

TIBCO Extends its Data Intelligence Vision

TIBCO is building out a comprehensive analytics portfolio through multiple acquisitions. Here's what's in store for Spotfire, Statistica and Alpine customers.

Sometimes one plus one really can turn into three. That's the formula TIBCO is hoping will apply to multiple acquisitions in the data-management and data science spaces over the last 18 months including Alpine Data Labs and Statistica (both now part of the just-announced TIBCO Data Science Platform), and the CISCO Information Server, previously known as Composite Software and now called TIBCO Data Virtualization.

To set the stage, formerly public TIBCO was acquired by private equity firm Vista Equity Partners in 2014. It has annual revenues of about \$783 million, according to Hoover's estimate. Founded in 1997, the firm has long prided itself on punching above its weight among industry giants. It started out as a messaging and integration middleware vendor up against the likes of IBM and Oracle. The company has landed many large, notable customers over the years, including Adidas, Bank of Montreal, Coca-Cola, FedEx, Johnson & Johnson, Pfizer and Unilever.

TIBCO's foray into analytics began with its acquisition of data discovery and visualization vendor Spotfire in 2007. The portfolio was extended in 2014 with the acquisition of Jaspersoft, a more reporting-centric, open-source business intelligence platform often used to embed analytical capabilities into commercial and custom software.

The analytics side of TIBCO is now billed as the "Augment Intelligence" portfolio, while the integration side is designed to "Interconnect Everything." The two sides

support each other, but I'll focus here almost entirely on analytics. TIBCO NOW 2018, the vendor's annual customer event, held September 4-6 in Las Vegas, gave some 1,600 attendees a great update and it gave me a better understanding of the company's many acquisitions and strategy.

Since 2014, TIBCO has acquired Statistica, Alpine and the CISCO Integration Server (formerly Composite Software), among other assets, and has added a data catalog.

Analytics Update

Before we get into the vision and strategy, the key analytics-related announcements at TIBCO NOW included the following:

Spotfire gets smarter. Set for general availability this fall, Spotfire X (as in the 10.0 release – skipping 8.X and 9.X) adds natural language querying (NLQ) and AI-driven recommendations. NLQ lets users explore data through typed, English-language questions rather than SQL code. The AI-driven recommendations feature employs an ensemble of machine learning techniques to scan data and spot interesting variables, exceptions and drivers. The recommendations are exposed as a list of notable data-exploration possibilities.

Spotfire gets faster. Spotfire X gains a new Data Streams function from TIBCO's StreamBase technology to support visual analytics on data streams. Spotfire Data Streams offers real-time intelligence for Internet of Things (IoT) applications and other (fast-growing) use cases for low-latency insight. Oil and gas, manufacturing, pharmaceutical and life sciences businesses are at the head of the list of potential customers. Spotfire Data Streams supports continuous monitoring as well as real-

time analysis and comparison to historical data to support trend- and root-cause analysis.

Data science options integrated. TIBCO Data Science is a start on the integration of various TIBCO analytics and machine learning capabilities, including last year's Alpine Data and Statistica acquisitions as well as the TERR (R server) and Statistics Server offerings long associated with Spotfire. Bundling brings together everything under a single license while initial integrations enable Statistica analytics to be run from Spotfire Analyst. Statistica can also call out to the Alpine engine to run analytics in cluster.

Jaspersoft gets cloud friendlier. TIBCO announced JasperReports IO to advance embedding of reports and visualizations in cloud-based apps. IO is a small-footprint rendering engine that wraps Java in RESTful endpoints for easier development. It's initially available as a Dockerfile. Highly scalable microservices orchestrated by Kubernetes are a planned next step.

MyPOV On Key Announcements

Spotfire is a solid data discovery and visualization option that's particularly popular with data-savvy analysts in manufacturing, oil and gas, life sciences, pharma and other demanding environments. TIBCO was ahead of the game in delivering integrated data-wrangling capabilities in 2015. Spotfire is not as intuitive for business users as is Tableau or Microsoft Power BI, but the addition of natural language query and AI-driven recommendations should help to make it more accessible.

The new streaming analysis capabilities in Spotfire X will give power users yet more sophisticated functionality for low-latency and IoT scenarios. Other upgrades in release X, including data visualization improvements and design-once, deploy-on-any-device responsive design, will please every Spotfire customer.

Every analytics vendor now seems to claim embedded analytics capabilities, but this is a market niche that Jaspersoft has focused on for a long time. Open source licensing gives developers a cost-effective, flexible foundation. With JasperReports IO, TIBCO is clearly intent on supporting state-of-the-art cloud architecture.

