Big Data Vendor Benchmark 2015
A Comparison of Big Data Solution Providers

Germany

An Analysis by
Experton Group AG
Munich, Germany

Authors: Holm Landrock, Oliver Schonschek, Prof. Dr. Andreas Gadatsch
Copyright

This analysis has been prepared by Experton Group. The information and data contained herein were gathered conscientiously and with the utmost care. Nevertheless, we cannot guarantee their exhaustiveness and accuracy. Nobody should take action, based on this information, without expert advice and an in-depth analysis of the individual situation.

All rights to the content of this analysis are reserved by Experton Group. For privacy reasons, the data and information remain the property of Experton Group. Reproduction – even of parts thereof – may only be made with the written permission of Experton Group AG.

Copyright Experton Group, 2014
# Table of Content

1 EXECUTIVE SUMMARY 6

2 PREFACE ON THE SECOND EDITION OF THE BIG DATA VENDOR BENCHMARK 8

2.1 Big Data - Complex Scenarios for New Insights 12

3 MARKET DEVELOPMENT 16

4 THE BIG DATA VENDOR BENCHMARK 19

4.1 Methodology and Analysis Design 19

4.2 Selection of Analyzed Vendors 21

4.3 Analysis Design 22

5 BIG DATA STRATEGY 24

5.1 Big Data Consulting 24

5.1.1 Benchmarked Big Data Consulting Providers 25

5.1.2 Big Data Consulting Evaluation Criteria 26

5.1.3 Benchmark of Big Data Consulting Vendors 26

5.2 Big Data Consulting & Solutions 29

5.2.1 Benchmarked Big Data Consulting & Solutions Providers 29

5.2.2 Evaluation Criteria for Big Data Consulting & Solutions 30

5.2.3 Benchmark of Big Data Consulting & Solutions Vendors 31

6 BIG DATA PROCESSES 36

6.1 Big Data Operations 36

6.1.1 Benchmarked Big Data Operations Providers 37

6.1.2 Big Data Operations Evaluation Criteria 38

6.1.3 Benchmark of Big Data Operations Providers 39

6.2 Big Data Analytics 42

6.2.1 Benchmarked Big Data Analytics Vendors 43

6.2.2 Big Data Analytics Evaluation Criteria 44

6.2.3 Benchmark of Big Data Analytics Vendors 45

6.3 Big Data Syndication, Reporting and Visualization 49

6.3.1 Benchmarked Big Data Syndication, Visualization & Reporting Vendors 51

6.3.2 Evaluation Criteria for Big Data Syndication, Visualization & Reporting 52

6.3.3 Benchmark of Big Data Syndication, Visualization & Reporting Vendors 53

7 BIG DATA SYSTEMS (SOFTWARE) 56

7.1 Big Data Aggregation 56
7.1.1 Benchmarked Big Data Aggregation Vendors 57
7.1.2 Big Data Aggregation Evaluation Criteria 57
7.1.3 Benchmark of Big Data Aggregation Vendors 58

7.2 Big Databases and Big Data Management 59
7.2.1 Benchmarked Big Database and Big Data Management Vendors 59
7.2.2 Evaluation Criteria for Big Databases and Big Data Management 60
7.2.3 Benchmark of Big Databases and Big Data Management Vendors 61

8 BIG DATA SYSTEMS (HARDWARE) 63

8.1 Big Data Appliances 63
8.1.1 Benchmarked Vendors of Big Data Appliances 64
8.1.2 Big Data Appliances Evaluation Criteria 64
8.1.3 Benchmark of Big Data Appliance Vendors 65

8.2 Big Data Storage 67
8.2.1 Benchmarked Big Data Storage Vendors 68
8.2.2 Big Data Storage Evaluation Criteria 69
8.2.3 Benchmark of Big Data Storage Solution Vendors 70

9 BIG DATA SOFTWARE 73

9.1 Big Data Protection 73
9.1.1 Benchmarked Big Data Protection Vendors 76
9.1.2 Big Data Protection Evaluation Criteria 76
9.1.3 Benchmark of Big Data Protection Vendors 77

9.2 Big Data Security Analytics 78
9.2.1 Benchmarked Big Data Security Analytics Vendors 80
9.2.2 Big Data Security Analytics Evaluation Criteria 81
9.2.3 Benchmark of Vendors of Big Data Security Analytics Solutions 82

10 OUTLOOK 85

11 INDEX 87

12 87

13 AUTHORS AND CONTACTS 88
## Table of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1</td>
<td>4-level model used for the Big Data Vendor Benchmark</td>
<td>10</td>
</tr>
<tr>
<td>Figure 2</td>
<td>The big data cycle of knowledge</td>
<td>14</td>
</tr>
<tr>
<td>Figure 3</td>
<td>Overall market development for big data solutions</td>
<td>17</td>
</tr>
<tr>
<td>Figure 4</td>
<td>Market development by 2019, divided by hardware, software and services</td>
<td>18</td>
</tr>
<tr>
<td>Figure 5</td>
<td>Experton Group Market Insight Quadrant</td>
<td>20</td>
</tr>
<tr>
<td>Figure 6</td>
<td>Benchmarked big data consulting providers</td>
<td>27</td>
</tr>
<tr>
<td>Figure 7</td>
<td>Benchmarked big data consulting &amp; solution providers</td>
<td>32</td>
</tr>
<tr>
<td>Figure 8</td>
<td>Benchmarked big data operations providers</td>
<td>39</td>
</tr>
<tr>
<td>Figure 9</td>
<td>Benchmarked vendors of big data analytics software &amp; solutions</td>
<td>45</td>
</tr>
<tr>
<td>Figure 10</td>
<td>Benchmarked big data visualization, reporting &amp; syndication vendors</td>
<td>53</td>
</tr>
<tr>
<td>Figure 11</td>
<td>Benchmarked big data aggregation vendors</td>
<td>58</td>
</tr>
<tr>
<td>Figure 12</td>
<td>Benchmarked big database and big data management vendors</td>
<td>61</td>
</tr>
<tr>
<td>Figure 13</td>
<td>Benchmarked vendors of big data appliances</td>
<td>65</td>
</tr>
<tr>
<td>Figure 14</td>
<td>Benchmarked vendors of big data storage solutions</td>
<td>70</td>
</tr>
<tr>
<td>Figure 15</td>
<td>Benchmarked vendors of big data protection solutions</td>
<td>77</td>
</tr>
<tr>
<td>Figure 16</td>
<td>Benchmarked vendors of big data security analytics solutions</td>
<td>84</td>
</tr>
</tbody>
</table>
1 Executive Summary

This 2014 edition of Experton Group AG’s independent and neutral comparison of big data hardware and software vendors as well as service providers is the second edition of the Big Data Vendor Benchmark.

The market for big data solutions and services is experiencing continuous growth rates, which is also reflected in the growing number of respective vendors and providers. For the purpose of this benchmark, Experton Group has identified more than 200 companies that have included big data into their product and service strategy or corporate communications, compared to 138 companies in the previous year. 121 vendors were identified as relevant for the Big Data Vendor Benchmark 2015 in Germany. These vendors provide a specialized product and services offering to address the challenges of big data scenarios. Experton Group has conducted 212 analyses across various categories.

Within the big data security analytics and big data protection categories, Experton Group analyzed companies with a clear and differentiated SIEM (security information & event management) or data protection offering that meets the specific big data challenges. Related requirements and demands for these two categories were particularly high.

The era of business decision-making processes based on simple reports generated from filtered and preselected data is coming to an end. The focus is no longer on simply counting and comparing KPIs, but rather on finding a statistically significant connection between these KPIs and all related information. A novel view of large structured as well as unstructured and semi-structured data can help gain really new insights.

While some areas of the market, including storage and appliances, have meanwhile reached a certain degree of maturity, this is not the case for all big data solutions and services in the market. The general project consulting business, for instance, still needs to catch up. Many competent project consulting providers have only started to recognize the great relevance of big data in 2013 and 2014. Challenges for these service providers to launch a specific offering and individual solutions to complement the offerings of hardware and software vendors are correspondingly high.

For the first time, Experton Group has also identified a "Rising Star" within each category; these are vendors that have not reached the leader quadrant, but have made significant progress, compared to last year's benchmark, have a very
attractive technological offering or have been included into the respective category for the first time and have reached a good position right away.

In some categories, the largest delta between the two axes is below twenty percent, which indicates a certain degree of market maturity. In these cases, users must thoroughly examine potential vendors and service providers accordingly.

Big data has passed the hype zenith. Some vendors that have taken enormous efforts (and have also set up a Potemkin village in some cases) were not able to maintain their momentum within the reference period for this benchmark - understandably so, since no company is able to constantly come up with innovations for a complex issue such as big data. On the other hand, the big data universe is based on existing hardware, software and network technology (especially existing business intelligence and business analytics solutions), and so, user organizations will soon implement big data scenarios into their business processes.

Experton Group has also found out that many vendors continue to provide international use cases and pilot projects as references, which may indicate that there are no German references yet or that the maturity of these projects is not deemed "presentable" yet. It may, however, also mean that German companies do not want to disclose information on their competitive edge.

Secondary sources, documentations and submitted information, on the other hand, hardly provide answers to the specific issues and demands of the German midmarket businesses and German global players. As a result, the German market differs from a global market perspective.

Users are very reserved when it comes to providing information on projects and project details. As our enquiries among user organizations have revealed, some projects do not pass the pilot project phase and have not gone live accordingly.
Preface on the Second Edition of the Big Data Vendor Benchmark

"Big data" is a much discussed topic, and definitions and markets are changing. While in the past, big data discussions had a technological focus, a shift can be observed to include the business perspective of the big data topic. Process models, analysis concepts and strategic concepts are published to show how companies can leverage the potential of big data solutions to increase their business success. This may be helpful to overcome current restraints and reduce information barriers. Experton Group’s Big Data Vendor Benchmark 2015 is the second comprehensive analysis of the German market for big data solutions. One characteristic of big data is its inherent cycle of knowledge. As the analyses performed for this benchmark have shown, some solutions provide already support for this cycle of knowledge.

Analyst Statement

Users have recognized that cultural change is an essential prerequisite for successful big data scenarios to provide a more in-depth view of huge amounts of data to not only generate better KPIs faster, but to gain insights and knowledge that may revolutionize the whole company. Vendors are on the right track and focus their solutions and services accordingly to reach these goals.

Holm Landrock
Senior Advisor Experton Group AG

While not all companies disclose their big data approaches to maintain their competitive edge, interesting use cases have been published, which reveal the potentials to be achieved with the "oil of the 21st century", i.e., the very fast and in-depth analysis of enterprise and environmental data in multiple representation forms (structured, semi-structured, unstructured). The focus is no longer on a brilliant technical solution (hardware, software) or the variety of data types to be processed, but rather on an innovative strategic approach for leveraging existing technologies.
The IT industry is known for its extremely fast pace of change. Analyses must be conducted and published in near real-time. Experton Group's second Big Data Benchmark documents market changes that have occurred since its first benchmark in 2013. This is the first benchmark that provides a comprehensive analysis of the German market and a neutral view of respective changes (progress, setbacks, stagnation). For the first time, this 2015 edition of the benchmark will also examine the situation in Switzerland.

This second Big Data Vendor Benchmark is a seamless continuation of the first edition, based on a multi-level model which compares the usage of big data solutions (user perspective) systematically with the vendors' offerings (see Fig. 1, 4-level model). The analysis model is based on the traditional IT management pyramid, i.e., strategy – processes – systems, and also accounts for the users' perspective, e.g. the chief information officer's perspective, the process manager's perspective that must optimize workflows or the technology experts' perspective that must provide adequate support for business requirements.
4-Leve Model

Each category analyzed within this benchmark is represented as a quadrant with a "portfolio attractiveness" and a "competitive strength" axis. For each axis, 10 criteria with 5 secondary criteria each are analyzed, which means that 100 individual criteria are evaluated for each category.

On the management level, vendors' strategic approaches and business process support options are analyzed in four categories. Vendors' strategic direction is examined within category 1, "Big Data Projects & Consulting", with two different quadrants: one for vendors with a specific big data consulting offering and one for other vendors where big data is part of the overall portfolio. Process support is examined within the categories "Big Data IT Operations", "Big Data Analytics" and “Big Data-Syndication/Visualization/Dashboards”.

Another four categories address the infrastructure level to examine systems from the hardware and software perspective: “Big Data Aggregation”, “Big Databases and Data Management”, "Big Data Appliances” and "Big Data Storage".

Figure 1: 4-level model used for the Big Data Vendor Benchmark
IT security is a specific aspect which is analyzed within two quadrants across multiple levels. The “Big Data Protection (Protection of Big Data)” quadrant examines how big data applications can be protected against unauthorized access and manipulation. “Big Data Security (Security due to Big Data)” examines how big data solutions can help to increase enterprise security (e.g., through fraud detection, to protect against credit card manipulation).

Overall, the market is divided in eleven quadrants and analyzed accordingly; detailed descriptions are provided in the respective chapters:

- Big Data Consulting & Solutions
- Big Data Consulting
- Big Data Operations
- Big Databases
- Big Data Appliances
- Big Data Storage
- Big Data Aggregation
- Big Data Analytics
- Big Data Syndication/Visualization/Dashboards
- Big Data Protection (protection of large data volumes)
- Big Data Security Analytics (security through big data)

Out of more than 200 companies that are visible in the German market as big data solutions vendors or service providers, 121 were included in the analysis, based on the Experton Group definition.
2.1 Big Data - Complex Scenarios for New Insights

The market is starting to recognize that the components of the Hadoop open-source framework are not a solution for big data scenarios, but rather just components for building big data solutions. Users are also starting to realize the competitive advantages gained by conducting complex analyses and using the results accordingly.

