

TDWI RESEARCH

TDWI CHECKLIST REPORT

SEVEN STEPS TO MAKING BIG DATA ACCESSIBLE TO EXECUTIVES

By Philip Russom



Sponsored by



tdwi.org

tdwi

MARCH 2013

TDWI CHECKLIST REPORT

SEVEN STEPS TO MAKING BIG DATA ACCESSIBLE TO EXECUTIVES

By Philip Russom



1201 Monster Road SW, Suite 250
Renton, WA 98057

T 425.277.9126
F 425.687.2842
E info@tdwi.org

tdwi.org

TABLE OF CONTENTS

- 2 **FOREWORD**
- 2 **NUMBER ONE**
Don't be a deer in the headlights.
- 3 **NUMBER TWO**
Treat big data as the egg, not the omelet.
- 4 **NUMBER THREE**
Practice data democracy.
- 5 **NUMBER FOUR**
Dress your data to the nines.
- 5 **NUMBER FIVE**
Let the kids play.
- 6 **NUMBER SIX**
Make hooking up easy.
- 6 **NUMBER SEVEN**
Get your head in the clouds.
- 7 **ABOUT OUR SPONSOR**
- 8 **ABOUT THE AUTHOR**
- 8 **ABOUT THE TDWI CHECKLIST REPORT SERIES**
- 8 **ABOUT TDWI RESEARCH**

© 2013 by TDWI (The Data Warehousing Institute™), a division of 1105 Media, Inc. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. E-mail requests or feedback to info@tdwi.org. Product and company names mentioned herein may be trademarks and/or registered trademarks of their respective companies.

FOREWORD

For decades, executives have complained that they're drowning in data, yet lack focused insights. If executives don't want to get lost in the sea of information, they need BI platforms that leverage big data in order to gain focused insights.

On the one hand, many executives today have relatively good technical skills when given the right tools for business intelligence (BI) and data warehousing (DW). This growing class of executives wants to do a certain amount of data exploration, sandbox development, data visualization, and other self-service BI tasks (as described later in this report). On the other hand, the majority of executives still depend on technical personnel, who must package and present big data and findings drawn from it in forms that executives are accustomed to.

In both cases, organizations need a simpler, more cost-effective approach to exploring big data and turning its raw information into business insights that are easily consumed by executive users—and nothing is more critical to that end than the executive dashboard.

This TDWI Checklist Report describes a BI environment and its attendant best practices that can greatly simplify executives' access to big data and enhance their consumption of insights and metrics that come out of big data and similar data sets. The sections of this report list and discuss the highest priorities for such an environment and the successful solutions built atop it.



NUMBER ONE

DON'T BE A DEER IN THE HEADLIGHTS.

Many technical and business users are shocked and overwhelmed by the immensity of big data, which gives them analysis paralysis. After all, most organizations lack experience with the multi-terabyte volumes typical of big data. But big data is big for reasons other than volume; it comes from many sources (both new and traditional) and it's in multiple structures (from relational to human language text). The size and complexity can be a barrier to finding just the right information needle in multiple haystacks and then consolidating these into a presentation that empowers executive decision making. Put all of this together, and it's natural that big data can freeze some organizations, like a proverbial deer in the headlights.

If you find yourself in such a precarious position, here are some things you can do to get out of the headlights:

1. **Find a safe but rewarding starting point.** To this end, TDWI has seen a number of organizations ask a pair of questions:
 - What are the important facts about our business that we have been unable to quantify properly?
 - Among these, which could be quantified using new sources of big data, combined with traditional data sources, so executives can take fast and accurate action?

Ideally, these questions lead to a straightforward, value-adding analytic application that incentivizes people to solve a problem that was previously unsolvable. Even better, there is a high likelihood that the outcome will improve the business.

For example, many organizations still lack credible metrics for customer satisfaction, customer sentiment, customer churn probability, and other customer metrics. The problem is that data for these is strewn across many applications (representing many customer touchpoints), as well as new, mostly untapped big data sources (e.g., the Web, social media, partners, third parties). Credibility for metrics comes only from bringing together all data from all sources. Doing so can get you and the other deer out of the headlights, and then lead you to similar successes elsewhere.

