

EXCLUSIVELY FOR
TDWI PREMIUM MEMBERS

TDWI RESEARCH

FOURTH QUARTER 2014

Ten Mistakes to Avoid

When Democratizing BI and Analytics

By David Stodder

tdwi.org

tdwi

Ten Mistakes to Avoid When Democratizing BI and Analytics

By David Stodder

FOREWORD

Data “democratization” is a high priority for many organizations as they seek to enable data-driven decisions and actions throughout their enterprises. Firms want to expand the user base for business intelligence (BI), analytics, and data discovery so more users can move beyond uninformed, “gut feel” decision making. Users are clamoring for tools that will improve upon canned reports, limited spreadsheet views, and the general information chaos that can make data more of a burden than a blessing.

Today, BI, data discovery, and analytics technologies are maturing to enable users to work with data without the usual hand-holding by IT developers and data managers. The danger, however, is that those easy-to-use, self-service technologies will lull organizations into overlooking important concerns as they deploy tools far and wide.

This *Ten Mistakes to Avoid* focuses on helping organizations ensure satisfaction as they democratize BI and analytics. It recommends a balanced approach for meeting user needs while addressing the necessities of data governance and management.

ABOUT THE AUTHOR

David Stodder is director of TDWI Research for business intelligence. He focuses on providing research-based insight and best practices for organizations implementing BI, analytics, data visualization, and related technologies and methods. He is the author of TDWI Best Practices Reports on mobile BI and customer analytics in the age of social media, as well as TDWI Checklist Reports on data discovery and information management. Stodder has provided thought leadership on BI, information management, and IT management for over two decades. He has served as vice president and research director with Ventana Research, and he was the founding chief editor of *Intelligent Enterprise*, where he served as editorial director for nine years. You can reach him at dstodder@tdwi.org.

© 2014 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.

MISTAKE ONE:

FAILING TO OBSERVE USERS' CURRENT STATE BEFORE GATHERING REQUIREMENTS

As the saying goes, you never get a second chance to make a good first impression. Organizations that seek to democratize BI and analytics need to get off to a good start. Although gathering user requirements is traditionally the first step, it would be wise to learn first about users' current state of data access, analysis, reporting, visualization, and sharing. By absorbing this knowledge, business analysts and developers will have a more nuanced understanding when they begin to gather project requirements.

Organizations are keen to break down the walls between users and developers to improve communication and development efficiency. Organizations can move faster if users and developers work together. Agile methods are becoming popular (and will be discussed later). However, even if organizations are not employing agile methods, it is critical before a project officially begins for users and developers to sit down together and look at the data, including who owns it and how it will be used in making decisions.

Most users who are not implementing BI, analytics, and data discovery tools work with spreadsheets for data access, analysis, and sharing. They depend on power users (aka "spreadsheet jockeys") who understand the data, can set up calculations and pivot tables, and can perform data preparation, filtering, scripting, and more. Be sure to get to know the power users and learn all you can from them.

Observe whether the users' culture (and, in particular, the executive leadership) seems ready to challenge "gut feel" decisions. You may discover that people and process obstacles to data-driven insights affect the success of BI and analytics projects. If the leadership is not open to the potential impact of more advanced data analysis on decision making, be prepared to move slowly. You will need to gain users' confidence through deliverables that provide immediate benefits. These first impressions will pay dividends as the project matures.

MISTAKE TWO:

NOT KNOWING ENOUGH ABOUT THE DATA

BI and analytics are nothing without good data. For democratization initiatives to succeed, it is as important to get to know the data as the users and what they plan to do with it. In many organizations, the *speed* data is growing is outpacing IT's *knowledge* of the data. Too many organizations neglect this task, making data chaos a barrier to success. BI and analytics projects grind to a halt because users are uncertain about data's quality and completeness.

To support democratization, organizations must implement more consistent and automated methods (such as profiling and data discovery) to increase their knowledge of the data and how data elements are related within and across sources. Certainly, regulatory requirements have been driving organizations to know more about their data so they can respond to audits and protect sensitive data. Now, BI and analytics are becoming drivers; by applying consistent methods and software tools where possible to learn about the data, organizations can save information management costs and increase user satisfaction.