And they understand that faster analyses, faster KPI determination and faster report generation is an intermediate step on their way to complex scenarios.

The market for big data solutions, software, hardware and services is mainly driven by the fast enhancement of underlying basic technologies, especially supercomputing technologies and procedures that are now used in enterprise computing environments, which enables companies to aggregate data from multiple sources and generated by devices and sensors to develop a completely new view of large data volumes.

Data have become the key asset in global competition. Information gathered from these data is a strategic resource within increasingly complex and dynamic market environments.

To achieve a new view of these data, new solutions, algorithms and systems are required, which in turn generate abstractive services that produce new data to be used for gaining new insights, generating even more data. This is the big data cycle of knowledge.

Obviously, a main obstacle to the implementation of big data scenarios are organizational barriers; overcoming technical obstacles, on the other hand, is relatively easy.

Users are becoming more interested in prepackaged platforms and services of ICT service providers, which can be used to test big data scenarios in the sandbox mode to show the potentials of enterprise data that can be enriched with external data.

The times when data were simply counted are over. The challenge is now to gain really new insights through a novel view of large structured, unstructured and semi-structured data.

This new view of data can be compared to forms and figures that we recognize in clouds or pictures illustrating completely unfamiliar situations, such as the "face on
But not all "insights" are really relevant, and so, validating and justifying such insights plays an increasingly important role. This is relevant for big data scenarios, because simply viewing statistical findings, for instance, may result in misinterpretations. For example, data gained through a traffic management system in a big city gave the impression that the sensors for a long section of the road were defective. All sensors showed an average speed of nearly 70 km/h although there is a 50 km/h speed limit for this road. However, nearly all road users travelling along the typical arterial road without traffic surveillance systems had speeds of clearly more than 50 km/h. We can thus conclude that potential and probable explanations for findings gained through big data analyses require domain and mathematical knowledge.

With information and communications technologies pervading practically all areas of our private lives and the business, big data are an inevitable ICT development. For corporate data volumes, which will easily reach the terabyte and petabyte range, new methodologies, algorithms and business processes are required to manage, process, analyze and distribute these data. In turn, information from these data creates an added value which can hardly be imagined today. In many areas, file sizes and the number of files will exceed current data volumes.

---

**Analyst Statement**

*Insights gained through big data analyses must be validated to determine their relevance. Statistical pitfalls must be recognized by data scientists to not be presented "insights" such as that the performance in PISA tests correlates with a person’s shoe size, as published on Facebook.*

*Holm Landrock*

*Senior Advisor Experton Group AG*

---

As such, big data enhance traditional business analytics applications and are characterized by an enormous increase of data sources (including film and audio documents), (internal and external) data users, processing speeds and data volumes.

Market considerations to reduce large data volumes by standardization, filtering or simplification are contrary to the big data concept. It is precisely the large amounts and granularity of structured, semi-structured and unstructured source data that open up completely new perspectives of existing structures and future product, service and business process designs.
Right from the beginning, Experton Group has included data protection and data integrity into its big data definition. Complex scenarios which can only be implemented by using new algorithms and procedures such as hardware encryption generate real competitive advantages. However, complex technological solutions and new procedures also require cultural changes and developments within the company, for instance, when it comes to eliminating internal barriers between divisions, data stores and access rights. On the other hand, such complex scenarios with novel kinds of technologies reward big data users with competitive success that goes beyond modern variants of location-based services.

---

Big data describes the collection of new information which must be made available to high numbers of users in near real time, based on enormous data inventories from multiple sources, with the goal of speeding up critical competitive decision-making processes.

Experion Group, 2011-2015
3 Market Development

Experton Group expects that the big data market will grow to reach nearly 3.2 billion Euros by 2019, which corresponds to a compound annual growth rate of 23.75 percent between 2015 and 2019 and makes big data one of the key drivers of the IT industry during the coming years.

Considering the definition of "big" data, the analyses, predictive exploration of situations and business intelligence workloads are beyond the capabilities of traditional ICT solutions. Therefore, users must invest into new IT infrastructure components such as big data appliance or involve service providers. Big data services differ from cloud services under qualitative as well as quantitative aspects. Another key investment area are database and analytics technologies.

Big data is an enhancement of existing technologies, and so, the market for solutions and services is subject to constant change. The new trend "Industrie 4.0", for instance, creates new data that can be included into complex usage scenarios. In the wake of Industrie 4.0, the big data market will also experience significant growth and development.
The Overall Big Data Market in Germany (in million EUR)

Source: Experton Group

Figure 3: Overall market development for big data solutions

This benchmark and the big data market does include established ICT vendors as well as numerous start-ups and specialized vendors, who constitute the core of specialized big data vendors with an exclusive big data focus. Often, these specialists act as technology and know-how partners of established and global technology and IT service providers, who often take over the responsibility of highly complex and investment-intensive projects. Such specialized big data vendors generate about eight percent of global spending and investments.

Experton Group has also analyzed the distribution of users’ investments across the hardware, software and services segments; as our findings confirm, big data is an issue that requires a lot of consulting and services support. Vendors are challenged to not only provide excellent technical solutions, but also on-premise trainings, workshops and support for users.
The Big Data Market in Germany (By Segments)

<table>
<thead>
<tr>
<th>Year</th>
<th>Hardware (in million EUR)</th>
<th>Software (in million EUR)</th>
<th>Services (in million EUR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015</td>
<td>297</td>
<td>325</td>
<td>723</td>
</tr>
<tr>
<td>2016</td>
<td>377</td>
<td>417</td>
<td>971</td>
</tr>
<tr>
<td>2017</td>
<td>459</td>
<td>506</td>
<td>1,246</td>
</tr>
<tr>
<td>2018</td>
<td>542</td>
<td>594</td>
<td>1,552</td>
</tr>
<tr>
<td>2019</td>
<td>626</td>
<td>679</td>
<td>1,893</td>
</tr>
</tbody>
</table>

Source: Experton Group

Figure 4: Market development by 2019, divided by hardware, software and services
4  The Big Data Vendor Benchmark

The Big Data Vendor Benchmark 2015 is based on a standardized matrix, comprising 100 characteristics, including 50 portfolio attractiveness and 50 competitive strength characteristics.

For the purpose of this analysis, Experton Group has focused on the enhancement of solutions presented in last year’s benchmark, proven market success and new developments of vendors.

4.1  Methodology and Analysis Design

This „Big Data Vendor Benchmark 2015“ is based on the “Experton Market Insight” methodology developed by Experton Group AG. This validated and internationally acknowledged methodology serves as basis for the evaluation and positioning of the individual vendors.

For each vendor, a detailed scoring based on 10 key and additional secondary criteria is provided for each product category. These criteria are weighted, based on the respective product category, resulting in an assessment of the individual big data offering’s attractiveness ("portfolio attractiveness") and the strength of the individual vendor ("competitive strength").

The benchmarked vendors had to supply answers to the questionnaire by June 12, 2014.

The “Experton Market Insight” Quadrant contains four segments where the vendors are positioned accordingly:

---

**The Big Data Vendor Benchmark**

The Big Data Vendor Benchmark 2015 is based on a standardized matrix, comprising 100 characteristics, including 50 portfolio attractiveness and 50 competitive strength characteristics.

For the purpose of this analysis, Experton Group has focused on the enhancement of solutions presented in last year’s benchmark, proven market success and new developments of vendors.

**4.1  Methodology and Analysis Design**

This „Big Data Vendor Benchmark 2015“ is based on the “Experton Market Insight” methodology developed by Experton Group AG. This validated and internationally acknowledged methodology serves as basis for the evaluation and positioning of the individual vendors.

For each vendor, a detailed scoring based on 10 key and additional secondary criteria is provided for each product category. These criteria are weighted, based on the respective product category, resulting in an assessment of the individual big data offering’s attractiveness ("portfolio attractiveness") and the strength of the individual vendor ("competitive strength").

The benchmarked vendors had to supply answers to the questionnaire by June 12, 2014.

The “Experton Market Insight” Quadrant contains four segments where the vendors are positioned accordingly:
Leaders:

The „leaders“ among the vendors have a highly attractive product and service offering and a very strong market and competitive position; they fulfill all requirements for successful market cultivation. They can be regarded as opinion leaders, providing strategic impulses to the market.

Market challengers:

„Market challenger“ are also very competitive, but there is still significant portfolio potential and they clearly lag behind the „leaders“.

Product challengers:

The “product challengers” offer a product and service portfolio that provides an above-average coverage of corporate requirements, but are not able to provide the same resources and strengths as the leaders regarding the individual market cultivation categories, often due to the size of the respective vendor or the weak footprint in the respective target segment.
Followers:

"Followers" are still lacking mature products and services or sufficient depth and breadth of their offering, while also showing weaknesses and improvement potentials in their market cultivation efforts.

"Rising Star":

For the first time, Experton Group has also identified a "Rising Star" within each category; these are vendors that have not reached the leader quadrant, but have made significant progress, compared to last year’s benchmark, have a very attractive technological offering or have been included into the respective category for the first time and have reached a good position right away.

No sponsoring agreements are concluded for Experton Group’s vendor benchmarks, and marketing content and time and content of marketing activities for the benchmarks are separated from the research phase. The evaluations constitute a qualitative positioning and not a ranking.

4.2 Selection of Analyzed Vendors

This Big Data Vendor Benchmark analyzes all vendors that are active in the German market with their own hardware, software or service offering for handling big data scenarios in user organizations. International vendors that are not visible in the German geography have not been included. This does, however, not mean that these vendors do not provide a big data offering. Rather, cooperation between users and these vendors would not be possible within the German or even European jurisdiction or the respective vendors say that they have no dedicated market approach for Germany. As thus, the benchmark provides a complete picture of the relevant vendors in Germany.

Vendors also associate the new hype topic "Internet of Things" with big data, and one single vendor may pack all respective buzzwords into two sentences; this concerns buzzword such as: „Consulting services, big data core competence, business intelligence, dashboards, data discovery, data warehouse, infrastructure, IoE (Internet of Everything), added value, operational processes, partners, predictive business intelligence, reporting, sensors and detectors, transactional processes, multiple data sources". Such lists, as provided by several vendors, illustrate that this topic is still very marketing-driven, while there is still scope for improvement of the technological approach.

For the purpose of this analysis, products, technologies and services described this way are regarded as individual portfolio components. Another criterion for the
analysis were projects that demonstrate that the vendors have achieved practical implementations between last year's Big Data Vendor Benchmark and the deadline for this year's benchmark (June 12, 2014 for Germany).

A thorough analysis was performed to determine whether vendors act within a "grey zone", consisting of other topics such as cloud computing, traditional business intelligence & reporting and high-performance computing. Vendors whose portfolio has a focus on such solutions were not included in the analysis. Also, sample projects, use cases and pilot projects of analyzed vendors within these areas were not accounted for within the analysis. On the positive side we have observed that the share of vendors that act within this grey zone has decreased, compared to last year’s benchmark.

4.3 Analysis Design

The analysis for this „Big Data Vendor Benchmark 2015” was conducted based on four phases:

Research:

Extensive secondary research was performed to ensure sound data for those evaluations that are not based on the vendors' own information; it included a review of the individual vendors' offerings and also an analysis of their Internet presences, product specifications and marketing materials.

Many interviews with product managers, technology experts and vendors' customers also contributed greatly to this benchmark. Based on numerous consulting engagements on the user side, Experton Group also has comprehensive experience when it comes to the assessment of the actual performance of the individual vendors.

Vendor survey:

A vendor survey, which was conducted, based on a questionnaire and interviews between our analysts and executives on the vendor side, formed another key part of this study.

The questionnaire comprised 58 questions for vendors and 59 questions for service providers within eleven categories that covered the complete scope, from strategy and technology to product and service offerings and support. The questionnaires were designed to address the offerings of IT service providers on the one hand and technology and solution vendors on the other hand.
Samples:

Samples were used to validate the vendors’ information through telephone calls to verify the information collected with the questionnaire. Also proposals were requested and sample checks of references were performed.

Benchmark:

The three previous steps formed the basis for the final evaluation and vendor positioning within this “Big Data Vendor Benchmark 2015”. Information and insights gained in the individual steps were consolidated and analyzed, based on the predefined criteria.
5  

Big Data Strategy

5.1 Big Data Consulting

To be able to use the various kinds of big data technologies, users must analyze their business strategy, processes, data structures, existing information systems etc. to determine potential fields of action and develop and implement respective use cases. In many cases, this requires external consulting expertise.

Compared to last year's benchmark, the group of companies that provide big data consulting services has grown significantly, which requires a differentiated analysis approach, since these providers pursue different consulting approaches, depending on their respective business strategy.

Big data can be viewed as an interdisciplinary function, and in this case, big data consulting is part of the provider's general business strategy and business process consulting portfolio. Alternatively, big data is handled as a separate topic, and in this case, the provider pursues a specific big data consulting approach, complemented by respective solutions.

The "big data consulting" category comprises providers that have positioned and integrated big data consulting services within their overall consulting offering as basic services to support their customers' strategy and business processes.

Analyst Statement

Big data consulting competence is not only about technological aspects. Customers do not only expect providers to help them with the strategic analysis of the respective decision-making situation, but also to develop sound individual use cases and implement concepts within operational business processes.

Prof. Dr. Andreas Gadatsch  
Senior Advisor Experton Group AG
Benchmarked Big Data Consulting Providers

The following 30 companies that have integrated their big data consulting services into their general consulting and project services offerings were included in this category:

- Accenture
- arcplan Information Services
- Ancud
- arvato Systems
- Atos IT Solutions and Services
- b.telligent
- BearingPoint
- CGI
- Comma Soft
- DATA MART Consulting
- Freudenberg IT
- FRITZ & MACZIOL Software und Computervertrieb
- Fujitsu Technology Solutions
- Informatica
- Inovex
- itelligence
- Mauerberg
- metafinanz Informationssysteme
- msg Systems
- odoscope Technologies
- ParStream
- PBS Software
• pmOne
• Realtech
• RhinoData
• SAS Institute
• SHS Viveon
• Software AG
• TCS Tata Consultancy Services
• YMC

Providers with a focus on management consulting were not included.

