

# Hadoop and Data Lakes

Use Cases, Benefits and Limitations

BARC Research Study

 **BARC**  
A CXP GROUP COMPANY





Timm Grosser  
Senior Analyst  
[tgrosser@barc.de](mailto:tgrosser@barc.de)



Melanie Mack  
Head of Market Research  
[mmack@barc.de](mailto:mmack@barc.de)



Jacqueline Bloemen  
Senior Analyst  
[jbloemen@barc.de](mailto:jbloemen@barc.de)



Jevgeni Vitsenko  
Research Analyst  
[jvitsenko@barc.de](mailto:jvitsenko@barc.de)

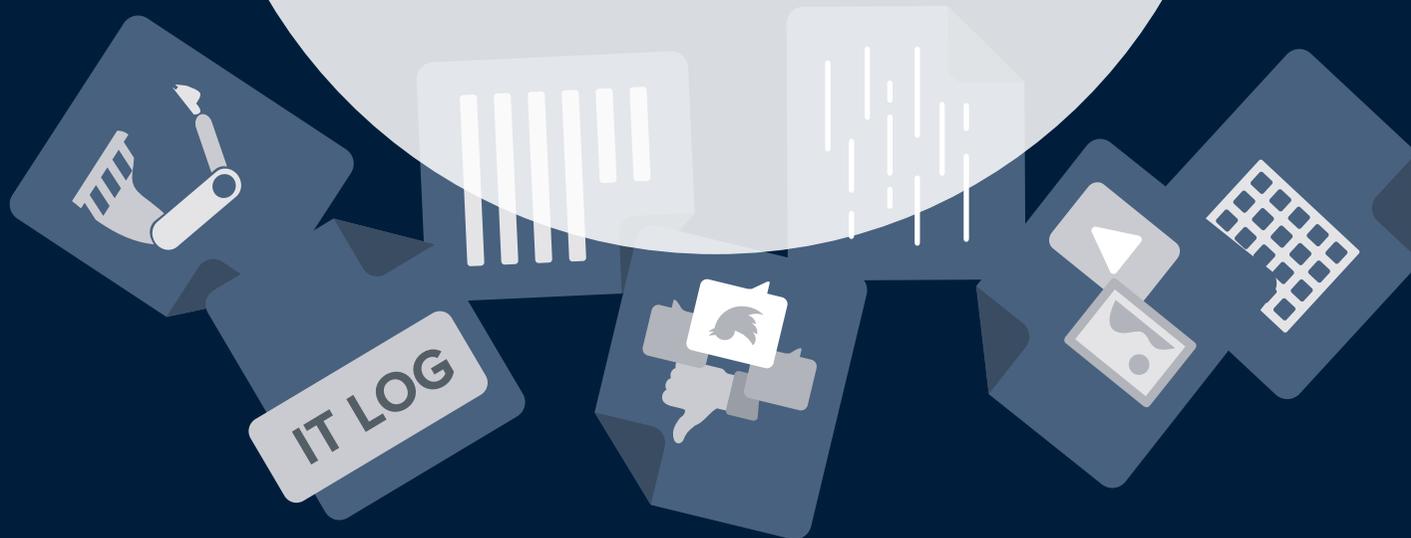
This independent study was conducted and written by BARC, an objective market analyst. We thank Cloudera, SAS, Talend and Teradata for sponsoring the survey. Thanks to SAS and Talend, this study is available in English language.





<b>4</b>		<b>Preface</b>	<b>31</b>		<b>Hadoop Theories Put to the Test</b>
<b>6</b>		<b>Demographics</b>	<b>35</b>		<b>Sponsors</b>
<b>8</b>		<b>Management Summary</b>	36		SAS
<b>11</b>		<b>Key Findings</b>	37		Talend
11		Usage	<b>37</b>		<b>About BARC</b>
17		Drivers			
19		Reasons for Use and Benefits			
24		Data Lakes			
27		Implementation			
29		Challenges			

# Preface





**D**iscussion surrounding Hadoop and data lakes is as relevant as ever. The Hadoop ecosystem is considered THE technological breakthrough for enabling companies to capitalize on the big data revolution. A data lake, in turn, is viewed as a broad data management concept and a prerequisite for data-driven companies. It promises a fast, efficient, low-cost way to manage, use and analyze any amount of data from different systems with varying structures. As a source for any type of analytic task, it can also claim to be the technological backbone of digitalization and the (big) datafication of the entire economy.

Hadoop, a top-level project of The Apache Software Foundation, is an open source Java framework for scalable, distributed applications. It includes a collection of components for administering, accessing and analyzing structured and unstructured data. Hadoop is capable of managing huge amounts of poly-structured data and adding value to new or established IT technologies. It is especially well-suited as a platform for implementing big data projects and is often viewed as a technology for data lake deployments. The

concept of a data lake, however, can extend far beyond that, depending on how broadly the term is defined. A data lake often focuses on data availability and providing downstream applications with schema-free data that is close to its raw format regardless of origin.

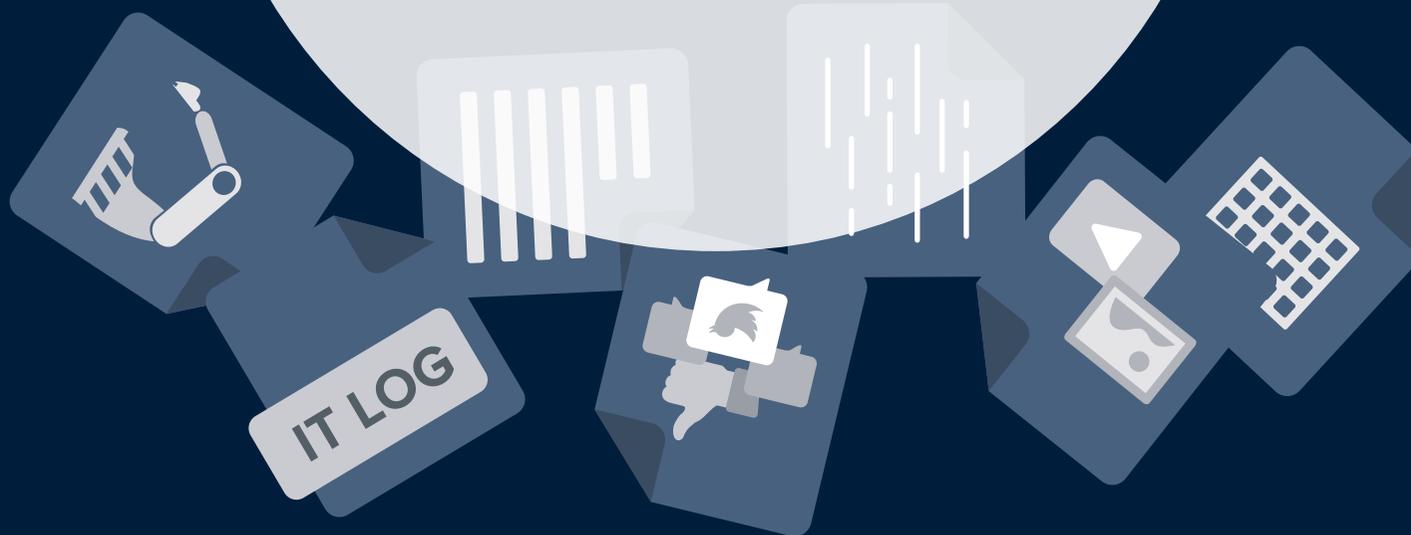
Enterprise projects with Hadoop technology and data lakes in production have only emerged in recent years. Accordingly, many companies find it hard to distinguish between media hype and the benefits that can realistically be enjoyed. There is limited experience in terms of how and where it really makes sense to implement, which obstacles can arise during implementation, and what potential benefits are delivered in real-world scenarios.

The following BARC user survey provides answers and insights. It explores the status quo of Hadoop and data lakes in general and real experiences from Hadoop use cases across the globe. It tackles important questions including:

