Deloitte.

Qlik

A new frontier in analytics

How to make the shift from traditional analytics to data discovery

Contents

The imperative for business data discovery

The modern-day digital revolution and the rapidly growing Internet of Things (IOT) are creating more data than ever seen before. The variety, complexity, and velocity of this data, and its many sources, are changing the way organizations operate. Perhaps the biggest factor in data growth is the advancements in technology that have made data capture and data sharing easier and more affordable. Modern users have also become more tech-savvy and are leveraging technology to generate and share more data and applications.

Another major factor is that sensors on devices are now generating data at a phenomenal rate. More and more devices are connecting to the Internet and these devices can capture and share data at a much faster pace than humans, creating more data to automate processes, deliver status updates and even drive machine-made decisions.

This massive amount of data can be very useful for your business if you have the tools and skills to extract its value. While data scientists have the skills for data analysis, they must collaborate with business users to ask the right questions. And even data scientists and information designers require easy-to-use tools for analytics and visualization so they can be productive when exploring data.

Analytics has become an important discipline because organizations that capitalize on the value of their data can differentiate their products and services and gain a competitive advantage. Technology advancements are only part of what's bringing analytics to a new frontier. Today's business users are accessing new types of data to support decisions and actions that span from operational details to strategic directions. Yesterday's data warehouse and business intelligence (BI) systems were simply not designed to provide the level of visibility and speed needed to respond in real-time to today's customer and operational needs.

Business users also want control over how they use their data. They want to engage with the data in an approach called data discovery. Data discovery means that the user is able to access the data when and where they need it, and then interact, analyze and visualize the data to explore its business value. When a user sees a visual expression of the data, this can lead to insights and spark a natural curiosity for further exploration. Data discovery puts the power in the hands of the business user, enabling an iterative process that is in contrast to the early days of static reports and data warehouses that were optimized for specific types of queries.

Let's examine some of the trends that are driving these changes.

Seven trends driving the shift to data discovery

The following trends are re-shaping data usage and transforming the discipline of analytics into data discovery.

#1: Unbound human and computer interaction

Usability is a crucial factor in delivering a good user experience and driving business benefits. Fortunately, the barriers between humans and technology are growing thinner. In the case of analytics, usability means going beyond static reports to give users the ability to access the data and interact with it through exploration, modeling, and visualization tools. It also means access from any device, including mobile and touch screen devices. These technologies inspire users to pursue new business opportunities and empower them to explore data and uncover the data's hidden insights.

#2: Ongoing, accelerating data boom

In its 2014 Digital Universe report, IDC predicts that the volume of data in the world will grow tenfold by 2020, from 4.4 Zettabytes in 2013 to 44 Zettabytes[1]. This exponentially growing data boom is a resource. Of course, users are demanding better, greater access to more and more data, regardless of where it comes from, but they need ways to sort the signal from the noise. That's where visual data discovery comes in. It gives users the self-service exploratory access they want, while also making it easier to spot trends and signals in the data.

#3: Rise of information activism

Information activism is about the behaviors and skills of users who are taking advantage of today's technology. We are witnessing a tech-savvy generation entering the workforce and older generations of workers are also learning new skills. Tech-savvy workers bring their own devices everywhere they go and expect to easily mash-up data, communicate, and collaborate with their peers. The overall trend is for knowledge workers to be much more actively engaged in finding business value in data. Data discovery is a natural outcome of this information activism. People with natural curiosity want to explore data to find insights and take action. They want the capability to have direct active access to many views of the data rather than simply read a standardized report.

[1] "EMC Digital Universe Infobrief" with research and analysis by IDC, April 2014. http://www.emc.com/collateral/analyst-reports/idc-digital-universe-2014.pdf

#4: Evolution and elevation of the role of Information Technology (IT)

In the early days of BI, the IT organization was the prime and often sole creator of reports. This meant that users had to be patient if they needed new reports or changes to existing reports. This has changed. Today's CIOs are driving their organizations toward an elevated role with a focus on business enablement. Rather than just providing infrastructure and acting as a gatekeeper, forward-thinking IT organizations are delivering tools and information products that empower business users to help themselves to what they need. Data discovery empowers business users to innovate and act quickly using an infrastructure of analytical tools and data sources provided by IT.

