



RESEARCH
STUDY

Data Preparation - Refining Raw Data into Value

Authors:

Robert Tischler

Timm Grosser



This study was prepared by BARC, an independent market analysis firm.
This study is available free of charge thanks to the generosity of Denodo, SAS, Tableau and TimeXtender.

denodo

sas
THE POWER TO KNOW.

tableau

TIMEXTENDER

Table of Contents

- Data Preparation - Refining Raw Data into Value 3
- Preface..... 3
- Management summary 4
- Survey Results 10
 - High expectations for analytics are driving data preparation to create value from data 10
 - Data preparation – a modern approach to data preparation as a source of hope for analytics and BI13
 - Expectations already exceeded – broad benefits from data preparation14
 - High potential for improvement in data preparation..... 16
 - The key personnel for data preparation are still IT..... 18
 - Business departments and IT are driving data preparation initiatives to a similar degree 19
 - Data preparation is predominantly used by experienced BI experts.....20
 - Productive cooperation is more important than rigid rules.....21
 - Data governance is a fundamental pillar for creating value from data..... 22
 - Data preparation between desire and reality..... 24
 - Spreadsheets were yesterday – now it’s time for the right technology..... 26
 - New data sources are only slowly being developed.....28
 - Data preparation does not happen by itself – the right conditions must be created 29
 - No one masters data preparation without problems.....31
- Conclusion 33
- Attachment A: Methodology and Demographics 35
- Sponsor Profiles..... 39

Data Preparation - Refining Raw Data into Value

Robert Tischler, Senior Analyst Business Intelligence

Timm Grosser, Senior Analyst Data Management

June 2017

RESEARCH
STUDY

Preface

There is high potential for creating value from the use of data. Being able to use it correctly requires an acute awareness of how to handle data and brings with it plenty of changes in terms of organization and technology. Business users demand a fast and flexible analytical landscape from their IT departments. A lack of resources, a lack of flexibility, and complex, historically grown systems are standing in the way of efficient and agile delivery. Smart data management is key for business success, especially in the field of digitalization where data and analytics are of increasing importance and have a growing influence on the business. Unfortunately, many companies are still learning this. For a long time now, data management has ceased to be only in IT's area of expertise. Traditional IT tasks related to data management are increasingly being taken over by business departments to ensure timely completion. First and foremost comes data integration and the provision of data.

Deployment scenarios for data preparation range from self-service BI tools directly accessing operational or analytical systems to specialized self-service data integration (DI) tools for supplying analytical models or explorative sandboxes with data. Self-service data integration tools enable business users to prepare relevant data in a flexible and self-reliant manner to be used for analysis. The self-service trend in DI has already reached the market, forcing leading DI vendors to offer options and interfaces as well as governance frameworks for data integration by business users. In addition, BI vendors (specialists and BI generalists) are adding data preparation capabilities to their analysis and data discovery tools.

This places additional strain on companies' BI organizations and BI governance as the responsibility for analytics becomes increasingly scattered. It is also a challenge for IT departments - who are in charge of operating the software tools, application servers and fulfilling data

Scattered responsibility for analytics places additional strain on companies' BI organizations and BI governance.

needs - as they have less control over user behavior. Therefore, striking the right balance between flexibility and data governance is a crucial element in the success of data preparation.

This study clearly demonstrates the benefits, timeliness and relevance of data preparation for analytics. It shows how and by whom data preparation is being driven and how the balancing act between governance and flexibility can be achieved by specifying the requirements for data preparation governance. In this BARC study, we also show how data preparation is used today, which challenges need to be overcome, and in which organizational framework this takes place.

Management summary

The increasing digitalization of business processes is making it necessary for companies to enable as many users as possible to gain insights from data (democratization of analytics). Many companies today view data preparation as the key to increasing their ability to efficiently use data in a distributed manner to optimize business processes, or to enabling new, innovative business models in the first place.

In today's economy, achieving efficient and agile data preparation is of utmost importance. Increasingly volatile and saturated markets create a complex business environment where the ability to differentiate by leveraging the power of analytics is vital. Organizations struggle to keep up with the demand for data for analytics to gain insight into changing market conditions. The pressure on analytical landscapes to provide data for in-depth analysis is high and addressing these needs requires skilled personnel and a modern approach to data preparation.

In today's economy, achieving efficient and agile data preparation is of utmost importance.

To find out more about current thinking on data preparation, BARC conducted an independent survey of over 695 BI professionals from a range of industries worldwide. The BARC Survey "Data Preparation - Refining Raw Data into Value" is one of the largest studies focusing on the conditions, benefits and challenges of data preparation. It has been made available to readers free of charge thanks to sponsorship from Denodo, SAS, Tableau and TimeXtender.

We have condensed the key findings of this study into seven hot spots.

da·ta pre·pa·ra·tion

/ˈdeɪtə pre.pa.raˈtʃjən/, n <comp>
the process of preparing and providing
data for data discovery and analytics

The goal of data preparation is to support business analysts and data scientists by preparing different kinds of data for their analytical purposes. The preparation of data can take place either in business departments or be performed centrally by IT. Data preparation is a sub-domain of data integration that can be executed with dedicated tools or traditional tools for data integration like ETL tools, data virtualization or data warehouse automation.

Data preparation is driven by concrete needs and is already widely practiced

Today's businesses face great challenges, as they have done throughout history. What is new is that the ability to use data systematically has become a decisive competitive advantage. Many companies have recognized this and are striving to solve many of their data usage problems by introducing or improving data preparation. The main drivers behind projects show that the hype around data preparation, which undoubtedly exists, is backed by „concrete“ requirements. Issues only indirectly linked to data management such as cost and usability are further down the list of drivers (see Figure 2).

When companies adopt trending technologies, they often have high expectations (see Figure 5). This is also the case with data preparation. Data preparation tools and methods are used to tackle major challenges, and our survey results show they indeed provide benefits. The already high expectations of users are consistently exceeded. If suitably embedded in the organization, data preparation offers a real opportunity to provide data for analytics in better shape and faster, and thus generate immediate benefits for the enterprise.

High expectations of the benefits of analytics and the need for agility are driving the use of data preparation. The share of companies already using data preparation to quench their thirst for information is correspondingly high (see Figure 8). Almost 70 percent of respondents reported they already use data preparation.

Even though its utilization rate is high and companies are drawing tangible benefits from data preparation, satisfaction with its implementation is not so high (see Figure 6), mainly due to challenges in the area of data management.

“Data preparation is the process of preparing and providing data for data discovery and advanced analytics.”

Hot Spot 1



Data preparation is key to refining raw data into value.

Challenges in the area of data management reduce the ability to leverage the full potential of data preparation.

We are only strong together – the best of two worlds

A promising approach emerging in the context of data preparation for analytics is the division of labor across departmental boundaries. The companies who report the greatest benefit and highest satisfaction with data preparation are those that have made data preparation a shared task between IT and business departments, where users with a strong sense of the business are able to prepare data independently in the main, but with the support of technical experts who ensure compliance with standards (see Figure 1). This enables business problems to be answered quickly when they arise. Developed solutions are then automated by technical specialists and made available to a wider audience. This can prevent the formation of separate solutions and silos caused by the isolated preparation of data.

Hot Spot 2



The companies achieving the greatest benefit and highest satisfaction with data preparation are those that have made data preparation a shared task between IT and business departments.



Figure 1: Which benefits do you generate through data preparation? By method of use (n=453)

In order to increase the achievement of objectives, not only do business users and IT have to cooperate, but management must also provide the appropriate resources for training (especially for business users) and suitable tools, and thus also strengthen their own position as a driver for improvements in the area of analytics.

Data governance is a fundamental pillar for creating value from data

Data governance uses standards and rules to enable efficient value creation from data. The prerequisite for creating value from data and therefore the task of data governance is to establish clear responsibilities, define transparent goals and structures, provide standard definitions, and ensure data quality and data security.

Our survey participants are aware of the relevance of data governance for data preparation, in particular to ensure data quality, data security and the use of standard definitions (see Figure 12). However, the need for action still seems to be high in order to implement these points satisfactorily.

