



# Big Data Vendor Revenue and Market Forecast 2012-2017

*Contributing authors: Jeff Kelly, David Floyer, Dave Vellante, Stu Miniman  
Original publication date: February 19, 2013*

The hype surrounding Big Data, which showed no signs of abating in 2012, now has big dollars backing it up. Factory revenue generated by the sale of Big Data-related hardware, software and services took a major step forward in 2012, growing by 59% over 2011(a).

The total Big Data market reached \$11.4 billion in 2012, ahead of Wikibon's 2011 forecast. The Big Data market is projected to reach \$18.1 billion in 2013, an annual growth of 61%. This puts it on pace to exceed \$47 billion by 2017. That translates to a 31% compound annual growth rate over the five year period 2012-2017.

## **Growth Drivers and Adoption Barriers**

The growth rate of Big Data revenue in 2012 was due to a number of factors, including:

- An increased awareness of the benefits of Big Data as applied to industries beyond the Web, most notably financial services, pharmaceuticals, and retail;
- The maturation of Big Data software such as Hadoop, NoSQL data stores, in-memory analytic engines, and massively parallel processing analytic databases;
- Increasingly sophisticated professional services practices that assist enterprises in practically applying Big Data hardware and software to business use cases;
- Increased investment in Big Data infrastructure by massive Web properties – most notable Google, Facebook, and Amazon – and government agencies for intelligence and counter-terrorism purposes.

In the enterprise space in particular, the combination of a better understanding of the use cases for Big Data and more mature product and service offerings resulted in a significant percentage of Big Data early adopters graduating from small, proof-of-concept projects to large-scale, production-level deployments. This evolution naturally required increased investment in Big Data hardware, software, and services. Feedback from the Wikibon community included multiple reports of \$100 million+ deals from both government and commercial buyers.

Additionally, a number of enterprises previously reluctant to undertake Big Data projects due to fuzzy ROI, lack of specific business use cases and/or concerns over product and services maturity, began

exploring Big Data in their organizations with small pilot projects, their concerns assuaged by the market potential underscored by the growth factors listed above.

The Big Data market is still within the confines of the early adopter phase and is poised for significant growth. For the Big Data market to reach its full potential, enterprises and vendors must overcome several obstacles. While a detailed discussion of these obstacles is outside the purview of this report, they are worth noting. They include:

- The well-publicized lack of analytic specialists and Data Scientists armed with both the technical skill and business acumen to derive insights from large, multi-structured data sets merged from disparate sources.
- A lack of understanding among enterprises on how to organize Big Data staff to best identify business requirements for Big Data projects and effectively communicate insights gleaned from Big Data to the business.
- Organizational resistance to adopting Big Data analytics-driven decision-making to replace “gut instinct”-style decision-making.
- Vendor marketing overly focused on “speeds-and-feeds,” product features and “Big Data-washing” rather than laying out a vision for Big Data in the enterprise, articulating a path to achieve this vision, and maximizing the potential for Big Data to disrupt well-established vertical markets.
- Development of Big Data platforms and tools by vendors that eschew open frameworks in favor of closed, locked-down solutions. This will limit interoperability with competing and complimentary products and reduce customer choice.
- A lack of best practices and related technologies for managing Big Data as a corporate asset, including data quality, data governance, and security platforms and tools.
- A dearth of Big Data application development tools and services that allow existing developers to build and customize Big Data applications using common and popular application development languages and processes.

## Big Data Vendor Revenue

As part of its market-sizing efforts, Wikibon tracked and/or modeled the 2012 Big Data revenue of more than 60 vendors. This list includes both Big Data pure-plays – those vendors that derive close to if not all their revenue from the sale of Big Data products and services – and vendors for whom Big Data sales is just one of multiple revenue streams.

