

Big Data Challenges in Capital Markets

Theodor S. Klemming
Senior Product Manager, Altair Panopticon

Using Visual Data Discovery to Meet Big Data Challenges in Capital Markets

Executive Summary

The world of financial services is not short on data, however, there is a shortage of tools that can handle everything the banks need to look at. From the very beginning, the Panopticon development team recognized the absolute need to work with what the industry is now calling “Big Data”.

The Panopticon team focused on a very different architecture compared to other visualization systems in order to accommodate requirements around what is now being referred to as Big Data, including the ability to visualize information from real-time streaming feeds.

Altair recognizes that capital markets is a real-time business; a truly effective visualization system must be able to work with real-time streaming data sources.

Altair’s Panopticon software connects to and retrieves data directly from the source — whether that’s a big columnar database, an in-memory store, a CEP engine, a message bus or even a spreadsheet. In addition to reducing system complexity and Total Cost of Ownership, this architecture eliminates the need for an expensive middle layer (like a data warehouse, data mart or internal proprietary database) that adds unnecessary latency to the system.



This screen shows a stream of market data from the last second extending back over a year. The underlying data consists of billions of records — bid and ask values from the market — recorded as often as once per millisecond.

It's a massive amount of information and there's simply no way to analyze it effectively — unless you use effective visual data discovery tools like these.

A trader can use a mouse to select a subset of all the available trading data. The chart will change and let him isolate an outlier. Is that something normal? Or is it an anomaly that needs immediate action?

Nearly all other systems do require a middle data layer, which means they waste precious time waiting for data to be loaded. Among other pitfalls, the inherent latency introduced by a middle layer in other systems makes visualization and filtering of real-time streaming data practically impossible. Any kind of data middle layer or internal data storage in a system for visualization and analysis will always run a high risk of becoming a bottleneck in the data flow. Using a middle layer approach to the architecture also creates scalability problems that become quite difficult and expensive to solve as data volumes grow larger.

This paper explores the requirements and use cases involving Big Data in Capital Markets and explains how a properly designed visual data discovery system allows traders and portfolio managers in large firms to visualize all the data resources available to them effectively.

Altair's approach of connecting directly to the source data repositories and data generators means there is no waiting for long ingest cycles or for OLAP cubes to be built. Decision makers know they are seeing the accurate and most up-to-date information available within the organization.

This is an exciting time in the world of visual data discovery. Demand is growing for real-time systems that enable a continuous view of operations in a wide variety of industries; besides capital markets, interest is growing rapidly in telecommunications, energy, manufacturing, and retail.

Data is the New Oil

Clive Humby, the co-founder of Dunnhumby, a retailing consultancy, coined the phrase "Data is the new oil" a few years ago. He and his colleagues built a very successful business and made their clients billions of dollars with that idea. But what did he mean?

Data is like crude oil: You can't use it until it has been refined. And, with data, it must be presented so people can make informed decisions on a timely basis. Only when it has been refined and analyzed does Big Data have any real value.

What Is Big Data?

The idea of "Big Data" has been around in the IT world since 2001, when Gartner defined it in terms of three dimensions:

- Volume: The amount of available data is increasing rapidly
- Velocity: New data is being generated and processed with increasing rapidity
- Variety: The range of data types and sources is also expanding

Recent advances in data storage and computing technology have brought the idea into the public consciousness. For example, most people are aware that online retailers are now mining their vast amounts of purchasing data to better understand buying patterns and to target advertising messages more effectively. Of course, the Big Data paradigm extends well beyond retailing to nearly every aspect of life and the new tools are making these resources more useful to people who must make decisions based on that data.

Big Data is Big Business

Companies are spending serious money on new tools to meet the Big Data challenge, and to increase the value — in operational terms — of their vast seas of data. IDC's research indicates that the market for Big Data technology and services is growing at a compound annual growth rate of 40% — about seven times faster than the overall IT market.

What's driving all this growth, besides the availability of the data itself?

IDC believes organizations that are best able to make real-time business decisions using Big Data solutions will thrive, while those that are unable to embrace and make use of this shift will increasingly find themselves at a competitive disadvantage in the market and face potential failure. Managers of Capital Markets firms, in particular, are embracing this idea and making major investments in Big Data tools in order to improve profitability and help reduce the occurrence of “black swan” events like the Flash Crash of 2010 and the more numerous market hiccups we've seen since then.