This is not always a simple undertaking as a successful implementation often requires adjustments to the organization. In the context of data preparation, it is primarily about finding the right balance between central stability and the desired level of decentralized flexibility in data preparation.

Hot Spot 3



Measures to improve data quality and the use of common definitions are two of the essential components in a solid data governance framework for data preparation.

Inconsistent organizational orientation – Is there a silver bullet?

The magic formula for anchoring data preparation within the organization is yet to be found. Although the use of data preparation methods and tools is now widespread, a clear division of tasks between IT and business users has yet to crystallize. Respondents to our survey indicate that organizations have adopted a broad range of approaches (see Figure 8).

IT plays an important role in the two most popular approaches (i.e. in at least half of the companies surveyed). The proportion of enterprises in which business users are actively working on preparing data is also high. Company size is less decisive in determining who and how data preparation is performed than the technical complexity of the requirements and the skills profile that exists in each department. Therefore, there is no sign of a one-size-fits-all magic formula, not least because the data preparation tools market is evolving rapidly.

Hot Spot 4



IT plays an important role in executing data preparation today, but business is increasingly assuming the lead role. Strengthen collaboration between both sides to reap the rewards.

Data preparation as an approach primarily for simple applications

The need for data management solutions is huge, especially for data preparation. When it comes to data preparation, the focus is currently not on complex processes for the preparation of data for advanced analytics or as the basis for big data, but rather on supporting simple applications such as the enrichment and further refinement of data for analyses. Functions for supporting data transformation and ensuring data quality or access to data are the most frequently used (see Figure 13). Thus, the majority of users view data preparation as an effective means of solving everyday problems. In contrast, fewer specialists require data preparation for advanced analytics or data science.

Hot Spot 5



Great benefits can already be achieved simply by tackling basic tasks.

The good news is that great benefits can already be achieved by solving basic tasks (see Figure 5). The extent to which the term “data preparation” is actually perceived in the market as the discipline of preparing data for data discovery or advanced analytics remains unclear.

Using the right tools for data preparation unleashes greater potential

The use of Excel is widespread and it also seems to be the first choice tool for data preparation today (see Figure 13). The fact that Excel does not support complex data preparation should be clear to everyone, and should also explain why predominantly simple applications are implemented using data preparation.

Hot Spot 6



Spreadsheets lack functionality for advanced integration tasks as well as the ability to map automated, high-performance and stable

processes for data preparation. The key question here is whether Excel is inhibiting the development of data preparation's potential due to its limited scope of functions, or whether there is simply a lack of use cases that require specialized tools for data preparation.

Overall, the performance of standard tools seems to be underestimated when it comes to data preparation, either as an instrument for IT-centric approaches or to provide direct support in business departments.

The value of data is recognized but the skills to make data useful are in short supply

Data preparation especially supports business departments and data scientists when it comes to preparing data for data discovery or advanced analytics and data science. The ability to gain valuable insights from data through efficient and largely independent data preparation in business departments is currently seen as patchy. Therefore, there is an urgent need for training and coaching in order to be able to implement sophisticated digitalization strategies. To do this, dedicated resources and budgets need to be allocated.

According to survey participants, these are among the greatest challenges in setting up data preparation initiatives (see Figure 18). This highlights the lack of focus on data preparation from management and the inadequate strategic anchoring of initiatives for the systematic use of data.

As with many aspects of data management, data preparation cannot be done „in passing“. It has to be viewed as a valuable step in the process of creating value from data: not a one-off project that can be outsourced and completed externally, but an ongoing endeavor that requires a high degree of competence. Therefore, data preparation needs to be deeply and extensively embedded in the organization.

Are Excel's limitations inhibiting the development of data preparation's potential, or is tool adoption low due to a lack of relevant use cases?

Hot Spot 7



The ability to gain valuable insights from data is currently patchy.

Experience is key: Use pilot initiatives to gather experience in data preparation and to demonstrate its value.

Data preparation needs to be deeply and extensively embedded in the organization. Acquire and nurture know-how and skills.

Survey Results

High expectations for analytics are driving data preparation to create value from data

The importance of data and analytics to business success is confirmed yet again by our survey results. Thanks to heightened awareness and discussion of big data and digitalization, responsibility among business users for the use of data has increased significantly. Data and its intensive use for analytics in business departments make a significant contribution to the success of today's companies. As Figure 2 shows, higher expectations in terms of performance, agility, and flexibility (49 percent) in particular are driving the topic of data preparation in companies. Similarly, strong drivers for the introduction of data preparation are leading to increased competitiveness through analytics (47 percent) and the ability to properly tackle current challenges in data management (46 percent). The goal behind the top drivers is usually to recognize changing conditions quickly, for example in the purchasing behavior of customers, in order to react faster than the competition. The top drivers in Figure 2 suggest that data preparation is therefore based on real requirements and is not just hype.

The goal behind the top drivers is usually to recognize changing conditions quickly in order to react faster than the competition.

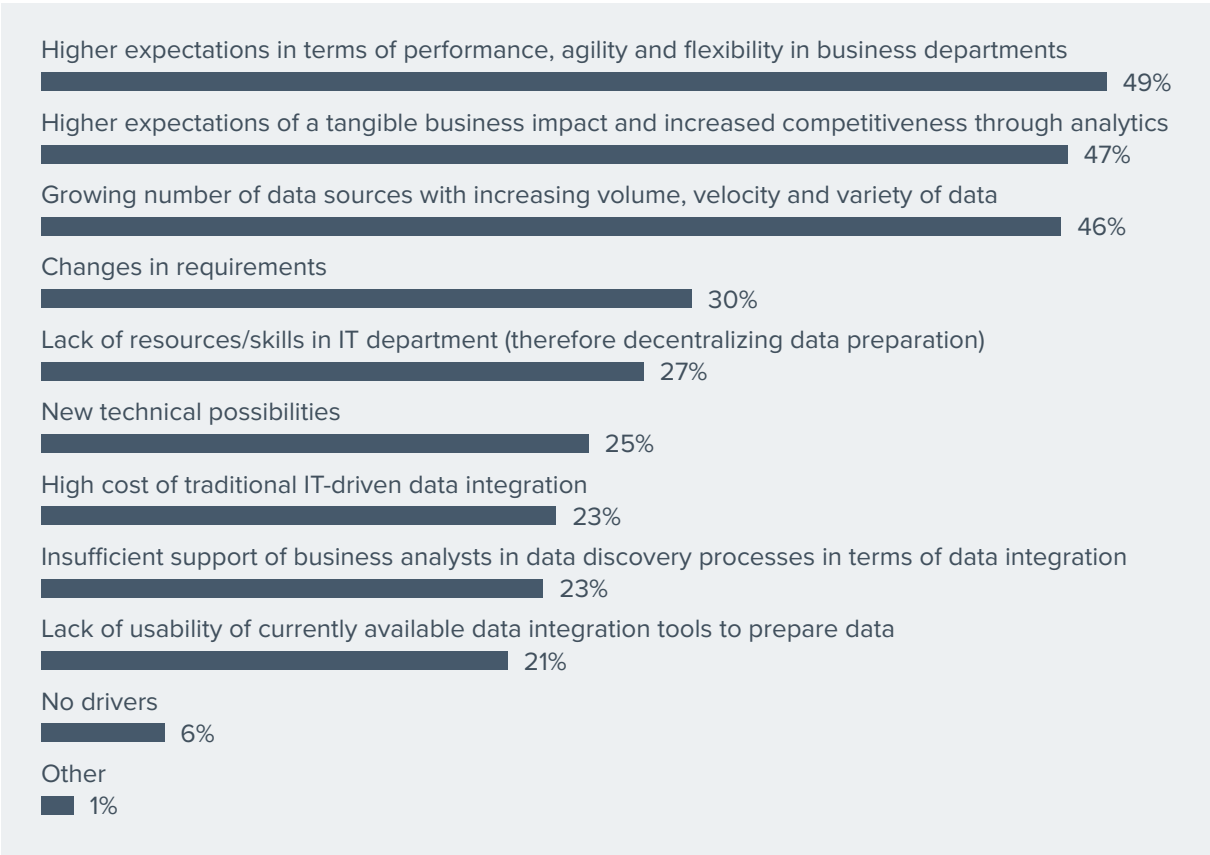


Figure 2: Which of the following are the most important drivers for data preparation in your company? (n=678)

The drivers of data preparation are primarily business-led in nature. The availability of new technical possibilities for implementation or easier-to-use tools for business departments are not among the most important reasons why companies are engaging with data preparation more intensively. The fact that these two points are way down the list of important drivers does not support the theory that data preparation has only become a mass phenomenon due to new, specialized tools that business users are able to handle. A lack of tools and the high cost of traditional landscapes also play a minor role only. This underscores the fact that there are „real“ requirements behind the high interest in data preparation, rather than the idea that technological hype is driving interest.