#5: Differentiate via information exploitation

As times change, so do differentiators. In the 1980s businesses differentiated themselves based on the quality and reliability of their business processes (think ERP). Increasingly, today's businesses are differentiated based on their use of information, making analytics an important tool for helping business users access the data they need and turn it into actionable insights. Analytics can be used to exploit information to create differentiated products and services that set businesses apart from competitors. Actionable information can also help organizations meet the challenges of today's global business climate, making this trend inextricably linked to the next trend — speed and agility.

#6: Need for speed and agility

Businesses lose out to more agile competitors when they fail to address a marketplace need or get left behind when innovation changes the rules of an industry. Gaining a competitive advantage from information requires that the analysis is in sync with the operational and market tempo of the business users. Traditional BI tools simply weren't designed to cope with today's demands for speed and agility. An OLAP-based architecture, for example, is designed to serve a process in which the questions are known in advance. Modification to techniques, dimensions, and adding new data take time in an OLAP system. OLAP cubes and tools are designed for specific types of queries, so data structures must be re-architected in order to achieve good performance for different types of questions. Today, users must ask and answer questions quickly and repeatedly. Then, after they have seen the results, they may want to explore different scenarios and models with different types of data and analysis.

#7: Transformation of analytics

Companies are introducing new products and marketing campaigns at an accelerated pace, and are even entering new markets at a fast pace. Faced with these business dynamics, business users must actively monitor new campaigns, understand what's happening today and take corrective actions. This requires a shift in analytics efforts to connect things that are seemingly unrelated and discover hidden insights in the data that helps people anticipate, or perhaps predict, what comes next. In other words, analytics itself is transforming and addressing modern requirements as organizations become more and more data-centered and data-driven.

The shift from baseline reporting to data discovery

Traditional BI reporting worked well for executive-level activities and macro decisions because the rate of change of information was relatively slow. Dashboards allowed executives to monitor trends over time so they could keep their business strategy in line with market trends, or monitor performance against quantitative goals and key performance indicators (KPIs).

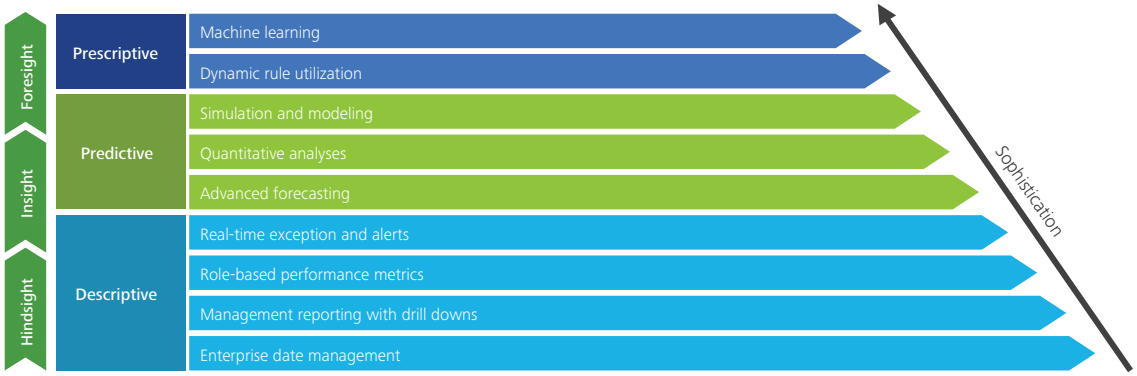

Figure 1. The evolution of analytics.

For Illustrative Purposes

Today's business culture operates much faster and information is required at many more points of contact. There are operational activities and daily decisions throughout the organization that can benefit from fast access to up-to-date data.

A customer service call center can be much more effective if the right information is at the fingertips of every customer service representative. When a representative can quickly review the customer's order history across retail outlets, reseller channels, and online storefronts, they can make special offers and provide better service to help elevate customer satisfaction levels.

While many of today's operational activities can benefit from data, there is a great diversity in the types of data needed. Each activity might happen at a different pace and use different sources of data, including both internal and external sources.

