

BI Data Governance

The Secret of Successful Business Decision Making

MARCH 2017

A White Paper by
Dr. Barry Devlin, 9sight Consulting
barry@9sight.com

As data becomes more voluminous and data sources more various, the opportunities to fail in Business Intelligence (BI) become ever greater. Such failures often arise from problems with data quality or issues with data preparation and delivery to the business. An over-emphasis on self-service and data discovery approaches can also lead to such shortcomings. Fail even once, and business begins to lose trust in the BI service and looks elsewhere for its data and decision-making support.

Comprehensive BI data governance is thus a vital part of any modern decision-making support system. Beyond the procedures, roles, and rules of general data governance, in the BI environment, data governance must be embedded and managed in the system itself. Based on an understanding of the path from data to decisions, this paper offers three cornerstones of BI data governance: (1) decision making as a process, (2) a single, integrated platform, and (3) an adaptive decision cycle.

It also reviews the data governance functionality offered by Yellowfin, and most recently updated in version 7.3+, which play strongly in support of the above cornerstones of BI data governance.

Contents

- 03** Data governance for smarties
- 05** From data to decisions
- 06** Three cornerstones of BI data governance
- 09** BI data governance as supported in Yellowfin
- 10** Conclusions

Introduction

When the Monthly Sales Management Meeting (MSMM) descends once again into a firefight of claims and counterclaims of ‘My figures are better than yours’, you know the old problem was **not** solved. This was not supposed to happen after you replaced the previous *ad-hoc*, spread-sheet-based system with a bright and shiny new self-service Business Intelligence (BI) tool.

The vendor assured you that spreadsheets were the problem: that they sprouted like weeds in every department. Each with its own inimitable way of calculating the sales figures. Each with its own peculiar formulae. And, more than a few with their own hidden calculation and transcription bugs. But, the same old spreadsheet-induced reporting problems have now resurfaced with the BI tool. So, what has gone wrong?

Your company’s BI strategy is well aligned with analyst opinion¹ that “*the business intelligence and analytics platform market’s shift from IT-led reporting to modern business-led analytics is now mainstream*”. You have adopted visually based data discovery, empowering business users to build and run their own analyses. You have seen a reduction in IT involvement in BI delivery—a role long blamed for stifling innovation and slowing reaction times in the business. BI and analytics is now clearly business led and agile. Your new self-service BI platform is well-respected in terms of capability and ease-of-use.

However, the actual behavior of business users has changed only marginally since moving away from spreadsheets. They do love the sexy new user interface and feel empowered and enabled to explore the data and hypothesize on causation. But, really, they’re still going to the same old sources for data—the friends they’ve made over the years who have access to raw data. They still discuss their analyses and results with the same colleagues they have depended upon for review since they started doing BI. They remain resolutely departmentally focused.

In many respects, nothing has changed except the tool used. Ironically, data quality and consistency have become an even bigger challenge as users are encouraged and enabled to engage in ever-deeper data discovery with enterprise-sanctioned self-service BI.

“Chronic data governance problem!” muttered the head of IT as she left the MSMM. She reemphasized—again—that the data warehouse should be the only (or at least primary) source of data for consistent BI and analytics. Her diagnosis rings true, but her solution doesn’t. Some of the most challenging data quality issues arise in the BI environment, irrespective of the data sources.

An idea takes root. Data governance is an organizational methodology, but, to apply it realistically and successfully in an analytics program, practical support tools are required. Why not embed data governance principles and approaches in the organization through the BI environment itself? Here it will be front and center for the business users and decision makers of the company—precisely where it belongs. Here, business and IT can collaborate closely in its enforcement and use, rather than the business thinking it’s cost-free magic offered by IT and the warehouse. Here, data governance might just take root and succeed.

Let’s explore the idea, starting with a quick overview of what data governance is, and what it isn’t.

The BI platform may offer the best place to show business the benefits of real data governance and to bridge the business-IT gap in pursuit of data quality.

Data governance for smarties

If your eyes glaze over at the mention of data* governance, you're not alone. Unfortunately, if you don't pay attention to the topic, a flying data shard from the information explosion will likely take your decision-making eye out or cut the jugular of your BI funding. Yes, it is that serious.

