

RESEARCH REPORT

The State of Streaming Big Data Analytics: 2014 Survey Results

April 2014

Executive Summary

As the speed of business accelerates, organizations produce increasingly vast volumes of high velocity data in a myriad of formats, commonly referred to as Big Data. The challenge for these organizations lies in being able to transform this deluge of data into instantaneous intelligence that can enable faster, better business decisions. Real-time Big Data Analytics initiatives are being discussed as a means to this end in boardrooms around the globe. But what does real-time Big Data Analytics really mean to these organizations? Does it include the ability to analyze and act on live, streaming data? What benefits do these organizations hope to achieve? And most importantly, are they ready for these initiatives?

In April of 2014, Vitria Technology, Inc. conducted a survey to gain some insight into real-time Big Data Analytics, and in particular, streaming Big Data Analytics perceptions.

The survey highlights:

1. Real-time analytics initiatives are relatively immature. Most firms equate real-time analytics to stored data analysis (e.g., Hadoop-based) – where the query and response is in real-time or near real-time; however the analysis is done on historical data, on-demand. However, 42% of the responses indicated that streaming analytics was essential for real-time analytics indicating a growing recognition of its importance.
2. 26% were considering open-source or home-grown solutions for real-time analytics. However, anecdotal evidence suggests that they recognize that these approaches can prove to be limiting down the line.

3. 26% planned to use stream processing or in-memory technologies for real-time analytics and 57% said that they either had an active streaming analytics initiative underway or planned to engage in one within the next six months. This indicates growing awareness of the value of such technologies for analyzing and acting on real-time insights while it still counts.
4. 32% ranked real-time customer experience monitoring as the top benefit. Real-time fraud detection and real-time infrastructure monitoring ranked #2 and #3 respectively. Real-time one-to-one marketing trailed at #4, further indication that firms aren't ready to reap the true benefits of real-time analytics - generating new revenue streams.

About the Survey Respondents

The survey was conducted at the Big Data Innovation Summit in Santa Clara, California. 91 conference attendees completed the survey.

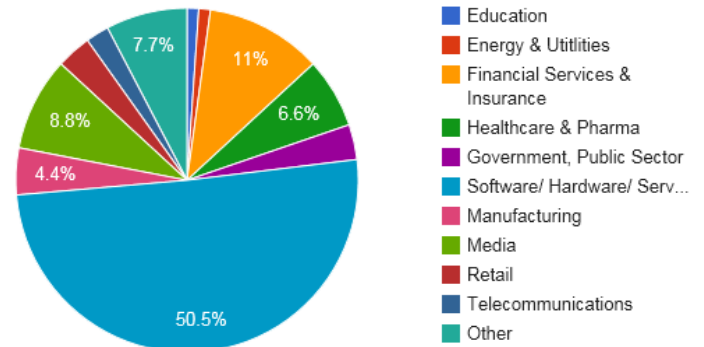


Figure 1: Industry segmentation of respondents

33% of the respondents were from five industry sectors - financial services, media, healthcare, manufacturing and retail. The financial services and insurance sector accounted for 11% of the respondents; the media and healthcare sectors accounted for 9% and 6% of the respondents respectively. 50% of the respondents were from the IT sector.

19% of the respondents identified themselves as data scientists, suggesting a rise in the number of such professionals that are actively engaged in Big Data initiatives. 17% of the respondents were data analysts or business analysts. 13% of the respondents were marketing or marketing analytics professionals. Architects and other IT professionals accounted for 22% of the responses. 9% of the respondents identified themselves as C-level executives.

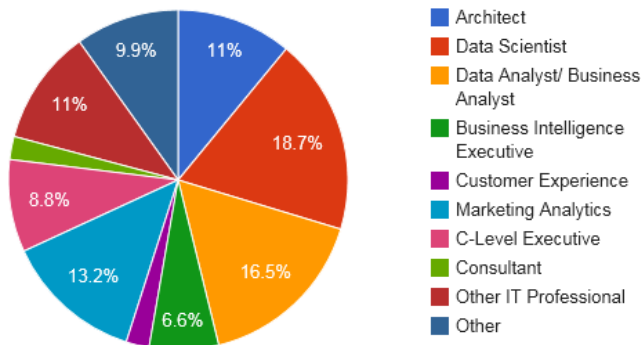


Figure 2: Role segmentation of respondents

The majority of the respondents identified themselves as decision-makers, technology evaluators or influencers for streaming analytics initiatives. 21% chose “decision-maker”, 37% chose “influencer”, and 22% chose “technology evaluator” as the option that best described their role.