The different view of companies depending on their size reveals another interesting insight. For example, large companies currently feel much higher pressure to adapt. For almost all the drivers mentioned, the challenges faced by large companies are seen to be significantly greater than those of smaller companies. These differences are particularly evident in the availability of resources (12 percentage points more than medium-size companies) and IT costs for data integration (14 percentage points more than small businesses). In large companies with more than 2,500 employees, data preparation therefore seems to be seen as an opportunity to solve the problem of a lack of flexibility by shifting data management skills from IT directly into business departments.

In view of the strong growth in the performance of data management tools, companies today should have already solved many of the challenges relating to data management for BI. However, old and new requirements are frequently stretching historically grown analytical ecosystems to their limits. This is not necessarily due to the technologies used. Figure 3 clearly confirms the need to optimize basic tasks and the methods for data preparation that are already easily controllable. For example, 49 percent of respondents experience problems combining data from different sources. 45 percent of respondents highlight problems with data quality and 39 percent have difficulties actually getting hold of relevant data for analytics.

An explanation for this is the growing need of business departments to intensively use data for analytical purposes. For years, observers have been highlighting a shift in the skills required for analytical tasks from IT to business departments. The expectation of being able to generate immediate value from data is encouraging this development and making data management and analytics a top priority today. This high priority either requires a viable solution to well-known and long-standing legacy challenges or provides the opportunity to formulate data requirements again and better.

Old and new requirements are stretching historically grown analytical ecosystems to their limits.

The expectation of being able to generate immediate value from data is encouraging this development and making data management and analytics a top priority today.

Only a relatively small proportion of respondents see challenges in preparing data for analytics in particular. Just 21 percent lack functions to search for and identify data, while 20 percent state that they do not have the right tools. The good news is that great benefits can already be achieved simply by tackling basic tasks. It remains unclear the extent to which data preparation is actually perceived as a discipline for preparing data for data discovery or advanced analytics in the market and wherein the main benefit is seen.

Great benefits can already be achieved simply by tackling basic tasks.

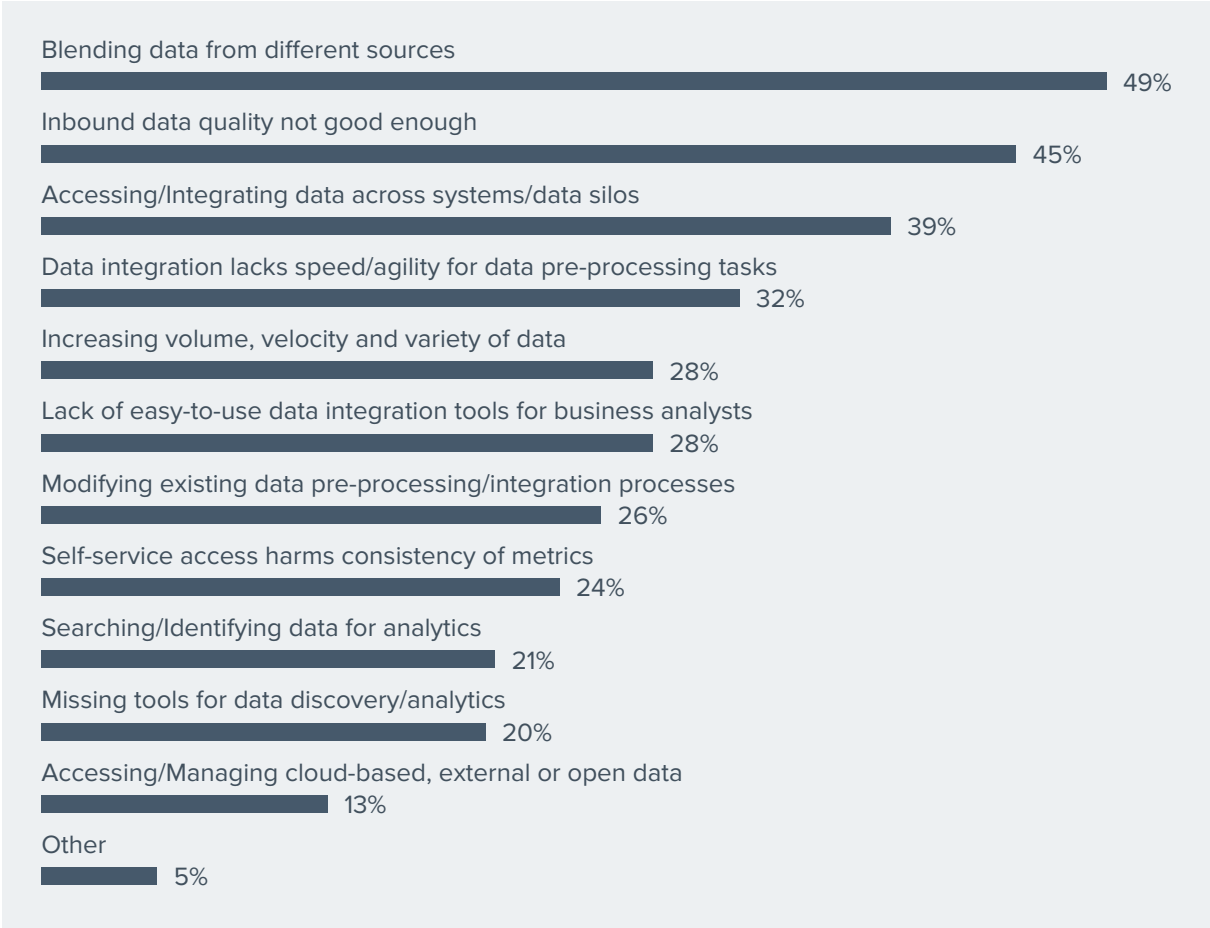


Figure 3: What are the most important data integration challenges in your analytics landscape? (n=677)

Data preparation – a modern approach to data preparation as a source of hope for analytics and BI

da·ta pre·pa·ra·tion

/'deɪtə pre.pa.ra.'tʃjən/, n <comp>
the process of preparing and providing data for data discovery and analytics

“Data preparation is the process of preparing and providing data for data discovery and advanced analytics.”

The goal of data preparation is to support business analysts and data scientists by preparing different kinds of data for their analytical purposes. The preparation of data can take place either in business departments or be performed centrally by IT. Data preparation is a sub-domain of data integration that can be executed with dedicated tools or traditional tools for data integration like ETL tools, tools for data virtualization or data warehouse automation.

Data preparation for analytics has a high priority in most companies. Already today, more than 75 percent of respondents consider data preparation to be important (see Figure 4). The proportion of companies who consider data preparation to be critical for their business rises from 12 percent today to 2.5 times that figure at 30 percent in the future. This escalation clearly shows that efficient data integration is significantly increasing in relevance. The challenge of data preparation is, therefore, much more visible and more highly prioritized, which can be explained by the higher pressure to take action due to digitalization or the opportunities newly opened up by big data, for example. Both topics call for business departments, in particular, to act.