Failures in data governance lead directly to information trust issues for the business. When trust is eroded, decision makers fall back on old, questionable behaviors for obtaining data and validating decisions. When trust is lost, regaining it is difficult. Getting data governance right from the get-go is the best—and, perhaps, only—way to build and retain trust. So, what is data governance?

One common explanation² states that “*Data Governance refers to the overall management of the availability, usability, integrity, and security of the data employed in an enterprise.*”. Note that data governance is not an IT or technology project. Rather, it is a set of rules, processes, behaviors, and organizational structures and responsibilities that apply to every aspect of data creation, integration, understanding and use, maintenance, archival, and deletion throughout the entire organization. That's big. However, the focus of this paper is on how these components can be embedded and operationalized in a BI environment to improve the quality and reliability of decision making.

Data governance in modern decision-making support

Data governance has long been a significant contributor to the success (or failure) of BI projects. Recent developments in business and IT further increase its importance:

1. The quality conundrum of external data

At some time during the mid-2000s, data (particularly as it relates to business decision making) crossed an important line. Previously, the majority of such data was sourced internally and its quality and reliability was in the hands of the internal systems and IT developers that created and maintained it. Since then, an increasing proportion of data comes from external sources. While internal data quality has often been questioned, it certainly far exceeds that of external data. The origin and prior processing of external data is dubious, because it comes from sources of wildly varying reputation via networks of suspect security. Even the basic veracity of its content must be questioned, as evidenced by the rise of ‘fake news’.

2. BI and analytics for the masses

Users' knowledge and prior experience of their data has long been key to avoiding data quality issues. However, as larger numbers and types of users take to friendlier BI and analytics tools, they often lack the skills to judge data quality. They thus make unjustifiable assumptions about how the data can be safely used. Data discovery and self-service BI allow users more freedom to manipulate data, even further beyond the limits of their data management skills. Data governance must therefore play a central and ever more important role in setting safe boundaries.

3. The need for speed in everything

Freddie Mercury's famous anthem “*I want it all, and I want it now*” underpins modern BI usage. Business has become increasingly dependent on real-time decision making as customers succumb to the dream of instant gratification. Combined with ever more external data, the perfect recipe for a data quality disaster emerges. Decision makers with limited experience rush into decisions, using whatever data is on hand.

Taken together, these three factors, which are at the heart of a data-driven business, point to the now vital role of data governance in ensuring that only the highest quality data is made available to business users.

The role and relevance of BI data governance have expanded greatly as business needs and IT tools have advanced in the past decade.

*Throughout this section, I use *data* and *information* interchangeably, in line with common practice. There is, however, an important difference between them which will be explained shortly.

When data governance fails

The immediate outcome of poor, or non-existent, data governance in the decision-making environment is the emergence of disputes among users, or departments across the business, about whose data is 'correct'. In fact, nobody's data may be correct, but the costly task of finding out whose might be wrong is often handed to IT. This exacerbates the problem, with effort that should go into ensuring data quality instead diverted to triage and investigation. And, oftentimes, the answer delivered is unacceptable to one or more parties.

With a growing distrust in the BI-provided data, users look elsewhere for answers. Power users and data scientists who have a few successes become heroes of the business and their roles are extended beyond the limits of their knowledge. Spreadsheet usage grows again. Cottage industries supplying data from private or departmental sources thrive. Funding that could have supported a more central approach to data governance is diverted to pet projects of departmental leaders.

Most disturbing of all, the business is exposed to a range of unnecessary or over-inflated risks as data quality deteriorates. Having the data governance frameworks, workflows and systems in place to prevent inaccurate data permeating throughout the business is critical. Failing to do so can be extraordinarily costly, as the following spreadsheet-induced errors show:

- At Fidelity Morgan in 1995, the accidental omission of a single minus sign in a dividend estimate spreadsheet resulted in a miscalculation of \$2.6 billion³
- Utah State Office of Education under-calculated the state's education budget by \$25 million in 2015⁴
- Canadian power company TransAlta accidentally over-paid for US power transmission hedge funds to the tune of \$24 million in 2003⁵
- Mouchel's pension fund miscalculation saw its profits misreported by more than \$13 million in 2011—the share price of the UK support services group plummeted and its CEO was forced to step down⁶
- One of Australia's big four banks, Westpac, had to halt trading and deliver its annual profit briefing early after sensitive⁷

But, data governance issues, and their repercussions, stretch far beyond spreadsheets and the finance department. Dirty data arriving in organizations' operational systems, reporting and decision-making processes has caused several severe customer service and marketing gaffes—not to mention PR emergencies. For example, the marketing team of a large financial institution accidentally sent out letters to its customers titled "Dear Idiot Customer John Doe" because disgruntled customer service staff had been entering derogatory comments into the customer service database⁸.