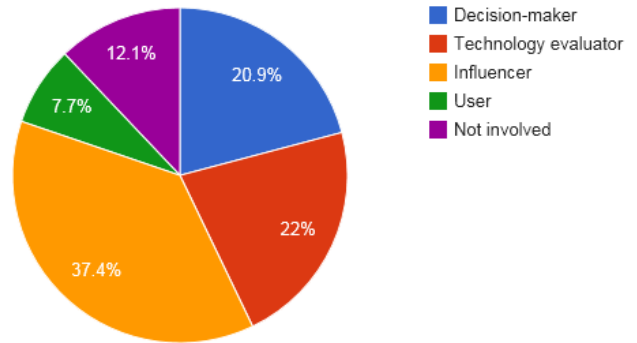


Figure 3: Level of involvement in real-time analytics initiatives

Key Findings

The four major findings from the survey were as follows:

#1: Streaming Data Analytics is Essential for Real-time Analytics, Says 42%

The survey respondents were asked to weigh in on what real-time analytics meant to them. Not surprisingly, the majority of the responses indicated that most people currently equate real-time analytics to either Hadoop-based analytics or some form of stored data analysis – where the query and response is in real-time or near real-time; however the analysis is being done on historical data, on-demand. However, there is a growing recognition for the need for streaming data analysis for true real-time analytics as is evident from 42% of the responses.

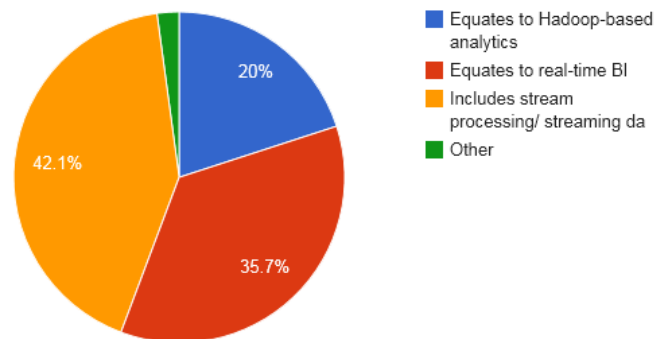


Figure 3: The meaning of real-time analytics

#2: Open-source and Home-grown Solutions are Popular but Firms Recognize Limitations

26% of the respondents said that they were considering the use of open-source solutions or developing solutions in-house for real-time analytics. However, several attendees that had started down this path also admitted to the fact that this approach can mean a trade-off on advanced functionality and that the cost of maintaining custom code was likely to get expensive over time. A respondent from a major telecom provider stated, “Building these capabilities ourselves can get old pretty quickly.”

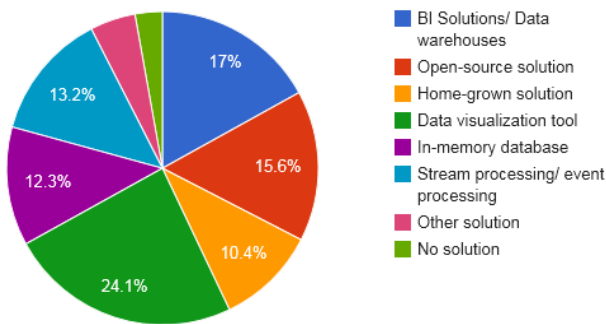


Figure 4: Solutions considered for real-time analytics

26% of the respondents indicated that they planned to use stream processing and in-memory technologies to help with real-time analytics further demonstrating a growing understanding of the need for such capabilities to glean real-time insight that can be immediately acted upon – in seconds and minutes. The fact that 24% said that they were considering data visualization tools and 17% were looking at BI solutions and data warehouses for real-time analytics is indicative of the relative immaturity of these initiatives. For a large number of firms, being able to consolidate data in Hadoop or a data warehouse, perform off-line queries and analysis and visualize the results using data

visualization tools currently equates to real-time analytics.