5.1.2 Big Data Consulting Evaluation Criteria

The evaluation criteria for this category are identical with the criteria for the "Big Data Consulting & Solutions" quadrant, where providers with a specific big data consulting offering were analyzed.

5.1.3 Benchmark of Big Data Consulting Vendors

The benchmark of vendors without a specific big data consulting offering shows a very clear profile: Very few (only four out of 30) providers could be positioned as leaders (Atos, Informatica, SAS and Software AG). This means that while their big data consulting business is integrated into their general services offering, the quality of their big data consulting services is as high as the quality of offering of providers with a specific big data consulting strategy. Ancud has been identified as "Rising Star".

Except for three challengers (itelligence, Fujitsu and arvato) all other providers were positioned as followers, although great differences can be observed between these providers. Altogether, this situation is due to the fact that many consulting houses still pay not enough attention to the big data consulting business, missing potential revenue opportunities.
Figure 6: **Benchmarked big data consulting providers**

Atos is one of the international consulting houses that has a strong presence in Germany and has included big data as a key aspect of their business strategy for quite some time. The target market includes midmarket businesses as well as the large enterprise segment. The Bull acquisition should help Atos to further improve their big data consulting potentials. Atos pursues a product-independent approach, allowing for a neutral support of customers, from a big data deployment analysis to strategy development and implementation.

US-based provider Informatica, whose offering includes a big data development environment, has several locations in Europe, including Frankfurt and Stuttgart. The offering addresses German requirements, with a focus on data management, in particular, master data management, data migration and data preparation; it is complemented by customer events, e-books and the Informatica Academy.

SAS is a global player that has been present and active within the big data segment for a long time, based on its core business. While more or less all staff members are working within the big data business, the provider is challenged to communicate its big data strategy consulting approach more effectively.
For the first time, Software AG, one of the most renowned and leading German software vendors, has been included into the „Projects & Consulting“ analysis. The "digital transformation" strategy, as formulated by CEO Karl-Heinz Streibich, is a challenge for Software AG’s consulting competence. Surprisingly, big data consulting is a component of the vendor's general consulting services; after all, the company pursues a dedicated big data strategy within the context of in-memory data management and real-time analytics and has included other traditional business segments such as business process management into its portfolio. This shows that big data has become an interdisciplinary topic that touches many other consulting segments, a fact that does certainly not have a negative impact on the overall leader position.

With its Deutschlandcard project, arvato Systems with headquarters in Gütersloh was among the 2012 finalists of a big data competition and was again included in our analysis within this category. The provider is challenged to better communicate its current big data portfolio in the market.

For technology providers such as Fujitsu big data is a core topic that has an impact on various products and, consequently, is supported through a broad portfolio of consulting services. It is an interesting question whether the vendor will further focus its strategy on the big data business.

For the first time, the Cologne-based start-up odoscope Technologies, winner of the Customer Experience Innovation Contest (source: Finanzen.net), has been included in the "Projects & Consulting" category. The company has built up a successful real-time operational intelligence business. Within the big data market, operational intelligence applications are among the younger technologies with quite some potential. The company's target group are customers from marketing or user experience departments that need an ad-hoc optimization solution for their web sites, based on real-time information.

itelligence, a system house with headquarters in Bielefeld, has comprehensive SAP project experience with a focus on German and European midmarket customers and is one of the newly added players in this quadrant. The company's big data activities have a focus on SAP HANA. Considering the reservedness of midmarket businesses when it comes to big data, itelligence is in a good position to acquire big data projects from their existing customer base.

RhinoData was founded, based on successful user projects of SMC IT Augsburg, and has been included in the Big Data Vendor Benchmark for the first time. RhinoData’s mission is to market their own service offering, based on their know-how gained through the implementation of big data solutions, for instance, with Exasol.
Ancud IT Beratung provides e-commerce, mobile app development and process consulting services. Increasingly, big data plays a major role within the provider’s consulting business. Ancud’s "Big Data Lab" can help users to get a first orientation to solve their big data tasks. The position as "Rising Star" should encourage the provider to address "higher-level" challenges.

5.2 Big Data Consulting & Solutions

To be able to use the various kinds of big data technologies, users must analyze their business strategy, processes, data structures, existing information systems etc. to determine potential fields of action and develop and implement respective use cases. In many cases, this requires external consulting expertise.

Compared to last year’s benchmark, the group of companies that provide big data consulting services has grown significantly, which requires a differentiated analysis approach, since these providers pursue different consulting approaches, depending on their respective business strategy. Either, big data is viewed as an interdisciplinary function, and in this case, big data consulting is part of the provider’s general service portfolio. Alternatively, big data is handled as a separate topic, and in this case, the provider pursues a specific big data consulting approach, complemented by respective solutions.

The "Big Data Consulting & Solutions" category comprises providers that have a specific big data consulting and also a specific big data solutions offering, for instance, a specific process or assessment model for big data projects or a specialized consulting competence center.

5.2.1 Benchmarked Big Data Consulting & Solutions Providers

The following 19 companies with a specific big data consulting and/or project offering in Germany were analyzed within this category:

- All for One Steeb
- Capgemini
- Computacenter
- CSC
- cundus
- eoda
- Hewlett-Packard (HP)
- IBM
- Microsoft
- Neofonie
- Pricewaterhouse Coopers (PwC)
- RELEX
- SAP
- Steria Mummert Consulting
- T-Systems
- The Unbelievable Machine (*UM)
- TIBCO Software
- TNS Infratest
- USU

Pure telecommunications carriers that provide basic platforms for big data applications, such as fast data networks, as well as companies with a focus on management consulting were not included in the analysis.

5.2.2 Evaluation Criteria for Big Data Consulting & Solutions

We have optimized the current evaluation criteria for the two consulting categories, based on our experience with the first Experton Group Big Data Vendor Benchmark, and have developed and evaluated specific criteria to adequately address the portfolio attractiveness and competitive strength of the evaluated service providers.

For instance, within the portfolio attractiveness evaluation, a special focus was laid on the scope of the providers' consulting services offering, including the development of a big data strategy, the identification of use cases, the analysis and adjustment of processes, technology selection support and strategy and concept implementation. Also, we have examined which kind of consulting services are offered and whether the providers have respective experience, including experience with traditional projects, coaching services or pilot projects, for instance based on agile methodologies. Other important aspects included
implemented projects and support of specific methodologies, such as standard models or specifically developed models, check-lists or benchmarks. Another criterion was the provider's offering of trainings for customers and its own employees, for instance, a data scientist training. Some companies have built up an extensive training offering, sometimes in cooperation with universities and academies.

To determine the providers' competitive strength, we have examined their technology or big data leadership. It was important to find out whether the management board and employees identify with the big data topic. While in some companies "all employees can handle the big data issue", it is often not given full attention and "one among many" tasks. Considering the multitude of hardware and software vendors, vendor neutrality was a key evaluation criterion. There is a great scope of vendors, from pure software vendors that only offer consulting for their own products, to typical consulting houses that have no own product offering and can therefore act as neutral providers. Other aspects included the respective company's big data core competence, as opposed to big data as a new "side business", and the development of own, innovative consulting approaches.

In addition to these specific criteria, many other aspects were included in the evaluation, such as the corporate vision and strategy of the respective providers.

Another important criterion to protect customers' investments is the providers' financial situation and stability. Other criteria included the availability of reference projects within the German market, the number and the quality of national and international partnerships.

5.2.3 Benchmark of Big Data Consulting & Solutions Vendors

Eight companies were positioned in the leader quadrant: Capgemini, Computacenter, HP, IBM, PwC, SAP, Steria Mummert and T-Systems. Microsoft is the „Rising Star“ of this category.

As has been the case last year, many providers position themselves across all industries of the market. However, great differences can be observed when it comes to the breadth and depth of their consulting offerings and the methodology support. On the positive side we have found out that some providers have made targeted investments to improve their big data consulting competence, which is, for instance, reflected in specific big data competence centers, a larger, well-trained consulting team, stronger partnerships and networking between providers.
As of today, many customers have not thought much about big data projects. Therefore providers must significantly expand their strategic consulting offering to identify potentials and use cases for their customers and develop respective implementation strategies. Developing specific consulting (e.g., big data assessments) and training offerings (e.g., data scientist) could help providers to improve their market position and also their positioning in the next edition of the Experton Group Big Data Vendor Benchmark.

For years Capgemini has had a large team and a comprehensive business information management offering for practically all industries. Big data has been integrated as a strategic topic. The provider’s target group are midmarket businesses and also larger enterprises. For frequently asked questions, for instance, concerning fraud detection or smart meters, prepackaged business analytics solutions, based on common standard products, are available.

Computacenter supports large customers from all industries to help them set up big data infrastructures with a focus on Hadoop and private cloud infrastructures to enable them to gain business-relevant information from their data, in cooperation with Woodmark Consulting, an experienced consulting house. The „Big Data Competence Center” provides extensive consulting know-how on the big data
appliances of leading vendors such as IBM or SAP. Computacenter plans to further strengthen and develop its existing big data strategy, which will be reflected in the next benchmark accordingly.

eoda has specialized in the analysis of structured and unstructured data and the visualization of the analysis results. The portfolio comprises consulting, software development, integration and training. eoda's Results as a Service offering for users can be purchased as a software product or a cloud service. Results as a Service is a modular analysis solution. eoda users from the manufacturing industry use this solution for condition monitoring and predictive maintenance tasks.

Hewlett Packard is a global player with a strong footprint and also active in the German big data market. Big data is a key pillar of the vendor's business strategy, and HP makes high investments (e.g., acquisitions) and develops innovative products (e.g., set-up of a sensor network) for its big data business. While the comprehensive consulting offering has a strong focus on HP's own products and services, it is nevertheless a convincing offering, due to its end-to-end approach. External certifications could help the provider to deliver a neutral perspective of its offering.

For IBM data constitute a natural resource, and so, big data are of strategic relevance. IBM’s big data platform provides the required hardware-software basis to also act as advisor to help implement their customers' projects. The consulting approach includes comprehensive customer support as well as specific offerings for software developers (BlueMix from the cloud). The provider should communicate its consulting competence for German companies more clearly.

PwC has been included in the Big Data Vendor Benchmark for the first time and was able to reach a leader position right away, thanks to the provider's comprehensive big data consulting offering. The consulting house provides customers methodologically sound and neutral consulting services and can leverage numerous installations within its own big data laboratory for customer projects.

Since Walldorf-based SAP AG has presented their HANA in-memory database, the vendor, who has been the dominant leader in the enterprise software market for many years, has also become one of the large big data players. Meanwhile, SAP has also enhanced the required consulting offering greatly, also through cooperation with external partners. Key components include the comprehensive trainings offering, part of which is also available as e-learning service, and specific SAP HANA cloud assessments.
Steria Mummert is an international consulting company with neutral big data consulting services, since the provider is engaged in technology partnerships, rather than selling their own hardware or software. While Steria Mummert has a convincing big data consulting offering, the provider could communicate this competence even better to address this topic more specifically among its customers. Steria Mummert pursues a convincing, method-based BI consulting approach, from strategy consulting to use case identification, feasibility analyses and implementation support, which also comprises big data.

T-Systems is an IT service provider and system integrator with a long-standing and broad big data product and service offering; the company was among the successful leaders of last year’s Experton Group Big Data Vendor Benchmark. Based on a specific model, T-Systems has enhanced and strengthened its already comprehensive consulting offering, in particular through standardized fixed-rate consulting offerings such as the "Innovation Workshop" or the "In-Memory Assessment" to guide customers from their initial idea to productive operations. Other positive services included „Hadoop-as-a-Service“. As a result, T-Systems’ position within this benchmark was improved considerably.

Three companies have just barely missed the leader quadrant of this category: All for One Steeb, CSC and The Unbelievable Machine.

All for One Steeb is one of those providers that have integrated big data into their business strategy, which is also reflected in the consulting offering. Many big data projects and targeted acquisitions (aventum Consult) show that big data is perceived as part of the core business and shall be expanded even further.

Analyst Statement

Specific consulting tools can be used to provide methodological support for big data technology implementations. Templates for typical usage scenarios help the customers to map their initial situation with potential target situations.

Prof. Dr. Andreas Gadatsch
Senior Advisor Experton Group AG
CSC was included in last year's benchmark. The company is one of the large big data analytics providers and has meanwhile expanded and refined their range of methodologies to provide method-based support for customer projects, including comprehensive strategy consulting and a phase model, including templates for project implementation.

The Unbelievable Machine (*UM) is one of the "newcomers" in this benchmark and category. The Berlin-based start-up is a specialized big data full-service provider with a high degree of flexibility to address individual customer requirements. *UM barely missed the leader quadrant and has achieved a remarkably good position.

Microsoft is an IT pioneer and global player in the IT industry. The company was not included in the projects & consulting category of last year's benchmark and has now been positioned as "Rising Star". Meanwhile, Microsoft's specialized big data consulting offering is visible in the market and it is interesting to observe how the vendor acts in this market. Microsoft does not only provide comprehensive big data products and services, but also uses these products and services within their own core business (e.g., the big data platforms Bing, Skype). The consulting offering has a focus on the analysis of sales and marketing data, social sentiment data, satisfaction analyses and predictive maintenance based on specific algorithms, with significant potentials.
6 **Big Data Processes**

On the management level, vendors’ strategic approaches and business process support options are analyzed in four categories.

