VIEW KEY INSIGHTS (HTTPS://WWW.GARTNER.COM/DOCUMENT/3074819/KEY-INSIGHT?REF=DDISP)

Eight Steps to Picking the Best Self-Service BI and Data Discovery Tool

FOUNDATIONAL **Refreshed:** 04 October 2016 | **Published:** 15 June 2015 **ID:** G00277712

Analyst(s): *Cindi Howson*

Summary

Business users are driving much of the new BI buying in search of the best self-service BI tool. This report provides an eight-step process to ensure buyers make an objective selection.

FOUNDATIONAL DOCUMENT

This research is reviewed periodically for accuracy.
Last reviewed on **4 October 2016**.

More on This Topic

This is part of an in-depth collection of research. See the collection:

SERIES OVERVIEW

How to Modernize Your Business Intelligence and Analytics Platform for Agility, Without Chaos (https://www.gartner.com/document/code/291630?ref=grbody&refval=3074819)

Overview

Key Challenges

The rise of data discovery and self-service BI tools is forcing BI leaders to revisit their BI tool portfolios, often meaning they augment BI platforms and existing investments with solutions from smaller vendors.

BI leaders may weigh requirements for governance and common semantic models higher than business users' requirements for agility, furthering the rift between business and IT and potentially sacrificing better capabilities for the sake of integration.

Data discovery tools look similar on the surface but have substantial differences in architecture and capabilities than traditional BI platforms and ad hoc query tools, making traditional RFP processes ineffective.

Data discovery deployments sprout at the departmental level with individual users often buying their own favorite tools, after which BI leaders struggle to set standards. Lack of standards and a plethora of overlapping tools raise BI cost ownership, lower support quality, stretch BI resources and introduce multiple versions of the truth.

Recommendations

BI leaders and buyers within business units should:

Mix and match modules from multiple vendors to provide users with the best tools.

Expand the BI tool portfolio beyond production reporting and ad hoc query tools to include data discovery.

Evaluate capabilities for governance and promotability of user-generated content, weighing the risks and benefits for greater agility.

Follow an objective selection methodology that focuses on the right tool for the right use case and user. Encourage user testing and evaluation to better understand product differences and potential.

Strategic Planning Assumptions

By 2018, data discovery and data management evolution will drive most organizations to augment centralized analytic architectures with decentralized approaches.

By 2017, 75% of IT organizations will have a bimodal capability; half will not handle this well, by either remaining overcontrolling or outright ignoring the risks of more agile approaches.

Introduction

As a business intelligence (BI) leader, you've struggled to keep pace with business demands for more data, more dashboards and modifications to existing reports — yet still there is a backlog of requests. Business users are buying their own BI tools, purchasing the likes of Tableau and QlikView as individual licenses and shadow IT initiatives. You can either ignore them or embrace them — ideally the latter — before the enterprise BI initiatives of the last decade degenerate into anarchy.

During the past 10 years, BI platform investments have centered largely on IT-led projects for large-scale systems-of-record reporting and enterprise data warehouses. These have tended to be highly governed and centralized, with IT-authored production reports or IT-designed semantic models to allow power users to author their own reports based on data in a well-modeled data warehouse and/or data mart.

Now, a wider range of business users are demanding access to interactive styles of analysis and insights without a lengthy IT-provisioning process. Data may come from multiple sources, including a data warehouse, independent data marts, a transactional system (directly), Web-based sources or external business partners.

Traditional ad hoc query tools allow a degree of self-service BI capabilities as long as IT has done some upfront provisioning. Data discovery tools (see Note 1), on the other hand, allow for self-service BI, but with less upfront IT involvement. For business users, data discovery is becoming synonymous with self-service BI, even though Gartner sees a range of self-service BI capabilities (see "How to Architect the BI and Analytics Platform" (https://www.gartner.com/document/code/265003?ref=grbody&refval=3074819)).