2. **Employ a platform designed for today's new requirements.** Most organizations in this situation have already deployed a business intelligence (BI) platform and a data warehouse (DW). These are killer platforms for traditional reporting, and therefore must be retained and nurtured. But such mature platforms were designed before the advent of new requirements

for big data, data visualization, advanced analytics, real-time operation, and agile BI.

One of the trends TDWI has tracked in recent years is diverse portfolios of tools for BI/DW and related disciplines. The reason for growing diversification is that retrofitting additional technology onto the core BI/DW technology stack is often difficult or unsuccessful. Instead, more and more organizations choose to complement the core stack with additional tools and platforms that were designed from the beginning for new requirements in real-time analytics, big data, visualization, and agility. Hence, if you need to take executive decision making to a higher level by satisfying new requirements, you should consider an additional tool or platform designed for that purpose.

- 3. Employ a platform that you can use as soon as possible.** A prolonged implementation of a traditional BI tool or platform runs the risk of putting you and the other deer back in the headlights. Look for platforms that address the new BI requirements, but out of the box with little or no system integration or development. A number of vendors offer standalone databases, departmental analytic applications, or cloud-based BI services that you can set up fast and get results from immediately; plus, these integrate with existing BI/DW stacks and other traditional applications.

NUMBER TWO

TREAT BIG DATA AS THE EGG, NOT THE OMELET.

The information contained in big data is the bulkiest ingredient of any recipe for big data insight, and the ingredients demand a lot of work before the dish is ready to eat. Like the eggs that go into an omelet, big data must be cracked, whipped into shape, spiced, and cooked before a consumable meal results. Making an insightful omelet from big data eggs is a multi-step recipe:

- 1. Discover new facts about the business as you explore data.** Data exploration leads to discoveries within the data. But this won't happen if you're not prepared. Go into exploration with goals in mind, such as discovering new clusters (e.g., young customers with high household incomes), hidden costs (extra shipping demanded of some products or customers), or locations (insurance claims for flooding per neighborhood).
- 2. Interpret your discoveries to form actionable insights.** For example, a new cluster of customers is a sales opportunity. Newly discovered shipping costs mean that your customers or suppliers are not as profitable or reliable as you thought. A shift in incidences for insurance claims may call for adjustments to actuarial tables.
- 3. Express your insights in terms executives can understand.** A new cluster of customers could become a new segment in your customer segmentation analytic application. Newly discovered costs could be represented as a shipping cost metric that rolls up into an overall cost KPI. Both cases would be visible in executive dashboards.
- 4. Rinse and repeat.** Follow the above steps regularly to monitor for business opportunities and problems or to scout for new data sources, or when users ask new business questions or request revisions to existing dashboards and other BI products.

✓ NUMBER THREE

PRACTICE DATA DEMOCRACY.

Too often, business analysts, data scientists, and similar power users share their analytic epiphanies with a very short list of peers and management. Those insights should be shared across departments if the enterprise is to get the fullest value from big data.

For example, analytics may show the root cause of a new form of customer churn. That’s when you develop and regularly quantify churn metrics to spot the trend if it returns. As another example, let’s say you’re tapping a new source of big data and, for the first time, you are able to reliably calculate customer sentiment. If you want executives to easily understand and track that sentiment, a business analyst or technical user should develop a series of metrics representing various aspects of sentiment that roll up into key performance indicators (KPIs).

Turning an ad hoc analytic insight into a repeatable BI consumable is called *operationalization*. This usually results in a polished deliverable, such as a dashboard, sandbox, metric, and so on (see Figure 1). Operationalization has distinct benefits.

Operationalizing the outcome of discovery analytics expresses the analytic epiphany in practical terms that many users can understand. Once that happens, the data and its meaning are democratized as they are shared across an enterprise. Democratized insights help the organization gain full business value from its investments in BI and related practices. Likewise, democratized insights (as opposed to those that are unique to a department or other group) keep diverse teams from contradicting each other with competing data sets or arguing over whose data set is best. And finally, operationalization takes a one-time analytic insight and turns it into a metric that can be tracked over time.

Ideally, popular operationalized BI deliverables should be on a platform that enables personalization. Such a platform helps the many diverse managers consuming a deliverable to tailor it to their own unique needs, in a self-service fashion. After all, a C-level executive needs a different view than a line-of-business manager or marketing director does. Democratizing big data this way spawns a social net of executive users who then learn from each other, sometimes in a viral fashion.