Data accuracy is especially important in the healthcare sector, where patient wellbeing and confidentiality are of primary concern. According to official records from Britain's hospital system, 17,000 pregnant *men* partook in "pregnancy-related services", including obstetric and prenatal exams, between 2009 and 2010 in the UK. The inaccurately entered medical codes resulted in significant billings, claims and regulatory compliance issues⁹. And, in 2015, it was found that poor data management prevented consistent creation, maintenance and reporting of essential data at the US Veterans Health Administration's Health Eligibility Center. The result? Delays to hundreds of thousands of veteran healthcare applications¹⁰.

But data governance is paramount to avoid costly mistakes in any industry or job function. NASA's Mars Orbiter was destroyed when it crashed due to data inaccuracies that "failed to convert English measures of rocket thrusts to newtons", resulting in the loss of a \$125 million spacecraft¹¹.

All these data-related disasters have one thing in common: unmanaged manual processes and ungoverned data management practices. With such examples, it's easy to see how bad data governance, applied to the area of BI, results in poor decision making. From small, daily decisions—each with limited impact on the bottom line—to regulatory reports, where significant error could send the CEO to jail, the business takes on risks it could have avoided. And how could it have avoided those risks? Through the timely and judicious execution of a data governance strategy and its implementation in the decision-making process. To build a successful BI program, it's therefore important to seek out analytics technology that supports and drives superior data governance practices.

When data governance fails, users revert to unsafe practices, putting the health and survival of the business at significant risk.

From data to decisions

Governance in the BI environment requires an understanding of the difference between data and information, as well as the role of context, as we follow the path from the data available in a BI system to valid and reliable decisions.

In *Business unIntelligence*¹², I defined information as the (now mostly) digital “recorded and stored symbols and signs we use to describe the world and our thoughts about it, and to communicate with each other.” Information thus comes from human sources and is the highly subjective record of human experiences, in the form of text, audio and video. From fiction to alleged fact, it is loosely structured and often ungoverned: it may or may not reliably tell the business what has happened in the real world. Today, such raw information comes directly to business from social media.

Traditional computing, on the other hand, starts with data. People’s wants and needs—information—is first converted to data by structuring and standardization (e.g. modeling, data entry, and validation) in operational systems, and cleansing as it moves to BI. Some data also comes from machines—ATMs, for example—within the enterprise. This is the formal and *well-governed* record of the business. It is highly structured, well suited to computer processing, and split into values and contextual descriptions (a relational table is a good example).

In addition, growing volumes of an increasingly important type of data comes from external devices via the Internet of Things (IoT). The reliability of this data is suspect and, like social media, stringent data governance practices are needed around its ingestion and use.

BI begins with data but needs information

The starting point for BI has always been data. Traditionally, it was all well-managed data from the operational systems of the business. Today, externally sourced IoT data is included. Externally sourced social media information is also involved, after it has been converted into data: it has been structured via analytics and statistics into a form suitable for summarization, averaging and other mathematical processing.

To make decisions, BI users first convert data back into information by applying **context** (including business metadata and personal knowledge), thus giving it meaning for the business. A key role of BI data governance is to ensure the existence and validity of such metadata. Without it, transforming data into useful business information for decision making is impossible. If the metadata is incorrect, limited, or out of date, the transformation will be faulty and decisions based upon such incorrect information will likewise be of dubious or worse quality.

In addition, knowing the sourcing and lineage of the original data and information is vital and a key aspect of BI data governance, for both internally and externally sourced data. Data governance shows BI users the origins and history of their data, enabling choices about how and where data can be used safely and appropriately. For internal data, for example, knowing that data comes from the new, well-documented CRM system can give BI users confidence to use it more than the data from an old, spreadsheet-based marketing application.