6.1 **Big Data Operations**

The big data operations category examines vendors of IT operational services and additional services for processing large amounts of data, in particular for operating complex big data scenarios. Today, the terminology is mixed up, and some vendors position their (often mature) services for the operations of standard business software or cloud services as services to be used within the big data context, which is only partly acceptable. For instance, a key question is whether users would be willing to transfer huge data volumes via the Internet to use such offering. Although "the cloud" is virtually unlimited, a cloud offering is not necessarily suitable automatically for complex heterogeneous data in the terabyte and petabyte range. On the other hand, cloud computing is one of the disruptive technologies and has certainly been a big data driver. The same is true for infrastructure- or platform-as-a-service. However, vendors tend to mix these terms and certainly various interpretations of big data create a certain "grey zone". Since 2011 Experton Group has communicated the same big data definition and also the big data cycle of knowledge concept. Within this context, the overall number of companies which have been collected for this benchmark was further reduced to select only those service providers whose offerings are suitable for big data scenarios, as defined by Experton Group.
6.1.1 Benchmarked Big Data Operations Providers

The following 21 companies were evaluated:

- All for One Steeb
- Amazon AWS
- arvato Systems
- Atos IT Solutions and Services
- Capgemini
- Century Link (Savvis)
- Computacenter
- CSC
- Freudenberg IT
- FRITZ & MACZIOL Software und Computervertrieb
- Fujitsu Technology Solutions/TDS
- Google
- Hewlett-Packard (HP)
- IBM
- itelligence
- Lufthansa Systems (LHS)
- Rackspace
- Realtech
- T-Systems
- TCS Tata Consultancy Services
- The Unbelievable Machine (*UM)

12, which is more than half of these vendors, qualified as leaders: All for One Steeb, arvato, Atos, Computacenter, Capgemini, Google, HP, IBM, itelligence, Microsoft, T-
Systems and Unbelievable Machines (*UM). FRITZ & MACZIOL was positioned as „Rising Star“.

6.1.2 Big Data Operations Evaluation Criteria

Companies whose big data offering is very strongly based on their existing cloud offering, did not achieve the same competitive strength within the big data context as companies that have developed dedicated big data infrastructures. This means that the resulting distribution explicitly relates to the respective provider's big data operations element within the overall offering. This also implies that a certain provider could certainly achieve another competitive position within the benchmark of general data center and consulting services.

Similarly, service offerings including not only typical Hadoop-related services, but also services for market-relevant concepts such as HANA, new databases and in-memory, achieved a better position than companies with an exclusive focus on Hadoop. The analysis also examined whether the data centers are operated in Germany and whether the telecommunications partner is a German carrier, which is important, also because user organizations are very sensitive when it comes to privacy and data protection issues.

Other criteria included IT infrastructure hosting offerings for big data usage scenarios and also the provider's ability to run the infrastructure on-premise. We have also examined the development of service and licensing costs, compared to the previous year, and whether the providers have passed price advantages resulting from strategic partnerships with ISVs (independent software vendors) on to their customers.

A key technology aspect was the provider's ability provide a convincing big data solution for backup, restore, high-availability and disaster recovery tasks. Nearly all providers are still lacking greatly behind. Most providers refer to Hadoop's inherent high-availability features (triple redundancy of nodes) and snapshot technologies.

Some providers have limited their concepts to the velocity and volume features, and have left out other features such as variety and, depending on the individual provider, voracity, value or, based on the Experton Group definition, the number of users of new insights, understandably so, since this allows them to describe existing IT infrastructure models as "big data"-enabled, and most cloud models also work as big data infrastructure models, obviously based on the rather static data behavior. Very large amounts of data can be processed faster with models such as MapReduce and its open-source derivatives and also through in-memory and flash
technologies. Since users can execute more queries in the same time, also quite playfully, providers refer to such scenarios as big data scenarios.

### 6.1.3 Benchmark of Big Data Operations Providers

As the analysis of vendors in this category shows, quite many providers have an offering that can be used for daily big data workloads.

![Benchmarked big data operations providers](image)

**Figure 8:** Benchmarked big data operations providers

Through its hosting daughter Softlayer IBM operates 13 (big) data centers in the US, Asia and Europe. As of June 12, 2014 (the deadline for submitting information for this benchmark), the data center in Germany was not available yet, but customers had the possibility to book the data center in Amsterdam. Softlayer plans to offer a total of 40 data centers by end-2014, including three new data centers in Europe. The offerings of IBM GBS (Global Business Services) were also evaluated. The data centers shall be connected through their own network infrastructure.

All for One Steeb is among the "newcomers" within the big data operations category and was included, mainly due to their big data services and solutions.
offering. For instance, the provider has enhanced its data centers in Germany to also address big data workloads, in cooperation with United VARs. User organizations have great confidence in Germany-based data centers, also due to the ongoing political debates. All for One Steeb’s technological portfolio which was enhanced through acquisitions in the first half of 2014 also contributed to the vendor’s positioning. In a first step, these acquisitions increased the provider’s customer base and also improved the licensing and service business. This growth is reflected in the big data consulting & solutions category. Between July 2013 and July 2014 All for One Steeb has acquired several companies, including ORGA, Grandconsult and avantum consult, which is further proof of the provider’s development towards independence from individual technology providers and multi-platform support.

Atos, too, has developed is German business and pursues a multi-vendor strategy to ensure its independence from specific hardware and software partners. For Atos, big data analytics is of strategic relevance and big data is perceived as a component and enhancement of the provider's established business intelligence and business analytics business. Atos' services are complemented by individual consulting services. Since July 2013, immediately after last year's Big Data Vendor Benchmark was published, Canopy was launched as a "one-stop cloud shop". The "as-a-service" offerings for infrastructures, platforms and software are bundled, together with additional services.

For the first time, Amazon AWS has been included into this benchmark. Amazon AWS is one of the providers with an excellent cloud offering, also for customer demands that require large data volumes, high processing speeds and the fast and easy set-up of uses cases and sandboxing tests. Amazon AWS has a strong emphasis on its role as IT infrastructure provider. In April 2014, Amazon AWS announced its most recent price reduction, which increases the provider's attractiveness for many companies. Amazon AWS can further enhance its offering through solution-oriented offerings and industry-specific services packages.

Several providers within this category are challenged to clearly differentiate their offering for complex big data scenarios from their cloud offering. Although cloud computing seems to provide virtually unlimited opportunities, many cloud offerings are not sufficient to address big data tasks, also because huge data volumes must be transferred.

arvato has made good progress and has also strengthened its partnerships with software vendors such as Microsoft, which has created great technical potentials which must now be leveraged, also through targeted service offerings.
Berlin-based IT service provider The Unbelievable Machine (*UM) is a newcomer and has made it into the leader quadrant right from the start. *UM has clearly separated its big data service offering from its cloud service offering and provides very cost-efficient infrastructure services. The provider also accounts for the sensitivities and concerns of German users regarding the public political debate. *UM has achieved significant market visibility through their clearly defined service packages and big data commitment, which goes beyond pure marketing activities. Together with many other criteria, this high degree of awareness has contributed to the very good position of The Unbelievable Machine.

T-Systems is among those companies that have mastered the challenges identified in last year’s benchmark in various categories and especially in the big data operations category. One major innovation, as compared to the previous benchmark, is T-Systems’ newly developed Hadoop-as-a-service platform that is technically mature and provides a nearly optimal service offering. Other key aspects that contributed to T-Systems' very good overall position include the "German" network of the provider's mother company Deutsche Telekom and the Germany-based data centers.

SAP service provider itelligence has achieved very good positions, due to its well-designed service offering and also due to criteria such as German data centers and HANA solutions offerings. Other key aspects include itelligence' industry-specific know-how and industry solutions and the high transparency of software licenses within the provider's IT operations offerings, which reduces users' capital investments and operational expenditures.

HP, too, has developed an attractive overall offering. Some providers have advanced their big data products and solutions, based on technical advancements of the basic technologies. HP has focused on the platform and is now able to provide a very comprehensive portfolio.

Computacenter has made significant progress, compared to last year’s benchmark, and has arrived in the leader quadrant. Key aspects that contributed to the good positioning include the high quality of the service offerings and an improved offering, compared to the previous year. Computacenter has made its big data “homework” in Germany and presents a very sound market approach for users big data workloads. For instance, Computacenter has not simply renamed its existing cloud offerings into big data offerings. As of June 12, 2014 (the deadline of this benchmark) some strategic decisions regarding additional cooperations and partnerships were not effective yet and were therefore not accounted for in this analysis.
Capgemini has made significant progress, a considerable achievement, especially considering the slightly different weighting, compared to last year's benchmark, and the clearly improved portfolio attractiveness position. The purely technical offering was complemented with additional services, solutions and training programs.

Google's good position is based on the provider's innovative strength and development efforts as well as the ease of use of Google's service offering for big data operations.

Traditional IT services providers such as FRITZ & MACZIOL and Freudenberg IT have also addressed big data and have entered this growing market. Therefore they were also included in this Big Data Vendor Benchmark for the first time. FRITZ & MACZIOL has demonstrated such great big data commitment that Experton Group has positioned the provider as "Rising Star" in this category.

6.2 Big Data Analytics

As this analysis has revealed, computational results gained from available analytics tools are mostly distributed to traditional user groups and in most cases to the corporate management and controlling. Application scenarios are therefore often consisting of graphs with attractively presented, intuitive key performance indicators and other typical business results. This situation has not changed. Explorative working is making slow progress only. However, user organizations would like to reach a broader group of users within the company. Existing big data solutions often need to be connected with separate media solutions such as enterprise wikis for this purpose, which increases related development costs.

Also no progress can be observed when it comes to linking big data with Industrie 4.0. Computational and analysis results are hardly used for influencing machines and facilities.
6.2.1 **Benchmarked Big Data Analytics Vendors**

Within the big data analytics category, Experton Group has evaluated the following 32 companies that offer solutions and products for the analysis of very large amounts of multi-structured data:

- Actian
- Angoss Software Corporation
- Blue Yonder
- Board Deutschland
- Cisco Systems
- Cloudera
- DATA MART Consulting
- Datawatch
- Datameer
- Dell
- Empolis Information Management
- Exasol
- Fractal Analytics
- Hewlett-Packard (HP)
- IBM
- MapR Technologies
- Microsoft
- MicroStrategy
- odsoscope Technologies
- Oracle
- ParStream
Blue Yonder, Empolis, HP, IBM, Microsoft, MicroStrategy, Oracle, ParStream, Qlik, SAP, SAS, Software AG, Splunk and Teradata have been positioned as leaders. Exasol has been positioned as "Rising Star".

6.2.2 Big Data Analytics Evaluation Criteria

Evaluation criteria included not only pure analysis functionality, but also the respective solution's ability to process data from multiple sources and database structures, as streaming formats or as machine data. Another aspect was the integrated visualization of computational results. We have also examined whether the respective solution also provides security functionality such as the anonymization or pseudonymization of incoming data streams. We have further differentiated between solutions that are an advancement of traditional BI/BA solutions and new kinds of solutions that were developed specifically for big data scenarios.

The speed of analytics solutions does not only depend on technical parameters of the overall systems, and so, the analysis also accounted for advanced concepts such as in-memory computing and vendor-specific, big-data-related technologies such as HANA as well as vendors' information on the real-time character of their solution.

Competitive strength criteria included the sales landscape, including channel partners and partners' big data enablement, also through complementing solutions.
and services. Vendors' specifications on the performance limits and suitability of their analytics solutions for public or private clouds were also included in this benchmark.

Considering the fact that big data solutions require much explanation, skills provisioning, including data scientist training, was another evaluation criterion.

As compared to last year's benchmark, we have slightly changed the weighting to strengthen the focus on local activities, i.e., the vendors’ market approach for Germany, and on the suitability of the analytics solutions with multiple data.

6.2.3 Benchmark of Big Data Analytics Vendors

Exasol has been positioned as "Rising Star". The analytics engine of this vendor with headquarters in Nuremberg can rely on very good support through partners. The company has received the "Rising Star" award for its highly interesting technological solution.

Blue Yonder provides one of the leading tools for predictive analytics and the vendor’s offerings are very successful. Blue Yonder also has a very good product
development history. The company has an active customer and partner landscape in Germany; with its predictive approach, Blue Yonder was one of the early and influential big data players in the market. Compared to last year's benchmark, Blue Yonder has improved its position significantly and has reached the leader quadrant.

Actian is one of those vendors that have advanced their traditional BI solutions into big data solutions. Actian uses vector processing to speed up relational databases and also column-oriented data stores, combined with a method to process the data within the processor cache, to achieve significantly higher speeds.

Thanks to the Jaspersoft acquisition, TIBCO was able to reach a significantly better competitive position, compared to the Big Data Vendor Benchmark 2013. In last year's benchmark, TIBCO was among the followers, and Jaspersoft was barely missing the market challenger quadrant. Considering the differently weighted criteria, a direct comparison reveals a clearly positive development of TIBCO (including Jaspersoft) in Germany.

USU's USU Service Intelligence solutions are based on Pentaho. The strong focus on ITSM (information technology service management) is an interesting approach for big data scenarios and also a challenge.

For the first time, Empolis has been included in the „Big Data Analytics Software & Solutions“ category; the vendor has advanced its solution portfolio and combined the solutions components into a platform. The Empolis analytics solution aggregates data from multiple sources, including no-SQL databases and Hadoop, and enrich them iteratively with new information, based on semantic ontologies, to detect interrelationships that could not be found through pure queries, not even many fast queries. The company has a qualified partner landscape and is present in the entire German market. The vendor's visualization solutions were evaluated within the respective category.

In this category, an evaluation of HP’s Autonomy, Vertica and Hadoop offering was performed; Vertica is HP's analytics solution. Since last year's benchmark, HP has integrated its analytics offering into a complete platform concept. HP's service offering through the provider’s service organization, comprising implementation, roll-out and operations of the solutions and available IT operational services, was also accounted for.