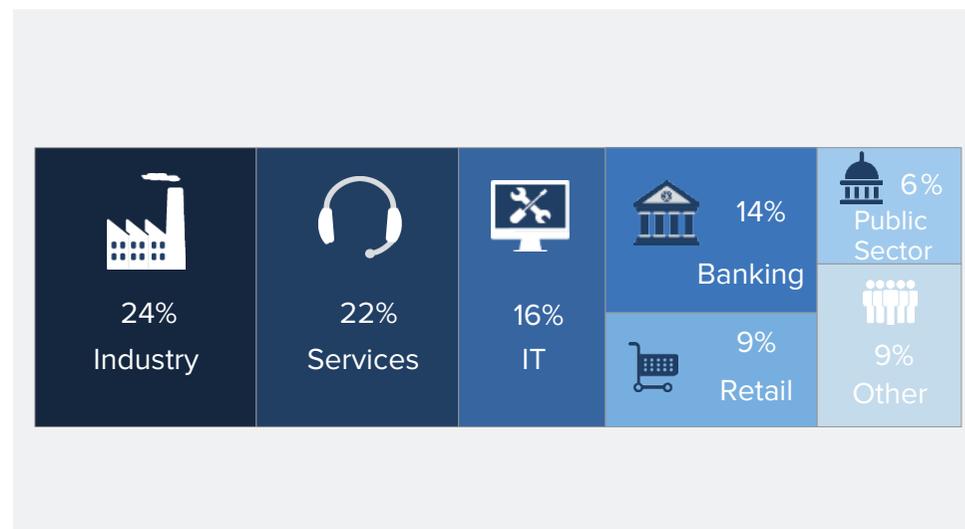
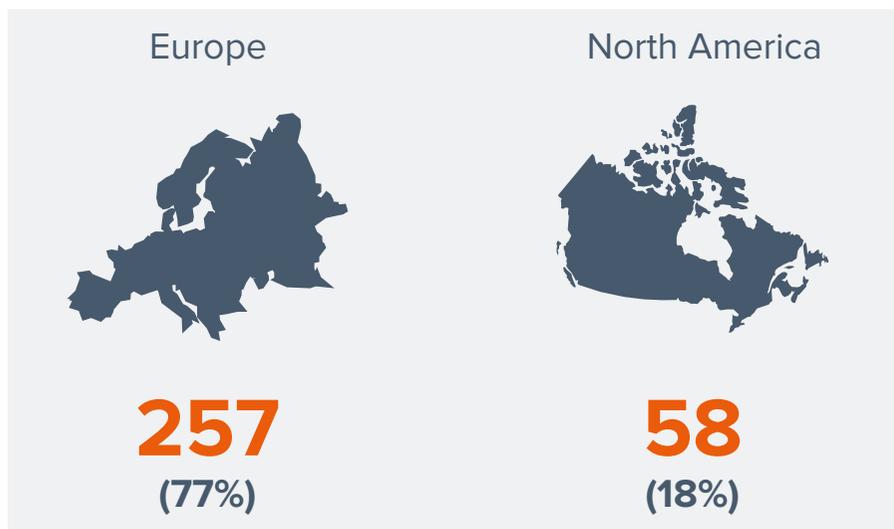
- How widespread is the current usage of Hadoop and data lakes in companies? What are their plans for the future?
- How do companies utilize Hadoop or plan to use it?
- How do companies currently use data lakes?
- What problems do they face?
- What real-world benefits does Hadoop bring? What projects have companies implemented already?
- How is the technological implementation set up?

This study was conducted independently by BARC. Thanks to sponsorship from Cloudera, SAS, Talend and Teradata, it can be published free of charge. BARC would like to thank you, our readers, in advance for your participation in future surveys as well. Your help and support are essential to fuel discussion through empirical data.

# Demographics



IT LOG



# Management Summary



IT LOG



## Hot Spot #1

Hadoop is a technology trend with great potential

The number of Hadoop projects in production is growing, especially in Europe. However, the number of companies that do not want to use Hadoop is also increasing. As more companies understand the potential benefits of Hadoop use cases, they can make better informed decisions on whether to use the technology. There is a surprisingly broad spectrum of Hadoop systems in production. Companies of all sizes with varying volumes of data, data types and demands for data currency are using Hadoop. That makes it a potential technology for a wide range of use cases in all types of companies. Hadoop is increasingly emerging from a simple file system to a runtime environment for analytic applications.

## Hot Spot #2

Opinion on data lakes is divided

Two sides have emerged in the debate on data lakes. One group views it as an important concept and a prerequisite for data-driven companies. The other considers it marketing buzz, a new name for an old concept, or a topic with no real relevance. It is possible that this gap is propelled by insecurities, which opens the opportunity to find a definition and create transparency in terms of benefits and best practices. The term 'data lake' is still not used consistently. There is a tendency, however, to use this term when referring to data preparation/storage or explorative environments.

## Hot Spot #3

BICC and data science teams are driving Hadoop and data lake projects

BI competence centers and data science teams are currently the main drivers behind Hadoop and data lake projects. They, in turn, are bringing more business relevance to these topics.



### Hot Spot #4

Companies primarily use commercial tools and Hadoop distributions for project implementations

Commercial tools and distributions received higher marks than Apache Hadoop components in most tool categories and are implemented much more frequently. The one exception is for streaming. Cost savings, functional performance/innovative capacity and operations improvement are the main reasons for choosing Apache Hadoop.

### Hot Spot #5

Customer intelligence and predictive analysis are clear Hadoop use cases

Customer intelligence and predictive analysis are by far the most common Hadoop implementations. Analyzing data from heterogeneous, divergent sources, predicting customer behavior, strengthening customer loyalty, and increasing flexibility are the most common benefits.

### Hot Spot #6

Hadoop generates analytical benefits

Hadoop's greatest benefits lie in the analysis of heterogeneous data from divergent sources, the prediction of customer behaviors/customer retention and in increasing flexibility. Hadoop not only takes on the role of the file system. It also serves as a platform and runtime environment with core data analytics and predictive analysis functionalities.

### Hot Spot #7

Hadoop enables use cases that were previously not possible

Users view Hadoop as a potential technology for implementing new types of use cases that were previously not possible with existing systems. Saving costs and implementing a technically better platform play a minor role when deciding whether to implement Hadoop.

### Hot Spot #8

Lack of skills and insecurities pose the greatest challenges

It appears that not much has changed since the last BARC Hadoop survey. A lack of professional know-how and technical skills still lead the list of challenges. European companies, in particular, are uncertain how to use it properly. Their counterparts in North America, in comparison, are more likely to complain about low usability, immature ecosystems and the high cost of training and development.

# Key Findings

## Usage



IT LOG



## The number of Hadoop projects in production is growing along with the number of critics

Hadoop is still being put to the test. The benefits of the framework, however, appear to be delivering good arguments in its favor.

The numbers show that two groups are emerging for and against Hadoop. Accordingly, Hadoop will not go down in history as the cure for all analytic requirements. It has advantages and disadvantages depending on the use case.

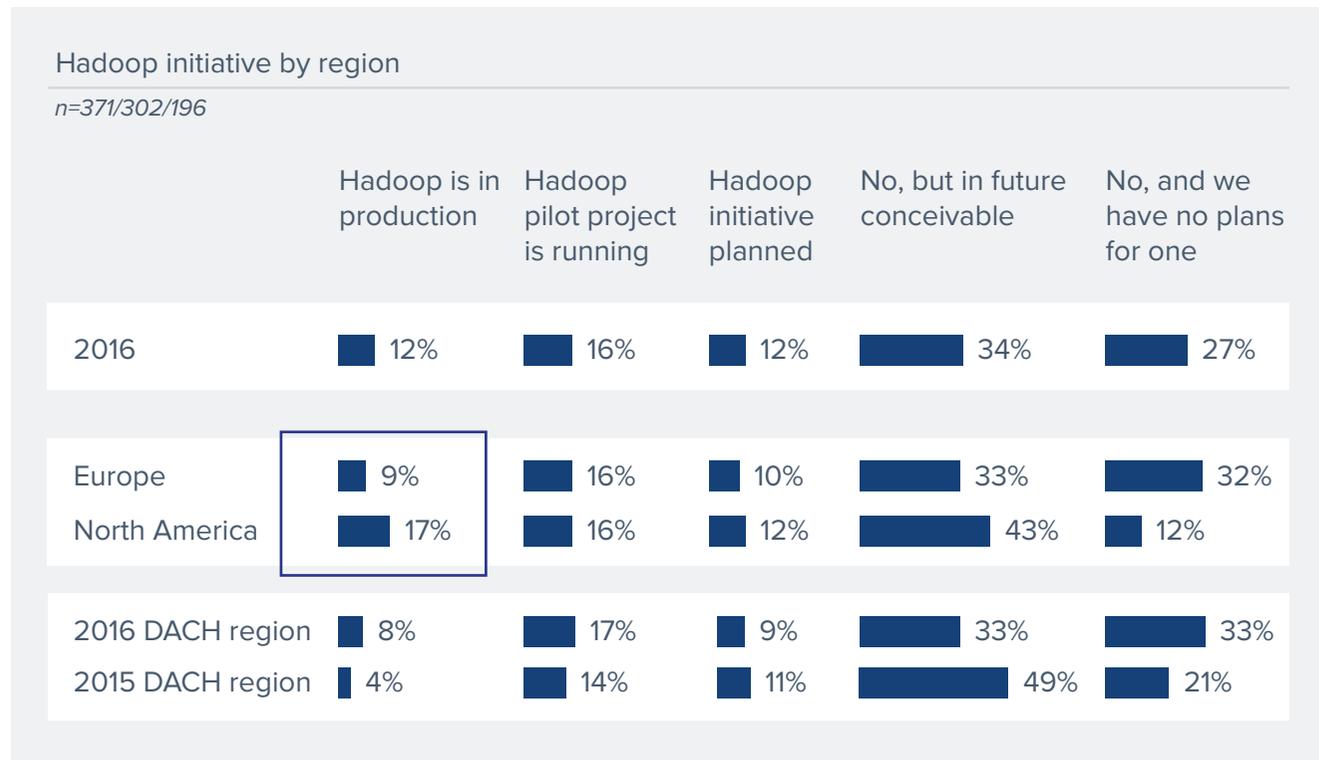
40 percent of participants worldwide are Hadoop supporters (e.g. with projects in production, a pilot phase or in planning). They show a clear interest in the technology. 12 percent of them already have Hadoop in production.

Many supporters view Hadoop as a potential component to help build analytic environments for special applications.

27 percent of respondents categorize themselves as Hadoop opponents. Another 34 percent currently do not have Hadoop but can imagine using it in the future.

### Clear jump in Hadoop usage throughout Germany, Austria and Switzerland

Germany, Austria and Switzerland, also known as the DACH region, recorded a clear jump in usage (4 percent to 8 percent) in comparison to last year. The number of Hadoop prospects (i.e. those that have Hadoop initiatives or are planning pilot projects) remained steady at 26 percent. However, the percentage of respondents who cannot imagine



using Hadoop rose from 21 percent to 33 percent.