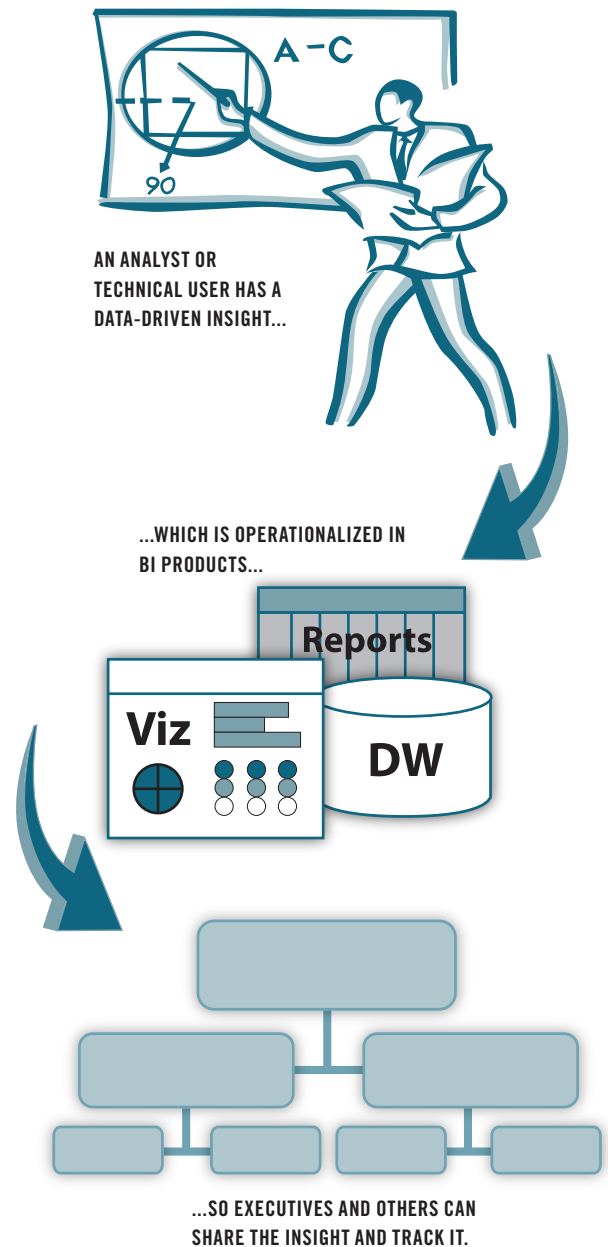


Figure 1. Operationalization.

 **NUMBER FOUR**

DRESS YOUR DATA TO THE NINES.

Executives typically don't have time to struggle with a difficult graphical user interface (GUI) or pore over every detail in a report. They need high ease of use in the GUI and targeted focus in the content. The executive dashboard has become *de rigueur* as the preferred style of reporting and analytics for most managers because the dashboard is the epitome of user friendliness and the poster child of targeted—even personalized—content. And the attractive, visual nature of dashboards enables you to dress your data to the nines, which boosts user adoption and productivity.

Note that executive dashboards vary a lot, and yet certain capabilities stand out as being more conducive than others to executives accessing big data:

- **Business performance management (BPM).** This practice identifies business entities or their characteristics that should be quantified as metrics, such as the metrics for customer attributes, sentiment, and cost mentioned earlier in this report. A good dashboard is lean, meaning that you can choose to display only the metrics that you need to monitor. This “manage by exception” method enabled by BPM is especially useful with big data, because you often need to ignore the bulk of big data and focus on truly actionable information.
- **Personalization per named user or group.** Every executive is different as a person and each has a unique management mandate. Instead of the distracting and time-consuming information glut of the average report (designed for multiple people), a personalized dashboard provides highly actionable information, thereby enhancing executive productivity and accuracy.
- **Advanced data visualization.** Dashboards are heavily visual, using graphic elements instead of actual numbers whenever possible. This ranges from simple pie, bar, and line charts to advanced forms of data visualization, such as hot maps, histograms, scatter plots, and hyperbolic trees.
- **Device independence.** Executives are on the move, whether traveling on business or playing a round of golf. So the truly modern executive dashboard should be accessible from a PC, smartphone, tablet, or any browser, and the functionally and GUI presented across these devices should be fairly uniform, if not identical.