As of today, IBM has the broadest offering of analytics solutions in the market, based on numerous acquisitions of BI/BA software houses during the last few years. IBM also has the strongest consulting organization, which was also included in the evaluation. IBM has a very strong big data market presence, including excellent reference projects, which have helped to increase an awareness of big
data analytics across practically all industries and company sizes. IBM also supplies the computer systems required for big data analytics, another aspect within this analysis. IBM is now challenged to substantiate its strong momentum from last year's benchmark with references for the German market. One of IBM's key differentiators are the vendor's highly advanced technical solutions. While other vendors often sell certain components separately, IBM integrates these components into its solutions. For instance, databases such as IBM's DB2 and also third-party databases are integrated with the analytics solutions. During the research phase for this Big Data Vendor Benchmark 2015 IBM announced that the company plans to sell the x86 computer systems division to Lenovo.

Microsoft has a broad and powerful product portfolio. Microsoft's big data strategy has a focus on developing and advancing the data warehouse and business intelligence stacks to integrate new big data technologies accordingly. Microsoft's analytics solutions comprise self-service BI as well as enterprise-class BI. The very broad range of partner companies allows user organizations of practically any size to implement individual and customized solutions. The actual analytics solutions – Microsoft Analytical Platform System and Azure SQL Database – enable users to work with relational systems as well as Hadoop clusters through SQL queries. Microsoft also offers alternatives to SQL-based solutions, including Analysis Services (for enterprise BI and OLAP), Azure Table Storage (as no-SQL solution) or PowerBI (for self-service BI).

MicroStrategy, too, has a complete portfolio for transaction-oriented applications (OLTP) and online analyses (OLAP). MicroStrategy's has a distinct BI history and has set up a big data strategy early enough. The vendor provides solutions for current user requirements, but must modify its portfolio successively to address new data sources and especially higher numbers of users, for instance, the whole sales organization of a large enterprise.

Teradata has one of the most comprehensive big data analytics offerings in the market. The company has strengthened its business analytics and business intelligence expertise to address the big data topic. For this purpose, Teradata's market strategy has a focus on big data analyses and provides support of a user community. While Teradata has achieved a very good position due to the completeness and technical maturity of its portfolio, competitors are closing in, and Teradata is challenged to communicate more directly with German user organizations and provide, for instance, concrete scenarios and information on the respective web pages. In this category it is especially important, not only for Teradata, to provide German references to convince the German user organizations. Basically, Teradata has achieved a very good competitive strength.
position; limiting factors include the stock price and the lack of local service, support and training activities for the analytics solutions.

Oracle has built up a very good portfolio, both through its own developments and acquisitions. The analytics solutions can work with content from multiple databases, including SQL and no-SQL databases and data from multiple storage architectures. The Oracle analytics solutions are complemented by a comprehensive portfolio of additional solutions including hardware solutions. Users can choose between various solutions, depending on their individual requirements. Oracle provides "Big Data Connectors" to integrate the technologies. Shortly after June 12, 2014 (the deadline for submitting information for this benchmark) Oracle made new big data announcements. Experton Group expects Oracle to further differentiate itself through a clearer focus on the specific requirements of user organizations, for instance large midmarket businesses, and respective positive feedback from the market.

As compared to previous year, ParStream has very much improved its German footprint for its analytics solution, which is based on supercomputing mechanisms. For instance, the company has strengthened its organization for Germany, which has a positive impact on the installed base and new customer business. The ParStream solution is hardware-independent and adds velocity to the big data cycle of knowledge, to query huge amounts of data with continuously new questions.

SAP’s key differentiator is its offering for HANA; HANA has evolved from a database into a platform for multiple application scenarios, in particular to address new big data scenarios in user organizations. SAP's use cases are convincing; market feedback regarding HANA’s big data suitability is mostly positive, and users can rely on getting support from SAP and their partners. As our analysis has revealed, there are improvement potentials for the concrete product, service and support offerings in Germany, especially to improve the vendors’ consulting, presales and service strength.

Some vendors advertise their solutions as being suitable for all kinds of data. However, if API (application programming interfaces) documentations are requested, they provide now answer or refer to the possibility of custom developments. This raises the questions how the prices specified for these appliances are made up, if variable components such as custom developments are included in the calculation.
6.3 Big Data Syndication, Reporting and Visualization

Based on Experton Group's big data definition, big data usage scenarios are also characterized by making the results of analysis, including predictive analyses, available to a much larger number of users, compared to traditional BI/BA scenarios. Therefore, this category of the Big Data Vendor Benchmark 2015 also includes solutions for visualizing, reporting and, ideally, syndicating (multi-channel distribution) information and insights.

The "face" of a big data scenario is made up of aspects such as the distribution of the gained information through multiple channels, the presentation of results from statistical methods in a comprehensible form for the broadest possible user groups for the lines of business and decision-makers and the fast generation of reports from multiple sources.

For Experton Group, the distribution of insights gained through big data analyses is a key component of complex big data scenarios. Big data expectations can only be fulfilled if new insights and knowledge is made available to as many users as possible within and, if necessary, also outside the organization. This requires new approaches for distributing analysis results.

To distribute complex insights also to new user groups, new forms of visualization are also desired, which means that applications must also give up traditional interfaces. A narrative user interface helps users understand the results gained from big data analyses.

Vendors' perception of "big data workloads" is also interesting. For instance, 500 lines of stock exchange data that can be explored interactively are presented within the context of a big data offering. While these 500 lines may have been extracted from 47 million records, which certainly are big data, the core idea of an overall view of all data is lost. Another vendor's example of a big data application is a table consisting of about 70,000 lines. Some vendors even say that a laptop is sufficient for their "big data" solutions.

Nevertheless, many vendors within this category reached a good or even very good position, mostly due to their good history within the market for visualization, reporting and dashboard solutions. Most vendors have enhanced and advanced their front-end solutions for BI and BA solutions, which is a good base for exploring data, also parts of very large databases. According to Experton Group's big data definition, big data is derived from traditional business intelligence and business analytics approaches, which is in line with the above explanation.
While one group of vendors markets multi-terabyte solutions, another group packs data to be processed (often a slice of a larger database) into the memory, saying that a well-equipped laptop is sufficient for their solution. Users are advised to thoroughly scrutinize all these buzzwords to determine the limitations of the respective solutions.

While the solutions and services within this category have a focus on the interactive slicing and dicing of large data volumes and the user delivers "the intelligence", the products and services within the analytics category have integrated the intelligence within the software.

Typically, many solutions that present data after a data warehouse extract rather playfully are called big data solutions; the actively distributed data are rarely reaching the gigabyte range, even if they have been extracted from millions of records or multi-terabyte data warehouses. This should enable employees from typical BI user groups such as finance and controlling to intelligently "play" and work with data to detect new interrelationships.

Currently, only few solutions for visualizing, reporting and distributing insights to larger organizational units have an industry-specific focus. Rather, the focus is on business figures. While nearly all of the vendors analyzed within this category say that they cover all industries with their solutions, many of them need partners with a specific industry competence to implement their solutions, especially if they want to address the German midmarket. Therefore, Experton Group’s methodology does not only examine the industry focus of vendors, but also their dependence from partners to ensure implementation.
6.3.1 Benchmarked Big Data Syndication, Visualization & Reporting Vendors

The following companies with a focus on data visualization or a specific partial solution for data visualization have been analyzed within this category:

- 1010data
- Actuate
- Angoss Software Corporation
- Blue Yonder
- Comma Soft
- Datawatch
- Empolis Information Management
- MeLLmo
- Microsoft
- pmOne
- Qlik
- Splunk
- Tableau Software
- TIBCO Software
- USU

Actuate, Blue Yonder, Empolis, Microsoft, pmOne, Qlik, Splunk, Tableau and TIBCO have qualified for the leader quadrant. Datawatch has been positioned as "Rising Star".
6.3.2 Evaluation Criteria for Big Data Syndication, Visualization & Reporting

Key evaluation criteria for this category include the overall solution, the company presentation, market approach and technology, including the big data character of this technology. Other aspects were the availability of data centers in Germany, if there is a cloud-based big data (analytics) services offering.

We also differentiated between open and proprietary solutions and determined the completeness of offerings, from the presentation of data to be analyzed to the distribution within an organization.

The portfolio attractiveness criteria included an increase in the number of active implementations of the respective solution in Germany, partnerships with independent software vendors and system integrators as well as sales partnerships in Germany. As is the case with many other categories, this results in clear differences, compared to the international market picture.

**Analyst Statement**

*The weak presence of strong international big data vendors may contribute to the impression that Germany lags behind when it comes to big data implementations. However, in many cases, companies are engaged in projects that involve more than just fast dashboards.*

*Holm Landrock*

*Senior Advisor Experton Group AG*
6.3.3 Benchmark of Big Data Syndication, Visualization & Reporting Vendors

Some vendors within this category have already set up a dedicated training and service offering, which is important, since this is a topic that requires much training. In case their own resources are not sufficient, software vendors need local training partners to provide preparatory workshops and also for the in-depth exploration of the big data subject.

![Diagram](image)

*Figure 10: Benchmarked big data visualization, reporting & syndication vendors*

Some international vendors lack a specific approach to conquer the German market, especially the midmarket. Within this category we have also accounted for competence centers and vendors’ local support of their customers. While from an operations perspective, it is sufficient to have a European instance, support and training services for employees that require employees to travel to London or companies to fly in trainers from Ireland cause significant cost which impacts the overall implementation cost.

Especially the vendors of solutions for the reporting and visualization for large data volumes must have a focus on adequate consulting, training and support service to
help users with this complex issue. "Self-service BI" requires very competent users. The distribution of new insights gained from big data within the whole (large) organization may be a challenge when it comes to the presentation of results.

Without imparting basic knowledge about the presentation of information from within big data by the vendors of respective solutions, there will be lack of qualified data scientist and data artists in Germany mid-term.

Comma Soft, a German vendor, offers a rather traditional BI tool and a tool for data visualization with interfaces to large database systems, has big data know-how and positions itself for big data in-memory processing workloads.

Since last year's benchmarks, TIBCO has acquired two companies that have significantly enhanced this service provider's solution portfolio, which contributed to the clear improvement of TIBCO's position, as compared to the previous year.

1010data is a specialist for big data slicing & dicing. The solution's strength include its ability to determine slices out of data volumes consisting of billions of records and prepare them for data exploring purposes, which enables users to work interactively in the detailed data. To achieve a better positioning, the company must develop a direct market approach for users from German companies and also a respective consulting, service and support portfolio. Also, data are stored within a rather undefined cloud, an additional challenge, considering current political debates.

Blue Yonder is a German company that can be described as rather "engineering-driven". The company's solutions and additional services provide a very good industrial depth without trying to meet the requirements of all kinds of industries, which is important within the newly evolving big data context. Blue Yonder also wants to enable more people, not only the experts from the lines of business that have been handling data in the past, to explore big data. Big Yonder was included in the category for big data syndication/visualization/reporting for the first time and has made it directly into the leader quadrant.

Microsoft uses Excel as important tool for the presentation layer of the underlying analysis solutions, databases and data warehouses. The Office tool is the industry standard for presenting relevant data, although certain apps or dashboard solutions in individual presentation functions are also attractive. Ultimately, there will again be a "deep dive into the data ocean", and this is possible with Excel, Sharepoint and PowerBI.

Tableau can be regarded as technology leader. The company has improved its German market presence and has built up a reseller and partner landscape.
Tableau has a distinct technological portfolio and is often used as functional reference.

Empolis provides an attractive visualization tool as the user layer of its analytics solution. Interaction with the semantic analytics solution provides fast visibility of interrelationships and allows for various kinds of presentation. The solution is not limited to typical diagram forms for key performance indicators. Empolis has customer references in Germany and receives positive feedback from the market.

Compared to some seemingly larger companies, "Rising Star" Datawatch has already developed significant activities in Germany. The company is now challenged to maintain the momentum gained in the last two years and strengthen its market presence, for instance, through stronger alliances with IT service and consulting companies and services that complement the software offering.
7 Big Data Systems (Software)

7.1 Big Data Aggregation

It is an exciting task to aggregate data and, in a first step, to collect these data. Variety does not only imply that an analytics solution can process multiple data types, but also that data suppliers deliver all kinds of data materials. For specific applications, data warehouse offloads (sections from large amounts of data) are required. The generated data must be aggregated with other data without changing or converting them (this would also require an additional and process step, which would constitute a point of interference). Currently it seems that users themselves are responsible for searching, provisioning and aggregating the data. Data aggregation is not a trivial task and nearly as complex as the task of developing suitable questions by data scientists; these are the questions used to gain new information from data.

While the market for solutions used to aggregate data from multiple sources is relatively limited, it is also characterized by a variety of very different approaches. Software and services with their various inputs and outputs require explanation, which affects both the competitive strength and the portfolio attractiveness ratings.

Some vendors within this category are young start-ups and are included in the Big Data Vendor Benchmark for the first time. Some of these new companies come from Germany and have focused their strategy on the specific requirements of German users, for instance with respect to their industry focus and their customers' company size.

The data aggregation category is a very young category of the Big Data Vendor Benchmark. Currently, big data are dominated by analytics solutions that often assume that data to be analyzed are already available within the company's IT systems or that data streams, for instance, from social media, are entered directly into an application software through cooperation agreements with respective providers. The companies benchmarked within this category are data collection and aggregation specialists.
7.1.1 Benchmarked Big Data Aggregation Vendors

The following 10 big data aggregation vendors have been analyzed:

- eoda
- Informatica
- Microsoft
- Neofonie
- Progress Software
- Qlik
- SAP
- Splunk
- Talend
- USU

Splunk, Informatica and Microsoft have been positioned as leaders. Progress is the "Rising Star" in this category.