Gartner estimates that more than half of net new purchasing is data-discovery-driven (see "Market Trends: Business Intelligence Tipping Points Herald a New Era of Analytics" (https://www.gartner.com/document/code/264016?ref=grbody&refval=3074819)). Of BI-related inquiries conducted in the last year, data discovery inquiries were 70% higher than inquiries regarding reporting. [1] (#dv_1_analysis_of)

Analysis

Mix and Match Modules From Multiple Vendors to Provide Users With the Best Tools

Data discovery vendors often use a "land and expand" strategy to win the loyalty of business users and data crunchers. After individual users buy a few licenses, deployment grows rapidly as other users follow suit. This selling approach makes it all the harder for a BI leader to later set standards or advocate buying from an incumbent vendor.

As a software category, BI first emerged in the early 1990s and was synonymous with ad hoc or business query tools. With these tools, users could create their own queries by pointing and clicking at elements in a business metadata layer; users no longer needed to know SQL to get to their data. Over the years, BI capabilities have expanded to include production reporting, OLAP, dashboards, mobile and, most recently, data discovery (refer to "Magic Quadrant for Business Intelligence and Analytics Platforms" (https://www.gartner.com/document/code/270380?ref=grbody&refval=3074819)).

CIOs and BI leaders would like to standardize on a single vendor for all the BI modules to reduce IT complexity and leverage existing relationships and buying power. However, as new BI capabilities emerge, the larger BI vendors are rarely the first to expand their capabilities. Dashboards, for example, is a module that mega BI vendors only added from 2005 to 2008; prior to this, customers relied on solutions from pure-play vendors. The same is now playing out with data discovery, cloud BI, smart data discovery and big data discovery.

While all BI vendors now have data discovery capabilities as part of their BI portfolios, specialty vendors such as Tableau, Qlik and TIBCO have had a decade head start. Just how robust the data discovery capabilities compare from the larger BI vendors to the early innovators varies greatly from vendor to vendor.

Any software selection should be driven by functional requirements and strategic considerations. Customers who select data discovery tools that are weaker in functionality may be selecting based on strategic factors such as incumbent vendor relationship, cost and product roadmap. However, as one customer who initially tried this approach and later changed tact explained, "Selecting a tool that is only good enough is the difference between a somewhat satisfied user and a truly delighted user." History has also shown us that, over time, products with excellent capabilities may eventually replace "good enough" products. Weaker products may get phased out of a vendor's portfolio, either through acquisition or innovation.

Some business users may argue that every user and department should be allowed to use their personal favorite BI tool, much as the shift in standards for mobile devices has evolved into bring your own device (BYOD). The challenge, though, is that multiple BI tools with overlapping capabilities mean that the authors, consumers and experts who support them must also learn multiple tools. BI users then spend more time on the technology rather than on the ultimate value: the business insights. Further, the average initial license cost for a data discovery tool is $1,700 per user, [2] (#dv_2_an_online) substantially higher than that of a smartphone. Companies that manage their BI portfolios show a higher business impact than those that set no standards at all. [3] (#dv_3_bi_scorecard) However, because there continue to be major differences among data discovery tools, business users and BI leaders can make a case for multiple data discovery tools within the company, but with clear usage guidelines by use case, user segment and subject area. The difference is in having multiple tools for justified reasons, rather than lack of a proactive approach.

Conversely, BI leaders may wish for a single solution from a single vendor that meets all their BI requirements equally well, but this is not the current state of the BI market. Even with the data discovery subsegment, each product has its "sweet spot" and capabilities — and these change over time.

Expand the BI Tool Portfolio Beyond Production Reporting and Ad Hoc Query Tools to Include Data Discovery

Gartner advocates a bimodal approach to IT (see "Predicts 2015: Bimodal IT Is a Critical Capability for CIOs" (https://www.gartner.com/document/code/271840?ref=grbody&refval=3074819)). Bimodal IT refers to having two modes of IT, each designed to develop and deliver information- and technology-intensive services in its own way.

- Mode 1 is traditional, emphasizing scalability, efficiency, safety and accuracy.
- Mode 2 is nonsequential, emphasizing agility and speed.

The more mature modules of BI platforms have provided Mode 1 BI needs, whereas data discovery supports Mode 2. As discussed in the 2015 Gartner BI Summit Keynotes, companies need to support both modes, but data discovery is what's new (see "Business Intelligence and Analytics Summit Opening Keynote: Bridging the Analytical Divide" (https://www.gartner.com/document/code/277128?ref=grbody&refval=3074819)).

