WHITE PAPER

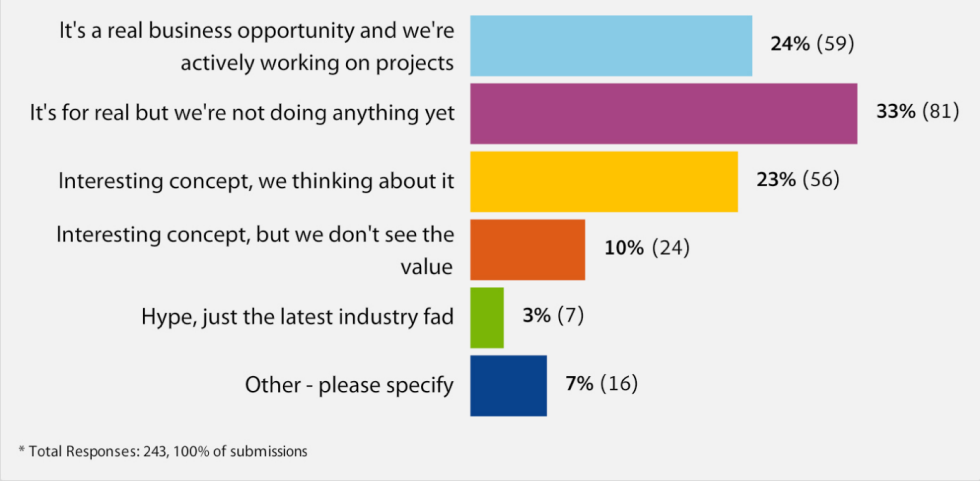
2012 Big Data Survey Results

TREASURE**DATA**
CLOUD DATA WAREHOUSING

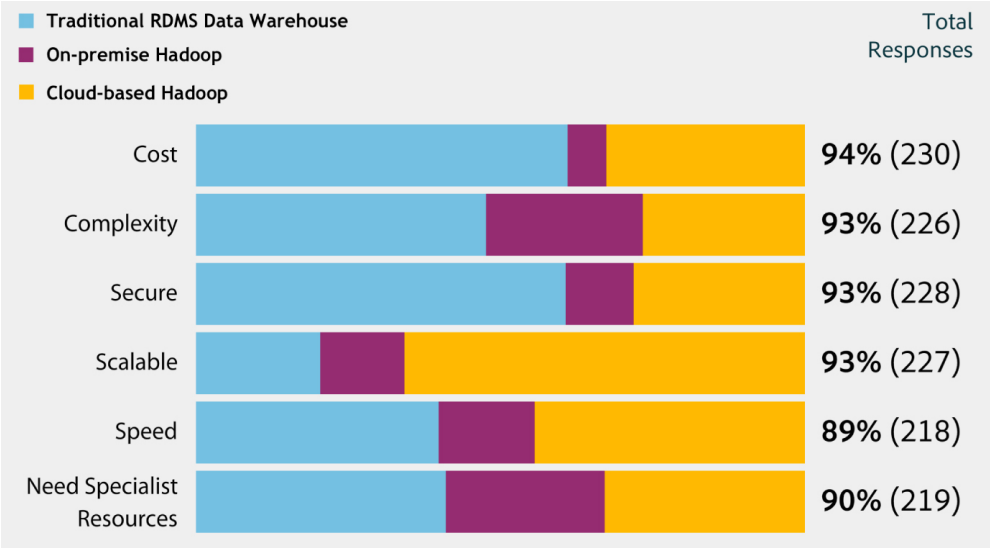
Overview

Meeting the challenges of big data has out-paced traditional warehousing technology and practices. Frustrated by the cost and pace of innovation associated with proprietary vendor solutions, many companies are turning to open-source alternatives to address their needs. To understand how people view these alternatives and to find out what data sources, analytic and presentation approaches they want to use in their big data projects, Treasure Data ran a survey targeting IT and business managers in fast-moving data-driven companies. Based on 243 responses, here are Treasure Data's 2012 Big Data Survey Results.

QUESTION 1:
HOW DO YOU VIEW THE BIG DATA OPPORTUNITY?



QUESTION 2:
WHAT ATTRIBUTES DO YOU MOST ASSOCIATE WITH BIG DATA ARCHITECTURES?



HOW TO INTERPRET THESE RESULTS

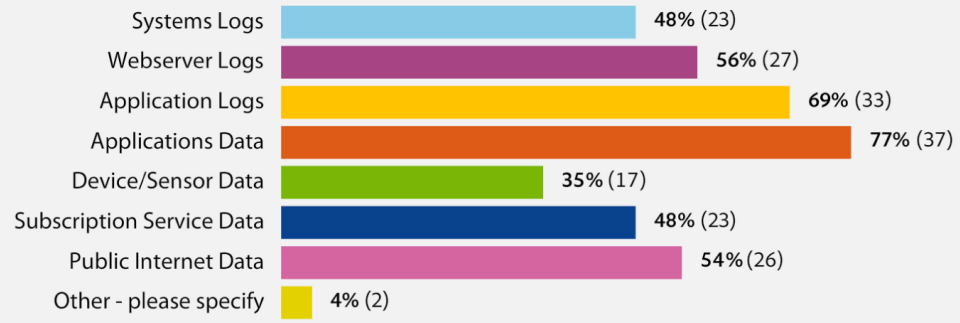
The longer the colored bar, the more respondents associated the attribute with the named architecture. Blue represents traditional data warehousing, purple represents Hadoop and gold represents Cloud architectures. A summary analysis of these results shows:

Traditional data warehousing approaches are seen as costly and complex but are regarded as secure by the majority of survey respondents.