### Widespread Hadoop use in North America

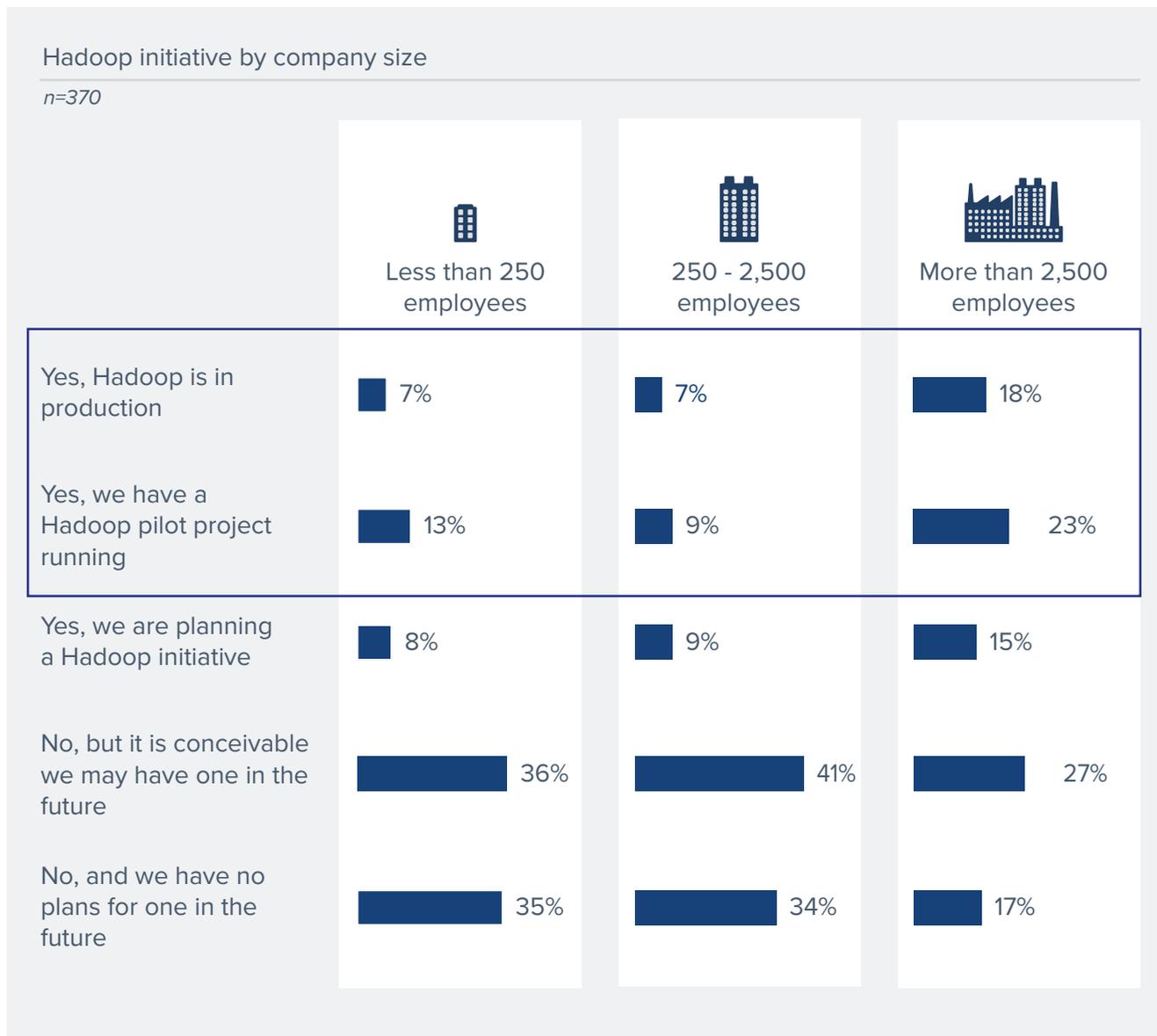
The number of Hadoop installations in production is nearly twice as high in North America (17 percent) as in Europe (9 percent). The North American market has a reputation for early adoption and higher

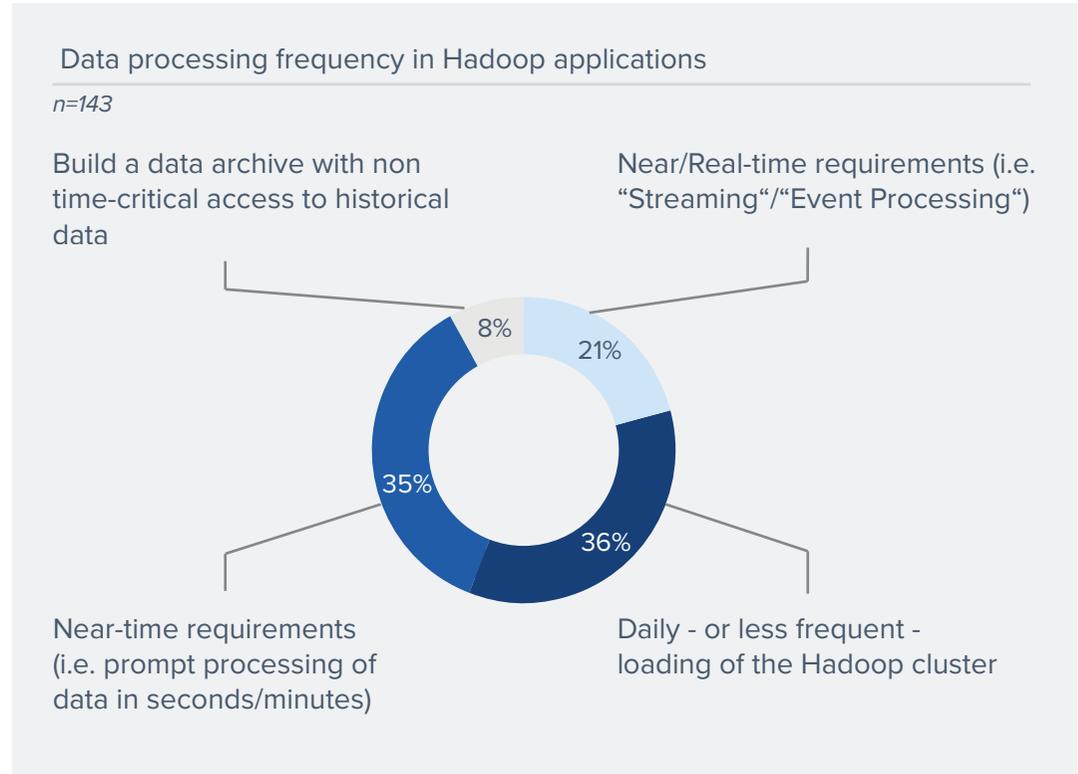
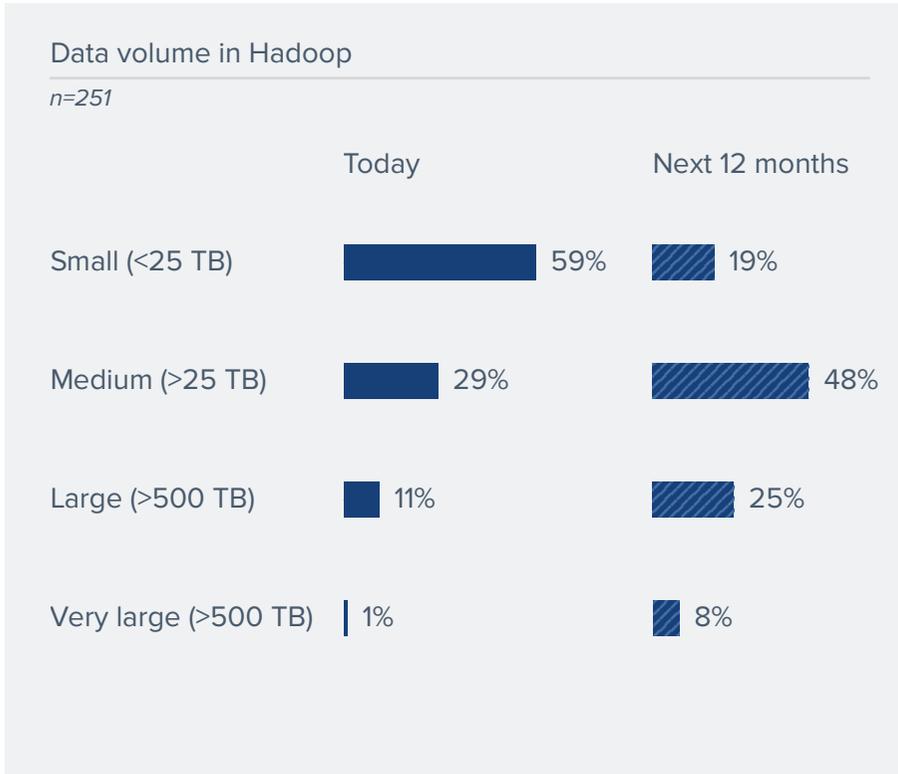
maturity rates in using information technology so this is not altogether surprising. However, when it comes to planned adoption of Hadoop, a similar proportion of North American and European companies have initiatives in the pipeline.

Hadoop usage depends on the use case.  
Company size is irrelevant.