The complete list is below:

2012 Worldwide Big Data Revenue by Vendor (\$US millions)						
Vendor	Big Data Revenue	Total Revenue	Big Data Revenue as % of Total Revenue	% Big Data Hardware Revenue	% Big Data Software Revenue	% Big Data Services Revenue
IBM	\$1,352	\$103,930	1%	22%	33%	44%
HP	\$664	\$119,895	1%	34%	29%	38%

Teradata	\$435	\$2,665	16%	31%	28%	41%
Dell	\$425	\$59,878	1%	83%	0%	17%
Oracle	\$415	\$39,463	1%	25%	34%	41%
SAP	\$368	\$21,707	2%	0%	67%	33%
EMC	\$336	\$23,570	1%	24%	36%	39%
Cisco Systems	\$214	\$47,983	0%	80%	0%	20%
Microsoft	\$196	\$71,474	0%	0%	67%	33%
Accenture	\$194	\$29,770	1%	0%	0%	100%
Fusion-io	\$190	\$439	43%	71%	0%	29%
PwC	\$189	\$31,500	1%	0%	0%	100%
SAS Institute	\$187	\$2,954	6%	0%	59%	41%
Splunk	\$186	\$186	100%	0%	71%	29%
Deloitte	\$173	\$31,300	1%	0%	0%	100%
Amazon	\$170	\$56,825	0%	0%	0%	100%
NetApp	\$138	\$6,454	2%	77%	0%	23%
Hitachi	\$130	\$112,318	0%	0%	0%	100%
Opera Solutions	\$118	\$118	100%	0%	0%	100%
Mu Sigma	\$114	\$114	100%	0%	0%	100%
TCS	\$82	\$10,170	1%	0%	0%	100%
Palantir	\$78	\$78	100%	0%	63%	38%
Intel	\$76	\$53,341	0%	83%	0%	17%
MarkLogic	\$69	\$78	88%	0%	63%	38%
Booz Hamilton	\$68	\$5,802	1%	0%	0%	100%
Cloudera	\$61	\$61	100%	0%	47%	53%
Action	\$46	\$46	100%	0%	63%	38%
SGI	\$43	\$769	6%	83%	0%	17%
Capgemini	\$42	\$14,020	0%	0%	0%	100%
1010data	\$37	\$37	100%	0%	0%	100%
10gen	\$36	\$36	100%	0%	42%	58%
Google	\$36	\$50,175	0%	0%	0%	100%

Alteryx	\$36	\$36	100%	0%	55%	45%
Guavus	\$35	\$35	100%	0%	67%	33%
VMware	\$32	\$3,676	1%	0%	71%	29%
ParAccel	\$24	\$24	100%	0%	44%	56%
TIBCO Software	\$24	\$1,024	2%	0%	53%	47%
MapR	\$23	\$23	100%	0%	51%	49%
Attivio	\$21	\$26	80%	0%	62%	38%
Fractal Analytics	\$20	\$20	100%	0%	0%	100%
Pervasive Software	\$19	\$51	37%	0%	59%	41%
Hortonworks	\$18	\$18	100%	0%	0%	100%
Informatica	\$17	\$812	2%	0%	78%	22%
QlikTech	\$16	\$321	5%	0%	74%	26%
DataStax	\$15	\$15	100%	0%	59%	41%
Basho	\$14	\$14	100%	0%	63%	38%
Microstrategy	\$13	\$595	2%	0%	59%	41%
Tableau Software	\$13	\$130	10%	0%	59%	41%
Couchbase	\$12	\$12	\$100%	0%	64%	36%
Kognitio	\$12	\$12	100%	0%	47%	53%
Datameer	\$11	\$11	100%	0%	79%	21%
Rackspace	\$11	\$1,300	1%	0%	0%	100%
LucidWorks	\$10	\$10	100%	0%	58%	42%
Digital Reasoning	\$10	\$10	100%	0%	51%	49%
Aerospike	\$8.8	\$8.8	100%	0%	80%	20%
Neo Technology	\$8.5	\$8.5	100%	0%	62%	38%
Think Analytics <sup>Big</sup>	\$7.9	\$7.9	100%	0%	0%	100%
Calpont	\$7.6	\$7.6	100%	0%	59%	41%
RainStor	\$7.5	\$7.5	100%	0%	67%	33%
SiSense	\$7.3	\$7.3	100%	0%	41%	59%
Revolution Analytics	\$7.2	\$13	56%	0%	56%	44%

