THE FUTURE OF VISUALIZATION
Big Data is meaningless without the right framework for understanding it. The volume, variety and velocity of data today have overwhelmed our ability to derive insights using traditional methods like Excel spreadsheets, KPIs and standard reports.

Companies have turned to visualization tools to tame the chaos of Big Data, and are asking better questions and making better decisions to gain a competitive advantage. And the visualization industry has done a fairly good job evolving the technology to keep pace with customer needs.

But we are now entering a critical, new, data-driven era in which visualization is becoming a bedrock technology for running a business, rather than an adjunct tool. The big question is: has the technology evolved enough to serve as a core enterprise system?

According to a survey by Ventana Research, “...48 percent of organizations have indicated that presenting data visually is an essential business analytics capability.” The research also revealed that “visual data discovery is one of the top three Big Data analytics needs not effectively delivered today, and visualization is more important for business than just a focus on velocity and volume.”

The technology eventually evolved toward data visualizations that not only present a dashboard view of the business, but allow users to drill down into the data to see more relevant and granular levels of detail. Users could now better identify outliers, anomalies and root causes of business drivers without having to be clairvoyant.

The visualizations also became much more complex. Applications that could present only a dozen or so points of data just a couple years ago could now represent tens of thousands of points of data.

MORE VARIETY

The next big development in visualization technology was its ability to reach beyond structured data in databases and data warehouses/data marts and ingest data from a variety of less conventional data sources. Until recently, data sources were a limited universe – SQL databases, documents and spreadsheets, data feeds such as EDI streams and industry-dependent specialized sources like machine-generated information from factory equipment or medical devices.

Now, in addition to these sources, data from PDFs, content management systems, social media, data feeds for managing equipment and log files are regularly pored over. And the data no longer has to be aggregated and homogenized in one place to do analytics.

This capability gave birth to the Big Data era. It was made possible thanks to new foundational technologies including faster, more powerful in-memory systems, columnar databases, new protocols like XBRL, open source technologies such as Hadoop and NoSQL, and analytic programming tools such as R and Python.

Visualization tools evolved to support this transition. They not only provided a visual, interactive interface with the data, but offered some data management capabilities in which many data sources could be organized and homogenized on demand.

SPEED OF BUSINESS

The third big change in the evolution of visualization was the rate at which data is consumed and refreshed. This is commonly referred to as data velocity.

In the past, data was refreshed and analyzed daily, weekly or monthly, because most of the BI applications developed were backward looking. They were designed to give you a historical perspective of what happened in your organization, not what’s currently going or just around the corner.

Today, data is flowing into organizations at unheard-of speeds from a variety of real-time sources. Fluid data technologies such as ticker feeds, sensors, message busses and complex event processing (CEP) engines, as well as external sources like social media, continually feed the corporate data trough.
And companies are increasingly deploying new advanced streaming infrastructures such as IBM® InfoSphere® Streams or Informatica’s Vibe Data Stream to drive the real-time components of their operations.

This data also travels along networks that are faster than ever before. Because much of it comes from outside the boundaries of the company, companies don’t have the ability to easily regulate the flow.

To fully exploit this data through analysis that helps optimize business performance, companies must ingest it at the velocity with which it arrives, thereby giving a real-time perspective of what’s happening throughout the organization. That means ingesting it, storing it and analyzing it, all without losing its real-time immediacy.

New visualization technologies had to morph from working in backward-looking, high-latency environments, to real-time visualization environments.

The shift began in the capital markets, where the requirements to manage risk continually throughout the day took the luxury of latency out of the business model. Firms in the big financial centers spearheaded development of a new class of analytical systems that enabled traders and analysts to visually analyze huge volumes of fast-moving market data as the market unfolded.

Soon, other markets with real-time requirements, like energy grid management, telecommunications, network management and manufacturing, adopted the new analytical technology because they could no longer wait a day or month to deal with problems; they had to tackle them as they occurred.

In-memory technologies helped solve the need for real-time processing velocity. But costs currently make it challenging to apply them to Big Data. Another obstacle is the need to aggregate real-time and historical data in order to gain valuable, in-the-moment context.

Take, for example, analyzing intraday performance of capital markets. It’s crucial to be able to aggregate both real-time and historical data to see not only the data, but the trending of the data. Companies need to understand what’s happening right now and what happened three seconds ago in the context of what happened three months ago.

What’s Next in Visualization

Because of these developments and challenges, Datawatch believes visualization technologies will soon evolve from being supplementary visual discovery tools for analysts, to becoming core applications that actively steer businesses.

