and be able to clearly communicate that requirement and dependency. Improving data for data's sake will never get approved. Improving data to progress a business program and objective will. Information management programs need a business justification to make them worthwhile. A business transformation project, for example, that will increase supply chain efficiency, organize clinical trials data, and/or improve the customer experience needs to have the business value outlined for each goal before starting the project, as does the data component of those projects.

An additional reason executives don't act like they are bought in is that the goals and objectives of certain personnel may conflict with the goals of data quality and the aggregated reporting requirements of senior management. For example, a call center manager is incentivized by the volume of calls and clients served, leading to potentially suboptimal data accuracy and duplicate records. If executives believe that data quality and accuracy is important, which they surely do, then they need to identify and encourage behaviors that lead to those results.

This means that, across the organization, we all need to become data driven. To be data driven means that data is considered an important part of the process, not just the end result of the process. It means that when we define a process, we consider how the steps in that process contribute to or diminish the accuracy, consistency, and timeliness of the data. For instance, do we search multiple ways for an existing customer record before entering a new one? Do we assess our existing data feeds before adding another? To be data driven also means that there is a real consideration of the trade-off between speed and quality. This implies that there is the ability to assess the impact of that trade-off and the willingness of all employees to raise it up the executive chain for consideration. Leadership is concerned about high productivity, but not if it sacrifices productivity in other areas. It is faster to enter a new customer, but identifying and resolving duplicate records adds days to quarterly reconciliations.

Being data driven also means having a view of that information and how it is created and consumed across the enterprise, not just within your operational unit. This is difficult because we don't all understand how data moves across an enterprise and the implications of additions and changes in data from one system or operational unit to the next. New approaches to Enterprise Information Management help by creating that enterprise

understanding and documentation of information that is not limited to a single operational unit or system. This enables you to answer the question, "Who are our top 10 customers globally?" rather than, "Who are our top 10 customers in the United States?"

Being data driven means creating data-centric development processes, not systems-centric or functionality-centric. We have been hearing about software development lifecycles and agile development approaches for years, and they have advanced the consistency and speed of development processes, but they don't necessarily consider the importance of data in that process. A data-centric development process explicitly considers the data quality, data model, data movement, and metadata as part of a development process, not just the functionality of the solution.

Most importantly, being data driven means understanding how data contributes to the success or failure of projects, programs, goals, objectives, and strategies. Those who are identifying and executing data programs need to understand those links, and those at the most senior level who are accountable for those corporate objectives need to understand data's involvement, too. ■

Kelle O'Neal is Founder and CEO of First San Francisco Partners. A veteran leader and accomplished adviser in the information management sector as well as a speaker and an author, Kelle is passionate about helping organizations apply data and intelligence to gain a true competitive advantage.

## Check out additional content on Data Informed

Find other articles like these and more at Data Informed: data-informed.com

Data Informed gives decision makers perspective on how they can apply big data concepts and technologies to their business needs. With original insight, ideas, and advice, we also explain the potential risks and benefits of introducing new data technology into existing data systems. Follow us on Twitter, @data_informed

DataInformed

Data Informed is the leading resource for business and IT professionals looking for expert insight and best practices to plan and implement their data analytics and management strategies. Data Informed is an online publication produced by Wellesley Information Services, a publishing and training organization that supports business and IT professionals worldwide. 