This aspect is particularly important for externally sourced BI data, such as social media or the IoT, with its much lower quality and reliability—a result of its poorly understood origins and processing. Data governance in this area must recognize that social media may be created to deliberately mislead—a characteristic seen clearly in ‘fake news’—and provide a means to determine real from false. Context—including sourcing, lineage, and external relationships—provides the basis for this judgement. For IoT data, contradictory and rapidly evolving definitions are often a challenge. Data governance clarifies these issues when BI users attempt to extract information from it.

BI users start with data and convert it to information by applying context before using it to drive decision making.

Context and quality are cross-enterprise concerns

In contrast to the more traditional and broader world of IT data governance and management, where the focus is largely on the quality of individual data elements, BI data governance re-quires a wider view, because decision making is a cross-enterprise process. The data used in BI crosses and re-crosses departmental and functional boundaries. Just because a data element is right for finance doesn't necessarily imply that marketing can directly use it with peace of mind. As BI data is widely shared across the organization, the context of its use changes. BI data governance must take this into account by showing the limits of validity of data from disparate sources and relating this to the roles and responsibilities of users in different business areas.

The road from data to information, and on to decision making, is paved with quality intentions. It is only with well-considered and broadly-implemented BI data governance that these intentions can be realized. We now explore the three cornerstones on which comprehensive BI data governance is built.

Three cornerstones of BI data governance

Decision making as a process

Today's emphasis on self-service BI and data discovery is, in large part, a response to the perceived bottleneck that exists in making all required data available to decision makers. The problem is often attributed to IT: it is said that they lack agility, knowledge of the business uses of data, and so on. This simplistic explanation misses the reality that decision making is a process that spans from the creation and consolidation of data in the IT domain to its manipulation and use by business users to make decisions.

Of course, in accessing, understanding, and acting on BI data, business users should be as self-sufficient as possible. In the metaphor of the salad bar, business users can pick and choose what they want to eat and combine ingredients as they prefer. But first, the salad bar must be stocked with fresh, quality, and complementary ingredients, which must be sourced from the best suppliers and transported to the salad bar. Furthermore, this type of 'self-service' is only the penultimate step of the process. After choosing the ingredients, the user must arrange and eat the salad!

Likewise, self-service BI and data discovery is only part of a decision process that involves data governance, IT, and business users, both expert and lightweight. Figure 1 shows this process:

- 0. Ingestion, consolidation and reconciliation:** when applied to internally sourced data, this preliminary step corresponds to the creation and management of a data warehouse. Legally relevant data demands governance in master data management and data warehouse initiatives run by IT before it arrives in the BI environment. For externally sourced data, this step represents loading and cleansing (often called wrangling) data into a data lake, often the province of data scientists and other experts.
- 1. Data access and preparation:** providing access to all potentially relevant data from any source, this step also offers some preparation of the data if needed. Separating access from preparation is difficult. Sometimes available data must be rearranged, reformatted or transformed in some way. When preparation is simple, business users may choose to do some of it themselves. But data experts and even IT are vital actors here, especially where preparation is more complex or data is new or of dubious quality. With such overlap in roles, a user-friendly and collaborative environment for all these parties is vital to support and drive successful data governance.
- 2. Evaluation, analysis and insight:** the heart, of course, of any BI work process, this is where business users excel and add most value. When data governance principles have been properly applied in the previous step, the validity of the results will be assured. However, new quality issues can certainly arise here too. Business users must also be able to loop back to the access and preparation step and collaborate with IT and data experts from here.

