
Buy vs. Build Considerations for Business Intelligence Platforms

By Jared Decker | President and Advisory Consultant, Expert Analytics

October 1, 2019 | Originally published in Enterprise Executive. Republished with permission.



When executives investigate the benefits of business intelligence platforms, they often see value but may wish to evaluate what functionality the software is providing that cannot be developed in-house in order to avoid BI platform licensing costs. Elementally, BI platforms are a combination of data handling and user interface components with features that vary in sophistication. When evaluating the buy vs. build question for the type of self-service analytics provided by BI platforms, sophistication is the key consideration. Here we discuss certain BI features which tend to be more sophisticated, along with custom-build considerations for each.

Administration

One common feature of market-leading BI platforms is governance capabilities. BI platforms provide administration capabilities out of the box to handle the following areas:

- **User and Data Security:** Management consoles allow administrators to manage which users can access which dashboards and reports, some down to which rows of data can be viewed.

-
- **Report Scheduling and Distribution:** Many BI platforms provide interfaces for scheduling reports and subsequent distribution to user groups via email, file drops, and other methods.
 - **Centralized Metadata Objects:** Some BI platforms provide the concept of centralized libraries of KPIs, hierarchies, and other data constructs.
 - **Auditing:** Most BI platforms include robust logging capabilities, allowing administrators and business leaders to monitor usage of the system, down to what information is being viewed and how often. This kind of auditing can lead to improvements, as management is made aware of what kinds of information is most relevant to different user groups.

BI Platform Administration Custom Build Considerations

Developing and setting up administration is not a particularly difficult task. There are many standards and plug-ins that are easily adopted to a variety of target deployment environments.

If proprietary authentication is required, this type of problem has been solved sufficiently so as to almost be incorporated as template code. On the other hand, some administration aspects of BI platforms such as centralized metadata objects are much more difficult to custom develop, because an additional repository and user interfaces must be developed to provide for the creation and management of these objects.

“When evaluating the buy vs. build question for the type of self-service analytics provided by BI platforms, sophistication is the key consideration.”

Data Handling

Data handling encompasses extracting data from source systems to processing and organizing it for optimal UI presentation and everything in between. Different BI platforms take different approaches to data handling, with some or all of the following features:

Source Data Extraction

Source data extraction provides tools for users to develop data extraction logic in a visual and simple interface while providing a mechanism for future scheduled data extractions. Some BI platforms have built-in capabilities to connect to almost any kind of database, in addition to providing support for connecting to web APIs such as those provided by Salesforce, Google, and many others.

Querying data from database systems is straightforward if the requirement is for the user to supply a query that is compliant with the target database system. However, many BI platform users are not necessarily well-versed in query languages such as SQL. More sophisticated BI platforms provide visual data selection tools which allows users to browse the data elements in the target system and select which tables and fields they would like to bring into the BI platform. When joins are required between data sets, these visual data selection tools mechanisms for this in a simple way that doesn't require technical knowledge of the source system.

Source Data Extraction Custom Build Considerations

Developing a user-friendly interface that abstracts the subtleties of many different types of source systems in order to provide a standardized, visual approach for business users to extract data would be a significant undertaking. This interface would need to incorporate different query language syntax varieties for different database platforms. For REST and SOAP APIs, a similar internal mapping would need to be developed in order to provide a consistent end user experience.

Analytic Model Construction

Some BI platforms allow for the creation of analytic models which may act as a superset of imported data, possibly incorporating new elements based on optional logic and formulae. Typical with analytical model construction abilities are mechanisms for creating hierarchies, sets, user-defined groups, and custom logic, all which serve as an “analytic overlay” on imported data, facilitating analysis and ease of use for end users that would not otherwise be possible.

Analytic Model Custom Build Considerations

The notion of allowing users to build their own analytic models would be an ambitious custom development undertaking because it requires a good deal of thought and analysis with respect to the methods and rules for how users may construct custom analytical models that expand or constrain original source data. For example, certain customizations, if allowed, could affect the integrity and performance of the system if rules and data are not properly constrained.

In practical terms, this is not a feasible requirement to be developed in-house except for in the most ambitious of projects, at least to the degree and sophistication of analytical model functionality in well-established BI platforms. However, some simple variations are certainly feasible for custom-built BI platforms, such as allowing users to define new calculated columns and custom groups.

Query Pre-Processing

Query pre-processing increases responsiveness and may occur at the time data is imported from source systems or at some other time prior to end user interaction. This can be an important feature, dramatically increasing response times as users browse reports and interact with data

visualizations. Response times in most marketing-leading BI platforms averages out to a few seconds for each interaction (i.e. page refresh or filter change).

BI platforms apply different types of query pre-processing that leverage software or hardware strategies.