Since operational activities can't wait for regularly scheduled reports, the speed of data access must match the speed and diversity of the business cycle. Figure 1 depicts the evolution of analytics from "descriptive" analytics, which provides the foundation for turning data into information, to "predictive" and "prescriptive" analytics. Predictive analytics is designed to provide advanced forecasts and the ability to model future results and outcomes. The top-tier of the diagram shows "prescriptive" analytics, which leverages machine learning techniques and dynamic rule engines to both interpret data and take action.

In short, descriptive analytics provides the hindsight view of what happened, predictive analytics offers insights for current processing, and prescriptive analytics enables foresight for actions and a vision of the future.

Today's analytical tools also support forecasting and can offer guidance for the next action steps — a prescriptive approach. In this scenario, the potential offers that a customer is likely to purchase are ranked and selected based on the customer's preferences, profile, order history, and the actions of similar customers.

An IT architecture for data discovery

Analytics tools must satisfy the increasing appetite for self-service access to data by knowledge workers throughout the organization. While faster reporting intervals might be able to match the pace of operational activities, this would not satisfy the need for knowledge workers to develop their own questions and interrogate the data to get answers. Business users cannot predict in advance what information they will need to make business decisions or support operational activities. The data discovery architecture must be able to help business users spend less time gathering and integrating data, and more time understanding, analyzing, and gaining actionable insights from the data.

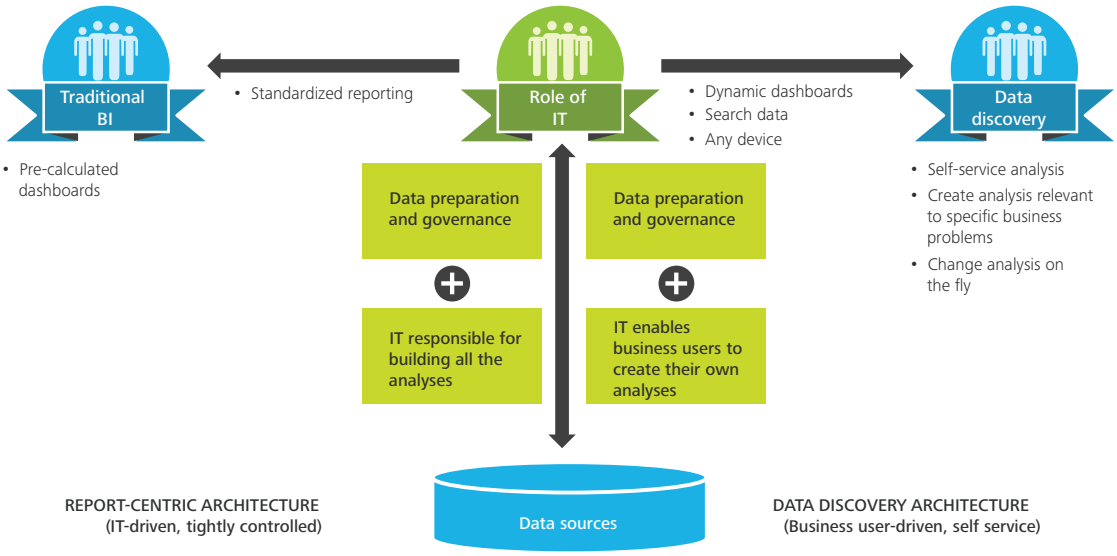

Figure 2. Architectural differences between traditional BI and data discovery.

For Illustrative Purposes

Implied requirements for analytics tools

To effectively support data discovery, analytics tools should support the following capabilities:

- **Easy and quick data retrieval** — Intuitive and fast data discovery.
- **Ability to ask more questions** — Explore data beyond the scope of a formal report.
- **Visualization of data**— Display previously hidden data patterns and communicate complex algorithms to decision makers.
- **Analysis of what-if scenarios** — Flexibility to change data factors and algorithms that influence the outcomes and then choose the inputs that can create the targeted outcome.
- **Forecasting** — Ability to reasonably assess future outcomes based on a trajectory from patterns and trends spotted in historical data combined with external data such as customer sentiment, economic forecast, capital markets, and even weather.
- **Varied data types** — Inclusion of structured and unstructured data to improve the value of the insight and expand the number of factors.
- **No user training** — Analytical tools should be easy and intuitive with minimal user training required. Users apply the same tests to analytical tools as they do for app downloads on their smartphones and tablets. If it's not intuitive, it will likely be discarded within minutes.
- **Schema-on-read design** — On demand integration of new data sources and no need to pre-process data to optimize for specific types of queries.