Data discovery tools help speed the time to insight in three main ways:

- Flexible access to new data sources in addition to a centrally provisioned data warehouse.
- Use of visual perception and advanced charts to highlight outliers and anomalies.
- Exploration and interactivity to discover root causes of problems, without the need for a report developer to explicitly define drill paths and prompts.

Faster time to insight has become a business imperative amid a global economy, the rise of digital business, the economic downturn and increased competition. Shifts in technology are also enabling and driving the need for more agile BI solutions, including big data, the Internet of Things, in-memory computing and cloud.

Customers who have deployed data discovery tools cite higher achievement of business benefits ("Survey Analysis: Customers Rate Their BI Platform Vendor, 2014" (https://www.gartner.com/document/code/262301?ref=grbody&refval=3074819)), user adoption and business impact. [4] (#dv_4_bi_scorecard)

Evaluate Capabilities for Governance and Promotability of User-Generated Content, Weighing the Risks and Benefits for Greater Agility

Business users want greater flexibility and agility to manipulate data and access new data sources. IT is in the difficult position of having to ensure data integrity, comply with Sarbanes-Oxley, and minimize the risk that decisions are made on misinterpreted or inaccurate data.

The business must decide what risks are acceptable. IT should act in a consultative, partner role on how the business can lower such risks. Integration with an existing semantic layer or existing BI reports is but one approach and possible requirement to consider. Business users should assess how valuable such integration and reusability is; if data has not been modeled according to their requirements in the past, such integration is not important. Similarly, if business users are trying to access new data sources not in a data warehouse, integration to existing semantic models also does not apply.

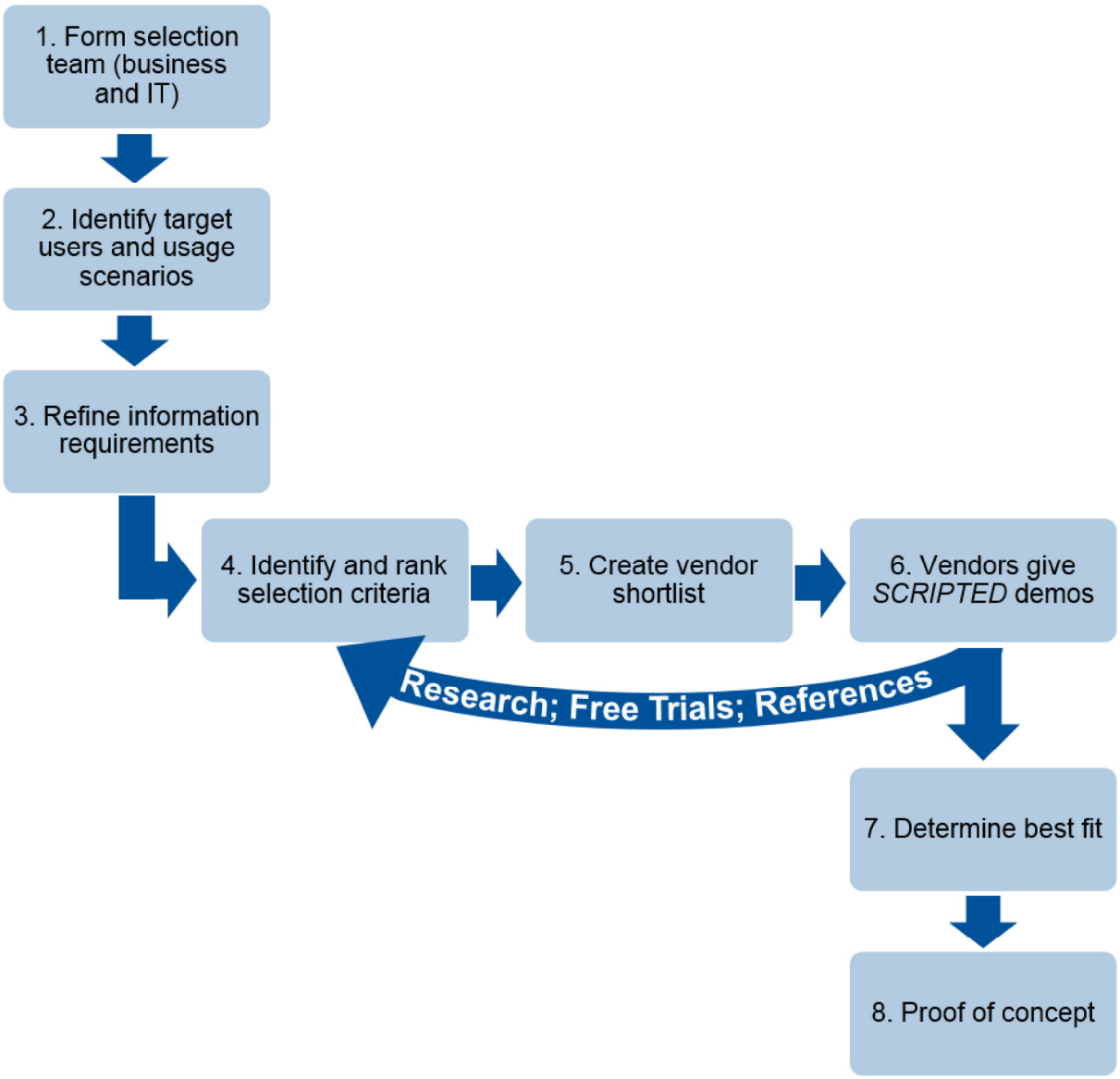
Alternatively, after-the-fact governance is another approach. With this workflow, the business user creates and manipulates the data according to their business requirements. As the data model is shared more extensively, it may be promoted back to the information portal with greater IT governance.