Talend	\$6.2	\$51	12%	0%	80%	20%
Jaspersoft	\$6.2	\$31	20%	0%	62%	38%
Juniper Networks	\$6.1	\$4,365	0%	71%	0%	29%
Pentaho	\$6.1	\$31	19%	0%	62%	38%
DDN	\$5.9	\$278	2%	63%	0%	38%
Actuate	\$4.6	\$137	3%	0%	63%	37%
Original Device Manufacturers	\$2,375	\$100,000	2%	100%	0%	0%
Other	\$1,593	\$197,170	1%	29%	22%	49%
<b>Total</b>	<b>\$11,448</b>	<b>\$1,223,373</b>	<b>1%</b>	<b>40%</b>	<b>21%</b>	<b>39%</b>

## Methodology

Regarding methodology, the Big Data market size, forecast, and related market-share data was determined based on extensive research of public revenue figures, media reports, interviews with vendors, venture capitalists and resellers regarding customer pipelines, product roadmaps, and feedback from the Wikibon community of IT practitioners.

Many vendors were not able or willing to provide exact figures regarding their Big Data revenue, and because many of the vendors are privately held, Wikibon had to triangulate many types of information to determine our final figures. We also held extensive discussions with former employees of Big Data companies to further calibrate our models.

Information types used to estimate revenue of private Big Data vendors included supply-side data collection, number of employees, number of customers, size of average customer engagement, amount of venture capital raised, and age of vendor.

## Big Data Definitions

It is critically important to understand how Wikibon defines Big Data as it relates to the market size overall and to revenue estimates for specific vendors in particular. Wikibon's definition of Big Data contains two equally important parts.

First, from a technology perspective, Wikibon defines Big Data as those data sets whose size, type, and speed-of-creation make them impractical to process and analyze with traditional database technologies and related tools in a cost- or time-effective way.

Second, Wikibon believes Big Data requires practitioners to embrace an exploratory and experimental mindset regarding data and analytics, one that replaces gut instinct with data-driven decision-making, and exchanges stubbornness for a willingness to question long-held assumptions. Projects whose processes are informed by this mindset meet Wikibon's definition of Big Data, even in cases where some of the tools and technology involved may not.

Based on the above definition, Wikibon includes the following products and services under the umbrella of Big Data:

- Hadoop software and related hardware;
- NoSQL database software and related hardware;
- Next-generation data warehouses/analytic database software and related hardware;
- Non-Hadoop Big Data platforms, software, and related hardware;
- In-memory – both DRAM and flash – databases as applied to Big Data workloads;
- Data integration and data quality platforms and tools as applied to Big Data deployments;
- Advanced analytics and data science platforms and tools;
- Application development platforms and tools as applied to Big Data use cases;
- Business intelligence and data visualization platforms and tools as applied to Big Data use cases;
- Analytic and transactional applications as applied to Big Data use cases;
- Big Data support, training, and professional services.

## **Key Findings: 2012 Big Data Market Highlights and Trends**

Following are key findings, market highlights and trends for the Big Data market in 2012.