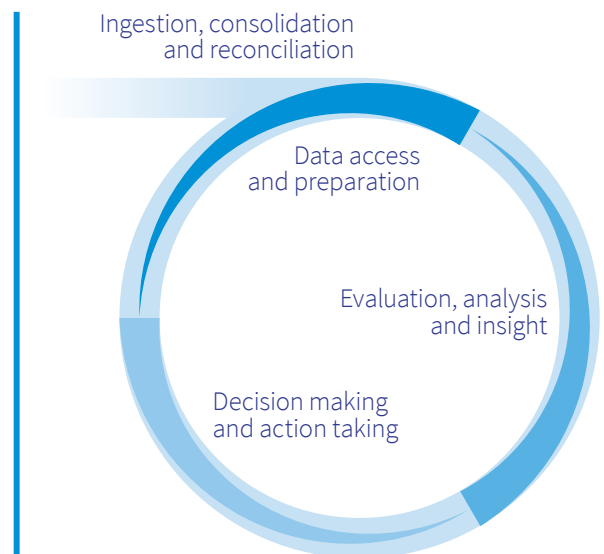


Figure 1:
Individual decision process

- 3. Decision making and action taking:** where true business value is generated, many BI tools often overlook this step. Here, insight is forged into decisions by assigning and monitoring tasks and actions throughout the organization. Here, BI data governance melds with change management, but the context and quality of the underlying data remain key: allowing actors to trace actions and manage tasks, all the way from the underlying data, through analysis, to committed and confirmed actions.

Steps 1-3 form the tight, agile, well-governed process loop of modern BI. Business users may traverse this loop many times in the course of a single decision. Therefore, all the data they may need should be available here in the BI environment. If not, work outside the BI platform is needed. In some cases, a visit to step 0 will be needed, to involve IT and data experts in complex data ingestion or consolidation tasks. Unfortunately, this will likely incur delays as more demanding data governance is required to ensure quality and consistency in the data warehouse or data lake. Generally, therefore, stepping back to the broader data governance environment is best avoided if possible.

A single, integrated BI data governance platform

The well-governed and managed process loop shown in Figure 1 demands implementation in an integrated platform. While a single platform is not mandatory, implementing data governance for BI across multiple platforms adds considerable complexity and possible gaps. Furthermore, with governance of ingestion, consolidation, and reconciliation (step 0 above) already in a separate enterprise-wide governance program, focusing all BI data governance into a single platform makes the most sense.

A single-platform approach offers a number of clear advantages:

- 1. Complete data lineage:** business users, as well as BI and data experts, may modify existing data and create additional data in the BI platform. How such data was calculated, by whom and when are all characteristics of interest to the business. This is especially so when results or decisions are disputed. Such lineage is a fundamental aspect of data governance and is easier to manage and ensure completeness if all this information is stored in and / or managed by a single platform.
- 2. Complete activity lineage:** closely related to data lineage (because many actions affect data), activity lineage also includes those actions that may not specifically change data. This may occur as work moves in the process from evaluation and analysis to decision making and action taking. Included here are records of collaborative contacts, messages, and actions, as well as actions taken to implement actual decisions. A single platform, where all BI-related activity occurs and can be tracked, offers the best possibility of governance.
- 3. Single or multiple consistent version(s) of the truth:** while recognized today that one “single version of the truth” (SVOT) is neither possible nor desirable as the one goal of BI, there is usually a need for a version of truth that is *primum inter pares* (first among equals) and to which others must relate. Such a favored version of the truth will typically be the one defined by finance for company and regulatory reporting. While others differ from it, they should not contradict or diminish it, leading to a consistent view of the data from specific points of view. This is actually a more complex, but also more realistic, requirement than the traditional SVOT. Governing and managing it demands comprehensive business and technical metadata, shared, reused business logic, and standardization of definitions. A single platform approach provides more useful support for such needs.
- 4. Secure and consistent environment:** the consistency, lineage and tracking needs of BI data governance can only be managed if a single logical security environment—consistently and comprehensively covering user, role, group, content, function and data levels—is in place. A single platform offers the most straightforward solution.

The driving need here is to provide a consistent, integrated path from the sources of data for BI, through the preparation and analysis, to the decision and action phase. Given the often-exploratory nature of BI, this consistency cannot be enforced by restricting user behavior. Rather, it must be embedded in the business and technical metadata of the environment, thereby guiding and influencing business usage. A single BI platform approach, while not the only way to achieve this, offers a simpler and more direct way to enable and drive BI data governance and management.

BI data governance benefits immensely from delivery via a single, integrated platform.

An adaptive decision cycle

A deeper and more nuanced understanding of BI, and the importance of data governance within it, comes from the recognition that two fundamentally different approaches to decision making occur in every organization. The first is the highly innovative, *ad-hoc* exploration that individual users undertake to address immediate problems or questions. Often performed in spread-sheets with minimal governance of data or analysis, BI has long tried—and mostly failed—to address this type of decision making. The second type is the formal or regular reporting and analysis based on pre-approved data and using agreed tools to ensure repeatability and quality of results, usually associated with a data warehouse and centralized BI implementation.