An elevated role for IT

One of the biggest challenges inherent in the centralized model of traditional BI is the burden on IT to maintain, modify and support the report portfolio. Report maintenance or generation can take days, weeks or months, and often the market opportunity is missed before the proper reports are ready.

The reporting backlog has pushed users to find their own way to get what they need, resulting in shadow IT reporting groups. This in turn has created a haven of "spread marts" maintained by "human middle-ware," a costly, inefficient and error-prone method of reporting that often leads to multiple versions of the truth.

A much better solution is the evolution and elevation of the role of IT. While IT may still generate some standard reports for mandatory reporting needs, the IT organization also provides a data discovery environment that supports user self-sufficiency. This decentralized approach allows the IT organization to focus on maintaining the data and analytics infrastructure while empowering users to explore data on their own with proper protection against security and privacy issues. IT essentially delivers data-as-a-service and analytics-as-a-service offerings that give users independence while reducing or virtually eliminating the bottlenecks that can occur when IT is responsible for all reports.

This requires a cultural shift in the IT organization. IT teams should collaborate with business users to ensure that analytical projects reach their potential and deliver high value to the business as well as to suppliers and customers.

People, process and technology are inextricably linked to successful outcomes

Like any project, successful analytics implementations require a combination of technology, people, and process. However, there are some specific nuances related to analytics implementations that are important to driving the project's success. The following sections identify several leading practices in the areas of technology, people, and process to help your analytics project achieve the desired outcome.

Technology

Technology is what gives people the power to collect, store, transform, integrate and interrogate data. Today's analytics technology supports an iterative approach to finding answers in the data and also offers rapid sophisticated analytical applications development for business users. These analytics tools provide more capabilities than traditional BI tools and also bypass the need to predefine data structures and optimized queries. They also make it easy to develop and deploy analytics solutions very quickly.

The technology infrastructure should be architected to support an enterprise approach to data discovery where all data sources are aggregated into a single physical or virtual enterprise data hub. Data should be moved only once when collected from its original source and saved in a centralized repository, allowing users to focus on the valuable activity of analysis instead of data management. The need for active metadata and sophisticated data management tools is another key requirement.

Data analysis must also be possible without pre-processing the data, and it must be easy to integrate multiple types of data, including unstructured data such as video and audio files. Keep in mind that the technology that can store all types of data in a unified structure may not always deliver the performance users require. Some situations may be best served by a technology architecture that includes multiple parallel in-memory systems to meet performance needs.

Lastly, technology plays an important role in driving user adoption. Characteristics of a technology environment that fosters user adoption include:

- **Speed** — Analytics results must be delivered as fast as an Internet search or users may opt out.
- **Usability** — Analytical tools must be self-explanatory with little or no training required.
- **Relevance** — It must be easy to navigate and flex the data and to understand how to retrieve it.

Building a highly visual and intuitive analytical environment for data access will result in users wanting more.

People

Regardless of how great the technology infrastructure might be, a project's success must overcome organizational and cultural barriers. Deciding whether an analytics center of excellence (COE) staffed with analytics specialists, data scientists and information artists will be centralized, federated or decentralized depends on your organization's culture, maturity, and goals. Common goals include raising the analytics IQ of your organization or transitioning analytical application development from IT to the business. The latter would require the appropriate development skills in the business side of the organization.

An analytics COE may include power users, educators, analytical evangelists, information designers and developers who deliver sophisticated application development. Deciding how deep and wide you want to enable analytical skills should be part of your process for designing the COE.