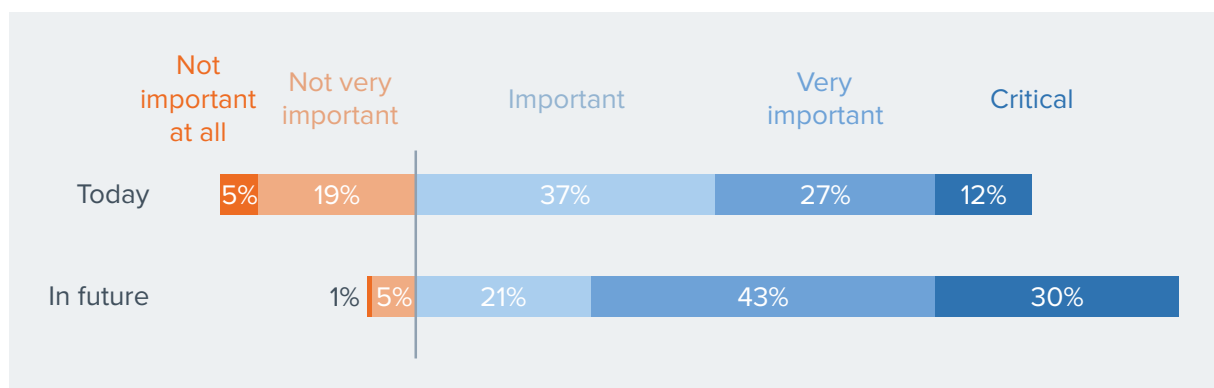


Figure 4: How important is data preparation for data discovery and advanced analytics in your company today?

Expectations already exceeded – broad benefits from data preparation

This statistic speaks for itself: only one percent of respondents see no benefit in data preparation! So data preparation offers benefits to almost everyone. Let us now qualify these benefits.

The analogy between drivers for data preparation (see Figure 2) and the benefit actually achieved is striking (see Figure 5). For example, the greatest benefit from data preparation is seen in the topics covered by the top three drivers. These are the higher expectations in terms of performance and flexibility of analytics as well as addressing the three Vs (volume, variety and velocity). This illustrates the high level of goal attainment through the use of data preparation. In addition, half of the respondents consider the shorter „Time to Insight“ to be of great benefit. Points that are more strategic in nature and have no direct business benefits, but which make critical contributions to successfully working with data – such as collaboration, democratization and reduction of information silos – are recognized, but do not represent the main benefits of data preparation according to the respondents. Understandably, the benefit of data preparation seems to lie primarily in the solution of specific data preparation tasks rather than in the strategic orientation towards creating value from data.

The benefit actually achieved through data preparation significantly exceeds companies' expectations. Only on one point are expectations marginally higher than the benefits actually achieved: in the ability to react to changing business conditions. A significantly higher benefit than expected was achieved by providing an overview of relevant data.

Data preparation can provide benefits for any company.

High level of goal attainment through the use of data preparation.

The benefit actually achieved through data preparation significantly exceeds companies' expectations.

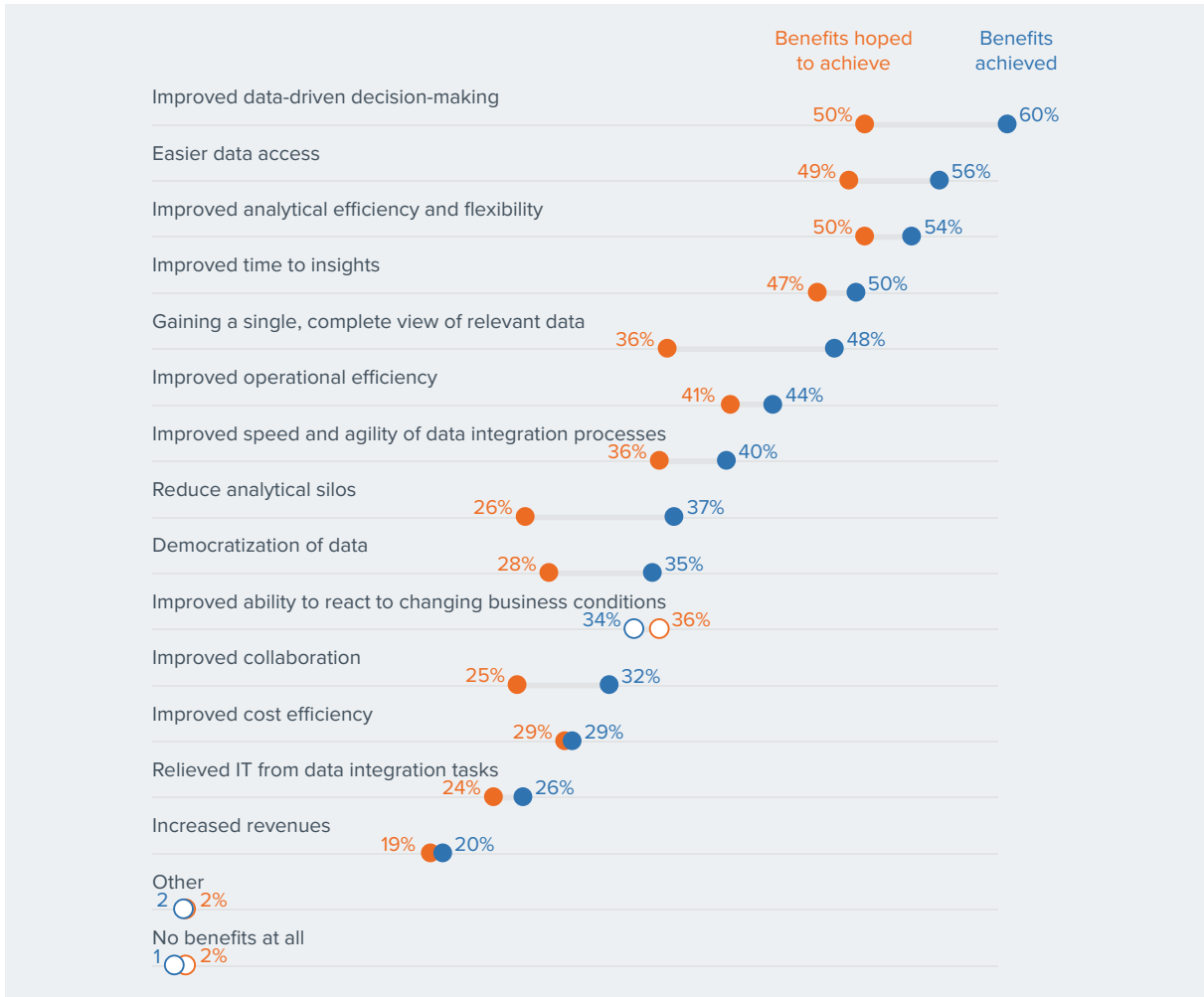


Figure 5: “Which benefits do you generate through data preparation?” (n=453) and “which benefits do you hope to generate through data preparation?” (n= 177)

Overall, the chart shows that data preparation generates significant improvements in many areas relevant to a company’s success and thus brings considerable benefits for these companies. Data preparation is thus perceived as a useful solution for a variety of today’s challenges in the analytical process.

Data preparation leads to significant improvements in many areas relevant to a company’s success and thus creates considerable benefits for companies.

High potential for improvement in data preparation

Whilst expectations of the benefits to be gained from data preparation are consistently high, the satisfaction shown in Figure 6 reveals considerable potential for improvement in many regards. Although a majority (53 percent) of companies are happy with the way they are handling traditional data integration processes, 50 percent have access to required data, and 46 percent are satisfied with performance, on average every third company struggles with the challenges of data preparation .

34 percent of companies surveyed are not satisfied with the quality of data. For years now, data quality has been one of the most important challenges and success factors, not only in traditional DWH environments, but equally for exploratory analysis. Particularly due to the growing importance of data in companies, data quality and master data management are gaining significantly higher priority as a prerequisite for creating value from data. In addition, 34 percent especially consider the identification and cataloging of data for analysis as an area in need of improvement . Only 37 percent of the participants in our survey are satisfied with today's solutions. One of the most important requirements for data preparation and therefore also a core raison d'être is the ability to react quickly and efficiently to changing requirements – the much touted agility. Modified requirements relating to the purpose of the analysis often manifest in the need to modify data structures. Therefore, providing and preparing data for ad hoc needs is vital to get value from data. On this point, dissatisfaction prevails at 40 percent compared to 35 percent satisfaction with the current state of implementation. The same is true for the management and integration of the data preparation workflow into the analytical landscape of companies with 36 percent dissatisfaction versus 32 percent satisfaction.