The secret to BI success, however, is to combine and enhance both approaches in an **organizational** process to encourage creative use of data and analysis, and—as appropriate—to bring the results of such creativity into a more formal and governed environment. The adaptive decision cycle, shown in figure 2, achieves this goal in three phases, each with identified user roles and data governance needs. These phases occur in sequence, moving from early innovation to full production-oriented use:

1. Exploring: business users working at an individual level begin by observing an event of interest, garnering data, innovating in an analytic exploration, and utilizing the conclusions in an attempt to reach a decision / action. At first, they are missing sufficient information, and iterate back to garner more. This leads to multiple iterations around the inner (blue) circle. In each cycle, they garner additional data from both common, certified sources and any others needed to perform the required analysis.

Explorers focus on their strength: understanding the business and its data. Data governance offers the best metadata available to guide explorers to the best data sources and to advise on valid usage.

2. Cultivating: when the exploratory work reaches a certain level of maturity, peer review is called for. Collaboration among peers is a key aspect of this (red) phase and ensures the work is fit for purpose in the department or group where it originated, and error free. This iteration follows the same steps as exploration and enables colleagues to formally seek advice, review data choices, analysis, calculations, and so on.

Cultivators (business users and data experts) share tasks to validate the data and calculations used by explorers. They are typically more experienced than average users and contribute to BI data governance because they appreciate its value to business goals. In this phase, governance focuses on task management between cultivators, documenting metadata about the analysis, and signing off the business design—data and analysis.

3. Grounding: as the analysis work achieves a level of popular use in the business, usage monitoring triggers a more formal involvement of data governance and IT functions to re-view the work for wider use in the organization via the gray iteration. This may lead to conditioning: the migration to a more production-level environment, the creation of formal data sources in the data warehouse or data lake and the implementation of formal governance and management procedures.

Grounders (data governance and IT) have a more formal role in creating and preserving a well-governed and managed BI environment. BI data governance provides formal and robust approval processes for data content, reports and views, and processes from data preparation to metadata population.

The adaptive decision cycle is an organizational process that aims to encourage innovative BI, as well as providing a mechanism by which governance and management can be easily applied, allowing enterprise-wide benefit from such innovation.

Although technology agnostic, the adaptive decision cycle can benefit substantially from implementation in a technological environment that supports its thrust and concepts. As already seen, BI data governance is most easily managed in an integrated platform. Many aspects of the adaptive decision cycle can benefit from collaborative technology, metadata management, and a process-driven approach to BI.

Many of the characteristics described in these three cornerstones of BI data governance are found in the Yellowfin BI platform, to which we now briefly turn.

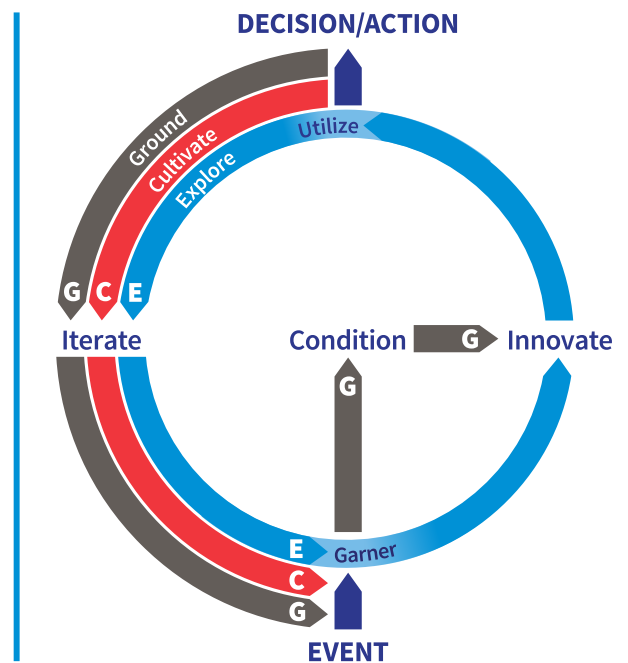


Figure 2:
Adaptive decision cycle

The adaptive decision cycle encourages innovative BI, and provides a mechanism to apply data governance and management.