- Market-leader IBM offers by far the largest product and services portfolio by both breadth and depth. The company also supports its Big Data practice with a well-crafted, high-level marketing campaign focused around its Smarter Planet initiative that often includes illustrations of real-world Big Data deployments. The biggest criticism of IBM from practitioners is that the company's portfolio is so wide and deep it causes confusion. IBM combats this confusion by initiating many Big Data customer engagements through its professional services division. A challenge and area of focus for IBM moving forward is to continue to articulate its Big Data vision in a way that focuses on industry solutions and not point products.
- HP achieved second-place status in the overall Big Data market by revenue in 2012. It did so mostly thanks to revenue derived from Big Data-related services, followed by sales of hardware to support Big Data deployments. HP by its sheer size is in a position to impact and participate in a number of Big Data deployments.
- Professional services was the largest segment of the Big Data market in 2012. Among firms that derive 100% of their Big Data revenue from professional and/or cloud-based services, which accounts for 39% of the overall Big Data market, the leader by total Big Data revenue was Accenture with \$194 million. Consolidated across vendors, professional and cloud services revenue accounted for \$4.4 billion of total 2012 Big Data revenue.
- Amazon continued and Google kicked off increasingly aggressive moves into the Big Data market. Each introduced new products and services to allow enterprises to leverage Big Data analytics and storage-as-a-service with the usual benefits associated with public Cloud services (elasticity, pay-by-the-drink, trading upfront CAPEX for monthly OPEX, etc.) Specifically, Amazon introduced RedShift, an analytic-database-as-a-service, to its portfolio and struck a deal with MapR to allow customers to run its Hadoop distribution on Amazon Web Service, among other announcements.

Amazon also continued to build its Elastic MapReduce business. Google finally got into the Big Data game by productizing Big Data tools and technologies, such as BigQuery, it has long used internally, and likewise introduced MapR as a service via Google Compute Engine.

- While M&A activity was relatively tepid, two important acquisitions took place in 2012 that have the potential to impact the long-term Big Data market. The first was VMware's acquisition of analytics firm CETAS. VMware had already begun efforts to apply virtualization technology to Hadoop, and the acquisition of CETAS gives the vendor a more comprehensive Big Data portfolio. The creation of the Pivotal Initiative further indicates that VMware and EMC are continuing to invest in Big Data for the long-term.
- The second deal worth noting was WANdisco's acquisition of Hadoop provider AltoStor. WANdisco specializes in data replication across the WAN, which it applies to Hadoop (both its own distribution as well as Cloudera's and Hortonworks' distributions) with the aim of making the open source Big Data framework reliable enough to support mission critical applications.
- Microsoft officially entered the Hadoop market in 2012 with the release of an on-premise Hadoop product - HDInsight Server for Windows – and a cloud-based Hadoop service - Windows Azure HDInsight Service. Both are based on Hortonworks' open source Hadoop distribution. Microsoft also announced PolyBase, which aims to allow the SQL Server Parallel Data Warehouse to execute SQL queries against data stored in Hadoop.
- A movement to bring SQL and NoSQL together in a unified platform was firmly established in 2012. Hadapt and Teradata Aster, which kicked off this movement in 2011 continued to lead the charge but were joined by competitors Cloudera, Microsoft and others in 2012.
- Facebook, Google, and Amazon as well several three-letter government agencies continued to invest heavily in commodity hardware to build out massive internal Big Data infrastructures. Facebook alone spent close to \$800 million on infrastructure in just three quarters in 2012. This spending is reflected in Big Data revenue for the original device manufacturer (ODM) category that appears at the bottom of the table. Specifically, Facebook and others like it purchase, configure, and deploy off-the-shelf hardware from ODM's such as Quanta, rather than purchasing commodity machines from vendors such as Dell or HP, to support the majority of their operations.

### **Spotlight on Hadoop and NoSQL Market Sub-Segment**

As mentioned in the introduction of this report, Hadoop-related software and services matured rapidly in 2012, leading to increased adoption of enterprise-level products by companies in industries beyond the Web. In many cases, companies that had previously deployed community (read: free) versions of vendor Big Data software bundles for proof-of-concept projects began upgrading to paid software and services to support production-level deployments.

As a result, leading Hadoop distribution vendors Cloudera and MapR enjoyed significant revenue growth last year. Cloudera grew revenue to \$61 million in 2012 from \$18 million in 2011. MapR grew revenue to \$23 million in 2012 from \$7 million in 2011. Hortonworks, in its first full year of existence, did \$18 million in revenue in 2012.