Follow an Objective Selection Methodology That Focuses on the Right Tool for the Right Use Case and User

Software selections can be fraught with hidden agendas and personal biases (see "Think Twice Before Changing Business Intelligence Tool" (https://www.gartner.com/document/code/270900?ref=grbody&refval=3074819)), particularly if pockets of users have dabbled with freemium versions of data discovery tools and formed opinions on their favorites. More challenging is when individual business units have purchased solutions and made substantial investments in licensing and skills that a BI leader or BICC is now trying to rationalize.

Follow this eight-step, objective evaluation methodology to remove the personal biases (see Figure 1) and ensure optimal investment for all stakeholders.

Figure 1. Data Discovery Tool Selection Process



Source: Gartner (June 2015)

1. Form the BI Selection Team

***Who*: Business and IT**

***Biggest mistake*: IT leading the initiative; business must drive**

***Deliverable*: Mission statement, scope, team members**

The selection team should comprise both business and IT representatives, from a cross section of business units and departments. Current BI application owners are prime candidates for serving on the core selection committee as they have unique insights on current successes and unmet needs. Giving them ownership in the selection process minimizes the risk of second guessing later. Business users who have experimented with and prototyped data discovery tools are also ideal candidates for the selection team. Instead of viewing shadow IT efforts as threats, BI leaders should tap these people and projects as experts who know what's needed in an ideal solution. At the same time, keep the selection committee small enough to be effective. The selection committee will elicit feedback from a larger user constituency to ensure buy-in. Whether or not a business user or a BI director should lead the team or have final decision authority will depend on a number of factors, including who has BI-buying budget and how well aligned the BICC is to the needs of the business (see "Business Intelligence Teams Need to Change With the Times" (https://www.gartner.com/document/code/270899?ref=grbody&refval=3074819)).