Regardless of the design of the COE model, understanding the business needs and its data are paramount. A partnership with IT for the data—and analytics-as-a-service requires clear roles and responsibilities for both business and IT team members. The protection of the data and privacy of customer information is a shared responsibility, and requires enterprise data governance with its policy and procedures embedded in every layer of the analytics architecture.

Additional members of the analytics COE might include virtual or physical members from the business community who bring a deep understanding of business domains and urgent business issues. Other members bring the analytical application development, data management and visualization know how. This combination results in delivery of analytics solutions that meet business demands because those who understand the business requirements are directly involved.

Additional important benefits of an analytics COE include raising the analytical IQ of the business, and enhancing the value and monetization of data. The analytics specialists working side by side with line of business and functional teams share their understanding of how to turn data into insights. As the business users learn more about how to apply analytics to their needs, their analytical IQ grows along with the value their data brings to the business.

One common barrier to successful analytics solutions is poor adoption among the user community. Usage of analytics tools is not always mandated, and often gut-feel or intuition are still the most common decision making approaches. A few forward-thinking leaders are beginning to require "test and verify" methodologies for all decisions that are made, but these situations are still rare.

Short of a mandate, one technique for promoting adoption is to share the "art of the possible" and show the power of analytics through a center that can demonstrate the highly visual and interactive capabilities of an analytical environment. In the end, gaining widespread adoption requires a cultural shift. Business users must be willing to accept the transparency that analytics provides. Sometimes data will show that previously executed actions or business decisions were not optimal. Many managers are somewhat attached to their old ways of doing things, or they do not believe the data. If the data is inaccurate, it's better to put in processes to improve the data so that trust can be restored.

A powerful sponsor is often the motive force to drive cultural change and user adoption. An executive mandate for the business to value and use data as the foundation for all business decisions and activities is a strong catalyst for such cultural change. Chief Analytics Officers or Chief Data Officers are taking on these leadership roles today. This gives data analytics a seat at the CxO table and makes it easier to drive awareness and cultural change.

Process

For process, there are several things to consider. The most obvious is the need for governance for data quality and to protect privacy and security. Data governance establishes the responsibility for business data ownership and sets standards to ensure that the analytics solution and the underlying data are reliable. This requires operational controls to monitor and manage the integrity of data collected, stored and distributed from the enterprise data hub. Implementation of role-based access controls is an imperative for managing access to data as set forth in the data governance standards. Other necessary controls should be established for such areas as data integration, metadata creation, KPI definition, hierarchy management, report, algorithms and dashboard creation. The types of

controls needed can vary widely and will depend on your business needs.

A potential challenge with a decentralized approach to analytics is effectively monitoring the accuracy of analytics and visualization results that are produced in a self-service fashion. Explore the risks for your organization when offering self-service capabilities and determine how to "water mark" results though automated governance and active oversight by members of your analytics COE.

For some operational tasks, it may be important to give free reign to certain users in exploring, analyzing and reporting on the data. In other cases, such as for strategic decisions, a rigorous validation process for new analysis will be required. The analytics COE can provide quality assurance, education and certification of key users to help manage risk.

Regardless of which organization structure or process is the best fit, your methodology for project delivery should also be scrutinized. An agile or lean analytics methodology works well for analytics projects because it supports iterative project execution and delivery of results in short four to six week cycles. Transitioning from a traditional system development life cycle (SDLC) can be a big leap. Set up a small group to design and adopt the right methods, tools and techniques and make your transition project by project. This approach helps get buy-in from IT and analytics specialists as they gain experience and understanding of an agile process.

The future of analytics is the hybridization of machine and human decisions

A benefit of predictive and prescriptive analytics is that they allow you to combine machine-made decisions with human interaction. Computer-based recommendations or prescriptions can be put into immediate action by the software system, or they can be directed to decision maker who then makes the final call. Combinations of machine-made recommendations with human interaction can be very effective in guiding operational tasks in business rather than replacing them.

Machines dependably execute repeatable processes and parse through large volumes of data quickly. They offer consistent quality and make no mistakes if programmed effectively using the right data. Time-critical, repeatable activities are well suited for a computer-generated prescription, such as the next offer to a customer.