Likewise, in the related NoSQL space a handful of vendors that offer commercial versions of popular open source databases enjoyed significant revenue growth as pilot projects blossomed into production deployments supporting real-time, Web-scale applications and services.

Among these vendors is 10gen, which offers a commercial version of the open source, document-oriented MongoDB; Aerospike, whose NoSQL database supports very low-latency online transactional applications; and DataStax, the company behind commercial Cassandra that counts Netflix among its marquee customers.

Leading the way in terms of revenue in the Hadoop/NoSQL subsegment of the Big Data market in 2012 was a 10-year-old firm, MarkLogic. The company’s NoSQL document store is in use at Bank Of America, the Defense Intelligence Agency and Warner Brothers, among other household names in the media and financial services industries.

Ultimately, however, the NoSQL market is largely up for grabs. Each NoSQL database has its related strengths and weaknesses, and no one NoSQL database currently “does it all.” Big Data practitioners must take a number of factors into consideration when selecting a NoSQL database to facilitate large-scale transactional workloads, including scalability, performance, security, and ease-of-development.

Below is a cut out of Big Data revenue associated with those vendors specializing in Hadoop and NoSQL software and services. Note that these vendors account for total Big Data revenue of \$272 million and are growing at a faster percentage rate than the rest of the Big Data market.

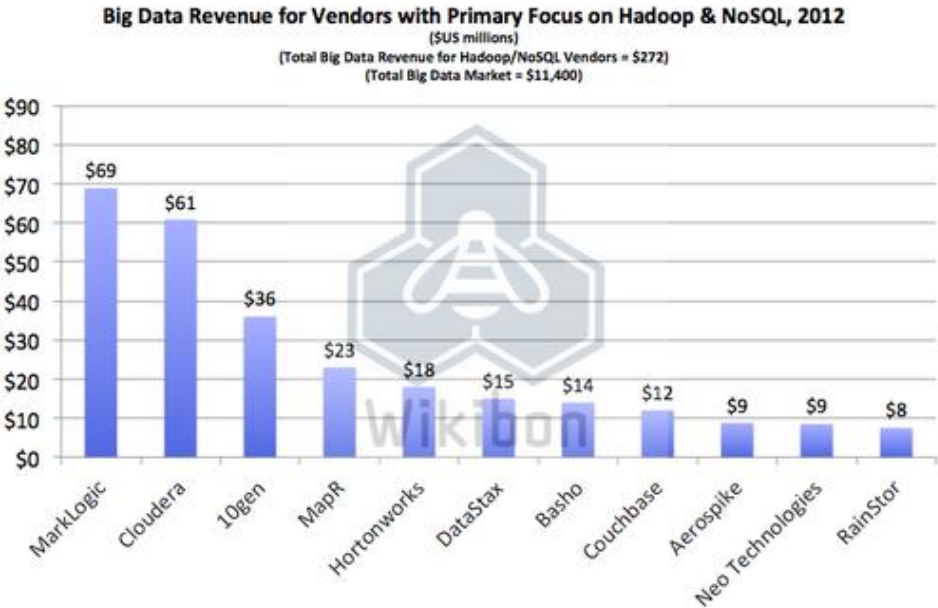


Figure 1 - Source: Wikibon 2013

**Big Data Revenue by Market Segment**

Below is a segmentation of the Big Data market by hardware, software and services.