The most common obstacle to realizing higher value from data is that organizations are faced with disconnected data silos and cannot easily bridge them to provide users with a single view of integrated data sources. Thus, the first step in establishing a single view is with "data discovery" to find out what data is located where and how it is currently being used. You can discover and document how data elements and their attributes are related within and across data stores, which is critical to building the single view. You should discover and document all extraction, transformation, and loading (ETL) operations; this step helps you weed out unnecessary or unauthorized ETL routines. All of this knowledge should be documented in a metadata repository, business glossary, or similar knowledge-capture system.

MISTAKE THREE:

FAILING TO EXPAND USER ACCESS TO NEW BIG DATA SOURCES

Organizations today are under pressure to enable access to multiple structured data sources as well as big data—semi-structured and unstructured sources such as text, social media data, and streaming machine data. This is particularly important for personnel in marketing, e-commerce, and customer sales and service functions who are interested in analyzing customer behavior across channels. If your organization is seeking to democratize BI and analytics to these users, you need to provide access to a wider range of data.

Business leadership should identify relevant sources and work with IT to determine which users and functions need to view which data sources. The tried-and-true strategy is to consolidate data from multiple sources into an enterprise data warehouse (EDW) that serves various users and business functions. Tasks for data mapping, profiling, discovery, ETL, and quality improvement then focus on what will go into the EDW.

However, this approach can be slow and impractical. Organizations should complement their consolidation strategy with data federation or virtualization. These modes use global metadata or master data to access “data in place” without having to move it to a central store. This is useful if data cannot be moved for regulatory reasons. The tasks to prepare the data and ensure its quality can be performed either through middleware or at the sources themselves.

Hadoop technologies offer another alternative. Organizations can implement Hadoop files to store “data lakes,” where the data is not restricted (as it is with an EDW) to just the acceptable types or structures. Working with raw data, users can be freeform in applying analytics. If users identify data that is worth moving into the EDW, IT can create ETL routines to make that happen.

MISTAKE FOUR:

FAILING TO BE PROACTIVE ABOUT SELF-SERVICE BI AND ANALYTICS

The heart of any democratization effort is putting the power of data into the hands of nontechnical users. These users need technologies that can take them beyond the limits of spreadsheets, canned reports, and static dashboards. With self-service technologies and methods, they can drive data access and analysis themselves and be free from having to ask IT for every new report and dashboard—or changes to existing ones. Common among self-service capabilities sought by users is personalized interactivity with the data so they can dig deeper and perform what-if discovery analysis on their own.

IT organizations often resist the notion of self-service BI and analytics, fearing that uncontrolled access and analysis will create problems with data management, performance, security, and quality. They prefer users stay within the EDW system and work with IT developers to add new data or functionality. Then IT can “keep bad things from happening,” to use the words of one IT chief, and stick to documented service-level agreements.

However, by not deploying self-service BI and analytics tools—or at least endorsing business users’ acquisition of them—IT is likely sowing the seeds of problems down the road. If users cannot get what they need, they will resort to other measures, including shadow or rogue IT systems. These (typically) standalone data marts will not conform to enterprise BI or EDW standards. It is better if IT is involved in managing self-service technology deployment.

Organizations should make deployment of self-service BI and analytics a core part of democratization. Technologies are advancing rapidly to enable users to personalize how they access and interact with data. By being proactive about self-service, organizations can manage deployments effectively and avoid some of the downsides caused by users installing tools without any oversight.

MISTAKE FIVE:

FAILING TO TAKE ADVANTAGE OF DATA VISUALIZATION

Graphical interaction with data is fast becoming the expected norm for users, from executives to frontline personnel. Visualization should be a key concern when democratizing BI and analytics. Good data visualization can be a huge contributor to making smarter decisions and improving productivity with data. Poor visualization, on the other hand, can mislead users and make it more difficult for them to overcome the data onslaught. Unfortunately, many organizations think of data visualization as mere “eye candy” and do not devote enough attention to it.

Many tools now offer diverse libraries of visualizations beyond the standard bar and line charts. Users can choose from scatterplots, heat maps, tree maps, geographic maps, variance charts, and more. They can also put data in motion, which can be an effective technique for “storytelling” with data and getting viewers’ attention. The latest versions of tools are improving how they work with underlying data integration systems to keep the complexity of data access hidden from users.

More advanced tools let users move easily from one visual object to another, offering an immersive experience for discovering data relationships, filtering views, and making correlations and comparisons without dropping out of the visualizations to write queries. Users can pick up visual objects and embed them in other applications, such as sales-force and contact-center management.