7.1.2 Big Data Aggregation Evaluation Criteria

Besides general evaluation criteria that applied for all categories we examined whether the software solutions or services provide capabilities to aggregate data from as many different data sources as possible. The service character of respective offerings was also evaluated.
7.1.3 Benchmark of Big Data Aggregation Vendors

Kassel-based eoda is included into this benchmark for the first time. The vendor's TableR tool can be used to prepare large amounts of data for the R statistics software. TableR was developed to aggregate information from multiple databases for R.

Within this category, Neofonie was evaluated for its Text Miner tool and also for its tools that can be used to prepare data for various downstream solutions. The company has developed out of the German MIA project (a portal for aggregating news streams from multiple sources) and has built up a remarkable portfolio consisting of consulting, support and services for its own solutions in no time.

Progress has been positioned as "Rising Star" for its Easyl product and its data connectivity solutions, which can be used to aggregate and analyze data from multiple applications.

Aggregation is a conceptual component of Splunk's big data offering. Splunk can process multiple data types received from multiple ports. The company has also
strengthened its presence in Germany and is successful among German user organizations. Splunk can also aggregate log files and other streaming data. Imported data are indexed accordingly. The Splunk offering can be used for structured, unstructured as well as semi-structured data (for instance, machine data). Splunk is one of the most versatile aggregation tools in the market.

USU AG's offering is mostly based on the Pentaho storage solutions and comprises services for aggregating data from multiple sources. USA AG is a relatively young player in the big data market and has a focus on social media data sources, which might be interesting for user organizations that want to use the Pentaho analysis and visualization solutions and need respective support.

7.2 Big Databases and Big Data Management

Within this category, we have evaluated SQL as well as no-SQL databases, analytical databases and other big data storage, indexing and management solutions.

7.2.1 Benchmarked Big Database and Big Data Management Vendors

The following 21 vendors of big database and big data management solutions were analyzed:

- Aerospike
- Cloudera
- Couchbase
- Hortonworks
- Hewlett-Packard (HP)
- IBM
- Informatica
- MapR Technologies
- MariaDB
- Microsoft
- MongoDB
- Neo Technology
- Oracle
- ParStream
- Progress Software
- Red Hat
- SAP
- Splunk
- SUSE
- Synop Systems
- Teradata

HP, IBM, Informatica, Microsoft, MondoDB, Oracle, SAP, Splunk and Teradata have reached the leader quadrant. Neo Technology has been positioned as „Rising Star“.

### 7.2.2 Evaluation Criteria for Big Databases and Big Data Management

Within this category, generic criteria included the specific vendor’s company size, big data history, financial situation, geographical coverage, strategy and strategy implementation.

More specific aspects to be examined included vendors' volume, velocity, variety information as well as information provided on the data security, disaster recovery, implementation and interoperability with other databases. Benchmarked products and solutions included big-data-enabled relational SQL databases, no-SQL databases, column-oriented databases, data management solutions with a big data perspective and MapReduce- or Apache Hadoop-based solutions.

The vendors' presence within the German geography also played a key role.
7.2.3 Benchmark of Big Databases and Big Data Management Vendors

New players benchmarked within this category are MariaDB, Neo Technology, Couchbase and Aerospike.

![Benchmarked big database and big data management vendors](image)

**Figure 12:** Benchmarked big database and big data management vendors

IBM has a comprehensive database offering, including an independent accelerator technology for queries of compressed data within in-memory data structures. The solutions feature a very advanced technology. Users can choose among various products. The very good consulting and service offering also contributed to the very good rating.

MariaDB is a vendor with origins in the MySQL universe and a strong German presence. Due to the vendor's strong commitment within the German market, MariaDB has been included in the benchmark for the first time.

For its big data platform, HP has enhanced a Hadoop distribution with its own extensions and has also added a comprehensive service package. HP, as a
generalist, can complement its databases with a large and powerful infrastructure and service portfolio.

Oracle has one of the technologically most advanced and leading-edge big databases and big data management offerings in the market. The Oracle 11 and Oracle 12 relational databases enjoy a high degree of adoption and feature many of the criteria required for big data scenarios. Oracle can leverage their existing data warehouse know-how and has also integrated Hadoop and "big data connectors" to further enhance their big data solutions. The Oracle portfolio also includes a Hadoop-based data warehouse and the TimesTen in-memory solution. Oracle also addresses workloads that can be better handled with no-SQL data management software. The Oracle NoSQL Database uses distributed storage nodes and can be scaled up to handle very large applications. Oracle's offering also comprises additional solutions for analytical or explorative workloads. There are appliances available for the database solutions, which has increased Oracles portfolio attractiveness within this category even further.

Neo Technology has developed a graph-based database with analytics functionality. The solution can, for instance, support the fast exploration and visualization of interrelationships between records from multiple databases. The graph theory developed by the mathematician Leonhard Euler several hundred years ago is applied to handle very large data volumes. For their interesting technological approach Neo Technology has been positioned as „Rising Star“. The young company is now facing the challenge to present and implement a clear go-to-market strategy for the German market. Neo Technology certainly has built up the foundation and also has a committed team to address this challenge successfully.

MongoDB has been newly included in the benchmark, because the company provides an interesting technological and also a popular no-SQL database and has also strengthened its activities in Germany considerably.

SAP is one of the leading vendors in the market and has achieved an excellent position, due to its overall offering and also for the development of HANA as a database solution that has in a way helped to shape the "big data" term. Meanwhile, HANA has evolved into a data platform for all kinds of applications, including business analytics and data warehousing. SAP has liberalized its licensing policy to enable partners to market the HANA platform.

SUSE and Red Hat have been listed as vendors of open-source-based operating systems that have added big data enhancements or Hadoop to their solutions to help users manage their data.
8 Big Data Systems (Hardware)

8.1 Big Data Appliances

Appliances are among the key product offerings in the big data solutions market: They enable users to perform their first big data analyses with a complete system, without major efforts, except for modifications to their business processes, as opposed to implementing their own open-source solutions on standard servers or using cloud-based solutions.

For the purpose of this analysis, appliances are defined as ready-to-use complete systems, consisting of:

- Computer systems
- Storage systems
- Network
- Database
- Middleware
- Applications

These appliances integrate the various components with a focus on big data, as defined by Experton Group – i.e., the retrieval of information which must be made available to very large numbers of users very quickly and retrieved from enormous amounts of data from multiple sources, while ensuring data integrity and the protection and privacy of data.

Some hardware vendors tend to combine servers and storage (often using the Blade form factor) within one rack and market this solution with a preinstalled Hadoop distribution. Such rack systems (for storage or computations) have not been included in this analysis.
8.1.1 Benchmarked Vendors of Big Data Appliances

The following 12 vendors were identified as vendors of big data appliances:

- Cisco Systems
- Dell
- EMC
- Exasol
- Hewlett-Packard (HP)
- Huawei Technologies
- IBM
- Microsoft
- NetApp
- Oracle
- SGI
- Teradata

Except for Dell, Huawei and SGI, all other vendors qualified for the leader quadrant. SGI is the "Rising Star" in this category.

8.1.2 Big Data Appliances Evaluation Criteria

Appliances are complete systems that are used to aggregate, process and distribute data, and our analysis determined the respective systems' suitability to address these tasks. Big data appliances are high-performance systems with very large integrated storage systems, and for this reason, their high-availability and disaster recovery components or related additional offerings were also relevant for this analysis. Additional criteria included their integration of open-source components such as Hadoop and proprietary big data solutions.
The degree of integration of the overall system was highly relevant in the appliances category. Companies whose appliance components come from one single source received a better evaluation than those vendors who also integrate third-party components within their appliances. Criteria also included the performance and the degree of integration of servers, storage and network equipment and also factors such as specific training programs and the track record. We have also examined the analytics and the visualization solution as the middleware or application layer of the respective appliances.

Compared to last year's benchmark, the vendors' presence in Germany, the offerings' suitability for regional requirements, customer satisfaction and industry focus played a more relevant role, which may have resulted in different positions, compared to the previous year.

Another criterion was the availability of competence centers in Germany.

8.1.3 Benchmark of Big Data Appliance Vendors

Nine out of 12 vendors have reached the leader quadrant, mainly because many appliances have been developed, based on industry standard server technologies,
which leads to very similar technological performance ratings. Main differences between the individual vendors concern the amount of proprietary developments, the completeness of offerings and the service, training and support offerings for the respective big data appliances.

Remarkably, all vendors have set up very good offerings, although big data is a rather "young" discipline. An individual analysis of the benchmarked vendors was performed, based on the criteria described above. The 100 criteria were weighted and fine-tuned individually and in groups. The results within this category are individual results for each company, rather than assessments relative to the other benchmarked vendors.

SGI has achieved the „Rising Star“ position, because according to Experton Group's findings, this company has experienced the most significant development. SGI's competitive differentiators are the new UV 2000 appliance and a gross single-system capacity of 64 terabytes. SGI should strengthen its presence in Germany, as should other vendors. Often, information provided to German customers is limited to an (automatically?) translated homepage that redirects users to English web pages. Many vendors provide no or hardly any information that addresses the specific requirements of the larger German midmarket and German global companies, which affected the ratings of these vendors for the German market.

The Oracle Big Data Appliance fits Experton Group's appliance definition very well, since it provides the complete technology stack within one system or one box. The core element is the "Oracle Big Data Appliance X4-2 Engineered System" solution. Oracle's appliance comes with a hardware layer that meets industry standard requirements at least, from power supply unit to server components, network and storage. With the databases and the supplied Cloudera distribution the Oracle appliance is a suitable tool for Hadoop-based big data scenarios. Oracle provides the Oracle Exadata Database Machine for very large data warehouses, which combines the Engineered Systems hardware with respective Oracle database and software products. Another Oracle technology variant is Oracle Exalytics, a business intelligence system for the in-memory processing of very large amounts of data. Thanks to this overall package, Oracle was able to improve its position, compared to last year's benchmark and could strengthen its position even further through a more positive overall feedback from the German user community. Better feedback could, for instance, be ensured through a service offering that is better suited to address the requirements and sensitivities of German users.

Similar to other vendors, Oracle has announced new products; while these products may be highly interesting for various categories within this benchmark,
their availability cannot be expected before the second half of 2014 and could therefore not be included into this analysis, due to the deadline of June 12, 2014.

HP has developed a clearly communicated platform strategy for their appliances and has started to implement this strategy accordingly. The core elements are an HP Converged System 300 for Vertica, an HP Converged System for Hadoop and an HP Converged System for SAP HANA, which are also used on HAVEn, HP's big data analytics platform. HP also provides about 700 adapters for its HAVEn platform to connect a great variety of data sources.

Exasol, too, is among the leaders of this category; its appliance, which is based on Dell hardware, features a convincing and powerful analytics solution and a highly mature platform. As we have found out, some German customers who formerly worked with other big data vendors have switched to Exasol.

Compared to last year's benchmark, nearly all vendors have greatly enhanced their appliances, making them ready for every-day operations, and could improve their position accordingly and to a similar extent. Last year’s confusing offerings have evolved into rather complete offerings.

On the other hand, the vendors have also closed ranks and could achieve a stronger differentiation through a stronger industry-specific focus and a stronger regional presence.

While some of the vendors within this category with a limited number of staff can certainly compete with large international players in the German market, the picture may differ on the international level.

### 8.2 Big Data Storage

Storage systems and storage appliances specifically designed for big data projects are a key component of users’ IT infrastructures. Criteria for this benchmark included various technical features and the scalability of the respective solution.

Many storage vendors also add storage capacities in private or public clouds to their portfolio, an aspect that has also been accounted for in this benchmark. Pure cloud storage offerings were not included.

Some established vendors offer dedicated big data solutions, consisting of hardware and software, and also storage appliances. They are accompanied and also attacked by new providers of big data storage hardware and software. Interestingly, some of these vendors, including established players, have announced their big data solutions quite some time ago, but have not launched
them in the German market yet; neither do they appear on these vendors’ price lists. Users can only start to parameterize their storage solutions within the context of a concrete project.

8.2.1 Benchmarked Big Data Storage Vendors

Seventeen vendors of big-data-enabled storage solutions were analyzed:

- Cisco Systems
- Dell
- EMC
- Fujitsu Technology Solutions
- Fusion-io
- GRAU Data
- Hitachi Data Systems (HDS)
- Hewlett-Packard (HP)
- Huawei Technologies
- IBM
- NetApp
- Oracle
- Quantum
- RainStor
- Scality
- SGI
- Symantec

Cisco, EMC, Fujitsu, HP, IBM, NetApp and Oracle have reached the leader quadrant of this category. Fusion-io is positioned as "Rising Star".
8.2.2 Big Data Storage Evaluation Criteria

Within few years, big data storage systems must be able to ensure the stable and permanent storage of data volumes in the Terabyte to Petabyte ranges. This requires new storage architectures that also provide fast access to these data. Another important point is which transport mechanisms are used within the respective storage system and between storage and server. Evaluation criteria also included the share of vendors' own development work for the respective solutions, which provides some differentiation. Other criteria included the vendor's dependence on partners for project implementation and the share of the vendors’ own complementary products, for instance, software.

On the technology side, the evaluation differentiated between traditional, hard-disk-based and flash-based storage systems and accounted for multi-level storage architectures and high-availability and disaster recovery functionality.

HANA is a key component of today's big data solutions. Therefore, we have not only examined the storage solution's suitability for open-source big data solutions and commercial business intelligence, business analytics and data warehouse applications, but also its suitability for SAP HANA or the availability of a SAP HANA certification. The storage offerings’ degree of integration with big data solutions also played a role. The resulting big data market picture for Germany may come as a surprise; vendors that provide or seem to provide superior technology find themselves in rather close competition, also with unexpected competitors.