Data quality remains a success factor and also one of the most important challenges for exploratory data analysis.

One of the most important requirements for data preparation is the ability to react quickly and efficiently to changing requirements.

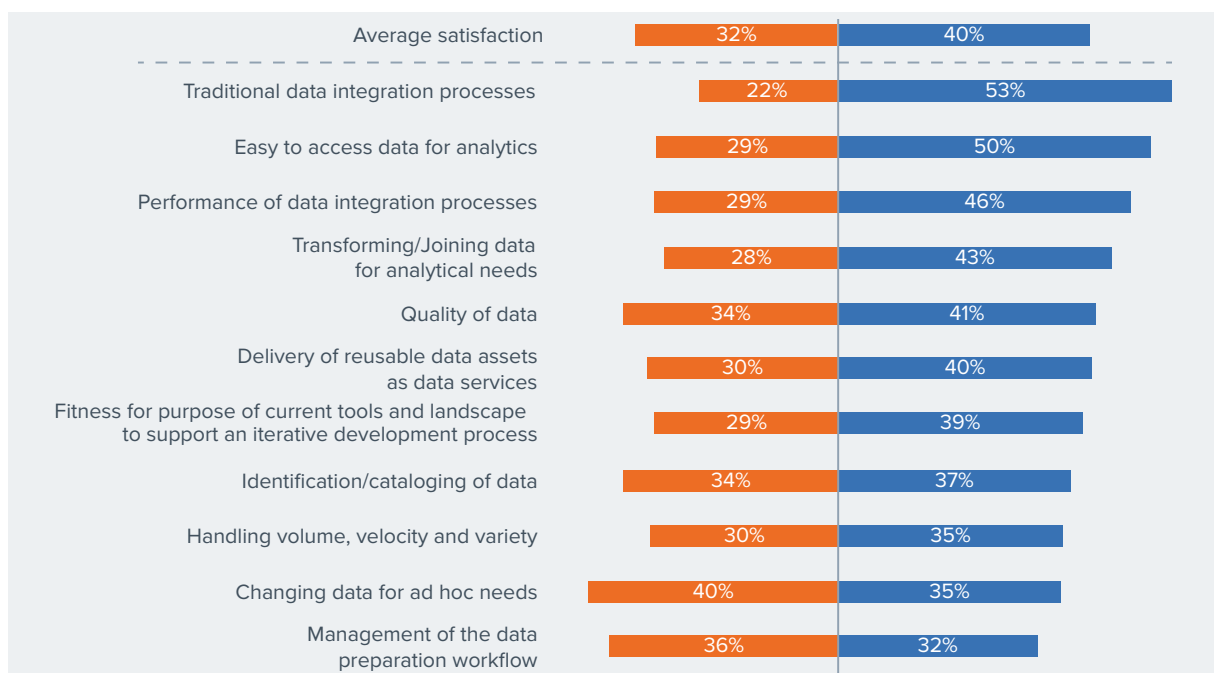


Figure 6: How satisfied are you with the following points regarding the way data is currently prepared for analytics? (n=686)

The desire for traceability and documentation of transformations stands out significantly with 45 percent (dissatisfied) of the responses versus 26 percent (satisfied). Traceability is often not self-evident and requires planning, consistent development as well as adherence to standards and rules, and ultimately also the correct use of tools for the analysis. Data lineage functions help to ensure the traceability of data flows in an automated manner and the provision of up-to-date documentation at all times. Today, this required transparency is often hampered by coded transformations, manual adjustments, and the use of different, poorly integrated tools. Uncoordinated processes, a lack of collaborative approaches, and poor architectures that lead to isolated solutions complete the mess found in many companies. Due to a shortage of time, the relevant documentation is often not created or updated. Transparency regarding the available data, its origin, and any transformations undergone are therefore lost for good or exist only in the minds of individual employees. The trend and necessity for cooperation between business departments and IT, especially when it comes to data preparation, create the potential breeding ground to solve this need jointly. It is clear that, with the increasing organizational distribution of data preparation, tools need to have functions for data lineage and automatic documentation (ideally role-oriented).

Data lineage ensures that data flows and transformations are easily traceable in an automated manner, and always up to date.

With the increasing organizational distribution of data preparation, data lineage and automatic documentation must be present.

Further interesting findings can be found in detailed analyses. For example, when looking at satisfaction according to type of implementation, it is shown that, in the case of anchoring in business departments and the use of adequate tools, the level of satisfaction (54 percent) is on average higher than with other forms of organizational anchoring (see Figure 7). This clearly demonstrates the need to involve business expertise in analytical questions at an early stage (i.e. during data preparation). Uncoordinated implementations, often carried out in spreadsheets, score poorly when it comes to satisfaction since the resulting formation of isolated solutions rarely achieves the goals that have been set.

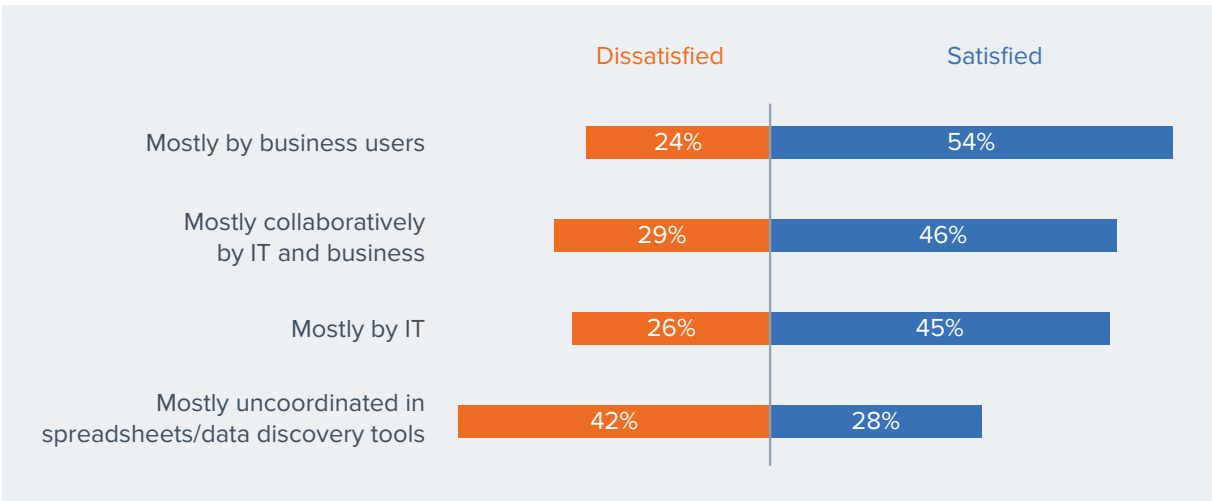


Figure 7: How satisfied are you with the following points regarding the way data is currently prepared for analytics? (n=661)

The key personnel for data preparation are still IT

70 percent of respondents already actively carry out data preparation for analytics. Only 3 percent consider the topic not to be relevant, as Figure 8 shows, and a further 11 percent plan to use or are already experimenting with data preparation. Patterns of use differ significantly between companies. In 25 percent of the companies surveyed, the preparation of data for analysis is undertaken jointly by business departments and IT. At 30 percent, this approach is again used more frequently in North America than the global average. From BARC’s perspective, this „interdisciplinary“ approach is the one which promises the most benefits and best results due to the combination of business know-how and technical excellence.

For a further quarter of the companies, data preparation is purely an IT issue. In this case, the need for technical expertise, usually due to higher technical or regulatory requirements, outweighs the benefits of an implementation in the business department. From market observations, an iterative implementation and incremental improvement of data preparation is difficult to implement in this model and therefore the accuracy of the data sets generated is not always optimal. On the whole, we also see a strong participation of IT in the collaborative approach and its role as an authority for the control and compliance of standards during implementation in business departments. Therefore, even today and despite all the marketing slogans, the technical expertise for implementing and operating data preparation lies mainly in the IT department.