BI data governance as supported in Yellowfin

Yellowfin was founded in 2003, when big, all-encompassing BI systems were already struggling with growing complexity and their strong IT-centric approach. The company's goal was a robust, scalable, metadata-driven platform that honored IT's strengths in data management and governance, as well as effectively delivering highly consumable BI to business users.

Yellowfin is well positioned to deliver dependable data governance for trustworthy, modern enterprise BI.

That dual focus has continued through the emergence of the BI self-service and data discovery movement. Better, more intuitive user interface elements continue to be added, but—in contrast to newer tools—Yellowfin continued to highlight the role of IT and data experts in delivering scalable solutions and quality data.

The collaborative Timeline functionality added in mid-2013, the upgrade of the administration console in Yellowfin 7 later that year, and the introduction of the Business-Analytics Workflow in the 7.2 release (April 2016) illustrate Yellowfin's growing commitment to the goals of BI data governance in a process-oriented approach to decision-making support. The most recent 7.3 and current 7.3+ release continue that trend.

The extensive BI data governance support of the current Yellowfin product can be traced through three different, but complementary, aspects of the product's Business Analytics Work-flow and Timeline.

Decision-making process based on collaboration and task management

Task Management allows users to raise, assign and track tasks throughout the platform. A business user may flag a data quality issue, annotate it with the exact context, and ask a data or IT expert to investigate. Progress is tracked and reported automatically. The same system allows business peers to preview reports or analyses before going live.

Expert Approval, extended in version 7.3+ to include data preparation, allows data governance staff to manage, curate, and approve enterprise content—including views, reports, dashboards, storyboards, and so on—before they are released for public consumption. Such tracking, auditing, and governing function allows the organization to ensure that users have access only to data and function that have been governed and validated by identified staff, such as data stewards and IT, with appropriate knowledge and skills. This is the foundation for data that can be trusted, analyses that can be believed, and decisions that are agreed upon across the board.

This level of transparent and managed collaboration between all parties to the BI system (1) streamlines communications between all parties, (2) provides checks and balances in data quality vs. agile delivery, and (3) ensures deliverables meet business requirements.

An Expert Approval process allows data governance staff to manage, curate, and approve enterprise content throughout the BI environment.

Streamlining decisions and actions

Alerts allow users to be automatically informed if data goes outside defined boundaries. Smart Tasks extend task management to allow users to prescribe actions, assign responsible parties, and set deadlines for resolution and tracking of assigned actions.

Together, this enables decision making by exception management, allowing business and other users of the platform to focus on what is important to their role, rather than detailed tracking of business processes that are working to plan. In terms of the adaptive decision cycle shown in figure 2, alerts and Smart Tasks are key aspects of bringing significant events to the notice of BI users and triggering the cycle into action.

In Yellowfin version 7.3+, these functions are extended to cover Expert Approval of data preparation tasks. When a user creates a data view, for example, with cleansed and transformed data, s/he must request approval before it can be put into production. The approval request appears in the Timeline of an Expert Approver as a Task.

Enabling auditability of the decision-making process

The Yellowfin Timeline records a users' BI activities and interactions in real-time, providing a searchable, personalized, chronological catalogue of the activities of, and anything related to, each individual user. This includes reports viewed, content created or shared, discussions en-gaged, approvals required or given, as well as alerts and notifications.

While clearly encouraging collaboration, and instilling a culture of data-based decision making, the Timeline system also contains a complete record of all the activities that occur in the Yellowfin platform, providing a searchable, auditable history of the decision-making process. This can form the basis for the organizational processes of cultivation (peer review) and grounding (promotion to production) described in the adaptive decision cycle.

Conclusions

BI data governance has come of age in the past few years, driven by an explosion of poor quality (mostly external) data sources and a mushrooming of interest in self-service and data discovery among business users. The business value available from both big data and easy business access to BI technology are indisputable. However, the benefits can be won only with a renaissance in data governance, particularly within the BI environment.