Survey findings show that not only large companies and those with large data volumes see the relevance of using Hadoop as part of a big data strategy. This is reflected in the distribution of our survey participants. The spread covers large (41 percent), small (20 percent) and midsize companies (16 percent).





### The majority of Hadoop applications process small amounts of data

The survey reveals that Hadoop usage does not depend on data volume and currency.

59 percent of the Hadoop scenarios described were implemented with "small" data volumes up to 25 terabytes (TB). Scenarios exceeding 1 petabyte (PB) are few and far between (1 percent). As expected, higher data volumes are anticipated in

the near future (i.e. the next 12 months).

Companies start with smaller-scale Hadoop initiatives in order to first see the potential for further development. The largest increase was reported among scenarios exceeding 25 TB.

### Hadoop usage is no longer limited to batch applications

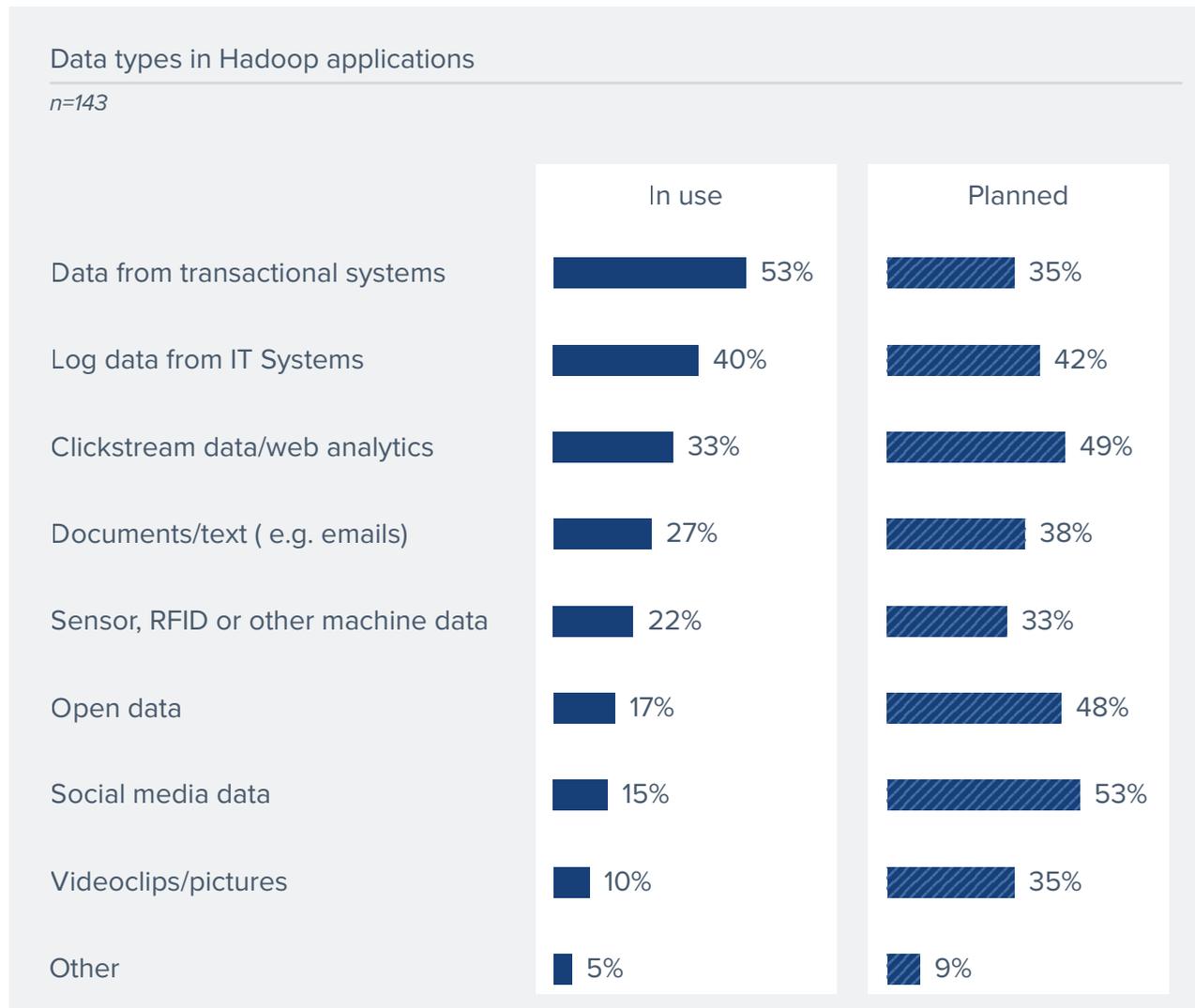
A closer look at data currency reveals high responses for streaming (21 percent) or near-time usage

(35 percent), especially in combination with customer intelligence applications. This shows that Hadoop and MapReduce are no longer limited to batch applications and are increasingly being used to address high demands on data currency.



The survey asked what types of data companies use in Hadoop applications. More than half of respondents (53 percent) use transactional data, followed by log data (40 percent) and clickstream data/Web analytics (33 percent). More than a quarter (27 percent) use documents and text files in Hadoop applications. This seems surprisingly high in comparison to the findings of BARC studies from years past. A closer look at the responses for planned usage shows ambitious endeavors in almost every data category.

The findings do not show that Hadoop is especially suitable for certain data types. Instead, it appears to be a potential technology for processing all types of data. It can also be assumed that Hadoop primarily uses data that can also be used in traditional platforms. If data volumes or types do not play a role, the reasons for using Hadoop still remain open. Cost, functionality for particular use cases, know-how and the availability of a technical infrastructure or existing IT processes, however, are also points to examine.





### Usage of Hadoop

n=144



Hadoop is used for many different purposes, especially as a runtime environment for new types of advanced analytics (65 percent), as memory for raw/detailed data (60 percent), and for data processing and integration (57 percent). Other scenarios, however, are conceivable and have been implemented as well.

A breakdown by region paints an interesting picture.

### North America uses Hadoop more intensively as a runtime environment for BI.

Using Hadoop as memory for raw/detailed data varies greatly between North America and Europe. 76 percent of respondents in Europe use it for this purpose in comparison to only 15 percent in North America.

One possible explanation is that with growing experience and maturity, Hadoop usage shifts from simple data storage more towards an analytic engine as a runtime environment for BI. This would also mean that North America has more

experience in analytic Hadoop usage as a whole and, therefore, focuses on other use cases.

Furthermore, using Hadoop as a runtime environment for advanced analytics is more popular in Europe (an 11 percent difference) than in North America. It appears that Hadoop is not “just” used in North America for advanced analytics, exploration or other new forms of analysis but rather as the technology itself to support both old and new analytic tasks. One could assume, therefore, broader usage and, in certain cases, better utilization of potential opportunities in this market as well.