Patterns of use differ significantly between companies.

The „interdisciplinary“ approach promises the best results due to the combination of business know-how and technical excellence.

The technical expertise for implementing and operating data preparation still lies mainly in the IT department.

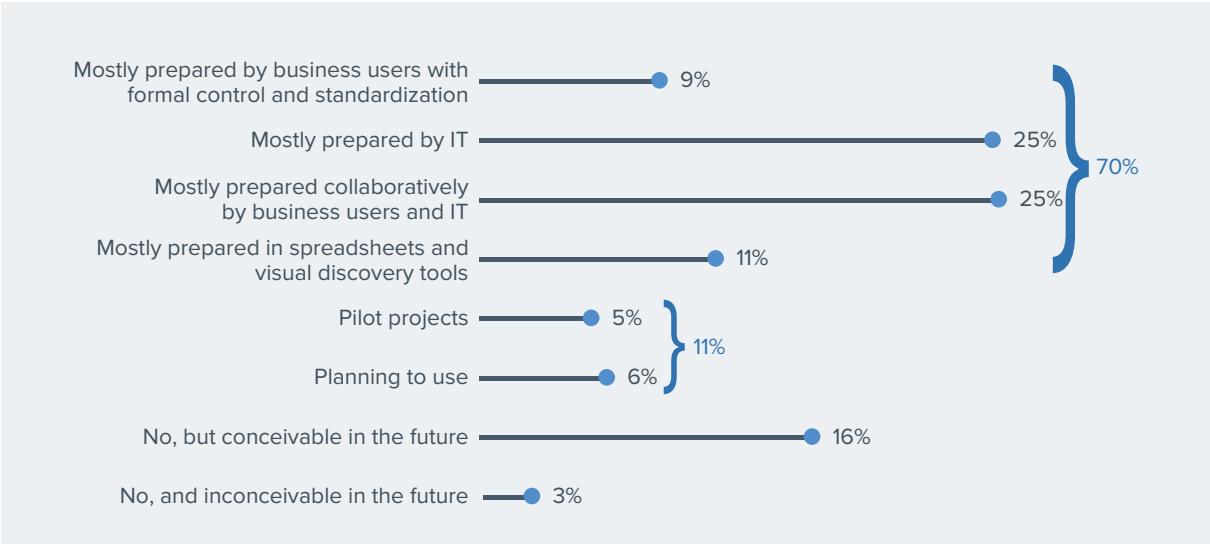


Figure 8: Is your company using data preparation? (n=677)

A look at the survey details, however, puts Figure 8 into perspective. For example, 34 percent of respondents from IT consider the topic to be primarily dealt with by the IT department. Responses from business departments estimate the role of IT in data preparation to be considerably lower at 15 percent.

Business departments and IT are driving data preparation initiatives to a similar degree

In more than half of the companies surveyed, it is the business departments and business users (53 percent) who are driving the implementation of data preparation initiatives. According to Figure 9, a similarly high share (48 percent) of IT is responsible for driving implementation. With 18 percent, the relevance of data science units is surprisingly low considering that data preparation is an essential task for them when it comes to data analysis. The comparatively low relevance can be explained by the fact that data preparation is already part of the repertoire of data analysis outside of advanced analytics and data science. Furthermore, dedicated data science units or labs are not established in all companies.

In more than half of the companies surveyed, it is the business departments and business users who are driving the implementation of data preparation initiatives.

For management, the issue seems to be more operational in nature. Thus, management directly drives this topic in only 25 percent of the companies surveyed and primarily leaves the step of data preparation to fall under the responsibility of business departments or IT. In doing so, in our view, too little consideration is given to the potential of data preparation to achieve strategic objectives, such as generating value from data by, for example, creating measures to improve the customer experience. In this situation, business users and IT often have to convince management of the importance of data preparation to get the appropriate mandates and necessary resources allocated. Analogous to BI, data preparation takes up most of the implementation effort, including in processes for advanced analytics or data discovery.

Business users and IT still have to persuade management of the importance of data preparation.

In Europe, there is more persuasion work to be done since only 22 percent of respondents from this region consider management to be a driver, while in North America, management drives data preparation to a much greater extent (33 percent). A further interesting finding from directly comparing countries is the role of data science units as a driving force. In North America, data science units are cited significantly more frequently (28 percent) as drivers than in Europe (just 15 percent). Dedicated data science units or labs are currently more established and more frequently encountered in the North American region than in Europe.

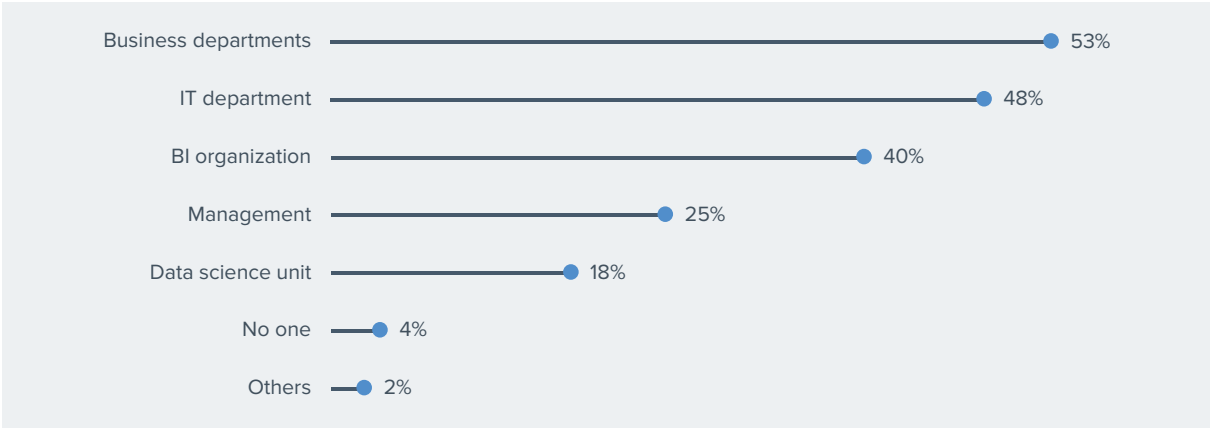


Figure 9: Who are the thought leaders and drivers of data preparation in your company? (n=686)

Even more strongly than in the assessment of usage and whether this is predominantly implemented in the IT department or in business departments, there are differences between self-assessments and external assessments. In BI organizations and IT departments, almost twice as many self-assessments are carried out than external assessments.

Data preparation is predominantly used by experienced BI experts

While data integration was still almost exclusively a topic for IT a few years ago, the need to draw findings from data for business purposes is driving a shift in the skills required for data management tasks from IT to business departments. Special data preparation tasks are increasingly being handled by people who also work on the relevant business issues. Figure 10 shows that 69 percent of respondents consider data preparation tasks in the context of advanced analytics to be carried out by business departments and 64 percent view this as a task for IT or BICC members. It is therefore not possible to clearly ascertain who has responsibility for data preparation. Rather it is managed differently from company to company. This is dependent on the drivers and goals, applications, and the skills and resources available.

Special data preparation tasks are increasingly being handled by people who also work on the relevant business issues.

The orientation towards less technically savvy users is a typical development and a sign of maturity in the life cycle of a technology. This has already been observed with cars which, shortly after being invented, could only be put into operation by specialists. Under the keyword Web 2.0, passive users and consumers have become active co-designers and thus „prosumers“ on the Internet. With the establishment of self-service BI and data discovery, reporting and analysis have also been democratized, following the life cycle from being technical solutions to becoming widely available for the masses.