It should be a top priority in democratization to improve data visualization and visual analysis, especially for users who have been limited to spreadsheets and simple reporting. Organizations should be careful to match visualization capabilities to the users’ roles and activities. Users can also benefit from training to help them avoid clutter and improve clarity, including how to personalize the appearance of visual objects in dashboards and other portals.

MISTAKE SIX:

FAILING TO BRING ADVANCED ANALYTICS TO MORE USERS

Along with democratization, one of the biggest trends today is advanced analytics. These types of analytics include data mining, machine learning, and predictive, geospatial, text, and customer behavior analytics. Often the focus is on employing algorithms to derive business value from sources such as online clickstreams, social media, and contact-center interactions. Increased collection and availability of big data, including emerging machine data, sensor, and Internet of things sources are also factors driving interest.

Predictive analytics in particular could be helpful to users in sales, marketing, supply chain management, and finance, among others. Predictive analytics enables organizations to examine data trends to explore what may happen next. Predictive analytics can enable users to look into the unknown future and into the likelihood of outcomes based on the performance of a predictive model's variables or measurements of reactions to certain stimuli. If supported by a strong analytical database system, predictive analytics has the power to deliver insights that traditionally could take weeks or months of analysis with just BI tools or spreadsheets. Decision cycles for marketing campaigns, fraud detection, risk management, and financial modeling could therefore be shortened and made more efficient.

Another key form of advanced analytics is text analytics. It covers a range of technologies and practices for analyzing text, extracting relevant information, and transforming sources by applying structure so that analysis can be repeated and adjusted over time. Text analytics enables users to expand their view to the volumes of information generated by human interactions that occur in the course of business. Insights based on content can provide rich context around the numbers that appear in BI reports.

If organizations cannot provide any advanced analytics capabilities for users, they should at least offer the ability to consume the output of advanced analytics performed by data scientists and analysts.

MISTAKE SEVEN:

FAILING TO MAKE SPEED A PRIORITY

As more users employ BI and analytics applications, demand for fresher data and fast query response will grow, especially from users in operations. To serve them effectively, organizations will need to shorten the time between an event or transaction and when that data can be accessed and consumed for reporting, analysis, and presentation. Some users want to push toward real time; they seek to set up predictive models and scoring algorithms to run on real-time, streaming data so they can discover patterns and anomalies for customer intelligence, fraud detection, and other needs.

Latency and query response in BI and analytics applications depend considerably on underlying data management and systems. It can be difficult to determine which component is most responsible, whether a query was written or optimized poorly, or if the data model is flawed. To reduce latency and improve query response, organizations will need to examine the entire stack.

Technology options today that are critical to shrinking the time to insight include:

- **Data virtualization (or federation).** Users can access multiple, distributed systems through a single layer without having to write queries to each system individually.
- **Columnar databases.** This alternative to the traditional row-oriented DBMS can accelerate query response and data retrieval because the system only needs to look at specific columns rather than retrieve entire rows.
- **In-memory computing.** BI and analytics applications can take advantage of a very large amount of memory to supply users with relevant data without having to go to disk for each request.
- **Cloud and software-as-a-service (SaaS).** Organizations can create analytic platforms or “sandboxes” in the cloud to service BI and analytics users. SaaS offerings can eliminate delays caused by traditional on-premises development.

MISTAKE EIGHT:

FAILING TO APPLY AGILE METHODS TO IMPROVE QUALITY AND REDUCE TIME TO VALUE

To realize the business benefits of democratizing BI and analytics, organizations need more agile and flexible development projects that can withstand continuous changes to user requirements, address complexity in data demands, and scale up to handle more projects and users. TDWI Research sees a strong trend today toward adoption of agile software development methods for BI, analytics, and data warehousing projects. Business and IT managers are questioning traditional “waterfall” development; they see it as a major reason why projects take too long and fall short of delivering value.

Agile development methods function with shorter, incremental cycles. The waterfall methodology typically delivers no value until the project’s end. With agile, users are not separate from developers; they are closely involved. Users work together in teams with developers, serving as “product owners” to define business needs. Agile’s scrum and similar iterative, team-oriented methods embrace a key principle of the agile manifesto: “Welcome changing requirements, even late in development; agile processes harness change for the customer’s competitive advantage.”