 **NUMBER FIVE**

LET THE KIDS PLAY.

As business analysts and technical users explore big data, they inevitably discover data subsets of interest, then extract these for further analysis by business executives and other users. At this point, analytic users no longer need access to the full mass of big data and the full processing resources of its hefty platform. Instead, analytic users need a different environment where they can drill into and analyze the data set just developed via data exploration. This is true whether the analytic user is a business analyst working independently or executives analyzing data sets created for them by technical personnel. All of these users need a special place where they can play with data.

The secondary environment just described is often called an *analytic sandbox*. A sandbox can take different forms. It could be a collection of data structures within a larger database or data warehouse, a standalone analytic database management system, an in-memory database, or some kind of data management service in a cloud. For executive users, a sandbox must also support the dashboards and metrics required of modern management methods.

The analytic sandbox offers a number of benefits for executive users. Most executives are monitoring an enterprise and making decisions based on data from multiple departments. The catch is that the data and its processing are flung across many departmental IT systems. The sandbox can be a collection point for pulling together disparate data in a more nimble fashion than the average BI/DW technology stack can.

A sandbox sounds a bit like a data mart, doesn't it? In a way, the analytic sandbox is the twenty-first-century equivalent of a data mart, but with significant improvements. Sandboxes vary, but the best ones provide the following:

- Governable environment, unlike standalone data marts
- Expiration date, to free up resources hogged by forgotten data sets
- Query governor, to isolate runaway queries
- Collaborative mechanisms, so that multiple executives and other users can play with a data set and share their views of it

NUMBER SIX

MAKE HOOKING UP EASY.

The success of data exploration, discovery, sandboxes, and dashboards depends on collecting and consolidating just the right data from multiple sources. This, in turn, demands a BI platform that can access and integrate with many operational platforms.

The four most important traits of an optimal platform are:

- **Interfaces to many data sources.** A sandbox and other analytic data sets contain a mix of data from multiple sources, such as big data sources (Web logs, machine data, social media) and traditional sources (customer data from CRM systems, recent call detail records from a telco application, supplier information from a procurement application). Sometimes data comes from external sources, such as third-party data providers, business partners, and government agencies (e.g., census data). Interfaces to cloud-based applications (such as Salesforce.com) become increasingly important.
- **Fast and scalable data loading.** With so much data coming from so many sources, scalable bulk loading is required to get big data into your BI environment. Inevitably, some of this data is time sensitive, so some interfaces need to collect and process data in real time.
- **Integration with traditional BI/DW data.** Data may also come from a data warehouse, which provides a more complete, historic view to complement the real-time feed or recent vintage common of other data sources.
- **Ease-of-use for data access.** A business analyst—much less an executive—doesn't have the time to wait for a technical aide to access and load data repeatedly. Modern BI platforms assume self-service for many user types (from technical to business). That self-service then requires ease of use from the platform's GUI for data access.

NUMBER SEVEN

GET YOUR HEAD IN THE CLOUDS.

Big data, analytics, data exploration, fact discovery, dashboard content, data visualization, real-time data—all of these new BI requirements are moving targets that also change size and shape constantly. The constant flux is due to the iterative nature of analytics and BI development and the accelerating speed of business.

For example, data exploration, by nature, leads the user to access one data collection after another. Discovering insights involves merging data from multiple sources, and the combination of sources evolves as the analyst searches for the right combination. This means that large volumes of data are collected and deleted, plus merged one way, then another. Some forms of big data stream 24/7, as with machine data and social media; other forms are updated daily or multiple times intraday, as with Web logs and applications data.

Accommodating the strenuous and evolving new requirements of BI and big data requires an elastic and agile platform. This is why various types of clouds are proving to be very useful for managing, analyzing, and reporting big data. In recent years, TDWI has seen a steady increase in so-called *cloud BI*, defined as any kind of BI implementation on any kind of cloud. That's because cloud-based BI services offer a number of advantages:

Clouds bring much-needed elasticity to BI technologies. The virtualized server farms of a cloud (regardless of cloud type) elastically provision CPU and storage resources as data exploration, analytic sandboxes, analytic processing, and real-time data demand them. As these workloads subside, the cloud “automagically” recoups platform resources and allocates them to other workloads. This assures linear scalability and high performance for BI processes without the need for tedious capacity planning or overbuilding for peak loads.