2. Define Target Users and Usage Scenarios

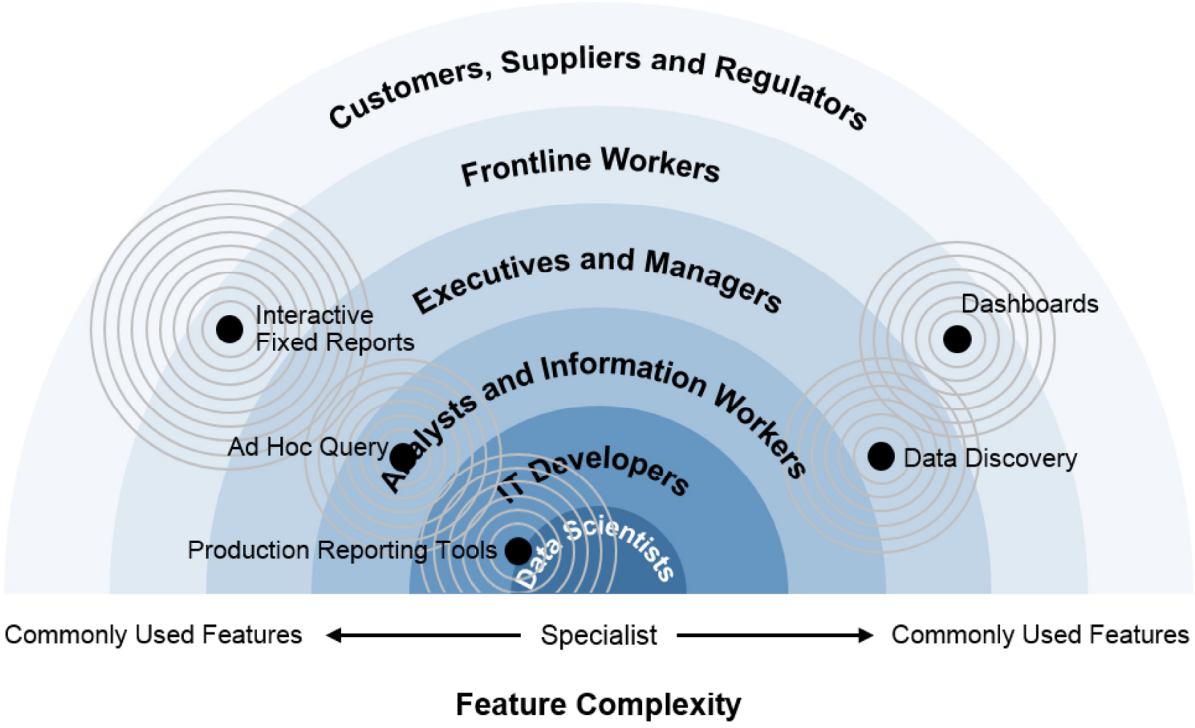
***Who*: Business**

***Biggest mistake*: Assuming everyone is a power user**

***Deliverable*: Usage matrix mapped to roles**

Despite BI's maturity, the industry does a poor job of understanding different user profiles and correspondingly matching product capabilities. Different user types require different tools or interfaces. Understanding these user segments is critical in managing the scope of your selection and resolving conflicting requirements. Figure 2 shows a spectrum of potential roles and the sweet spot (the dot) for each BI module. As you deploy certain modules beyond the sweet spot, you may encounter limitations and have decreased user satisfaction with the tool. Data discovery tools may have capabilities that span user segments that will evolve over time. For example, SAP Lumira initially lacked the ability to create dashboards but added that in a subsequent release; QlikView has limited capabilities for user-assembled dashboards but Qlik Sense does support user-defined dashboards. Also consider who will do the authoring; while most data discovery tools allow business users to author a dashboard, vendors may have an additional module for IT-authored dashboards.

Figure 2. Which Users Need Which Tools?



Source: Gartner (June 2015)

In the broadest sense, you will have two groups of users: producers, who author content, and consumers. However, you can further define these groups according to how, where and when they use the tool. Some data discovery users will want to extensively manipulate and mash data from multiple data sources; others may want only light interactivity via a tablet device. Refer to the "Critical Capabilities for Business Intelligence and Analytics Platforms" (https://www.gartner.com/document/code/270381?ref=grbody&refval=3074819) for additional use cases.

3. Refine Information Requirements

***Who*: Business and IT**

***Common mistake*: Assuming the data is all in one place or that all data models work the same with all data discovery tools**

***Deliverable*: High-level scope document or storyboards**

In our rush to see modern BI tools with their visual appeal, refining the information requirements is an often-overlooked step. Companies that have multiple BI tools know the harsh reality that every tool handles data and schemas slightly differently. Unless you simply want another pretty tool with "bad" data, you have to incorporate information requirements into your BI selection process. This does not detail source system to target mapping, but rather, considers how the data will be analyzed and potentially from which sources. For example, users may express the requirement to view sales with inventory to calculate day sales inventory (DSI) by various product groupings, time periods and regions. This single business question translates into a host of technical features such as:

Multipass SQL to query two fact tables in a data warehouse, or, if sales and inventory are in two source systems, the ability to blend data from multiple data sources

Semiadditive measures to aggregate inventory across product groupings but not across time periods