The Traditional Reporting Model: Failing in the Real World

In the past, managing the complex Business Intelligence reporting systems commonly used in the corporate world required teams of specialists whose jobs it was to fully understand the intricacies of the system and to translate business requirements into useful reports.

This approach worked to an extent, but also created significant problems. Building and maintaining reports was time consuming; most senior managers will have heard the phrase “it will take us two weeks to build your report” many times from their IT people. In addition, actually running a report could take hours or even days in some cases. Getting IT people to truly understand the underlying business questions that the report was intended to address was fraught with pitfalls and potentially damaging misunderstandings.

All in all, it meant that managers were making decisions based on reports that either did not use current data, filtered out the most interesting outliers or hid important patterns and trends in multipage columns of numbers.

Visualization: The 4th Dimension of Big Data

The idea of using visual aids to analyze data goes back to the 17th century, when Rene Descartes, the French philosopher and mathematician, invented his method for representing quantitative data in relation to two-dimensional coordinate scales — what we now call a Cartesian plane.

This visual approach was extremely useful and in the hundreds of years since then, the idea of using pictures to represent numbers has become commonplace. The reason this works so well is that the human psycho-visual system is one of the most sophisticated parts of our bodies; we have evolved over millions of years to be extremely sensitive to visual cues. The idea behind visual data discovery is to tap that resource in our brains to help us literally see, very quickly, what is really going on in a sea of numbers.

Big Data presents many technical challenges to visualization systems. Most of the tools that had been available were simply unable to cope with Gartner's “Three Vs of Big Data” as the magnitude of data in all three dimensions increased exponentially. A new generation of visual data discovery tools was needed to visualize all the data resources available to them effectively.

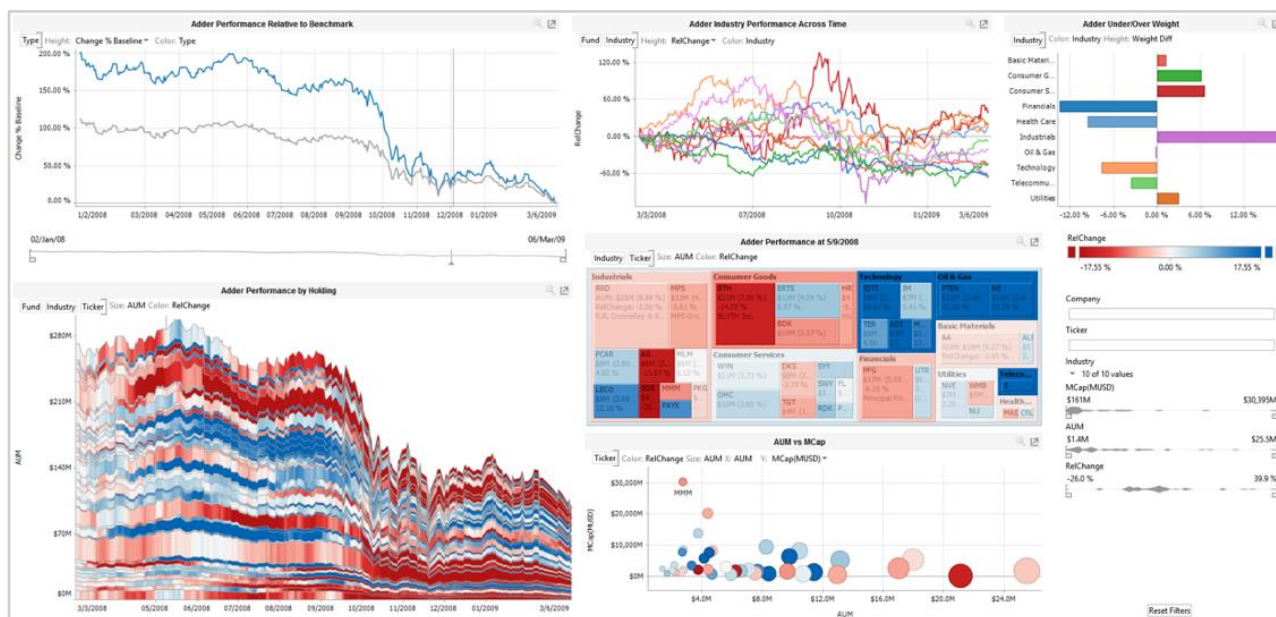
A properly designed visual data discovery system allows people to cope with:

- Data sources of any size
- Data that is changing in real time
- Data stored in multiple types of systems and in multiple formats