As for my take on the data science moves, let's look at how the broader portfolio comes together.

The Data Intelligence Portfolio

TIBCO executives say underlying data intelligence capabilities including data wrangling, data cataloging and virtualization will be a boon to analytics and data science teams alike.

To get a better understanding of TIBCO's analytics strategy, I sat down with Brad Hopper, vice president, analytics product strategy, who drew out a simple pyramid to explain what he informally calls the data intelligence portfolio. It starts at the top of the pyramid with the data wrangling capabilities built into Spotfire. These features enable individual analysts and data scientists to combine, manipulate, clean up, standardize and otherwise munge data for visual analysis in Spotfire or in data science work.

Data analysts and data scientists don't tend to methodically build out workflows with every data-manipulation step planned out in advance, as would an ETL professional. The focus is on letting these users wrangle the data, yet each step they take is captured and recorded in the background. Even if the steps weren't applied in a linear, planned-out way, Spotfire's Data Canvas retains a logically ordered, documented and trackable workflow of all the joins, transformations and cleansing

and editing steps behind each new source that can be edited, perfected, reused and audited.

The next layer down in the Data Intelligence pyramid is the TIBCO Spotfire Data Catalog, which is built on technology from Attivio that supports machine-learning-based crawling and indexing of all data in the enterprise. This includes data lakes with unstructured information, which the catalog parses and enriches with metadata. Fuzzy matching helps to find relationships among data.

This middle layer is designed to aid search, discovery and data prototyping, so it's aimed at teams using Spotfire, but Hopper also sees it coming into play for data scientists. Teams might develop marts for standardized dashboarding and visualization through Spotfire or reporting through Jaspersoft. The catalog and data wrangling tools will help data science teams spot and develop new data sets.

At the bottom of the Data Intelligence pyramid is TIBCO Data Virtualization (formerly Cisco Information Server/Composite Software). This is where IT can draw on the data workflows captured on the Spotfire Data Canvas and turn them into hardened data marts and sanctioned, production-grade data sources. The heavy lifting is done through virtualization rather than traditional ETL and data-movement. Also described, variously, as data fabric, logical data warehousing or virtual data layer, data virtualization approaches are gaining traction. TIBCO Data Virtualization competes directly with Donodo (which I recently wrote about in combination with Looker [in this recent case study](#)).

MyPOV on TIBCO's Data Intelligence Strategy

Analytics and data science can only thrive where there is ready access to solid, well-managed data. Thus, I do like how the data catalog and data virtualization assets complement Spotfire and, potentially, Tibco Data Science. I've been

following [Composite Software](#) for a long time, so it's good to see stepped up investment in what's now known as TIBCO Data Virtualization. I didn't get to see enough at TIBCO NOW, so I've asked for deeper briefings on how DV works with modern cloud architectures, unstructured sources and object stores such as Amazon S3.

The breadth of TIBCO's portfolio and the one-stop-shop/one-throat-to-choke opportunity it affords is likely to appeal to IT buyers. I'd make a point of getting analytics buyers in the loop, too. You might use DV to deliver IT-sanctioned data sources for data governance purposes, for example. One Spotfire customer lamented, during a DV session at TIBCO NOW, that users want direct, not virtualized, access to the tables and sources they already know. Adoption of TIBCO DV may require a bit of change management, even if it's ultimately a boon to data access and governance.

As for TIBCO's data science portfolio, I see the moves to bring together the Statistica, Alpine and Terr/Spotfire assets on one platform as appealing to existing customers, but I doubt the collection will win over greenfield converts. For one thing, they remain separate products with very basic points of integration at this point. For another, data scientists starting from scratch are embracing all things open source, including Jupyter and Zeppelin notebooks and direct coding of open algorithms. Alpine Data certainly tried to cater to this next-generation crowd, but it wasn't winning in the marketplace.

For now, TIBCO's access-to-everything licensing will appeal to existing Statistica and Alpine Data customers (as will the complement of catalog, data wrangling and DV capabilities, if they're also Spotfire customers). Long term, however, I'd want to see deeper integration and refactoring that truly brings these assets together as a single product worthy of the "platform" billing, with connected workflows and consistent user interfaces.

In short, I like what TIBCO is doing with Spotfire, catalog capabilities and TIBCO DV. Only time will tell if the data science portfolio can become more appealing than the sum of its parts.