#3: Real-time Customer Experience Monitoring Tops the Benefits List

The survey respondents were asked to weigh in on the anticipated benefits from real-time analytics. 32% ranked real-time customer experience monitoring as the top benefit. Real-time fraud detection and prevention ranked #2 with 22% of the respondents indicating it as a major benefit. Real-time network and infrastructure monitoring was a close third at 21%, followed by real-time one-to-one marketing at 18%. The fact that the making real-time offers trails at #4 suggests that firms aren’t quite ready to reap the true benefits of real-time analytics - generating new revenue streams. Most firms appear to be testing the waters with asset protection and loss prevention (e.g., fraud detection) initiatives as well as optimization initiatives (e.g., network optimization and customer experience monitoring).

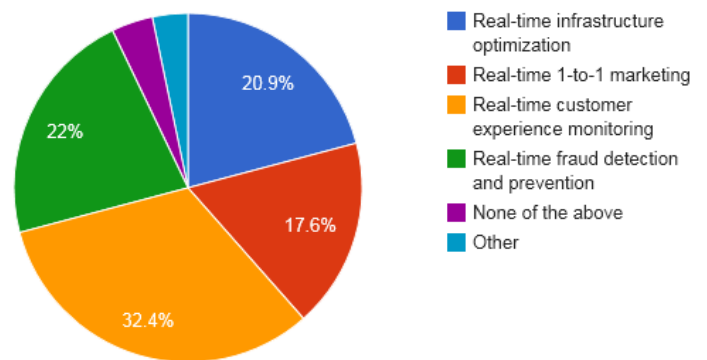


Figure 7: Anticipated benefits from Real-time Analytics

#4: Streaming Analytics Initiatives within the Next Six Months, says 57%

Of those surveyed, 57% noted that they either had an active streaming analytics initiative underway or planned to engage in one within the next six months. This is further indication that companies are beginning to recognize the value of streaming analytics.

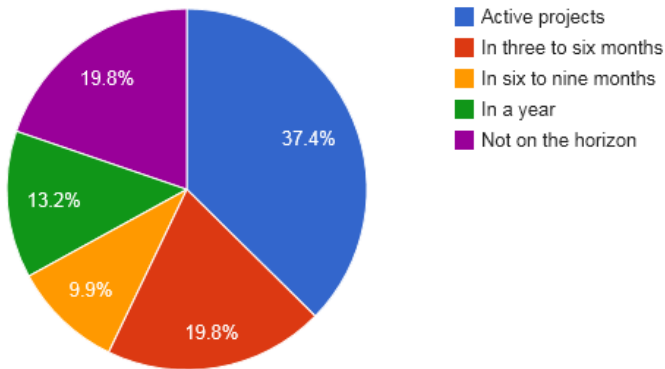


Figure 8: Streaming Analytics Initiatives Timelines

Recommendations

IT organizations that are striving to deliver actionable insight in seconds or minutes to their business counterparts should take stock of their analytics architecture.

Consider complementing your existing enterprise business intelligence (BI), data warehouse and other stored Big Data infrastructure such as Hadoop and NoSQL databases with solutions that can ingest, correlate and analyze streaming Big Data as well to produce continuous, real-time operational intelligence.

Consider streaming data analytics solutions that provide advanced functionality that go well beyond what open-source solutions can provide you with. Evaluate solutions that also incorporate capabilities that let your organization take immediate action on the discovered insights –e.g., immediately send out alerts and notifications, or even kick off a remedial process. When rethinking your analytics infrastructure, look for approaches that afford a quick win around a specific business problem but factor in being able to scale for multiple initiatives.

Appendix

The Vitria survey was conducted in April 2014 using the offline polling technologies from iSurvey at the Big Data Innovation Summit in Santa Clara. The survey was promoted to attendees via conference-specific promotions, Twitter and at the Vitria booth.

Surveys were completed at the Vitria booths using iPads. Survey respondents were also entered in to a drawing to win an iPad. A total of 91 respondents filled out the survey, and the complete results of responses are available upon request.