Capital Markets Applications for Visualizing Big Data

The Capital Markets industry has been an early adopter of advanced analytics. Many of the largest investment banks in the world have selected Panopticon, Altair’s real-time visualization software, to support several mission-critical applications. They chose Altair based on several criteria, including:

- **Connectivity:** Banks need a visual data discovery system that can handle a wide variety of data inputs from their established risk and trading systems.
- **Performance:** Traders and portfolio managers need a system that makes it easy to identify outliers, clusters, trends very quickly. Basically, they use Panopticon to reduce the amount of time it takes for a trader or manager to identify a problem or an opportunity, understand all the implications and take appropriate action.
- **Simplification:** Banks have configured Panopticon to support the operations of many different departments. In addition, the technical design of the system means that users can access it over standard web browsers and even from mobile devices like iPads and smartphones.



The Capital Markets industry is a major user of advanced data visualization technology. Their people use dashboards like this one to analyze risk, monitor trading performance and identify trading opportunities even in extremely fast-changing market conditions.

When Panopticon is deployed as part of a Big Data initiative, risk managers in large banks are able to see, in graphical form, extremely complex data in ways they can interpret very quickly. They can use our interactive filters to “slice and dice” and zoom in and out on the data — essentially they can change their perspectives on all that information easily — to isolate anomalies that may represent something they need to look at more closely.

Typical applications for visualization-enabled Big Data systems within the major banks include:

Market Risk

Many banks are seeking to replace or augment traditional tabular reports on market risk and replace them with interactive analytical dashboards. Market risk data is quite complex. Banks use Panopticon to develop dashboards specifically optimized to display DV01 (also called “Price Value of a Basis Point”) calculations, Value at Risk (VaR), options deltas, vega, gamma, and so on — across all asset classes (equities, equity options, FX, fixed income and derivatives).

Counterparty Risk

Banks understand that their data visualizations of counterparty risk must cope with multi-layered, aggregated and normalized data sets, including domiciles, limits, parents, parent-to-child roll-ups, limits applied at each level, and transaction data superimposed at each level. It is a complicated task that, in most cases, will also utilize data from many different data repositories in the bank. Panopticon’s ability to federate data from multiple sources is a critical function in such deployments.

Credit Trading

Banks need a visual discovery system that provides their front office people with a usable reporting tool on credit trading activities and also supports all the new regulatory reporting requirements affecting banking operations. With Panopticon, a bank is able to see multidimensional market value versus notional value displays and perform in depth analysis of historical trends. They can visualize overall, aggregated position data (for example, bonds + CDS) by sector, book, trader or individual issuer. In most cases, a Panopticon treemap is embedded in credit trading dashboards to facilitate market value, capitalization and DV01 analysis. Panopticon enables a bank’s people to see meaningful sets of dynamic risk attributes in multiple partitions, hierarchies and time periods with a scope and depth they’ve simply never had access to before.

Case Studies in Financial Services

Banks and other financial institutions are among the earliest adopters of Big Data technology. They generate massive amounts of information as part of their daily business, and their people are also charged with making good, profitable decisions based on a flood of data coming in from the markets and other sources.

Visual data discovery is playing an increasingly central role in the Big Data infrastructure for the best-run banks. It enables managers to keep a close eye on risk, performance and trading activity based on millions of daily transactions without being overwhelmed by the amount or complexity of the information.

Let’s look at some real-world examples of how visualization of Big Data is playing a role in two major institutions:

Pre- and Post-Trade Analysis at a Big Four Bank

The bank employs an enterprise-wide application that supports over 400 buy-side firms and more than 1,000 users globally. Traders use the system to maintain portfolios containing as many as 5,000 securities each. With the advanced data visualization tools built into the system, they can see the different attributes of the portfolio, including composition, the weight of each security and the impact cost- contribution of a security in the portfolio.

Users report that they particularly like the ability to clearly see two parameters like Percentage of Average Daily Volume and Expected Impact Cost at the same time. They can immediately figure out which securities require further investigation without having to pore over reports for thousands of instruments.