# Key Findings

## Driving Factors

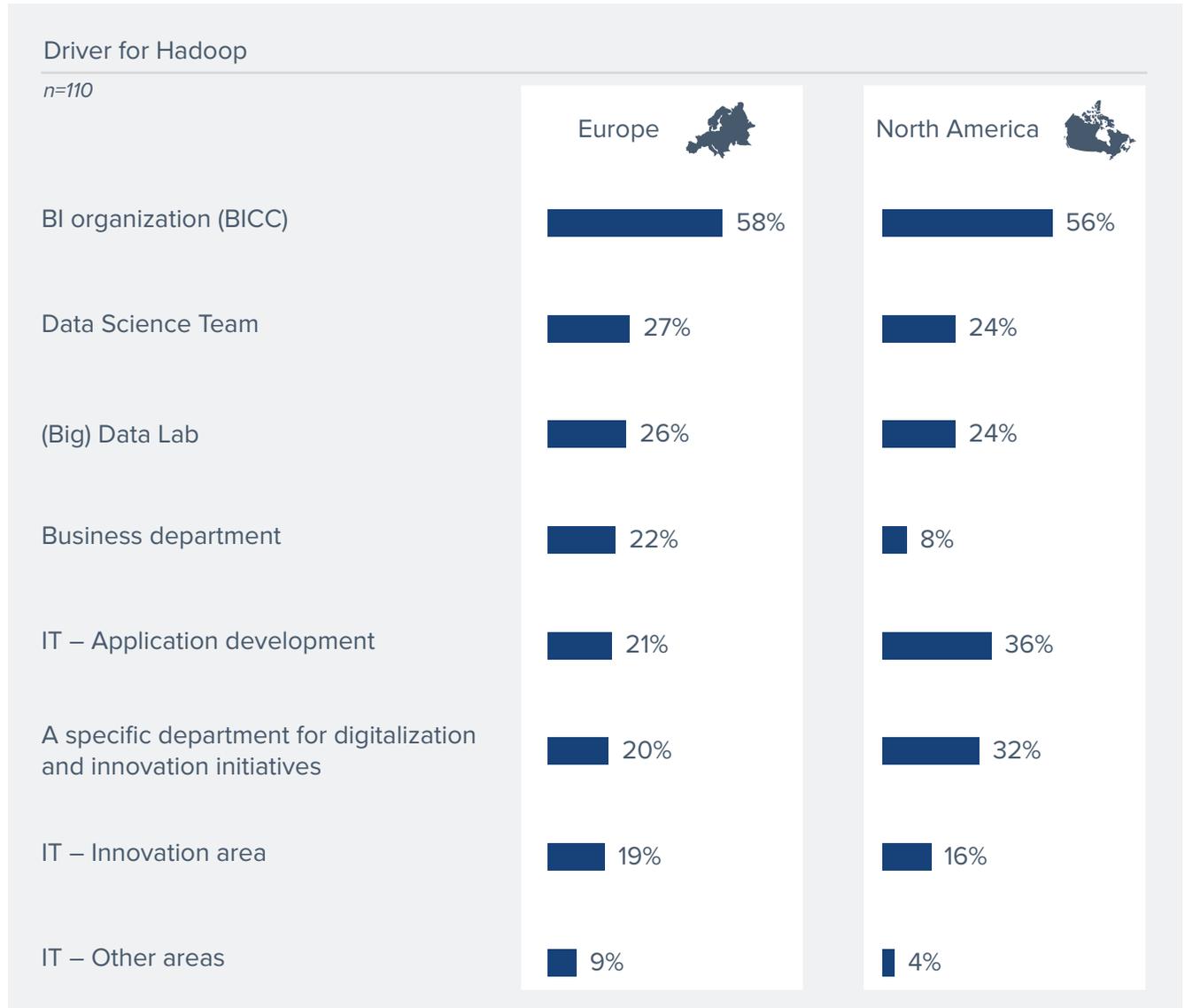


IT LOG



The BARC Hadoop Survey 2015 showed that IT departments were the drivers and visionaries behind Hadoop technologies and data lake initiatives. But times have changed. With 58 percent in Europe and 56 percent in North America, business intelligence competence centers have now emerged as the driving force in companies by some distance.

Special organizational units, such as data science teams and big data labs are becoming more commonplace, especially in Europe. In contrast, IT application development and independent departments for digitalization and innovation are driving Hadoop initiatives in North America. On the whole, pure IT departments are losing ground as a key driver. One reason could be the high demand for accessing data in business departments. This development primarily involves organizational units that have closer ties to business departments or specialize in analytics.



# Key Findings

## Reasons for Use and Benefits

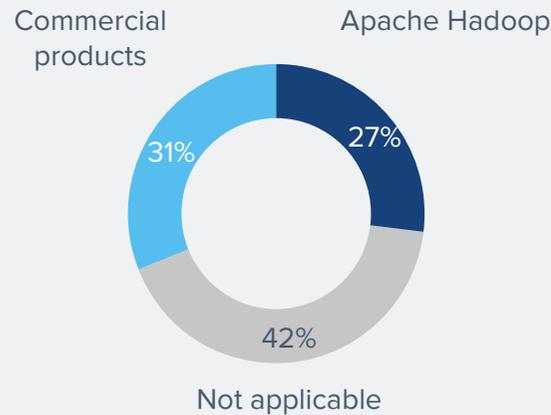


IT LOG



### Tools for implementing Hadoop

n=348



Almost a third of participants use commercial products or Hadoop distributions to implement Hadoop projects. Over a quarter (27 percent) build their own Hadoop ecosystems with Apache components. That number is surprisingly high because it takes in-depth knowledge to harmonize and manage the components as part of the implementation, maintenance and operations.

Over 40 percent of respondents stated that they do not know or cannot clearly identify the components for the implemented application scenarios. One can assume that the technical implementation is not always clear-cut, for example, due to the many different products in a Hadoop ecosystem.

### Tools for implementing Hadoop by software category

n=141

	Apache Hadoop	Commercial products	Not applicable
Data storage & access	40%	45%	14%
Data integration & quality	28%	48%	23%
System management	25%	41%	34%
Streaming	25%	22%	53%
Governance and security	23%	31%	46%
Advanced analysis & visualization	20%	64%	16%

### Users prefer on-premises installations

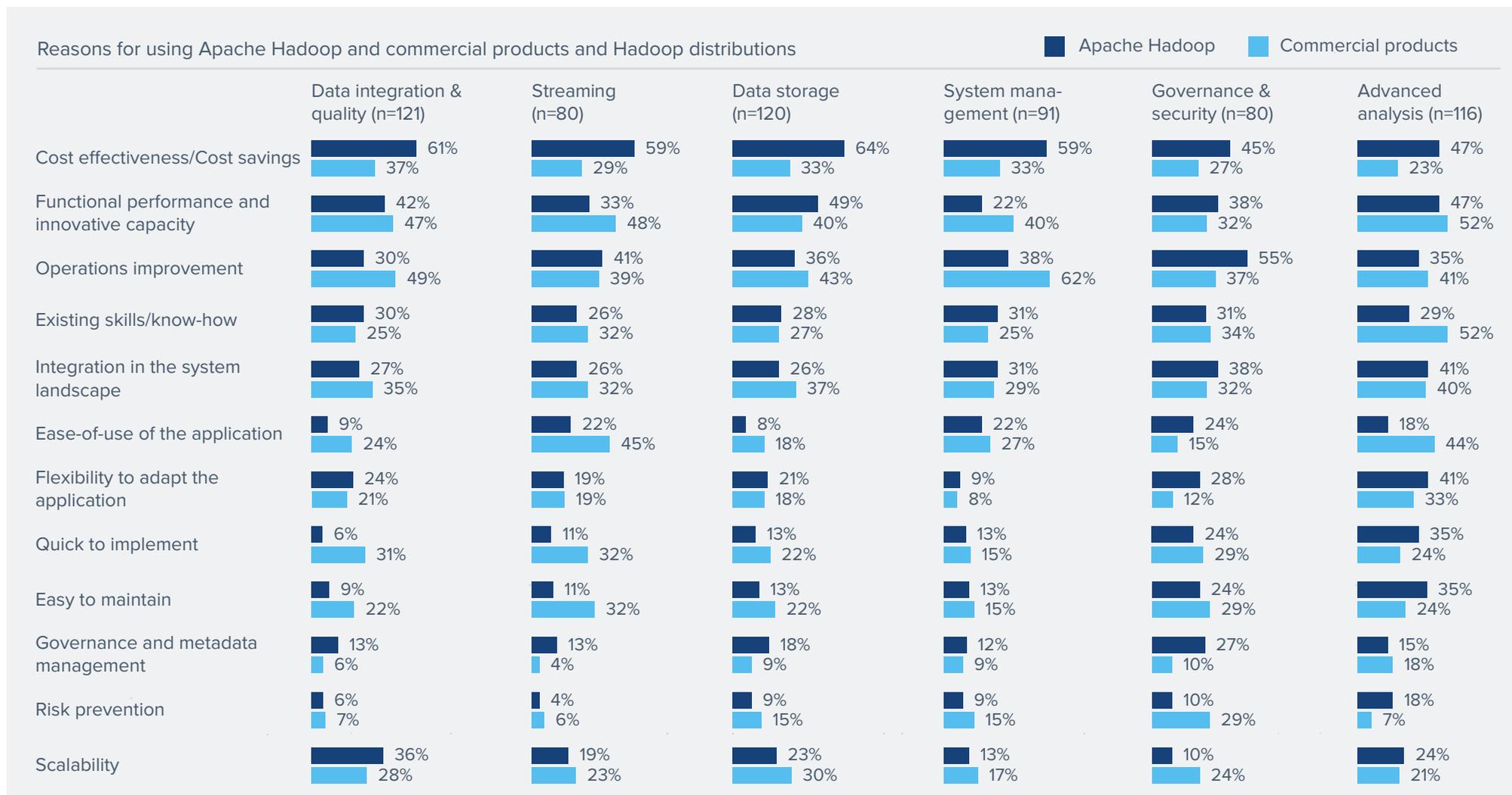
Further analysis of the software selection process shows that 61 percent choose in-house deployments of Hadoop distributions. Only a few rely on managed services (11 percent), cloud platforms (9 percent) or appliances (10 percent).

Analysis and visualization is clearly the domain of commercial tools (64 percent).

A closer look at the tool categories reveals other interesting insights. Commercial tools and Hadoop distributions clearly dominate the categories for data integration and data quality (48 percent), system management (41 percent) and, especially,

advanced analytics and visualization (64 percent). Data storage is the only area where the usage of Apache Hadoop as an open source framework is very high in comparison to commercial tools or Hadoop distributions. Data storage with the Hadoop Distributed File System, (HDFS) is one of the original basic functions of Hadoop and is therefore better known.

About half of participants use no tools for streaming (53 percent) or governance and security (46 percent). There appear to be no clearly defined products for these categories.



Participants cited cost effectiveness/savings but also functional performance/innovative capacity and operations improvement as the main reasons for choosing the open source components from Apache. In fact, these same reasons were quoted

more frequently for Apache components than for commercial tools and Hadoop distributions across all tool categories. The reasons for using commercial tools vary by tool category. Functional performance/innovative capacity and operations

improvement are popular reasons for choosing commercial products, as is ease of use in the areas of streaming and advanced analytics.



Customer intelligence/experience projects (32 percent) closely followed by predictive analytics projects (31 percent) top the list of Hadoop use cases. Customer intelligence already has many use cases since customers were at the heart of the discussion at the start of the hype around big data, and vast amounts of data on customers, their behavior and channels is available. Many of these applications, such as next-best offers in Web portals or POS data analysis, are already in production.

Predictive analysis has many use cases as well. Viewed as the epitome of explorative analysis, it helps uncover the hidden potential in data.

An analysis of use cases by company size uncovers more findings:

- Small companies focus much more on technical use cases such as data warehouse offloading.
- Customer intelligence is a major topic in midsize companies while predictive analysis is the focus of large companies.
- Midsize and large companies use significantly more clickstream, sensor and social media data.