Hardware Strategies

In the case of hardware strategies, data is repositioned to either spread the query workload across more hardware or to leverage the most efficient hardware resources such as CPU and RAM, while minimizing reliance on less-efficient resources, namely disk spindles.

Software Strategies

Software approaches to improving response times typically employ data indexing, pre-caching, or pre-aggregation.

With indexing, the BI platform identifies which fields to index and what type of index to use on each field – no small task if done manually - as any DBA will attest to. A tool that can systematize the process of indexing based on the nature and structure of the data rather than anticipation of the specific queries that will be executed (i.e. the standard approach to indexing) is an impressive feat, and this is precisely the feat that some BI platforms accomplish.

With pre-caching, necessary data sets are resolved in advance so that when users open or change filters on reports and dashboards, the necessary data to populate those reports and data visualizations is retrieved directly from a stored location, rather than from the underlying data store where the data is in its lowest level of detail. Pre-caching necessarily entails a finite set of data sets, which also means that the analysis possibilities are somewhat close-ended.

As the primary method employed by OLAP databases, pre-aggregation strategies seek to pre-calculate higher-level intersections of data that may be required to populate reports and data visualizations at non-transaction presentation levels. For example, source data may be loaded at the home address level but pre-calculated at the city, state, and country levels. Pre-aggregation strategies are more open-ended than pre-caching strategies, but completely open-ended, given the amount of human interaction and computational resources required to define and pre-populate



the most optimal aggregations that will enhance a most users' experience with rapid response times as they drill into vast regions of data.

Query Pre-Processing Custom Build Considerations

Though some simple pre-processing options such as report caching are easily custom-developed, more sophisticated approaches are difficult to accomplish. For example, custom-developing an automated indexing feature would require the development of a rules-based data design platform that is rigid enough to enforce rules regarding how the data is structured such that it can apply indexing opportunities automatically, which is difficult if allowing the flexibility to incorporate a range of data structures in order to be truly useful. Though such a thing may have been custom-developed by some ambitious enterprise for internal reporting, we have never come across an auto-indexing custom-built reporting tool. Custom-developing an automated aggregation-designing and construction system would be even more challenging.

Conclusion

This article provides a glimpse into some of the complications that should be considered in a buy versus build decision involving the kind of functionality provided by BI platforms. There are more factors to consider such as other types of features not addressed in this small sample, product improvements, and support, but we do have a general rule of thumb to offer: The higher the number of users and use cases, the more sophisticated the BI platform will need to be, leaning in the direction of buy. Likewise, the more features required, the more sophisticated, leaning in the direction of buy. In other words, custom-built BI platforms are most feasible when the features and use cases are limited, as depicted in this figure. In summary, we recommend thinking very carefully about the decision to custom-build what would otherwise be out of the box functionality from a BI platform.

Key Takeaways

- Enterprises with sophisticated IT departments may seriously consider building a custom enterprise reporting and analytics platform but should be aware of the risks and tradeoffs of this route vs. purchasing a BI platform.
- Established BI platforms incorporate decades of experience
- Post-deployment support should not be discounted – vendors offer an edge over IT departments here
- High Diversity of use cases and features leans in the direction of buy
- Speed of interactions and queries can make or kill an analytics initiative; vendors generally offer an edge

-
- BI tools such as Qlik, Microsoft, and Tableau are competent, feature-rich, and well-supported BI platforms that offer ample customization options. It would be a very expensive and time-consuming effort to attempt to replicate such functionality and user experience as is provided by these tools.

About the Author

Jared Decker is the founder and president of Expert Analytics, and provides advisory consulting services to clientele regarding their strategic data platforms and initiatives. He has more than 16 years in BI and data analytics consulting, with 15+ years in data architecture roles, designing data warehouses and BI platforms for clients in verticals such as commercial real estate, capital management, credit card, and consumer retail. He is the co-author of several analytics books published by Wiley, is a frequent big data columnist for executive tech publications, and has provided onsite training engagements for numerous *Fortune 500* companies, including Halliburton, Humana, PepsiCo, PPG Industries, and consulting companies that service *Fortune 500* and government clientele.

About Expert Analytics

At **Expert Analytics** we provide data experts and expert teams to carry out your strategic data initiatives expeditiously and cost-effectively. Our experts are experienced in a wide range of data tools and platforms and are highly credentialed with decades of successful client deliverables, advanced degrees, and certifications.

Our experts help you achieve success with your most strategic initiatives and improve your teams' performance across a range of projects through mentoring and advanced expertise.

Expert Analytics, based in Dallas, Texas, was created by data warehousing and analytics veterans and thought leaders. From the beginning, **expertise is the company's culture** and is deeply woven into every client delivery.

www.ExpertAnalytics.com

(800) XPRT137 | (800) 977-8137

Info@ExpertAnalytics.com