Public clouds are gaining adoption for BI usage. TDWI surveys show that many BI users are concerned about the security of their data on a public cloud (sometimes called off-premises or third-party clouds), plus the technical challenge of moving data onto and off clouds. Yet TDWI surveys also show user organizations moving beyond these concerns, as early adopter users increase in number and as BI best practices adapt to cloud platforms.

Adopting a cloud-based BI service is quick and easy compared to traditional implementations. In just a few days, you can acquire a software-as-a-service license from a third-party cloud BI provider, integrate data with it, and start using BI functionality you didn't have before. A traditional implementation for similar functionality typically takes weeks or months, at best. This is the kind of kick-start that organizations need when trying to get out of the big data headlights.

Cloud BI services usually satisfy BI's newest requirements.

Some are designed for diverse big data, exploding and contracting data sets, advanced analytics, device-independent dashboards, and real-time operation. Instead of retrofitting these functions onto an existing BI/DW technology stack, it might make more sense to complement that stack with more modern cloud-based BI services.

A cloud service for BI is a solid outsourcing strategy. It can be cost-effective compared to building your own cloud for BI or extensions to your BI/DW technology stack. TDWI regularly sees cloud BI services paid for via modest departmental budgets. Adopting a cloud service for BI can reduce risk by avoiding the complex system integration and development work typical of many in-house BI implementations. The cloud or its provider handles administrative functions, so you need not handle your own high availability, data backup, load balancing, and most database administration work.

ABOUT OUR SPONSOR



www.domo.com

When Josh James was longtime CEO at Omniture, he had data sources galore, but he didn't have a way to bring them together to tell the full story of what was happening in the company. Business intelligence was cumbersome, time-consuming, and outdated. He wanted a single dashboard that connected all of his data sources and reported up-to-date metrics. And since he was always on the go, he expected the same experience on any mobile device he used.

Haven't seen one of those lying around recently? Neither did we. Until now.

With more than \$125 million in funding and backing from the biggest players in SaaS, Josh has created a new form of business intelligence for the executives that actually use it. The result is a dashboard that enables business users to find the information they need, at the moment they need it. We call it Domo.

Domo's executive management platform gives you all of your critical business data, in one place, the moment you need it. It's all there—every stat you want to see from every part of the company, completely updated, and on any device.

Suddenly, the numbers start making sense. And so does Domo.

ABOUT THE AUTHOR

Philip Russom is director of TDWI Research for data management and oversees many of TDWI's research-oriented publications, services, and events. He is a well-known figure in data warehousing and business intelligence, having published over 500 research reports, magazine articles, opinion columns, speeches, Webinars, and more. Before joining TDWI in 2005, Russom was an industry analyst covering BI at Forrester Research and Giga Information Group. He also ran his own business as an independent industry analyst and BI consultant and was a contributing editor with leading IT magazines. Before that, Russom worked in technical and marketing positions for various database vendors. You can reach him at prussom@tdwi.org, [@prussom](https://twitter.com/prussom) on Twitter, and on LinkedIn at [linkedin.com/in/philiprussom](https://www.linkedin.com/in/philiprussom).

ABOUT THE TDWI CHECKLIST REPORT SERIES

TDWI Checklist Reports provide an overview of success factors for a specific project in business intelligence, data warehousing, or a related data management discipline. Companies may use this overview to get organized before beginning a project or to identify goals and areas of improvement for current projects.

ABOUT TDWI RESEARCH

TDWI Research provides research and advice for business intelligence and data warehousing professionals worldwide. TDWI Research focuses exclusively on BI/DW issues and teams up with industry thought leaders and practitioners to deliver both broad and deep understanding of the business and technical challenges surrounding the deployment and use of business intelligence and data warehousing solutions. TDWI Research offers in-depth research reports, commentary, inquiry services, and topical conferences as well as strategic planning services to user and vendor organizations.