BI data governance is but a subset of the larger topic of enterprise-wide data governance. This broader scope is also vital, but BI data governance is of particular importance because it is through BI and analytics that much of the new externally-sourced—and often poorly governed—data arrives in the enterprise and is used in decision making. Furthermore, the speed and agility demanded of decision making today requires more closely integrated governance processes and functions.

Decision making in today's BI environment can and should be considered as a process, at both individual and organizational levels, so that speed and agility can be well balanced with quality data and reliable decisions. These processes are most easily and successfully implemented and managed within a single BI platform spanning from data sources to eventual decision making and action taking.

Yellowfin offers such a single BI platform. Its long-term emphasis on providing a well-managed data environment—together with an agile, easy-to-use analytics platform—puts it in a good position to meet these modern BI data governance needs. Yellowfin further provides an extensive set of collaboration, task management, and metadata functionality that directly and equally support business users, data experts and IT staff. Paired with a comprehensive set of data governance policies and processes, Yellowfin offers a strong base for the implementation of a viable and valuable BI data governance approach.

Treating decision making as a process is key to the creation of high-quality data and the delivery of reliable decisions with real business value.



Dr. Barry Devlin is among the foremost authorities on business insight and one of the founders of data warehousing, having published the first architectural paper on the topic in 1988.

With over 30 years of IT experience, including 20 years with IBM as a Distinguished Engineer, he is a widely respected analyst, consultant, lecturer and author of the seminal book, “Data Warehouse—from Architecture to Implementation” and numerous White Papers. His new book, “*Business unIntelligence—Insight and Innovation Beyond Analytics and Big Data*” (<http://bit.ly/BunI-Technics>) was published in October 2013.

Barry is founder and principal of 9sight Consulting. He specializes in the human, organizational and IT implications of deep business insight solutions that combine operational, informational and collaborative environments. A regular tweeter, @BarryDevlin, and contributor to numerous publications, Barry is based in Cape Town, South Africa and operates worldwide.



Brand and product names mentioned here are trademarks or registered trademarks of Yellowfin and other companies.

Endnotes

- ¹ Gartner, “*Magic Quadrant for Business Intelligence and Analytics Platforms*”, 16 February 2017.
- ² TechTarget, “*Guide to managing a data quality assurance program*”, <http://searchdatamanagement.techtarget.com/definition/data-governance>
- ³ LinkedIn, “*Top 10 Most Expensive Spreadsheet Mistakes of All Time*”, <https://www.linkedin.com/pulse/top-10-most-expensive-spreadsheet-mistakes-all-time-carl-nixon>
- ⁴ Utah Desert News, “*Calculation error leaves schools short by \$25 million*”, <http://www.deseretnews.com/article/865553839/Calculation-error-leaves-schools-short-by-25-million>
- ⁵ CIO, “*Eight of the Worst Spreadsheet Blunders*”, <http://www.cio.com/article/2438188/enterprise-software/eight-of-the-worst-spreadsheet-blunders.html>
- ⁶ AccountingWeb, “*Spreadsheet error leads to Mouchel meltdown*”, <http://www.accountingweb.co.uk/practice/general-practice/spreadsheet-error-leads-to-mouchel-meltdown>
- ⁷ CIO, “*Eight of the Worst Spreadsheet Blunders*”, <http://www.cio.com/article/2438188/enterprise-software/eight-of-the-worst-spreadsheet-blunders.html>
- ⁸ LinkedIn, “*Four Dirty Data Horror Stories – When Good Data Goes Bad*”, <https://www.linkedin.com/pulse/dirty-data-horror-stories-when-michael>
- ⁹ Data blueprint, “*Data Horror Stories: Pregnant Men and a Psychic Retailer*”, <http://www.datablueprint.com/data-horror-stories-pregnant-men-and-a-psychic-retailer>
- ¹⁰ HealthIT Interoperability, “*Data Governance Failures Exposed in VHA Eligibility System*”, <http://healthitinteroperability.com/news/data-governance-failures-exposed-in-vha-eligibility-system>
- ¹¹ TechTarget, “*Data quality management: Problems and horror stories*”, <http://searchdatamanagement.techtarget.com/feature/Data-quality-management-Problems-and-horror-stories>
- ¹² Devlin, Barry, “*Business unIntelligence*”, (2013), Technics Publications LLC, http://bit.ly/BunI_Book