Competitive strength criteria included the availability of blueprints and project plans as well as feedback from the German market, also in the form of references and practical experiences. Since the focus was on the German geography, vendors' position varies significantly from their global position.

As our research has revealed, many companies, including many of the leading vendors, still have only international use cases and pilot projects as references. This may indicate that Germany is lagging behind, compared to global big data developments, although neither the German economy, i.e., the users, nor the vendors within the German market have done something wrong.

Considering the widely used term and the multitude of solutions and services as well as the many different definitions, companies are forced to explain their offering and market approach on a German website or a microsite as part of the company's other websites. Sometimes it seem vendors only present English content quite deliberately to supply only general information and provide more detailed explanations not before the presales phase and only when concrete sales activities have been initiated. To fill the gap, many vendors have a German
homepage or microsite where the offering is summarized on about two screen pages or less.

8.2.3 Benchmark of Big Data Storage Solution Vendors

Compared to last year’s benchmark, some vendors of big data storage solutions have made little progress only, and so, we can assume that between 2013 and 2014, vendors have been busy to further consolidate initiatives started during the hot hype phase until 2013, have made important acquisitions and have developed individual components into coherent offerings. Only few vendors, including Fusion-io, IBM and Oracle, have launched new products.

These new technologies include solutions based on NAND flash storage for extremely low latency during storage access.

NetApp is one of the global storage leaders and provides a comprehensive portfolio of storage solutions, including solutions for large data volumes. NetApp’s leader position is further substantiated by the vendor’s reliability and good reputation in the market as well as the versatility of data center storage solutions.
Cisco offers storage products from EMC, HDS and NetApp and combines these offerings with infrastructure components and integration services into an independent big data storage offering.

Oracle has made particularly great progress, compared to last year’s benchmark. The big data products announced in 2013 have meanwhile been updated and positioned in the German market. The Oracle Exalytics In-Memory Machine X3-4 storage system can process big data at high speeds.

HDS (Hitachi Data Storage) is one of the key independent storage technology vendors. The HDS solutions are components of other vendors' OEM solutions. HDS' excellent partnerships have also contributed to the analysis results, and on the other hand, the HDS storage technologies have been accounted for in the evaluation of respective OEM partners. HDS could use its platform to strengthen its activities in Germany to provide mature technologies, also to German midmarket businesses.

According to HP, HP was the first vendor to deliver a performance of 5128 seconds for data volumes of 10 terabytes, which corresponds to 1.99 Gigabyte per second\(^2\), based on the Terasort benchmark, which was developed for determining the Hadoop performance of the benchmarked solutions.

Only few vendors have published Terasort benchmark results within their marketing campaigns.

Fujitsu was benchmarked for its Externus DX solution for business warehouses and Eternus JX solution for Hadoop configurations. These two systems enjoy a very good position in the market, but Fujitsu must work to emphasize their specific suitability for big data workloads.

For the first time, Huawei has been included in the Big Data Vendor Benchmark. The Huawei OceanStor product family is a big data storage offering that has evolved out of cloud storage systems. However, technologies who are basically suited to address current requirements must prove there suitability for more complex big data scenarios. Huawei is challenged to leverage its strong international entry in the enterprise computing market also in Germany. In any case, Huawei is a noteworthy new player in the market.

RainStor provides very fast, scalable storage systems and has interesting international references, but needs to strengthen its German market presence.

---

\(^2\) Hardware Configuration: 18 HP ProLiant DL380 Gen8 servers; Dual 6 core Intel® E5-2667 2.9GHz processors; 64 GB memory; 16 x 1 TB SAS 7.2K disks per node; 4 x 1GB Ethernet. Software Configuration: Red Hat Enterprise Linux 6.2; Java Platform, Standard Edition, JDK 6 Update 29-b11. (Source: Hewlett-Packard)
Fusion-io has contributed greatly to the development of the relatively new flash storage technologies. This commitment is rewarded with the "Rising Star" position in this category.
9 Big Data Software

9.1 Big Data Protection

Often, big data is discussed within the context of privacy and data protection debates, since there are concerns about "transparent customers" and "transparent users". Respective privacy and data protection measures must be taken to prevent the abuse of large data collections. Traditional data protection must evolve into big data protection.

A look at normal planning activities for setting up data protection measures reveals the difference between "normal" data protection and the protection required for a big data project.

The first step to protect personal data within a project is to identify these data. It must be determined which category of personal data is concerned, since the required protection depends on this category. So-called special types of personal data are particularly critical. According to the German Data Protection Law (Bundesdatenschutzgesetz, BDSG), these include data concerning information on racial or ethnic origin, political opinions, religious or philosophical convictions, union membership, health or sex life. For projects where such data are involved security needs are especially high.

The same is true for personal data that are subject to professional secrecy or are related to actual criminal offences or administrative offences or the suspicion of punishable actions or administrative offences, and for personal data concerning bank or credit card accounts. Under certain circumstances, these data are subject to the obligation to report unlawful access to data (§ 42a BDSG), which often causes damage to the respective company's image.
Traditional approaches provide for the classification and protection of obtained data. For this purpose, the data are searched for specific keywords that indicate a certain data category. A word such as "account number" indicates that bank account data are concerned, which fall into the category of data that require special protection. Traditional keyword search tools soon reach their limits if large amounts of data must be searched or need much too much time for such analyses. New, powerful data search and data classification methods are required to address these issues.

The anonymization of personal data, a key data protection aspect, is another major data protection challenge when large data volumes must be masked accordingly. Personal data must be changed in a way that does not allow for any conclusions to be drawn as to the identity of individual persons, which also requires that personal data must be detected and processed automatically. Such efforts are not always successful, even if only limited amounts of data are concerned.

To ensure that data masking has been successful, users need new approaches that are more reliable and faster than commonly used solutions. Manual control of big data and limitation to few samples is not possible.

First IT security solutions with a specific focus on big data are available on the IT security market. This includes tools for identifying, classifying and anonymizing personal data and also powerful data encryption, data backup and data integrity solutions.
Experton Group recommends that organizations that want to engage in big data projects answer the following privacy and data protection questions before initiating such projects:

- Does the respective solution provide powerful data search and data classification functionality?
- Does the solution provide functionality to anonymize/pseudonymize/mask large amounts of data?
- Does the solution provide high-performance encryption functionality?
- Does the solution provide high-performance integrity control?
- Does the solution provide high-performance data backups?

Good, specific big data protection solutions help to answer some of these questions positively. The market for big data protection is still in its infancy, and comprehensive solutions that provide answers to all of the above questions are not available yet.

**Analyst Statement**

The market for big data protection is still in its infancy, and comprehensive solutions that provide answers to all of the above questions are not available yet.

*Oliver Schonschek*
*Research Fellow Experton Group AG*
9.1.1 Benchmarked Big Data Protection Vendors

15 of the available data protection solutions in the German market were classified as big data protection solutions. The current versions (at the time when we conducted the research) of the following solutions were evaluated (in alphabetical order):

- Aircloak Anonymized Analytics System
- Cloudera Gazzang
- Dataguise DgSecure Platform
- Hewlett-Packard (HP) Data Protector
- IBM InfoSphere Guardium
- Informatica Data Masking
- Innovative Routines International Data Masking
- Oracle Exadata
- Protegrity Big Data Protector
- Rohde & Schwarz SITLine ETH Ethernet-Verschlüsseler
- SafeNet ProtectFile
- Symantec NetBackup
- Uniscon Sealed Freeze
- Voltage SecureData for Hadoop
- Vormetric Data-centric Security for Big Data Environments

9.1.2 Big Data Protection Evaluation Criteria

Experton Group’s analysis has revealed that only few out of the multitude of data protection solutions are currently able to meet the specific big data challenges. The analysis has included the general criteria applied to all quadrants of this Big Data Vendor Benchmark and also the following specific requirements and aspects:
- Data encryption, integrity and deletion
- Multi-tenancy capability, access control for big data structures
- Data masking, anonymization and pseudonymization

Considering the high complexity of big data projects, a special focus within this category was on existing test versions, comprehensive support (in Germany and in German language, if possible), training offerings (in Germany and in German language, if possible) and a broad, high-quality partner structure. Besides functional differences, results for these criteria also varied, sometimes greatly, for the analyzed solutions.

### 9.1.3 Benchmarked Big Data Protection Vendors

IBM, HP, Symantec and Rohde & Schwarz SIT qualified for the leader quadrant. Uniscon has been positioned as "Rising Star".

---

*Figure 15: Benchmarked vendors of big data protection solutions*
IBM’s InfoSphere Guardium is a convincing solution, especially due to its powerful security and data protection functions and the breadth of functionality as well as the corporate vision and strong big data technology, the solution’s visibility in the market and the large partner infrastructure.

HP’s Data Protector, too, features leading-edge technology and security and data protection functionality for big data, a high degree of awareness among users and good support through partners.

Symantec NetBackup addresses big data availability and backup issues and provides a strong technological functionality, combined with security functions to protect against unauthorized access. Symantec has a broad partner landscape to cultivate the market.

The Rohde & Schwarz SITLine ETH Ethernet encryptor has a focus on the safe transmission of large data volumes and provides a convincing solution with a high security standard and performance. Within the German market, the vendor can build upon the high degree of awareness among customers.

"Rising Star" Uniscon’s solution Sealed Freeze is an interesting solution for protecting the privacy of large data volumes; near-term, the vendor should demonstrate the performance of his solution in big data projects to increase his visibility and pave the way into the leader quadrant. Already today, the vendor has a convincing vision, data protection concept and functionality.

Experton Group expects an increase in professional, powerful offerings and improved depth of functionality from the big data protection vendors that populate this market. The big data protection market is a young market, and demands for respective solutions will increase significantly.

9.2 Big Data Security Analytics

In ongoing privacy and data protection debates, big data analyses are oft perceived as a potential threat. However, if user data are anonymized as much as possible and are not analyzed except for the purpose of their collection, as required by data protection laws, big data do not bear risks, but significant potential to increase the security of data: Big data analyses enable users to set up an early warning system for IT attacks.

Attacks on IT systems and sensitive data are getting more and more sophisticated and are conducted through multiple levels and numerous access routes. Detecting such access requires comprehensive searches for traces under consideration of
many "sensors" – networks, servers, devices and also clouds and social media, which are increasingly abused for attacks.

Traditional SIEM solutions (security information & event management) are not able to master big data challenges:

- Security-related company information is literally amounting to mountains of data, the more so when "historical" data from previous months must be analyzed to track potential attack routes.

- Responses to potential attacks on IT systems and data must be extremely fast to minimize damages through potential attacks.

- The multitude of relevant data sources is very high and can easily amount to hundreds of sources. Security-related information is not only supplied through networks, but also all kinds of (mobile) devices, operating systems, local and Internet-based applications, IT security solutions such as identity management, vulnerability management and malware protection as well as IT system and operations management systems. Depending on the individual company, multiple locations must be included in the analysis of data sources.

Advanced solutions for security information and event management leverage big data functionality to provide enhanced and improved detection of risks for personal and other sensitive data. Big data evolves into an enabler of intelligent security, attack prevention & detection and state-of-the-art data protection, resulting in a high and growing demand for big data security analytics as today’s SIEM solutions.

As we know, big data has two faces when it comes to data protection and data security. One the one hand, large data volumes from multiple sources raise concerns that detailed user profiles may be created, making the "transparent" user a reality. On the other hand, big data analytics solutions also enable users to perform analyses that can help to increase data security significantly. Threat intelligence solutions analyze security-related information from multiple "threat sensors" and can make intelligent forecast threats.

However, big data security analytics provide much more than just access to a new generation of security information & management (SIEM) solutions. Big data security analytics can also support identity & access management (IAM) initiatives to better handle new kinds of threats and the increasing connectedness of multiple kinds of devices and applications and the Internet of things (IoT).
Actually, big data security analytics, combined with traditional identity management, results in a new kind of identity & access management. It is easy to draw a logical connection between big data and identity management: The quantity of digital identities is increasing enormously, due to the developments of mobile devices, networked sensors, social networks and other cloud services. Each device, each networked machine in Industrie 4.0 projects has a digital identity. These identities must be related and linked accordingly, for users use these devices to access services and machines; the various digital identities are connected.

However, the fact that defined policies and access rights exist for each identity and each link between identities (such as a defined user uses a specific device) reveals the high complexity and large volumes of data involved. Identity & access management comprise even more big data aspects: State-of-the-art access controls are context-related, i.e., access rights depend on various current parameters, including the user's location, the date and time and the content of data to be transferred with a specific application. Intelligent and near real-time analyses of current risks and required protection, security measures to be taken and permitted rights must be determined. It must be obvious immediately whether the user is permitted to transfer these data from this location at this time with this device or not.

This is where the big data security analytics component of state-of-the-art identity management solutions comes into play and where big data can support data protection and data security initiatives through an intelligent and predictive risk analysis and recommendations of security measures and permissions for identity & access management purposes.

9.2.1 Benchmarked Big Data Security Analytics Vendors

19 out of the SIEM solutions available in the German market were classified as big data security analytics solutions. The current versions (at the time when we conducted the research) of the following solutions were evaluated (in alphabetical order):

- Arbor Networks Pravail Network Security Intelligence
- Bluecoat Security Analytics Platform (by Solera)
- Datameer Big Data Analytics
- EMC RSA Security Analytics
- Hewlett-Packard (HP) Arcsight
9.2.2 Big Data Security Analytics Evaluation Criteria

In the course of the research performed for this Big Data Vendor Benchmark Experton Group has examined many SIEM solutions that have not met big data requirements. Our analysis only includes solutions that actually provide big data security analytics functionality and services. Evaluation criteria included event collection and event processing rates, event or log data storage capacities, the speed of analysis and response times as well as the kind and number of potential responses.