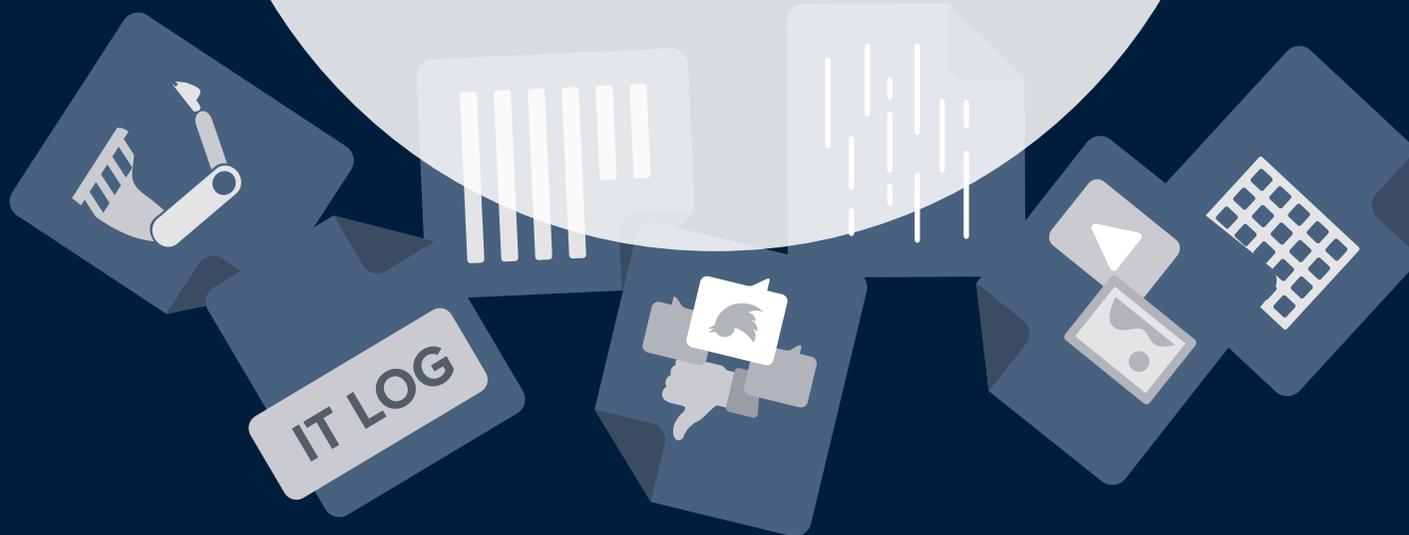
Using Hadoop has advantages almost across the board. Among the named benefits of Hadoop initiatives, comprehensive analytics and improved data integration (59 percent worldwide) ranked first followed by building a platform to better predict customer behavior and improve customer retention (53 percent worldwide). Increased analytic flexibility (47 percent worldwide) came in third. Regional differences can be found in these responses. Aside from predicting customer behavior, the North American sample uses Hadoop more heavily for predicting the success of sales or products (42 percent), monitoring and optimizing IT systems (38 percent), and increasing the effectiveness of operational processes (35 percent).

Another interpretation is that North America uses Hadoop on a broader scale and analyzes areas beyond customers – for example, products – in order to make operational processes more efficient. This last point, in particular, is where BARC sees the greatest potential for data. In order to become a data-driven company, it is important to act on the results of data analysis in operational processes instead of just following predefined workflows and rules. Users should be able to access additional functionality within an individual process and take actions based on insights. This closes the classic management loop from information to action.

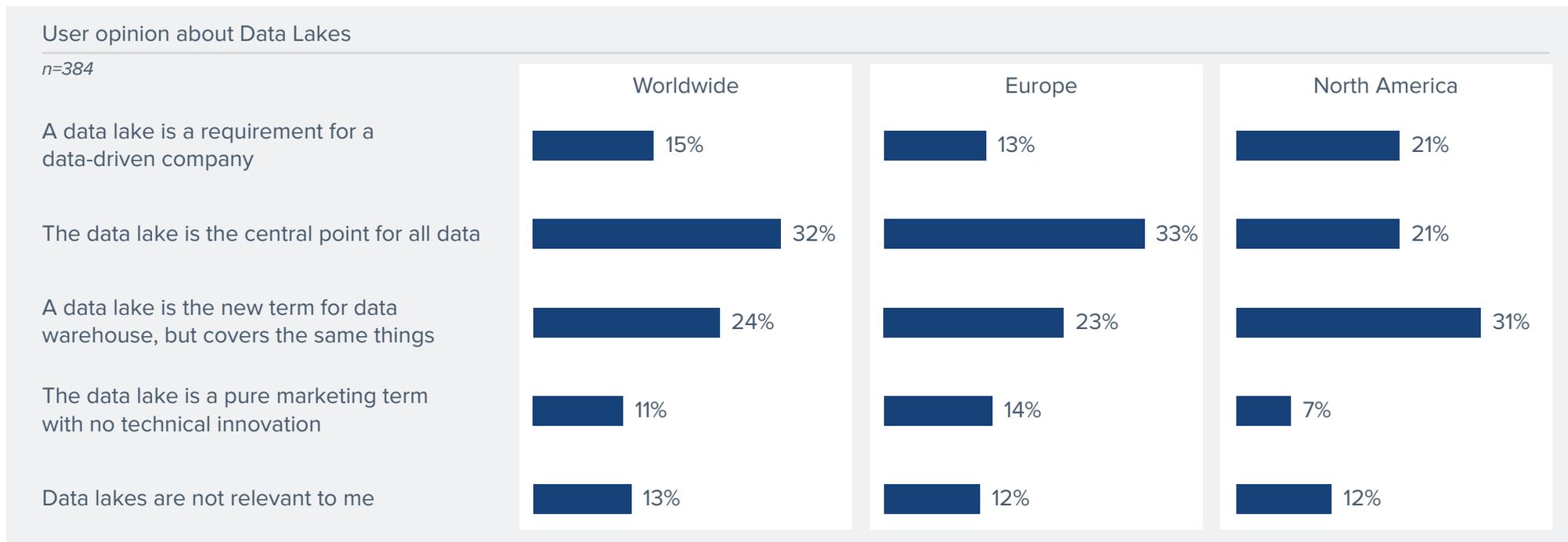


# Key Findings

## Data Lakes



IT LOG



Data lakes are a highly debated concept yet still lack a clear definition. Many view them as places to store structured, semi-structured and unstructured data – schema-free and close to its raw data format. The structure comes with usage, when the data is needed. This definition, however, still leaves many unanswered questions:

- Is the term “data lake” a synonym for Hadoop and big data technologies or is it a collection of all data storage concepts available within a company?
- Is a data lake physical storage or a logical concept?

- Which governance requirements apply to data lakes?

Regardless of the answers to these questions, the concept is relevant. 47 percent of users worldwide confirm the benefits of data lakes.

35 percent of respondents still view the data lake as a new term for an old concept or a pure marketing term. 13 percent feel that the data lake concept is irrelevant.

### Variations by region

In a regional comparison, over one fifth of respondents in North America view the concept as a pre-

requisite for a data-driven company (compared to 13 percent in Europe).

In Europe (46 percent supporters, 37 percent opponents) and North America (42 percent supporters, 38 percent opponents), respondents are split into two factions.

On the whole, there still appear to be insecurities regarding the benefits of data lakes. Its unclear definition and/or lofty promises from vendors can make it even harder to make a realistic assessment of this concept.