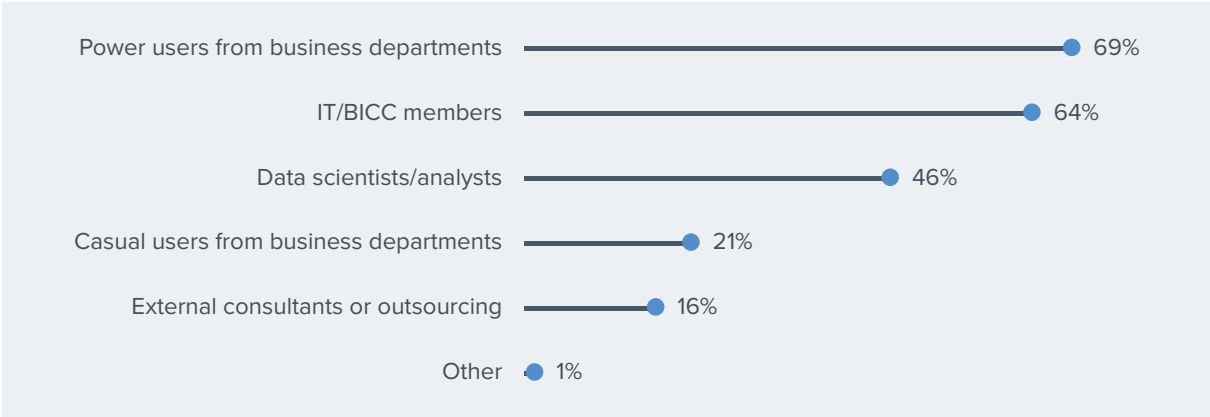


Figure 10: Who in your organization uses or will use data preparation to provide data for data discovery and advanced analytics? (n=540)

Data preparation is used more frequently by existing structures and competences in the form of power users or IT/BICC members than by data science specialists. Exploratory analysis experts are considered to be users in less than half (46 percent) of companies. Thus, according to this BARC survey, the know-how required for data preparation already appears to exist in the companies. We can confirm a high level of expertise in data preparation for BI and simple data discovery applications. We cannot confirm the knowledge about data preparation for more complex analyses by means of data discovery or advanced analytics. Bringing in external consulting power to get things started and for coaching to build an in-house team of experts also often proves to be an effective means when it comes to data preparation, not only to train people in the use of tools but also in proven approaches to implementation. The fact is that in both cases, the close involvement of a company's own business experts is required in order to ultimately generate insights from data.

Business users and IT still have to persuade management of the importance of data preparation.

Productive cooperation is more important than rigid rules

As the trend towards the shared use of data preparation by business users and IT already shows, almost 80 percent of the survey participants believe that this collaboration must be intensive and successful in order to ensure consistency and efficiency in data preparation in the first place (Figure 11). Although data preparation is more frequently taking place in business departments, and this was viewed as an absolute no-go and a risk not too long ago, only one fifth of companies today see this development as a threat to the quality and security of company data.

Intensive collaboration between business and IT is necessary to achieve the required consistency and efficiency of data preparation.

If data preparation is carried out in different areas of a company, IT should assume the role of quality assurance (65 percent). IT does not have to implement all the requirements itself but should be involved throughout in order to enable coordination and harmonization for greater efficiency and quality assurance. The implementation of clear guidelines, structures and responsibilities (63 percent) is a further building block to ensure these two goals are met in an information landscape with a high division of labor.

Today, the view has already taken hold that different rules and methods are required in the context of data governance (62 percent) for heterogeneous applications for work with data – from standard reports to big data discovery. Companies today have already internalized the need for different priorities when it comes to speed and reliability during exploratory work and during the standardized preparation of information by BI factories.

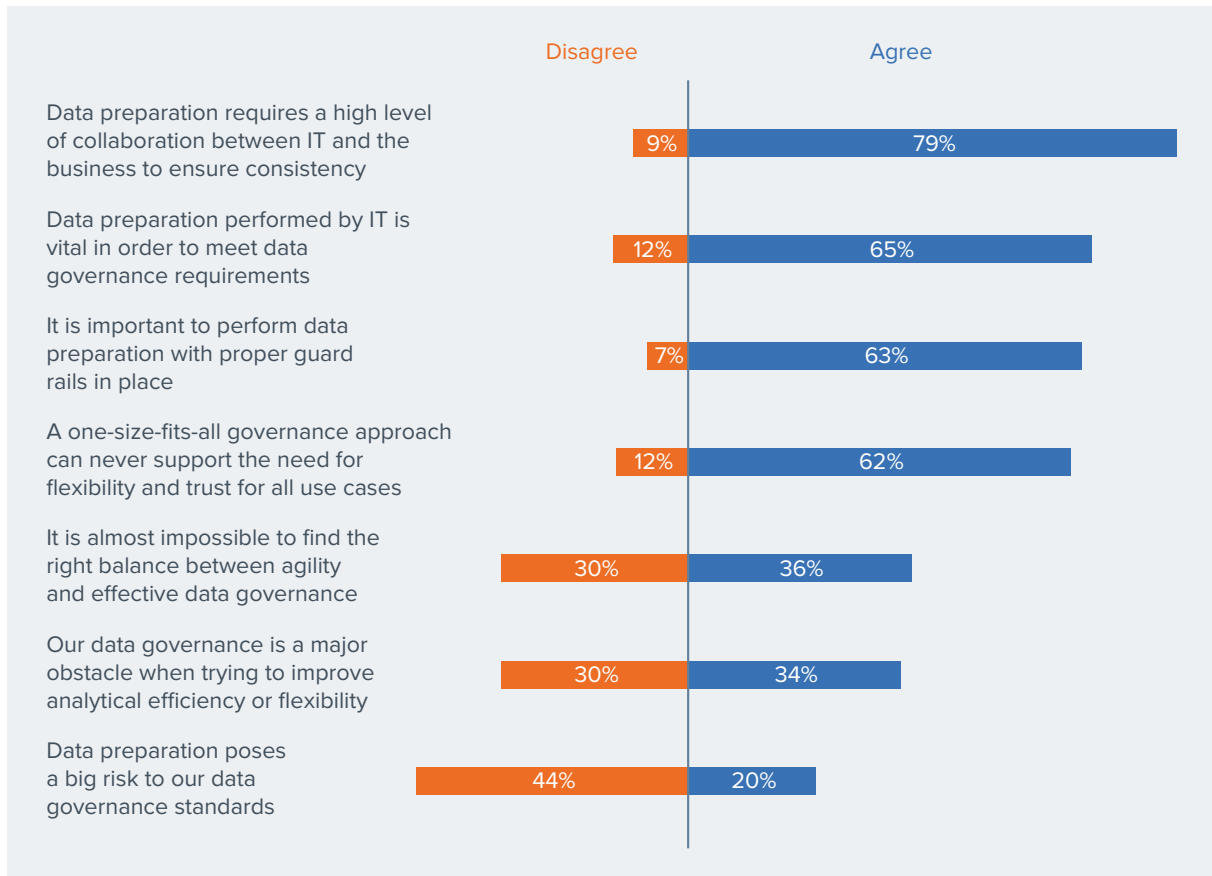


Figure 11: To what extent do you agree with the following statements? (n=671)

One challenge is implementing data governance as an enabler for analytics and not as a bureaucratic overhead that prevents flexibility. For example, the survey participants are still quite divided in this opinion with 30 percent (disagree) versus 36 percent (agree) and 30 percent (disagree) versus 34 percent (agree) respectively. There is still a lack of best practices in this area and there is a need for clarification in order to assess the benefits conclusively. The majority (44 percent) of respondents disagree with the statement that data preparation poses a big risk to standards. This may be due to the fact that many businesses employ data preparation in exploratory or sandbox environments, because use in such secure environments minimizes potential risks.

Data governance is a fundamental pillar for creating value from data

Data governance uses standards and rules to enable efficient value creation from data. The prerequisite for creating value from data and therefore the task of data governance is to establish clear responsibilities, define transparent structures, provide standard definitions, and ensure data quality and data security. Data is valuable capital. Its maintenance and management is becoming a prerequisite

for the digital transformation, regardless of the context in which it is used. It must be possible to use data across departmental boundaries and silos in order to gain new potential from data. Awareness of this need is growing, as the overall importance of typical data governance tasks shows in Figure 12. There is still a long way to go to get there for many companies. For example, the figure also shows that dissatisfaction still prevails in almost all task areas of data governance, or that no satisfactory solutions have been found yet.