Hadoop was not strongly associated with any of these attributes, however, Hadoop is most strongly associated with complexity and the need for specialist resources.

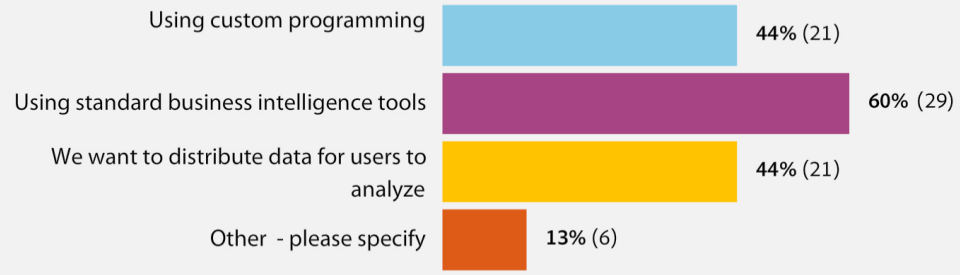
The Cloud is strongly associated with scalability and, more surprisingly, with speed. On the downside, Cloud solutions are not associated with security.

QUESTION 3:
WHAT DATA SOURCES DO YOU WANT TO USE IN YOUR BIG DATA PROJECTS?



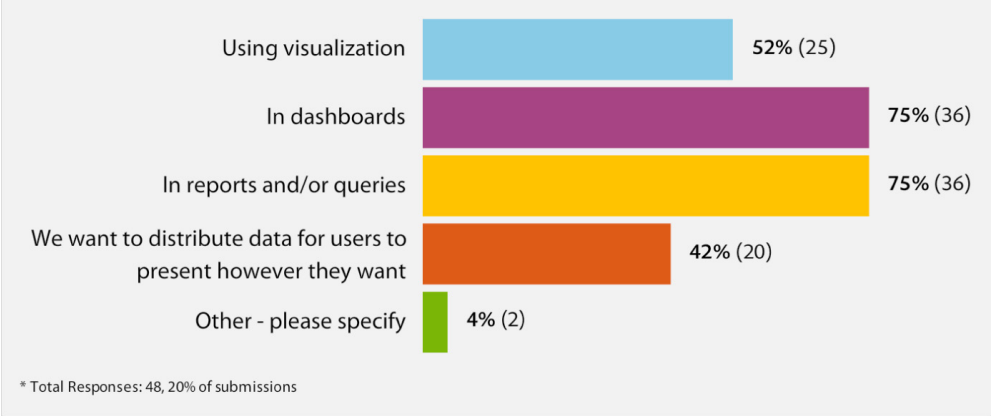
* Total Responses: 48, 20% of submissions

QUESTION 4:
HOW DO YOU WANT TO ANALYZE BIG DATA?



* Total Responses: 48, 20% of submissions

QUESTION 5:
HOW DO YOU WANT TO PRESENT YOUR BIG DATA?



Summary

While most people (90%) regarded big data as an interesting concept, only 24% have active big data projects. However, an additional 56% of people think big data is for real and/or something they are thinking about. 20% of people either didn't see the value of big data to them or didn't feel they understood enough to have an opinion. Interestingly, only 7 out of 243 people (3%) regarded big data as hype.

With regard to big data architectures, traditional data warehousing pproaches are seen as costly and complex. Hadoop is most strongly associated with complexity and the need for specialist resources while the Cloud is strongly associated with scalability and, more surprisingly, with speed. All three approaches were strongly associated with the need for specialist resources.

Of those people that have active big data projects, most are using log files as a primary source of their data but a surprising number are also using Internet feeds (54%) and subscription services (48%). Structured application data (77%) remains the most common single data source.

In terms of analysis and presentation of big data, most live projects use multiple approaches. Using Standard BI tools (60%) to drive reporting, queries and dashboards is the most common approach (75%).

ABOUT TREASURE DATA

Treasure Data combines the power of open source with the scalability and economics of the Cloud and their own patented software innovation, to deliver a big data analytic service that can be running in days not months without specialist IT resources and for a tenth the cost of other alternatives. With customers ranging from social selling start-ups, through online gaming leaders to global corporations, the Treasure Data Cloud Data Warehouse is already hosting over 100 billion rows of live data and is processing 10,000 new messages per second. Treasure Data was founded by some of the smartest engineering minds in the open source world and is backed by leading angel investors from the venture capital, open source and enterprise software markets. For more information, please visit www.treasure-data.com.