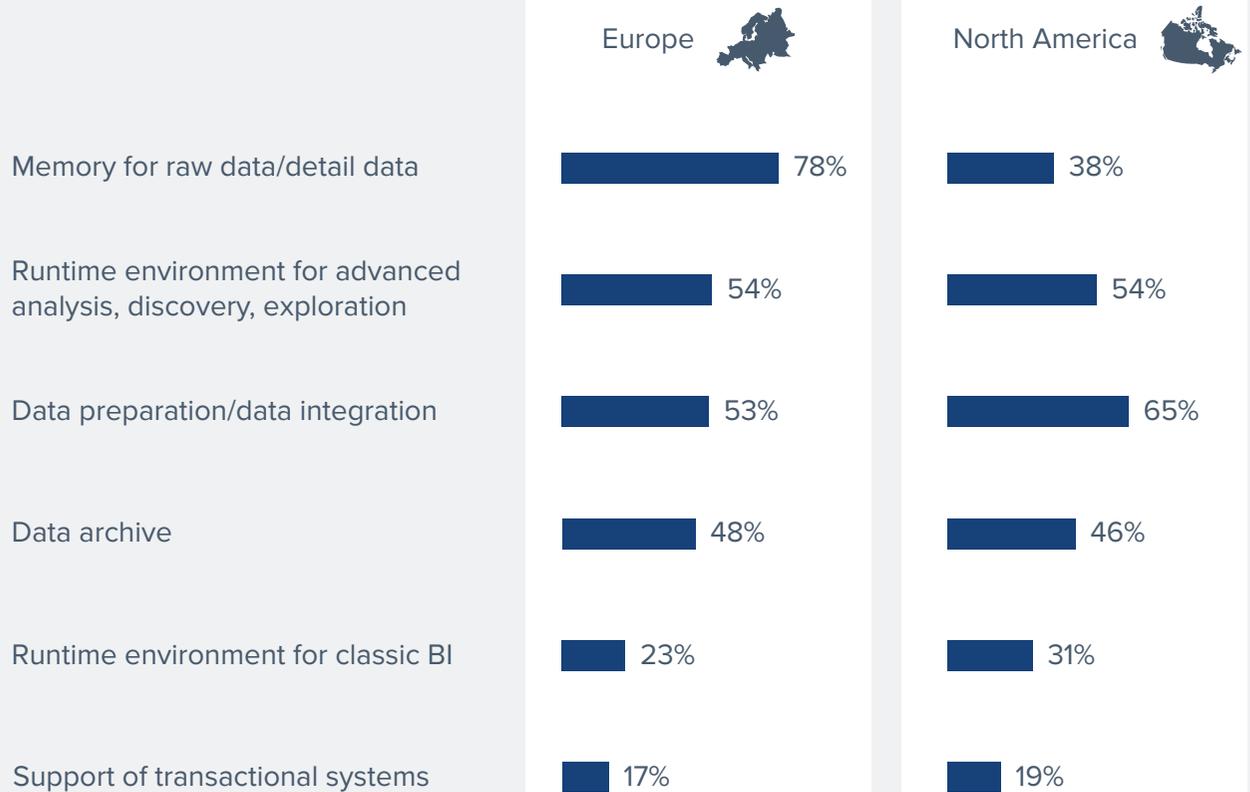


A precise definition of the term “data lake” is yet to be established, but it is clear that it is a broad concept that is not confined to a single usage scenario. There is, however, a tendency to use the term to refer primarily to data preparation and storage, or to an explorative environment.

A differentiated view of data lake usage by region reveals further insights. Data lakes are more likely to be deployed for preparing and using data in North America than in Europe. In Europe, however, a response of 78 percent clearly shows that the data lake concept is primarily used as memory for raw and detail data.

### Usage of Data Lake

*n=100*



# Key Findings

## Implementation

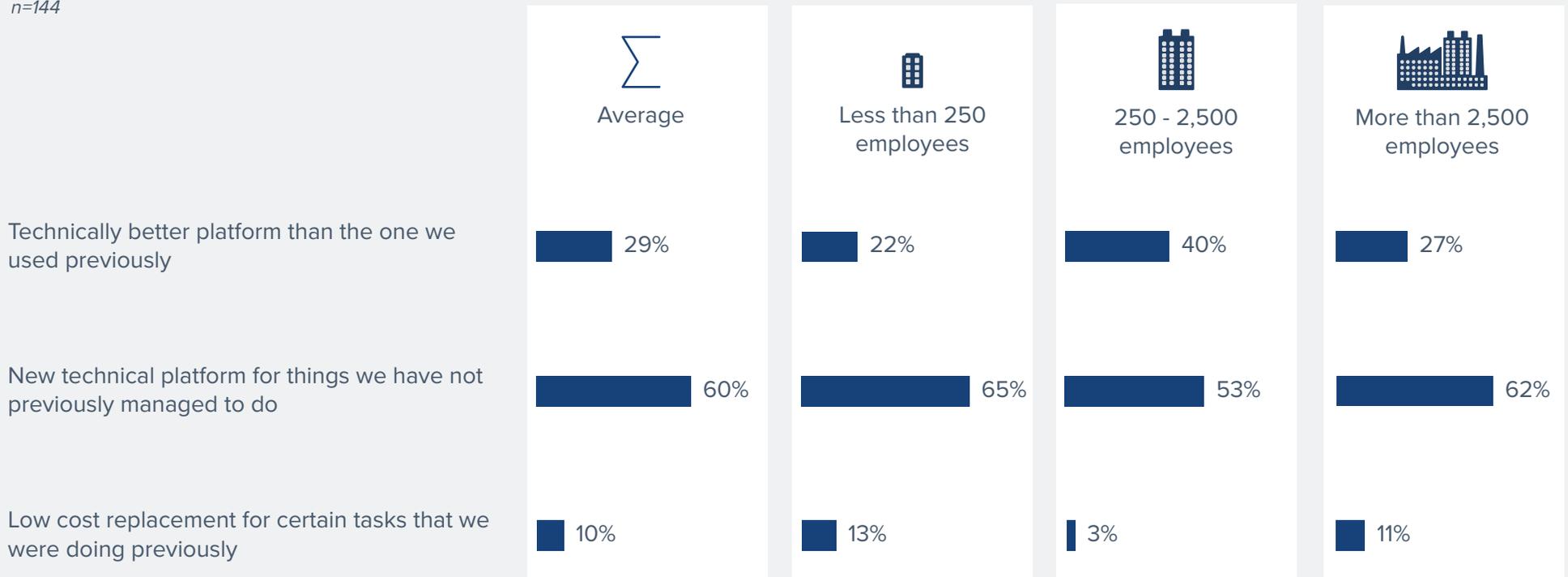


IT LOG



### Main driver for using Hadoop by company size

n=144



The question of the most important drivers for using Hadoop shows a clear trend. Six out of ten responded: “Hadoop is a new technical platform for things we have not previously managed to do.” Three out of ten answered that Hadoop is a “technically better platform”.

Only in midsize companies were these two drivers (“optimization through a technically better plat-

form” with 40 percent and “implementation of new types of use cases” with 53 percent), relatively close together.

Only one in ten respondents listed costs as the main driver. This number is surprisingly low because costs were identified as a reason for using Hadoop components (see p. 21).

BARC’s findings show that Hadoop usage is

primarily driven by new analytic requirements. Costs are not viewed as a main driver. Although cost savings have been considered a decisive advantage of Hadoop for a long time, there seems to be a more differentiated view today. Hadoop can be - but is not necessarily - less expensive depending on the use case.

# Key Findings

## Challenges





BARC's previous Hadoop Survey identified usage insecurities and a lack of technical and professional know-how as the main challenges for Hadoop. It is a similar story in this survey. Lack of professional (54 percent) and technical (50 percent) know-how topped the list of challenges.

A regional comparison offers further insights. A lack of know-how to use new findings from data research (50 percent) was more frequently viewed as a challenge in Europe. North American participants, in contrast, had more problems with a lack of sponsors and support from executive level (36 percent) and the usability of Hadoop (26 percent).

It seems that while many talk about Hadoop, few use it. And those who do apparently do not want to talk about it. That is understandable. After all, Hadoop is about gaining a competitive advantage, which the results of this survey confirm. The market still lacks experience, security and, especially, knowledge when it comes to Hadoop.

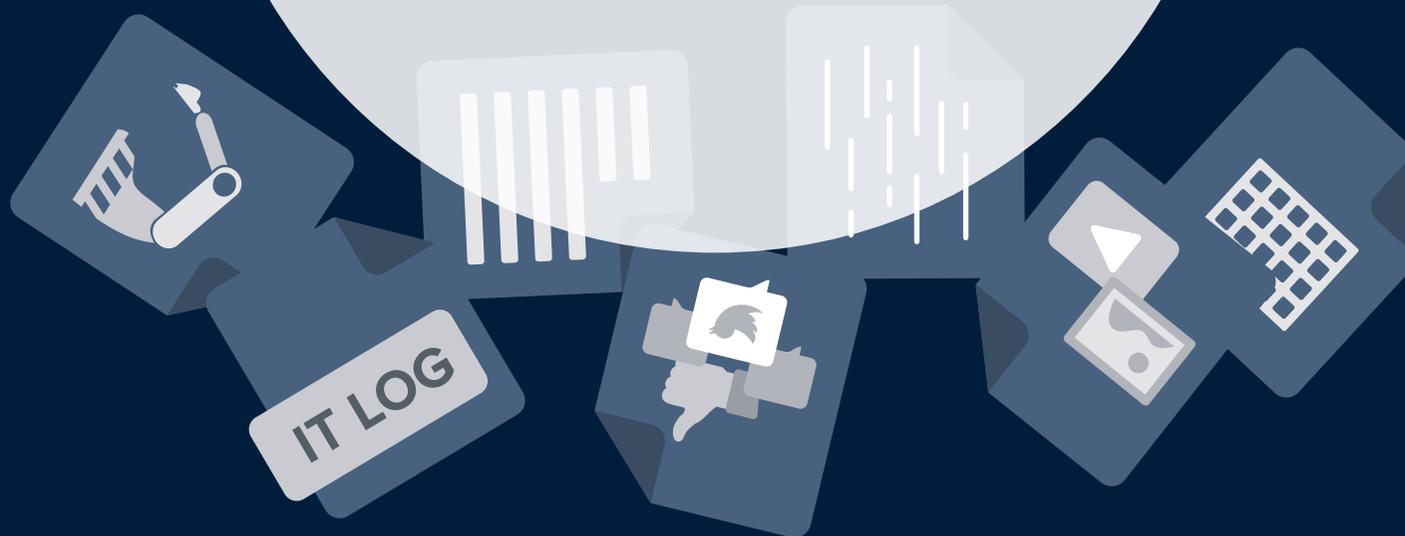
Criticisms of the Hadoop system, such as data protection and security, were only listed as a challenge in every fourth or fifth case. Broken down by industry, data security was a concern primarily in finance (33 percent) and the public sector (38 percent) in comparison to the manufacturing/processing industries (18 percent) and retail (11 percent).

### Challenges when implementing Hadoop

n=309



# Hadoop Theories Put to the Test



IT LOG



After the hype comes disillusionment and the growing realization that Hadoop and data lakes are not the answer for all analytic tasks. On the other hand, the technology and concept has great potential. There is still a need to further clarify this

area and make the usage and benefits of Hadoop and data lakes more transparent and tangible based on real experiences. This still appears to be one of the greatest challenges in the debate. To help address these insecurities, BARC analysts

listed some theories about Hadoop and the data lake concept prior to this survey, compared them with the survey findings, and added their closing comments.