They will incorporate a much greater set of capabilities. They will become true platforms for managing data. They will incorporate new algorithms and predictive, analytic capabilities so that business users can not only “see” new discoveries, but intelligently know how to act on those discoveries. For example, R and Python programming capabilities are starting to be integrated into the software to incorporate logic into the visualizations.

It will be a business user-led revolution, rather than remain the domain of specialist analysts. According to Gartner, “Companies will shift their future investment away from IT-developed reporting solutions toward business-user-led analysis solutions. IT will focus more of its effort on system-of-record data modeling and governance. As a result, data discovery will displace IT-authored static reporting as the dominant BI and analytics user interaction paradigm for new implementations by 2015.”

Retail is a good example of visualization’s dramatic impact when it’s transformed from a tool to a core platform for managing day-to-day business. Monitoring and auditing millions of dollars’ worth of transactions annually is a considerable challenge for any retailer with a POS (point-of-sales) machine. The transaction logs generated by POS machines are true representatives of ‘Big Data’—voluminous, complex and largely incognizable.

Manual extraction of data from POS transaction logs and arrangement into a tabular output is a time-consuming endeavor for which most retailers have no extra resources. The lack of a complete, timely picture of sales performance has
far-reaching financial effects. Inventories cannot be properly managed, accountants cannot ascertain profitability, and business owners cannot project realistic expansion goals. Most importantly, retail operations cannot focus on activities that promote products that sell well and eliminate those that do not.

An enterprise-wide visualization platform could incorporate a sales auditing system that extracted POS reports in real time, producing visualizations that are easily read and analyzed by those involved in the sales process. Users across the organization could visually discover store metrics from these large POS datasets to view store performance. They could drill down from store-level summaries to daily breakdowns by brand and product. Ultimately, they would make better business decisions using accurate and up-to-date information.

Online marketing is another industry ripe for transformation by next-generation visualization. In the age of e-commerce, there can no longer be any delays between the time a potential customer considers a purchase or contact with a company and the time a marketer sees it happening. Marketers need up-to-the-second visibility into who their website customers and prospects are, what call-to-actions or content they engage with, and what ads and incentives will trigger them to convert. They also need to assess how their affiliate networks are driving traffic, as well as gather more qualitative, unstructured data from social media, mobile applications, Amazon reviews and the like. Wait too long to sort out all these data points and competitors will step in and steal customers away.

With real-time data visualization, marketers can act faster and do things like instantly compare current customer behavior to predict their future behavior, create new messages to win them over. They can do A/B testing in real time to quickly identify which content, offers and messages truly engage customers, and adapt their site or ad buys to meet those needs.

They can also identify who their most profitable customers are by creating demographic triggers for sought-after characteristics such as education, income and geographic location. They can instantly capture the commentary, concerns and issues of customers’ conversations in social media to take corrective actions, keep existing customers from defecting, or preemptively steal customer mindshare away from a competitor’s online campaign.

To become a true enterprise platform like these examples demonstrate, next-generation visualization technologies will have to encompass three core capabilities:

- **Transformation**: the ability to capture and transform data from virtually any source at the speed of business
- **Distribution**: the ability to securely and automatically channel information to departments that need it
- **Optimization**: the ability to securely share, archive and analyze data discoveries that can impact the business

Next-generation visualization technologies will also incorporate the latest scientific research in machine-human comprehension. The goal is to enable a visual data discovery environment designed to work with the way the human brain processes information.
One interesting area being incorporated into new visualization interfaces is pre-attentive processing — the visual properties that people process almost subconsciously without the need for focused attention. The human mind has an ability to intuitively and effortlessly always know a big shape from a smaller one, for instance, or understand that two shades of red are different, and feel that items displayed as grouped in a box are of the same kind.

Using this power, a load is taken off short-term memory, which otherwise would have a lot more to keep track of through conscious, intellectually costly processes. By assigning appropriate information value to visual characteristics, including size and color, the user’s perception and understanding of data is greatly enhanced.

CONCLUSION

Visualization technologies have evolved adequately to meet business demands over the years, but the industry now has reached an inflection point, where it must transform from a tool to a platform for running business. To stay competitive, organizations need to harness all relevant information regardless of its type, volume or speed of delivery. Combining real-time data ‘in flight’ with unstructured and semi-structured data in interfaces optimized for pre-attentive processing and similar human cognitive abilities, next-generation analytics will dramatically improve how Big Data gets visualized to isolate and resolve problems as they occur, perceive hidden patterns, track emerging market trends, and identify opportunities for competitive advantage and improved business processes.

1 Four Types of Discovery Technology for Using Big Data Intelligently, Ventana Research,
2 Location Analytics: Putting the Evolution of BI on the Map, Aberdeen Group, 2/28/2014