Geographic mapping

4. Define and Rank Selection Criteria

Who **: Business and IT**

Common mistakes **: Trying to replicate reporting capabilities already available; not envisioning or learning new possibilities with data discovery**

Deliverable **: Evaluation matrix and scoring template**

This is the core of the selection process, and there are multiple methods to capturing user requirements: individual user interviews, gap analysis and brainstorming sessions, to name a few. Most data discovery vendors have free trial editions that provide powerful, tangible insights into what is currently lacking in your BI portfolio. To the extent that you have established an innovation lab as part of the BICC, leverage the knowledge of those evaluators in identifying requirements. Key at this point is translating a requirement into a BI tool capability. For example, users will not say, "We want a BI tool that supports outer joins," but they may say something like "we want to look at patient treatment options, what's been tried and what's not been tried." In order to develop a list of requirements, you need to know what's possible. For this reason, it's useful to ensure a working knowledge of the data discovery market before gathering detailed requirements. Use Gartner and vendor-specific webinars to learn the fundamentals.

Product features are often easier to rank than strategic considerations. Agreeing the relative importance of something like "BI market leadership" versus a capability like "offline iPad support" is a difficult and often contentious task. For this reason, BI leaders should have two broad categories of evaluation: functional and strategic. Strategic considerations include cost of ownership, existing relationships and skills investment, availability of resources, vendor roadmap and vision.

Use the "Toolkit: Data Discovery RFP and Toolkit" (https://www.gartner.com/document/code/253034?ref=grbody&refval=3074819) as a checklist of selection criteria, applying your own weightings to the criteria. These weightings and requirements may be adjusted after scripted demos, when users have a better understanding of the possibilities and implications. Also consider emerging capabilities such as search, natural-language processing and smart data discovery ("Hype Cycle for Business Intelligence and Analytics, 2014" (https://www.gartner.com/document/code/262070?ref=grbody&refval=3074819)). In this regard, developing and ranking your selection criteria is an iterative process.

5. Request for Information

Who **: IT**

Common mistake **: Long laundry list of requirements in a form that does not facilitate scoring and comparisons**

Deliverable **: RFI**

It would seem like a glaring omission to skip this step, yet increasingly, RFIs seem to provide diminishing value. RFIs create a lot of work for the vendor and not much value for the customer. Vendors often say "yes" to each requirement, even when a more accurate answer is "not really, but possible with lots of workarounds."

However, some requirements are showstoppers that an RFI can help weed out. For example, if your company's standard operating system is Linux and the vendor doesn't support Linux, that requirement might cause the vendor to be removed from your shortlist. Other requirements are not so clear-cut. Follow these steps to improve the value of an RFI:

1. Define your requirements well to avoid misunderstandings between you and the vendor.
2. Ask for specific product names, feature names and explanations for how the requirement will be fulfilled. Know when the answer requires a "yes/no" response or is more a matter of "how."
3. Keep the RFI short, emphasizing the critical requirements that will be decisive in your selection or standardization and that are differentiators.
4. Complement vendor RFI responses with a heavy dose of your own research from customer references and discussion groups, as well as Gartner Market Guides and Critical Capabilities reports.

6. Scripted Demos

Who **: Business, IT and broader user base**

Common mistakes **: Letting vendor control demo topics; not focusing on critical requirements and differentiators**

Deliverables **: Updated evaluation matrix; demo feedback scores**

Vendor demos should be scripted so users can objectively compare vendors and products. Prepare a consistent agenda for each vendor to follow. In the agenda, allow time for a discussion of strategic considerations as well as specific product capabilities identified as critical. Be sure to invite an extended user base to the demos to elicit qualitative feedback and ensure users have a stake in the decision-making process. Based on your priorities defined in Step 4, ask demo participants to score the vendors on their ability to meet the various requirements.

Scripted demos can be based on either the vendor's sample data or your own. Using internal corporate data further identifies how each vendor's tool is different, yet requires a large investment from both you and the vendor as well as a formal confidentiality agreement. Such an investment is advisable with a handful of vendors but impractical with many vendors. Collect user feedback immediately following the scripted demo session, using a scoring template. If you wait to collect feedback for days or weeks after, users are likely to have forgotten the details and may mix up vendors.