Besides the general criteria used for the various quadrants of this Big Data Vendor Benchmark, a specific focus was on the following big data security analytics criteria:

- Number and kind of data sources
- Analytical performance
- Integrated detection policies (to detect threats and attacks)
• Supported reaction templates
• Interfaces for the vendor's own and third-party solutions

With privacy and data protection playing a key role for big data analyses, we have also examined whether the analyzed solutions provide functionality to anonymize and pseudonymize personal data.

Implementing a big data security analytics solution may be a highly complex task (for instance, to address multiple data sources), and so our analysis also accounted for available demo versions, trials, comprehensive support (in Germany and in German language, if possible), training programs (in Germany and in German language, if possible) and a broad, high-quality partner structure.

Experton Group recommends that user organizations ask the following questions when selecting a big data security analytics solution:

• Does the solution support all of the company's existing IT resources, devices and services?
• Does the solution provide a clear and transparent view of incidents and alerts (dashboard)?
• Is it possible to track and analyze incidents across multiple data sources?
• Does the solution enforce internal security policies?
• Is it possible to select business-critical alerts and incidents and prioritize them accordingly?
• Is it possible to perform an inventory analysis of all connected systems within the network at the push of a button in order to ensure the completeness of analyses?
• Is it possible to analyze user activities (e.g., log-ins) in compliance with privacy regulations?
• Does the solution respond to threats and security incidents immediately and in the desired form (for instance, alerts)?

9.2.3 Benchmark of Vendors of Big Data Security Analytics Solutions

IBM, T-Systems, HP, McAfee, Splunk, LogRhythm and EMC/RSA achieved a leader position. NetIQ is the "Rising Star" in this category.
In the big data security analytics segment, IBM with its IBM Security QRadar SIEM has demonstrated foresight through its sound big data vision, comprehensive functionality and high performance for state-of-the-art security information & event management (SIEM), a high visibility in the market and relevant reference projects.

T-Systems demonstrates how an optimized and round offering can be set up for the German market, based on good solutions such as Splunk and RSA Security Analytics. T-Systems does not only provide a broad and mature functionality, but also has a good reputation on the German market, a focus on data protection, good access to customers and provides supporting services.

HP Arcsight is also among the leading big data security analytics solutions, since the analytics functionality addresses the challenges related to big data and multiple data sources; the solution also provides convincing data protection functionality and enjoys a high visibility in the market, which is reflected in respective references.

McAfee Enterprise Security Manager provides a convincing technology and functionality. The big data character of the solution is further supported through numerous interfaces to data sources and a high performance. While the solution has a high degree of visibility in the market, the vendor should work to improve its big data position.

Splunk is a well-known big data analytics solution, also in the German market, and features good functionality for security applications. Compared to specific security offerings, there is still some improvement potential regarding the security functionality and market visibility. The cooperation with T-Systems shows how Splunk can be placed successfully in the Germany market.

LogRhythm provides strong big data functionality and performance and a convincing technology. Additional partnerships could further improve the vendor's competitive strength in the German market.
EMC’s RSA Security Analytics is one of the leading solutions for big data security analytics with a convincing functional breadth and performance. The partnership with T-Systems also had a positive impact on the evaluation.

NetIQ has been identified as “Rising Star”, due to the convincing functionality and technology. Additional impulses, for instance through additional partnerships would be useful to improve the vendor’s competitive strength. Interoperability between NetIQ Sentinel and identity & access management (IAM) is a particularly positive aspect.

Big data security analytics has evolved from two sides and this development will continue: Big data analytics solutions that address security issues, and SIEM solutions that have been enhanced to handle large data volumes and multiple data sources and provide high processing speeds. Short-term, big data security analytics will gain even more strategic relevance through interoperability with other security solutions such as IAM and also due to the growing threats through intelligent attacks.
Outlook

With the big data wave moving from USA to Europe, rather hesitantly first and then very quickly, big data continue to be a topic in ongoing debates. But how long will big data remain such a trend topic? While we cannot provide a final answer to this question, we will try to clarify a few aspects.

Benefits

Companies with a web-based business model have leveraged big data technologies very early, since investments and benefits are close together (e.g., Amazon, Facebook). Companies with a traditional business model (e.g., automotive industry, finance, healthcare) require differentiated analyses to develop individual business models, based on big data technologies. After initial reluctance, these companies, too, are starting to recognize the benefits of these new opportunities and integrate big data in their projects.

Similar to the "e-commerce wave" more than 10 years ago, the typical "big data project" of the future will evolve to become the standard (by 2020 probably), and then, "big data" will be a term that will not even be worth mentioning. Big data will become the "normal project business", as is the case with all "e-commerce technologies" today.

What can big data service providers do to improve their portfolio and market position?

Visibility of offerings

Quite often, vendors' big data offerings can only be found by searching the respective company website. Many of these vendors are challenged to improve the visibility of their own competence, and the best way to do so is to use the language and wording of potential customers, including relevant reference projects from all kinds of industries.

Methodologies

Big data products are very complex and need respective explanations. Customers must recognize the potentials and benefits for their own business, which requires frameworks, models, templates and other tools to support the project across the whole project life cycle.
Standardized products

Big data product vendors should standardize their offerings to ensure a high degree of adoption of their products. Big data projects should no longer be handled as individual projects, which is still the case quite often. Similar to an ERP implementation, they must be based on a model or an industry-specific model, if applicable, and work with templates for multiple use cases.

Training offerings

There is quite some demand for data scientist trainings and similar qualifications. IT vendors can benefit from and improve customer and employee retention rates through such offerings.
11 Index

A

Accenture · 25
Actian · 44, 47
Actuate · 52
Aerospike · 60, 62
Aircloak · 77
All for One Steeb · 29, 38
Amazon AWS · 38, 41
Ancud · 25, 26, 29
Angoss Software Corporation · 44, 52
Arbor Networks · 81
arcplan Information Services · 25
arvato Systems · 25, 26, 28, 38, 39, 41
Atos IT Solutions and Services · 25, 26, 38, 39, 41

B

b.telligent · 25
BearingPoint · 25
Blue Yonder · 44, 45, 47, 52, 55
Bluecoat · 81
Board Deutschland · 44

C

Capgemini · 29, 31, 32, 38, 39, 43
Century Link (Savvis) · 38
CGI · 25
Cisco Systems · 44, 65, 69, 70, 72
Cloudera · 44, 60, 67, 77
Comma Soft · 25, 52, 55
Computacenter · 29, 31, 33, 38, 39, 42
Couchbase · 60, 62
CSC · 29, 35, 36, 38
cundus · 30

D

DATA MART Consulting · 25, 44
Dataguise · 77
Datameer · 44, 81
Datawatch · 44, 52, 56
Dell · 44, 65, 68, 69

E

EMC · 65, 69, 70, 72, 81, 83, 85
Empolis Information Management · 44, 45, 47, 52, 56
eoda · 30, 33, 58, 59
Exasol · 29, 44, 45, 46, 65, 68

F

Fractal Analytics · 44
Freudenberg IT · 25, 38, 43
FRITZ & MACZIOL Software und Computervertrieb · 25, 38, 39, 43
Fujitsu Technology Solutions · 25, 26, 28, 38, 69, 70, 72
Fusion-io · 69, 70, 71, 73

G

Google · 38, 39, 43
GRAU Data · 69

H

Hewlett-Packard (HP) · 30, 31, 33, 38, 39, 42, 44, 45, 47, 60, 61, 63, 65, 68, 69, 70, 72, 77, 78, 79, 81, 83, 84
Hitachi Data Systems (HDS) · 69, 72
Hortonworks · 60
Huawei Technologies · 65, 69, 73

I

IBM · 30, 31, 33, 38, 39, 40, 44, 45, 48, 60, 61, 62, 65, 69, 70, 71, 77, 78, 79, 81, 83, 84
Informatica · 25, 26, 27, 58, 60, 61, 77
Innovative Routines International · 77
Inovex · 25
itelligence · 25, 26, 28, 38, 39, 42

J

Juniper Networks · 81

L

LogDrill · 82
Logpoint · 82
LogRhythm · 82, 83, 84
Lufthansa Systems (LHS) · 38

M
MapR Technologies · 44, 60
MariaDB · 60, 62
Mauerberg · 25
McAfee · 82, 83, 84
MelLmo · 52
metafinanz Informationssysteme · 25
Microsoft · 30, 31, 36, 39, 41, 44, 45, 48, 52, 55, 58, 60, 61, 65
MicroStrategy · 44, 45, 48
MongoDB · 61, 63
msg Systems · 25

N
Neo Technology · 61, 62, 63
Neofonie · 30, 58, 59
NetApp · 65, 69, 70, 72
NetIQ · 82, 83, 85

O
odoscope Technologies · 25, 28, 44
Oracle · 44, 45, 49, 61, 63, 65, 67, 68, 69, 70, 71, 72, 77

P
ParStream · 25, 44, 45, 49, 61
PBS Software · 26
Pentaho · 45, 47, 60
pmOne · 26, 52
Prelert · 82
Pricewaterhouse Coopers · 30
Progress Software · 58, 59, 61
Protegrity · 77

Q
Qlik · 45, 52, 58
Quantum · 69

R
Rackspace · 38
RainStor · 69, 73
Realtech · 26, 38
Red Hat · 61, 72

RELEX · 30
RhinoData · 26, 28
Rohde & Schwarz · 77, 78, 79

SafeNet · 77
SAP · 28, 30, 31, 33, 42, 45, 49, 58, 61, 63, 68, 70
SAS Institute · 26, 27, 45, 72
Scality · 69
Seculert · 82
SGI · 65, 67, 69
SHS Viveon · 26
Software AG · 26, 28, 45
Splunk · 45, 52, 58, 60, 61, 82, 83, 84
Steria Mummert Consulting · 30, 31, 35
SUSE · 61
Symantec · 69, 77, 78, 79
Synop Systems · 45, 61

Tableau Software · 52, 55
Talend · 45, 58
TCS Tata Consultancy Services · 26, 38
Teradata · 45, 48, 61, 65
The Unbelievable Machine (*UM) · 30, 35, 36, 38, 39, 42
TIBCO Software · 30, 45, 47, 52, 55, 82
TNS Infratest · 30
Trustwave · 82
T-Systems · 30, 31, 35, 38, 39, 42, 82, 83, 84, 85

Uniscon · 77, 78, 79
USU · 30, 45, 47, 52, 58, 60

Voltage Security · 77
Vormetric · 77

Y
YMC · 26

Z
Zscaler · 82

13 Authors and Contacts
Holm Landrock is a specialized journalist and Senior Advisor at Experton Group.

His main areas of coverage include big data and supercomputing as well as technical-scientific information processing. Within the Experton Group team he adds big data competence and contributed significantly to Experton Group's first "Big Data" multi-client study. He is head of the Big Data Vendor Benchmark team.

Since 2006, Holm Landrock has also worked as a freelance journalist. Before, he worked as a journalist and PR consultant for various global IT companies and has published more than 200 technical articles.

Holm Landrock has completed an IT specialist training in Dresden; since 1982, he has studied and worked with enterprise-class IT systems.

Oliver Schonschek is a Manager Advisor at Experton Group.

In addition to his work as advisor, he also works as publisher, author and journalist for various specialized publishing companies and editorial offices.

His analyses and publications cover information security and privacy issues, with a focus on the interfaces between IT, business, law, compliance, privacy and data protection and data security. He is interested in issues such as the business usage of data under compliance with privacy regulations and the protection of IT infrastructures without compromising user privacy.

After he had completed his studies at the University of Bonn with a degree in physics, Mr. Schonschek started to work at a medium-sized software company in 1997, which was also commissioned by trade companies to process their sensitive data. Often, this raised questions about security and data protection, and so, Mr. Schonschek specialized in related issues.

Since 2006, Mr. Schonschek has published books, articles and analyses on privacy, data protection and IT security in renowned publications and publishing companies. At Experton Group, he contributes to analyses and studies on IT security and security- and privacy-related aspects in other areas of IT.
Prof. Dr. Andreas Gadatsch is a Senior Advisor at Experton Group.

Prof. Dr. Andreas Gadatsch is a professor of business management, with a focus on business information technology, at the Institute of Economics at the Bonn-Rhein-Sieg University of Applied Sciences in Sankt Augustin. In his role as a Senior Advisor at Experton Group, he is currently involved in studying big data from a business perspective and examining the impact of big data technologies on enterprise information and process management.

Since 1986, Prof. Dr. Gadatsch has worked for various companies as consultant, project manager and IT manager. In 2000, he was appointed professor of business management, with a focus on organization and data processing, at the University of Applied Sciences in Cologne. In 2002, he joined the Bonn-Rhein-Sieg University of Applied Sciences. Since 2011, he has been head of the part-time master program for innovation and information management.

Prof. Dr. Gadatsch’s main areas of coverage in research and education are information management (IT controlling), business process management (process modelling) and ERP systems.

He is the author of more than 250 publications, including 16 books, some of which were published in several languages and editions.

Prof. Dr. Gadatsch holds a degree in business management of the University of Applied Sciences in Cologne and a degree in business administration of the Open University Hagen. He holds a Dr. rer. pol. degree as external post-graduate from the Institute of Business Information Technology chaired by Prof. Dr. Hermann Gehring.
About Experton Group

Experton Group is a leading IT research, advisory and consulting company. The company has 30 experienced analysts in Europe who support mid-sized and large organizations with their IT strategic planning and implementation. In Germany, Experton Group has offices in Munich, Frankfurt and Kassel.

More information on our research can be found under:

http://experton-group.de

http://blog.experton-group.de/

https://twitter.com/ExpertonGroup

Contact
Experton Group AG
Carl-Zeiss-Ring 4
D-85737 Ismaning
Tel. +49 89 923331-0
Fax +49 89 923331-11

Version 1.0 of September 12, 2014