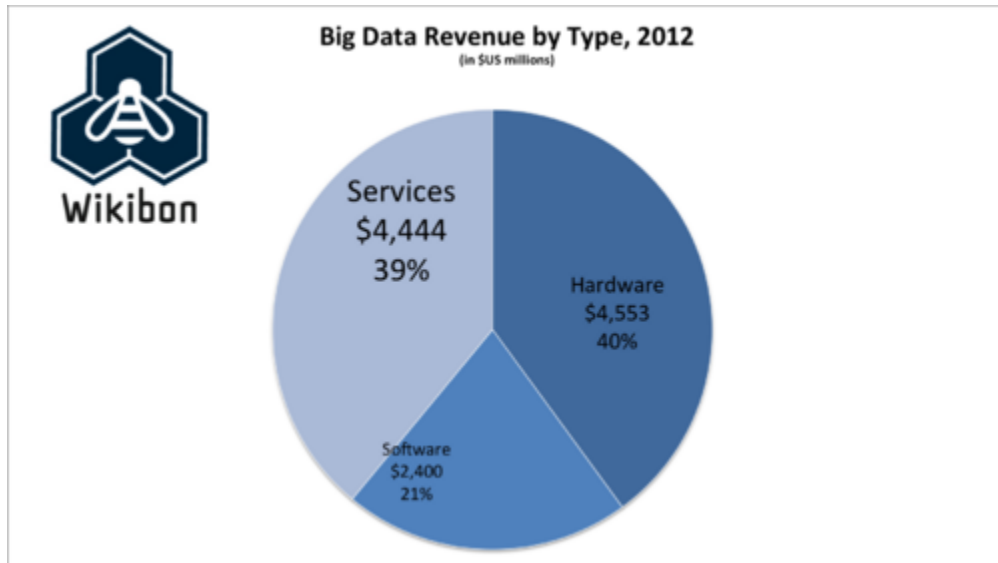


Figure 2 - Source: Wikibon 2013

Wikibon further dissected Big Data revenue by type down to a more granular level.

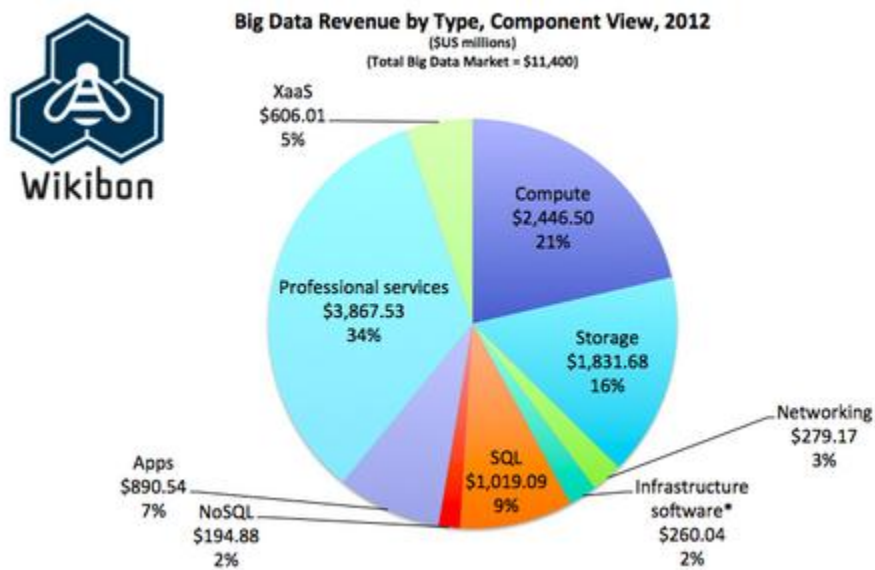


Figure 3 - Source: Wikibon 2013; \*Big Data infrastructure software includes Hadoop, data integration, data quality and other data management software.

### Wikibon's Big Data Forecast

Wikibon projects the Big Data market to top \$18 billion in 2013, a growth rate of 61%. Looking beyond 2013, Wikibon forecasts the total Big Data market to approach \$50 billion by 2017, which translates to a

31% compound annual growth rate over the five-year period 2012-2017. While the global economic outlook is for slow to stagnant growth over this period, Wikibon believes the Big Data market will not be severely impacted and may, in fact, benefit from enterprises needing “to do more with less,” which effective Big Data analytics facilitates.