Humans are good at considering the context of a decision and verifying whether a machine-generated recommendation or action is a good choice for a particular customer. A mismatched recommendation from a machine can happen when there are factors that are outside the logic of the machine's algorithm. In this case, decision makers can readily detect the mismatch and make an alternative recommendation.

Consider an example where a retail outlet uses mobile technology and analytics to deliver an offer to every recognized customer's mobile device that is brought into the store. If a customer's purchase history has been women's business attire, the analytics tool would probably recommend a similar brand or style, and perhaps offer a discount. An in-store associate, however, might inquire if the customer is looking for herself or for a gift before making an offer, thus virtually eliminating the risk of an irrelevant offer to the customer.

Case study—The business value of data discovery

A large retailer wanted to reduce operational costs across their business operations by using retail point-of-sale and inventory data to identify potential areas for cost savings. With their existing data warehouse environment, it took months to bring new data online and create the appropriate reports to address new types of questions. This meant that cost savings opportunities were not being spotted in time to make the necessary changes to improve business results.

The retailer shifted to a Hadoop-based enterprise data hub and began using analytics tools, such as Qlik, to discover and visualize data. This new infrastructure made it possible to quickly and easily analyze the data, find cost savings opportunities, and implement corrective action.

A business-led team, with the support of IT, used the new tools to analyze gross margin return on investment (GMROI) across all stores. The data discovery process first identified specific items that had low inventory turns at a given location. Then the discovery process analyzed similar items not moving in other stores. In some cases, the trends were regional. The items might not be selling in certain geographical areas or populations because the demand there was different compared to national norms.

The previous data warehouse environment did not make it easy to find this kind of information because item profitability was reported on an aggregate basis across stores. There was no ability to visualize the variance across thousands of items and stores. By finding the trends, the retailer quickly addressed the issues and adjusted inventory on a regional or local basis rather than nationally. As discoveries were made in the agile environment, they were built into the traditional data warehouse environment for continuous exception reporting.

Leading edge analytics tools can blend the wisdom of information workers with technology's ability to process big data. Information workers can be made more productive and more effective by using machine recommendations as a starting point. They save time and take advantage of the power of big data while preventing machine interpretations from being put into action without a decisions makers' filter.

The hybridization of machine and human decisions forces organizations to answer challenging questions:

- How do we measure the business risk of using machines to make decisions and recommendations?
- What types of decisions or activities are appropriate for machine-based prescriptions and actions?
- Could increased dependence on machines affect customer satisfaction or customer loyalty?
- How will machine-based recommendations and decisions affect our competitive standing, and what will happen if we do nothing?

While analytics tools provide the opportunity for significant automation, it's important to find the right balance between human and machine decision making. Analytics is a way to augment your existing operational and managerial staff, not replace it.

Conclusion

We are in the midst of a transition in how analytics tools are being used in today's information age economy. The old days of relatively stable data marts and data warehouses are giving way to a new era in which the data flows like a river and must be analyzed as it changes — no matter how turbulent.

Data discovery not only makes it possible to scrutinize large volumes of rapidly changing data, but also enables faster, better decisions and actions. Advances in analytics technology are empowering users to explore data through intuitive user interfaces that invite the user's natural curiosity to understand the hidden messages in the data. Achieving the desired results with data discovery also requires cultural and organizational shifts as well as appropriate governance and control.

Decision makers must be willing to accept the transparency that analytics creates in exchange for the power to interrogate data quickly. With the enterprise data hub and analytics application development infrastructure in place, the role of IT changes to that of delivery of valuable data for users to explore and visualize— a data-as-a-service offering. An analytics center of excellence (COE) staffed by business users, data scientists and information designers will allow the organization to deliver more value from its data.

The evolution of analytics tools has propelled us into a new frontier, giving businesses the ability to transform operations at virtually every level— not just the managerial ranks. Data discovery holds the promise of making us all more productive as we create more and more business value from growing volumes data. Organizations that can optimize the process of finding insights in their data may achieve greater agility and can build a more sustainable competitive advantage.

Contacts

Jane Griffin
Managing Director, Data Analytics
Deloitte Canada
jgriffin@deloitte.ca

James Richardson
Business Analytics Strategist
Qlik
james.richardson@qlik.com