Long waterfall cycles carry the danger of baking into the applications design mistakes and bad code that can remain undiscovered until users encounter them in the final release. Agile’s shorter, more incremental cycles allow projects to fail fast so that errors can be discovered and corrected more quickly. Although agile presents some challenges for BI, analytics, and data warehousing projects, the methods fit for many applications where continuous refinement is to be expected. BI and analytics applications need to accommodate new data and new requirements.

Organizations should examine whether traditional waterfall development methods are why BI, analytics, and data warehousing projects are moving too slowly and not delivering value. They should consider whether agile might be a better fit for their objectives to expand BI and analytics across the organization.

MISTAKE NINE:

LEAVING MOBILE OUT OF YOUR DEMOCRATIZATION STRATEGY

TDWI Research finds that most organizations anticipate their users will spend an increasing amount of time on their mobile devices—tablets in particular—implementing or accessing BI and analytics applications. Many organizations are concerned about data security, which will limit what users can do on mobile devices. However, as the devices become sufficiently robust to support native visual BI and analytics, pressure for functionality will grow.

TDWI Research also finds that mobile BI and analytics are most successful in organizations where mobile is part of a bigger strategy. The alternative is that users will download applications on their own and use their own data, plus whatever snapshots they can pull from spreadsheets and desktop BI systems. This can lead to data quality, consistency, and security problems.

Organizations that have been successful in deploying mobile BI and analytics have discovered that users have increased visibility into business performance. Rather than deliver passive data reports in graphs or charts, they are using alerts, checklists, and call-to-action buttons to draw users' attention to changes in conditions or other important information. Leading tools are employing responsive design methods to tailor screens to the form factors of mobile tablets and smartphones. Developers must contribute by helping users select visualizations that work best on the particular device.

Organizations should develop a strategy to deploy mobile BI and analytics to customer sales, service, and support personnel. TDWI Research finds that meeting the needs of these frontline users is high on the mobile deployment priority list for the majority of organizations. The potential is great for upgrading the business value of customer interactions with personnel on the road at appointments, trade shows, and other engagements outside the office.

MISTAKE TEN: **NEGLECTING TO ESTABLISH A BALANCED DATA GOVERNANCE STRATEGY**

Organizations with little governance over what data users are accessing and using can encounter problems, including poor data quality and consistency and lax security. However, data governance that is unnecessarily tight will reduce the business value of data and most likely drive frustrated users outside the enterprise system to find their own solutions. Thus, organizations need to take data governance seriously as they democratize BI and analytics and establish a balanced approach.

Creating a data governance and/or center of excellence (CoE) committee can facilitate a collaborative relationship between business and IT leadership about data governance and how to apply it to the expansion of BI and analytics. Many organizations are employing such committees to encourage best practices and guide technology adoption. Business leadership from across the organization should be represented along with IT, data management, and development. The committee can articulate the business value of shared BI and data assets, including for regulatory data governance. This group can develop road maps for establishing company-wide data access and directing consolidation of unnecessary data silos.

Data quality should also be a focus of the data governance or CoE committee. In many cases, BI and analytics users and developers have little control over data quality unless the “garbage in, garbage out” issues are addressed at the sources. A joint committee can raise such issues and determine how to remedy the problems with the backing of business and IT executives.

To avoid mistakes with BI and analytics democratization, business and IT leadership should provide joint, visible, and sustained leadership of projects. Good leadership will instill confidence in users, allowing them to feel that they can rely on the quality of their BI and analytics applications.

ABOUT **TDWI**

TDWI, a division of 1105 Media, Inc., is the premier provider of in-depth, high-quality education in the business intelligence, data warehousing, and analytics industry. TDWI is dedicated to educating business and information technology professionals about the best practices, strategies, techniques, and tools required to successfully design, build, maintain, and enhance business intelligence and data warehousing solutions. TDWI also fosters the advancement of business intelligence and data warehousing research and contributes to knowledge transfer and the professional development of its members. TDWI offers a worldwide membership program, five major educational conferences, topical educational seminars, role-based training, on-site courses, certification, solution provider partnerships, an awards program for best practices, live Webinars, resourceful publications, an in-depth research program, and a comprehensive website, tdwi.org.



555 S Renton Village Place, Ste. 700
Renton, WA 98057-3295

T 425.277.9126
F 425.687.2842
E info@tdwi.org
tdwi.org