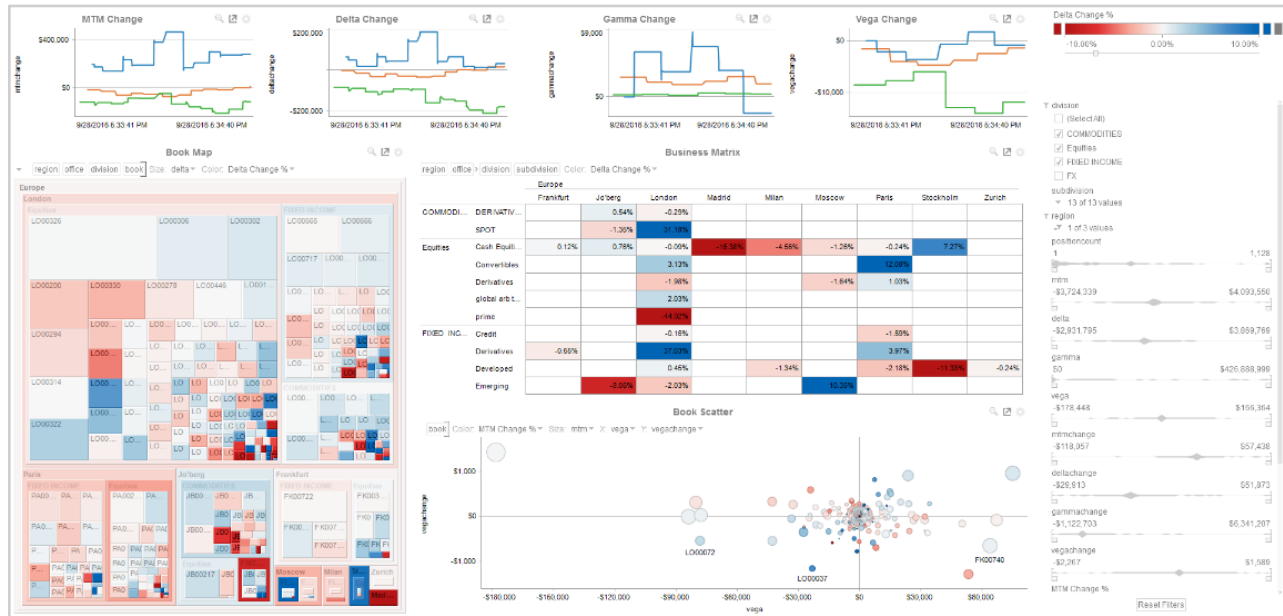


Portfolio managers use dashboards like this to analyze intraday performance — down to the millisecond if needed — for all the constituents in a portfolio. They can also look at aggregated data for multiple portfolios and spot problems and opportunities at a glance.

The system provides intuitive analysis and exploration of the data in ways that are simply impossible with traditional reports. It allows traders and the bank's clients to group and analyze their securities by country or sector easily. In post-trade analysis, the trader can see instantly how well she performed compared to benchmarks like Full Day Volume Weighted Average Price. They can also compare their performance to the universe of other traders using the system.

With their data visualization tools, the bank's traders do their own analysis without depending on an IT team to produce reports. They can ask their own questions and explore the performance data about their portfolios in ways that are simply impossible without such a tool. They save enormous amounts of time by using the interactive filtering and reordering tools built into the system.

The system has been very well accepted by the bank's traders and client firms since it saves so much time and allows users to better understand the data underlying the performance of their portfolios.



This dashboard allows traders to visualize market moves and changing risk parameters for thousands of equities on a single screen. Interactive filters let them quickly isolate outliers and expose patterns and trends that would otherwise be hidden in mountains of numbers.

Portfolio Analysis at a Global Asset Management Firm

One of the world’s largest asset management firms uses advanced data visualization to analyze position and market data. The main screen for the portfolio management application uses a treemap (sometimes called a “heat map”) to display a wide variety of real-time performance parameters for all the portfolios each manager is responsible for. The managers use interactive filters to expose outliers and hidden patterns and use that information to make buy and sell decisions.

In particular, the managers need to look at intraday performance at the portfolio level and also for all instruments held in groups of portfolios. For example, a portfolio manager may need to look at position and risk data for General Electric, which he holds in multiple portfolios. He needs to understand the impact GE is having in every portfolio that holds it. With their visual data discovery tools, the portfolio managers can see which stocks are helping and which are hurting in terms of intraday performance.

The firm’s senior managers need the same capabilities, only they are responsible for the performance of many different portfolio managers, with each of them managing multiple portfolios. It’s easy to see how traditional reports simply could not keep up with the markets, especially under volatile market conditions.

The firm’s managers report that one of the most useful views they have lets them see performance data and analyst ratings on a single screen.

The real power of visual data discovery lies in its ability to help managers pick and choose from a wide variety of parameters very quickly. Managers can visualize risk, trading activity, and position data based on the selected parameters and then easily filter out less interesting elements to focus on the instruments that require closer study.