7. Determine Best Fit

Who **: Business and IT**

Common mistakes **: Overemphasizing incumbent vendors; not allowing users to reweight requirements or add items**

Deliverable **: Updated scores with product and vendor summaries**

Using your requirements matrix defined in Step 4, rescore the RFI responses based on the demos. Incorporate strategic considerations, qualitative research and customer feedback to determine which vendor(s) most closely matches your company's short- and long-term BI needs. If you have one clear leader, do not fully dismiss the runners-up. You may find out during the proof of concept (POC), contract negotiations or pilot that your first choice has insurmountable issues.

8. Proof of Concept

Who **: Business users, IT and broader user base**

Common mistakes **: Trying to build a production system; inconsistent user testing of multiple products**

Deliverables **: Updated scores with product and vendor summaries; identifications of key findings and showstopper items**

You may have only one or two vendors that move onto the POC stage. This stage is your chance to test the tool in your environment or the cloud — with your data. Even if your company has pockets of deployments of data discovery tools, the POC provides a consistent testing and evaluation environment for all stakeholders. Some companies have set up war rooms with side-by-side testing stations to complete the POC.

A POC is only a test, though to confirm key capabilities. It's important to keep the selection team focused on the critical requirements defined in Step 4 rather than endlessly playing with the software or attempting to create production-ready dashboards. The POC may be a throwaway: its sole purpose is to confirm that the product works as you expect it to for the intended use cases and users. Carefully manage the scope to use one subject area and a handful of visualizations or dashboards. The sample dashboards should be based on information requirements defined in Step 3 and be of moderate complexity (don't waste your time with simple tabular lists that all BI tools can handle or, conversely, the killer mashup that would take your best developer a full month to design). The POC will give insight into how you may need to adapt the rest of your BI architecture and delivery team, but it is not the point at which you solve all the implementation problems.

This eight-step process brings the BI selection team to a final decision. Procurement, implementation and sharing of best practices are the next steps.

Acronym Key and Glossary Terms

BI	business intelligence
BICC	business intelligence competency center
OLAP	online analytical processing
POC	proof of concept
RFI	request for information

Gartner Recommended Reading

"Critical Capabilities for Business Intelligence and Analytics Platforms" (https://www.gartner.com/document/code/270381?ref=ggrec&refval=3074819)

"What Data Discovery Means for You" (https://www.gartner.com/document/code/253035?ref=ggrec&refval=3074819)

"Toolkit: Data Discovery RFP and Toolkit" (https://www.gartner.com/document/code/253034?ref=ggrec&refval=3074819)

"How to Deliver Data Discovery Products" (https://www.gartner.com/document/code/270537?ref=ggrec&refval=3074819)

"Predicts 2015: Power Shift in Business Intelligence and Analytics Will Fuel Disruption" (https://www.gartner.com/document/code/270932?ref=ggrec&refval=3074819)

"Business Intelligence and Analytics Summit Opening Keynote: Bridging the Analytical Divide" (https://www.gartner.com/document/code/277128?ref=ggrec&refval=3074819)

"How to Be Agile With Business Analytics" (https://www.gartner.com/document/code/270535?ref=ggrec&refval=3074819)

"Think Twice Before Changing Business Intelligence Tool" (https://www.gartner.com/document/code/270900?ref=ggrec&refval=3074819)

Evidence

[1] Analysis of approximately 1,710 Gartner inquiries from April 2014 to April 2015 described as reporting or data discovery related — 1,020 concerned data discovery versus 590 for reporting.

[2] An online survey of vendors' customers conducted in October 2014, which yielded 2,083 responses overall. Of these, 187 specific to data discovery were used to calculate average license costs.

[3] BI Scorecard 2014 Successful BI Survey, 534 responses.

[4] BI Scorecard 2014 Successful BI Survey, 534 responses.

Note 1

Data Discovery Vendors and Modules

The below list is not exhaustive.

- Advizor Solutions Analyst
- Datawatch
- IBM Watson Analytics
- Information Builders WebFocus InfoDiscovery
- Logi Vision
- Microsoft Power BI
- MicroStrategy Visual Insight
- Oracle Visual Analyzer
- QlikView and Qlik Sense
- SAP Lumira
- SAS Visual Analytics Explorer
- Tableau
- TIBCO Spotfire