Big Data Predictions for 2016

A roundup of big data and analytics predictions and pontifications from several industry prognosticators.

By Andrew Brust for Big on Data | December 30, 2015 -- 19:56 GMT (11:56 PST) | Topic: Big Data Analytics

At the end of each year, PR folks from different companies in the analytics industry send me predictions from their executives on what the next year holds. This year, I received a total of 60 predictions from a record 17 companies. I can't laundry-list them all, but I can and did put them in a spreadsheet (irony acknowledged) to determine the broad categories many of them fall in. And the bigger of those categories provide a nice structure to discuss many of the predictions in the batch.

Predictions streaming in

MapR CEO John Shroeder, whose company just added its own MapR Streams component to its Hadoop distribution, says "Converged Approaches [will] Become Mainstream" in 2016. By "converged," Schroeder is alluding to the simultaneous use of operational and analytical technologies. He explains that "this convergence speeds the 'data to action' cycle for organizations and removes the time lag between analytics and business impact."

The so-called "Lambda Architecture" focuses on this same combination of transactional and analytical processing, though MapR would likely point out that a "converged" architecture co-locates the technologies and avoids Lambda's approach of tying the separate technologies together.

Whether integrated or converged, Phu Hoang, the CEO of DataTorrent predicts 2016 will bring an ROI focus to streaming technologies, which he summarizes as "greater enterprise adoption of streaming analytics with quantified results." Hoang explains

that "while lots of companies have already accepted that real-time streaming is valuable, we'll see users looking to take it one step further to quantify their streaming use cases."

Which industries will take charge here? Hoang says "FinTech, AdTech and Telco lead the way in streaming analytics." That makes sense, but I think heavy industry is, and will be, in a leadership position here as well.

In fact, some in the industry believe that just about everyone will formulate a streaming data strategy next year. One of those is Anand Venugopal of Impetus Technologies, who I spoke with earlier this month. Venugopal, in fact, feels that we are within two years of streaming data becoming looked upon as just another data source.

Internet of predicted things

It probably won't shock you that the Internet of Things (IoT) was a big theme in this year's round of predictions. Quentin Gallivan, Pentaho's CEO, frames the thoughts nicely with this observation: "Internet of Things is getting real!" Adam Wray, CEO at Basho, quips that "organizations will be seeking database solutions that are optimized for the different types of IoT data." That might sound a bit self-serving, but Wray justifies this by reasoning that this will be driven by the need to "make managing the mix of data types less operationally complex." That sounds fair to me.

Snehal Antani, CTO at Splunk, predicts that "Industrial IoT will fundamentally disrupt the asset intelligence industry." Suresh Vasudevan, the CEO of Nimble Storage proclaims "in 2016 the IoT invades the datacenter." That may be, but IoT technologies are far from standardized, and that's a barrier to entry for the datacenter. Maybe that's why the folks at DataArt say "the IoT industry will [see] a year of competition, as platforms strive for supremacy." Maybe the data center invasion will come in 2017, then.

Otto Berkes, CTO at CA Technologies, asserts that "Bitcoin-born Blockchain shows it can be the storage of choice for sensors and IoT." I hardly fancy myself an expert on

blockchain technology, so I asked CA for a little more explanation around this one. A gracious reply came back, explaining that "IoT devices using this approach can transact directly and securely with each other...such a peer-to-peer configuration can eliminate potential bottlenecks and vulnerabilities." That helped a bit, and it incidentally shines a light on just how early-stage IoT technology still is, with respect to security and distributed processing efficiencies.

Growing up
Though admittedly broad, the category with the most predictions centered on the theme of value and maturity in Big Data products supplanting the fascination with new features and products. Essentially, value and maturity are proxies for the enterprise-readiness of Big Data platforms.

Pentaho's Gallivan says that "the cool stuff is getting ready for prime time." MapR's Schroeder predicts "Shiny Object Syndrome Gives Way to Increased Focus on Fundamental Value," and qualifies that by saying "...companies will increasingly recognize the attraction of software that results in business impact, rather than focusing on raw big data technologies." In a related item, Schroeder predicts "Markets Experience a Flight to Quality," further stating that "...investors and organizations will turn away from volatile companies that have frequently pivoted in their business models."

Sean Ma, [Trifacta](#)'s Director of Product Management, looking at the manageability and tooling side of maturity, predicts that "Increasing the amount of deployments will force vendors to focus their efforts on building and marketing management tools." He adds: "Much of the capabilities in these tools...will need to replicate functionality in analogous tools from the enterprise data warehouse space, specifically in the metadata management and workflow orchestration." That's a pretty bold prediction, and Ma's confidence in it may indicate that Trifacta has something planned in this space. But even if not, he's absolutely right that this functionality is needed in the Big Data world. In terms of manageability, Big Data tooling needs to achieve not just parity with data warehousing and BI tools, but needs to surpass that level.

The folks at Signals say "Technology is Rising to the Occasion" and explain that "advances in artificial intelligence and an understanding [of] how people work with data is easing the collaboration between humans and machines necessary to find meaning in big data." I'm not sure if that is a prediction, or just wishful thinking, but it certainly is the way things ought to be. With all the advances we've made in analyzing data using machine learning and intelligence, we've left the process of sifting through the output a largely manual process.

Finally, Mike Maciag, the COO at AltiScale, asserts this forward-looking headline: "Industry standards for Hadoop solidify." Maciag backs up his assertion by pointing to the Open Data Platform initiative (ODPi) and its work to standardize Hadoop distributions across vendors. ODPi was originally anchored by Hortonworks, with numerous other companies, including AltiScale, IBM and Pivotal, jumping on board. The organization is now managed under the auspices of the Linux Foundation.

Artificial flavor

Artificial Intelligence (AI) and Machine Learning (ML) figured prominently in this year's predictions as well. Splunk's Antani reasons that "Machine learning will drastically reduce the time spent analyzing and escalating events among organizations." But Lukas Biewald, Founder and CEO of Crowdflower insists that "machines will automate parts of jobs -- not entire jobs." These two predictions are not actually contradictory. I offer both of them, though, to point out that AI can be a tool without being a threat.

Be that as it may, Biewald also asserts that "AI will significantly change the business models of companies today." He expands on this by saying "legacy companies that aren't very profitable and possess large data sets may become more valuable and attractive acquisition targets than ever." In other words, if companies found gold in their patent portfolios previously, they may find more in their data sets, as other companies acquire them to further their efforts in AI, ML and predictive modeling.

And more

These four categories were the biggest among all the predictions but not the only

ones, to be sure. Predictions around cloud, self-service, flash storage and the increasing prominence of the Chief Data Officer were in the mix as well. A number of predictions that stood on their own were there too, speaking to issues as far-reaching as salaries for Hadoop admins to open source, open data and container technology.

What's clear from almost all the predictions, though, is that the market is starting to take basic big data technology as a given, and is looking towards next-generation integration, functionality, intelligence, manageability and stability. This implies that customers will demand certain baseline data and analytics functionality to be part of most technology solutions going forwards. And that's a great sign for everyone involved in Big Data.

This post was updated on January 4th, 2016, to attribute CA Technologies' prediction to Otto Berkes, the company's CTO. Originally, the prediction was attributed to Aruna Ravichandran, a DevOps expert at CA, which was incorrect.

Original article – http://www.zdnet.com/article/big-data-predictions-for-2016/