Wikibon further expects the balance of revenue generation and value to shift from Big Data infrastructure and middleware to value-add services and software over the next five years. As noted, hardware revenue accounts for 40% of Big Data revenue and a large portion of software and services revenue is associated with infrastructure software and technical services that tie Big Data platforms and data together.

Wikibon believes Big Data infrastructure, middleware, and technical services will become increasingly commoditized as they mature and common standards are adopted. Practitioners will increasingly look to NoSQL and in-memory database software, streaming analytic platforms, vertically focused analytical and transactional applications and application development platforms (both on-premise and Cloud-based) and associated consulting and professional services to address specific, high-value business problems and opportunities.

Below is Wikibon’s Big Data market forecast broken down by market component through 2017.

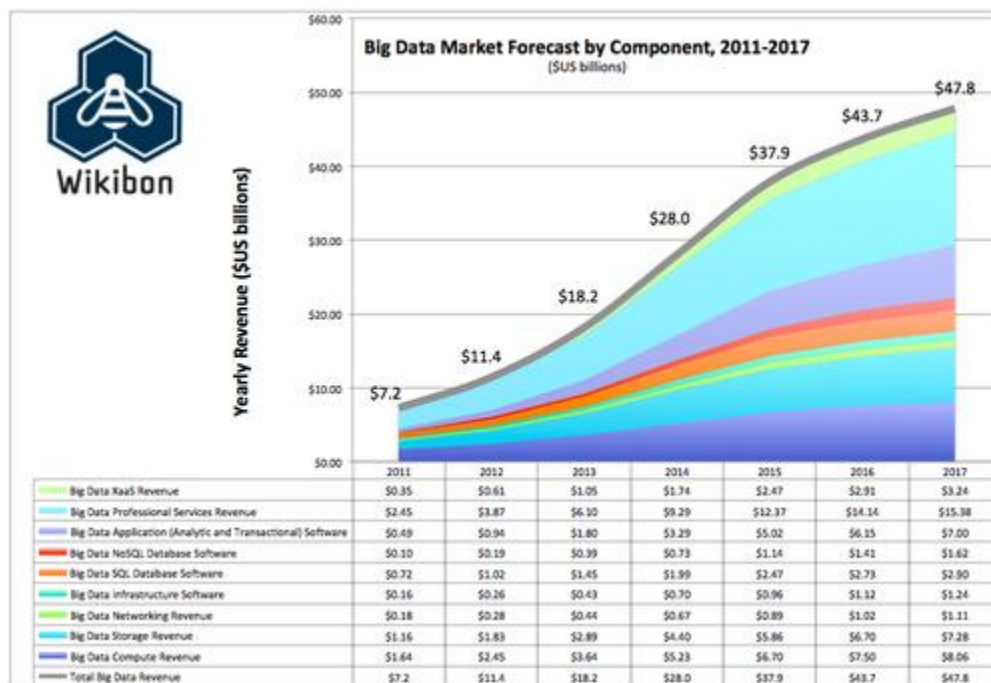


Figure 4 - Source: Wikibon 2013

Wikibon will be looking in more detail at the components that make up the Big Data market, shown in Figure 4. For additional analysis of components of this forecast, see [Big Data Database Revenue and Market Forecast 2012-2017](#) for an examination of SQL/NoSQL and [2012 Big Data Revenue by Vendor](#) for a breakdown by segments.

Action Item: While Big Data vendor revenue is forecast to grow significantly over the next five years, Wikibon believes that Big Data practitioners will create much more value than technology and service providers in the long-term. When selecting vendors to support Big Data initiatives, therefore, CIOs and Big Data practitioners must evaluate the products and services on offer in the context of how best to monetize Big Data to achieve competitive advantage. This includes evaluating “speeds and feeds” and other product features but should also include evaluating how well vendors can assist enterprises in adopting a sustainable culture of data-driven decision-making.