### Survey findings

The findings show no clear answer. Respondents view Hadoop as a more or less suitable technology to implement a data lake concept.

### Theory

Hadoop is the preferred technology for implementing a data lake.

### BARC analysis

There are no clear guidelines for building a data lake. There are many open questions when it comes to designing a data lake, especially regarding metadata management and requirements for a virtual/logical data lake. As a result, Hadoop cannot be called a “preferred” technology in general.

### Survey findings

Survey respondents viewed this neither as a main advantage nor disadvantage of Hadoop. Basic suitability can be assumed just as in commercial tools and Hadoop distributions.

### Theory

Hadoop has functional advantages over “classic” BI/DW tools.

### BARC analysis

Hadoop offers advantages in terms of programming flexibility. However, standard software tools all have their own benefits so the decision on which approach to adopt should depend on the project and the resources and skills available. The prerequisite is to have the basic equipment regardless of the actual requirements.



## Survey findings

The survey shows that flexible application design is a strong reason for choosing Apache. A fast, simple implementation (implementation efficiency) was more of a reason for choosing commercial tools and Hadoop distributions.

## Theory

Hadoop is fast, flexible and easy to implement.

## BARC analysis

Hadoop offers advantages in terms of programming flexibility, and it can also be used for predictive analytics. However, standard software tools all have their own benefits so the decision on which approach to adopt should depend on the project and the resources and skills available. It also depends on the available knowledge of massive parallel processing (MPP).

## Survey findings

This has been confirmed both in this and previous surveys.

## Theory

Hadoop supports data with many different structures.

## BARC analysis

Yes, in the sense of a simple file system saving different formats. The schema comes with usage.

## Survey findings

This is generally true, but not the main reason for using Hadoop.

## Theory

Hadoop is cost-efficient.

## BARC analysis

It can be, but not always. Many simply think of license costs. Costs, however, also depend on the implementation, hardware and operations.



### Survey findings

Survey participants viewed this neither as a main advantage nor as a disadvantage of Hadoop. Basic suitability, therefore, can be assumed.

### Theory

Hadoop easily scales with growing data volumes and workloads in parallel environments.

### BARC analysis

Generally speaking, yes, Hadoop offers advantages in terms of programming flexibility. However, standard software tools all have their own benefits. A basic set of equipment should be guaranteed regardless of the actual requirements.

### Survey findings

Hadoop is mainly used as a technology for analytics and is rarely deployed for online/real-time processing.

### Theory

Hadoop can be used for analytics as well as online/real-time processing.

### BARC analysis

Generally speaking, yes, Hadoop offers advantages in terms of programming flexibility. It needs to be considered that Analytics and transactional applications require different designs, components and system configurations.

# Sponsors





### SAS business profile

[www.sas.com](http://www.sas.com)

With more than \$3 billion in sales, **SAS is one of the world's largest software companies and the leading vendor of big data analytics solutions.** At more than 80,000 locations around the world, enterprises rely on SAS analytics solutions for a competitive edge in **strategic and operational decisions** by tapping a wide range of business data – both separately and in conjunction with external data of any scale – for solid business insights.

Big data analytics is the key to not only managing, but profiting from the **digital transformation** and successfully putting the **disruptive processes** it entails in place. Thanks to 40 years of experience in the field of data analysis, SAS not only has sweeping vision, but also technology that is pragmatic, proven, secure and built for swift, productive deployment.

SAS systems can be found throughout the business world and in public administration. Its **core industries** are banking, insurance, trade and manufacturing. Banks use SAS to control their processes and ensure regulatory com-

pliance. Insurance companies use it to detect fraudsters. Retailers rely on SAS to optimize customer communication and campaign management, and to improve online shoppers' customer experience. Industrial enterprises use it to manage their service and maintenance processes – for instance to ensure that components are replaced before they cause unplanned downtime.

SAS big data analytics helps enterprises **extract maximum value from their data.** No matter how large and how complex the data sets are – SAS software identifies the relevant structures and relationships. Data become insights, and thus a foundation for solid and prescient business decisions.

**SAS high-performance analytics** takes full advantage of Hadoop and in-memory computing for fast, economical big data processing. SAS also offers enterprises a platform to analyze, enhance and review data – a major contribution to data quality and governance.

All SAS solutions are also available as **mana-**



**ged services** and can be deployed in the public **cloud**, the private cloud or in hybrid cloud environments. One focus here is on solutions for self-service and mobile business analytics and data visualization that enables individual departments and the management level to gain valuable insights from data without having special knowledge of statistics or requiring support from the IT department.

**Background:** SAS arose out of a research project at North Carolina State University. Established in 1976 and headquartered in Cary, North Carolina, SAS employs around 14,000 people in 59 countries worldwide. Heidelberg has been the home of SAS' German headquarters since 1982. The subsidiary has offices in Berlin, Frankfurt, Hamburg, Cologne and Munich and currently employs 520 people. German customers include Allianz, Continental, Commerzbank, HUK Coburg, Fraport, DER Touristik, Nestlé, Galeria Kaufhof, BASF and Meyer Werft.



## Talend business profile

[www.talend.com](http://www.talend.com)

Talend ([NASDAQ: TLND](#)) is a next generation leader in cloud and big data integration software that helps companies become data driven by making data more accessible, improving its quality and quickly moving data where it's needed for real-time decision making. By simplifying big data through these steps, Talend enables companies to act with insight based on accurate, real-time information about their business, customers, and in-

dustry. Talend's innovative open-source solutions quickly and efficiently collect, prepare and combine data from a wide variety of sources allowing companies to optimize it for virtually any aspect of their business. Talend is headquartered in Redwood City, CA. For more information, please visit [www.talend.com](http://www.talend.com) and follow us on Twitter: [@Talend](#).



### Contact information

Talend Inc.

800 Bridge Parkway

Suite 200

Redwood City California 94065

United States

Tel: +1 (650) 539 3200

[sales@talend.com](mailto:sales@talend.com)

[www.talend.com](http://www.talend.com)

# About BARC



# BARC — Business Application Research Center

## A CXP Group Company



BARC is a leading enterprise software industry analyst and consulting firm delivering information to more than 1,000 customers each year. Major companies, government agencies and financial institutions rely on BARC's expertise in software selection, consulting and IT strategy projects.

For over twenty years, BARC has specialized in core research areas including Data Management (DM), Business Intelligence (BI), Customer Relationship Management (CRM) and Enterprise Content Management (ECM).

BARC's expertise is underpinned by a continuous program of market research, analysis and a series of product comparison studies to maintain a detailed and up-to-date understanding of the most important software vendors and products, as well as the latest market trends and developments.

BARC research focuses on helping companies find the right software solutions to align with their business goals. It includes evaluations of the leading vendors and products using methodologies that enable our clients to easily draw comparisons and reach a software selection decision with confidence. BARC also publishes insights into market trends and developments, and dispenses proven best practice advice.

BARC consulting can help you find the most reliable and cost effective products to meet your specific requirements, guaranteeing a fast return on your investment. Neutrality and competency are the two cornerstones of BARC's approach to consulting. BARC also offers technical architecture reviews and coaching and advice on developing a software strategy for your organization, as well as helping software vendors with their product and market strategy.

BARC organizes regular conferences and seminars on Business Intelligence, Enterprise Content Management and Customer Relationship Management software. Vendors and IT decision-makers meet to discuss the latest product updates and market trends, and take advantage of valuable networking opportunities.

Along with CXP and Pierre Audoin Consultants (PAC), BARC forms part of the CXP Group – the leading European IT research and consulting firm with 140 staff in eight countries including the UK, France, Germany, Austria and Switzerland. CXP and PAC complement BARC's expertise in software markets with their extensive knowledge of technology for IT Service Management, HR and ERP.

### Other Surveys

BARC research reports bring transparency to the market



**The Planning Survey 16** is the world's largest survey of planning software users. Based on a sample of over 1,200 responses, it offers an unsurpassed level of user feedback on 13 leading planning products.



The BARC **Big Data Use Cases** Survey explores the usage of big data in companies worldwide. 559 business and IT decision-makers completed the survey in the first quarter of 2015.



The **BI Trend Monitor 2016** from BARC reflects on the trends currently driving the BI and data management market from a users' perspective. We asked close to 2,800 users, consultants and vendors for their views on the most important BI trends.





**Business Application Research Center – BARC GmbH**



A C X P G R O U P C O M P A N Y

**France**

BARC France (Le CXP)  
8 Avenue des Ternes  
FR-75017 Paris  
+33 1 530505 53  
[www.cxp.fr](http://www.cxp.fr)

**Germany**

BARC GmbH  
Berliner Platz 7  
D-97080 Würzburg  
+49 (0) 931 880651-0  
[www.barc.de](http://www.barc.de)

**Austria**

BARC GmbH  
Goldschlagstraße 172 / Stiege 4 / 2.OG  
A-1140 Wien  
+43 (1) 8901203-451  
[www.barc.de](http://www.barc.de)

**Switzerland**

BARC Schweiz GmbH  
Täfernstrasse 22a  
CH-5405 Baden-Dättwil  
+41 76 340 35 16  
[www.barc.ch](http://www.barc.ch)

**Rest of the World**

+44 1536 772 451  
[www.barc-research.com](http://www.barc-research.com)