A look at the details of the survey reveals that the issue of data quality is seen as critical or very important for data preparation for 70 percent of respondents. Standard definitions (66 percent) and data security (64 percent) are also considered essential in order to use data beneficially through data preparation. The implementation of security policies as one of the tasks of data governance is now a matter of course for 60 percent of companies. Fortunately, the importance of data quality as a prerequisite for data preparation seems to have been recognized, since without quality-assured master data and a common understanding of data and terms, companies will find it difficult to combine and use the multitude of new data in any meaningful manner. However, as the evaluation also shows, it is especially in this area that most of the homework is still to be done. This is something that the market for master data management (MDM) is also noticing.

Data quality is key for data preparation

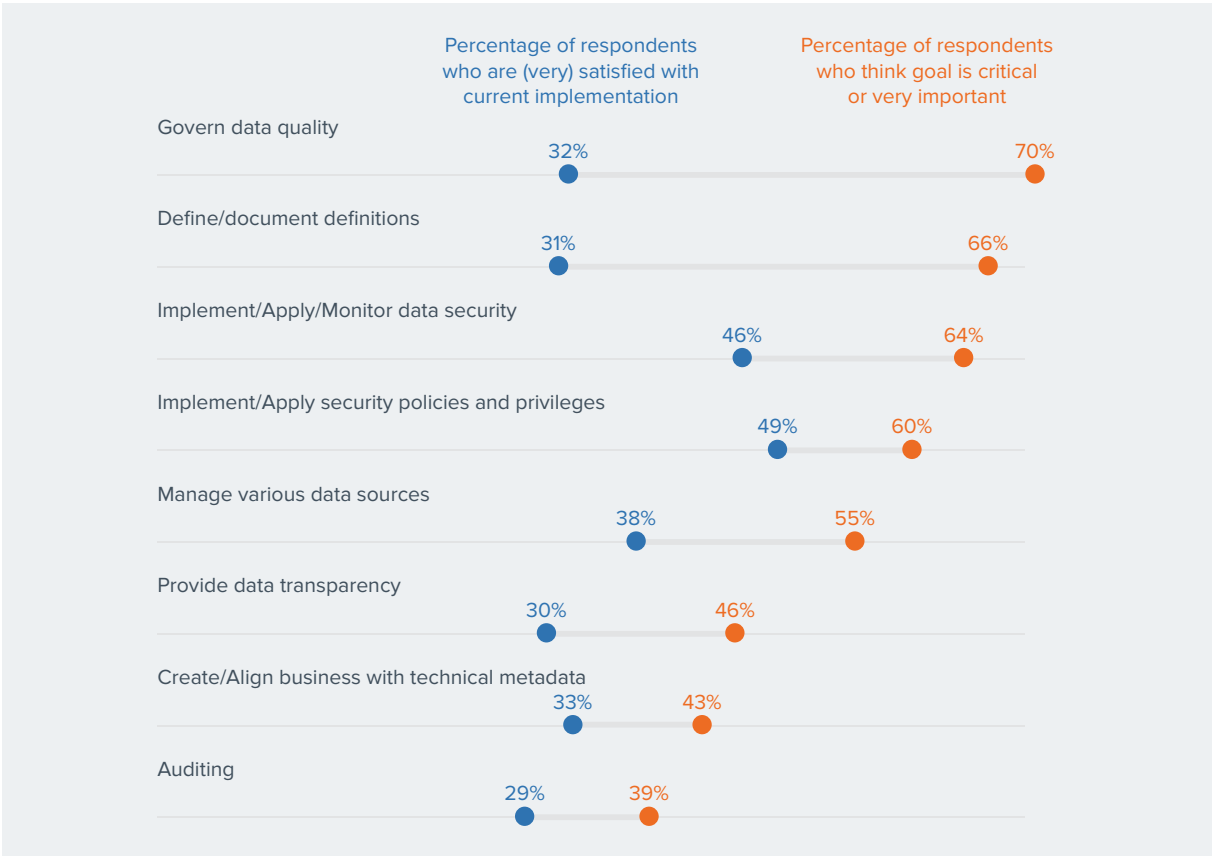


Figure 12: "How important are the following data governance objectives to maintaining overall data quality, consistency and trust in data when using data preparation?" (n= 662) and "How satisfied are you with the current degree of implementation of the following data governance objectives in terms of target achievement in your company?" (n= 662)

There is a great need for action in many companies when it comes to defining key figures in a uniform manner as well as defining the meaning of data generated and used, as only 31 percent of the participants are satisfied with the current state of the solution here. Accordingly, many companies are finding it difficult to define data clearly. For example, how to clearly define who is counted as a customer and how to determine which valid views on the customer exist. Boundaries between individual information silos in the companies are too big. As a result, only limited and isolated valid definitions have developed. Creating a uniform image of data as a prerequisite for creating value from data is the purpose and goal of current projects in these companies.

There is a great need to get a common view on data and its meaning in order to use it in a way that will generate insights.

The lack of a connection between business information and descriptions of the data used in the companies surveyed is particularly noteworthy. For example, only 33 percent of respondents are satisfied with the alignment of business metadata with technical metadata. Data creates value in the right context with the right significance for the business application. Transparency, traceability and the controllability of data are the focus of consideration for business departments and also IT, in order ultimately to be able to operate profitably using the data in terms of value creation. Above all, support for metadata analyses across system boundaries is not possible in many companies today. For example, many companies already define and maintain their business and technical metadata, but only few succeed in making this metadata available in all the various systems found within large, heterogeneous analytical landscapes in order to gain a better understanding.

Data preparation between desire and reality

Despite having proved to be unsuitable, spreadsheet applications are used for data preparation in three quarters of all companies (Figure 13). It is particularly noteworthy that this usage is not only high in business departments, as expected, but also among users from IT and BI organizations. The high use of spreadsheets is an indication that data preparation often deals with simple applications and manageable data volumes. As we shall see, this assessment is also supported by the fact that, in most cases, rather simple functions are used for the preparation (Figure 14). Even in large companies, simple use cases are the rule rather than the exception, since Excel and co are used more frequently there (79 percent) than in small companies (73 percent).

It seems that spreadsheets are used for everything. Most companies are using them for data preparation today.

Data integration platforms or ETL tools are also at the top of the board in terms of usage (63 percent). Sophisticated tools with a large functional range can be used for a variety of applications. Through efforts to make the tools easier to use and more intuitive, providers have also made this category of tool accessible for use by experienced business users, even if many tools still have potential for improvement.

The use of dedicated data preparation tools and data warehouse automation (both 38 percent) are already surprisingly high. The main focus of these tools is to automate the creation and adaptation of ETL processes for the development and maintenance of a data warehouse. Tools for exploratory data analysis, which often also provide good and easy-to-use data preparation functions, are used much less frequently. Data discovery tools are only used in 3 out of 10 companies for data preparation. Comparing the currently already wide distribution of these tools with their use for data preparation, we come to the conclusion that the functions available for data preparation are insufficient for many users, and they rely on other tools, thereby accepting interruptions in the iterative process of data analysis. The equally low use of advanced analytics and data science tools for data preparation (23 percent) reflects the generally lower use of these tools. From the survey participants' point of view, data science and data discovery tools as well as tools for data virtualization and big data integration have high potential, as the figures for planned usage show.

The widespread use of spreadsheets, even compared to traditional data integration, illustrates a clear desire of the users: they no longer want to model data preparation in often highly abstract data flows, but rather apply the rules to the data themselves. A by-product of this change will be a recipe that ensures reproducibility for automation and transparency. For example, the advantages of spreadsheets can be combined with the strengths of data integration tools.

The lack of coverage shows that there is currently great potential for improvements in the tooling landscape for data preparation, because even if many applications can be „just about“ implemented using spreadsheets, this is outweighed by the advantages of professional and specialized systems. For example, the participants in our survey state that they have problems accessing data considerably more often when working in spreadsheets than with all other tools. Companies can thus achieve significant improvements in all areas, as the data on satisfaction and benefits clearly shows.

There is a lot of potential to rethink tooling. There are many professional tools available for data preparation that suit various individual needs and produce much better outcomes than Excel can.

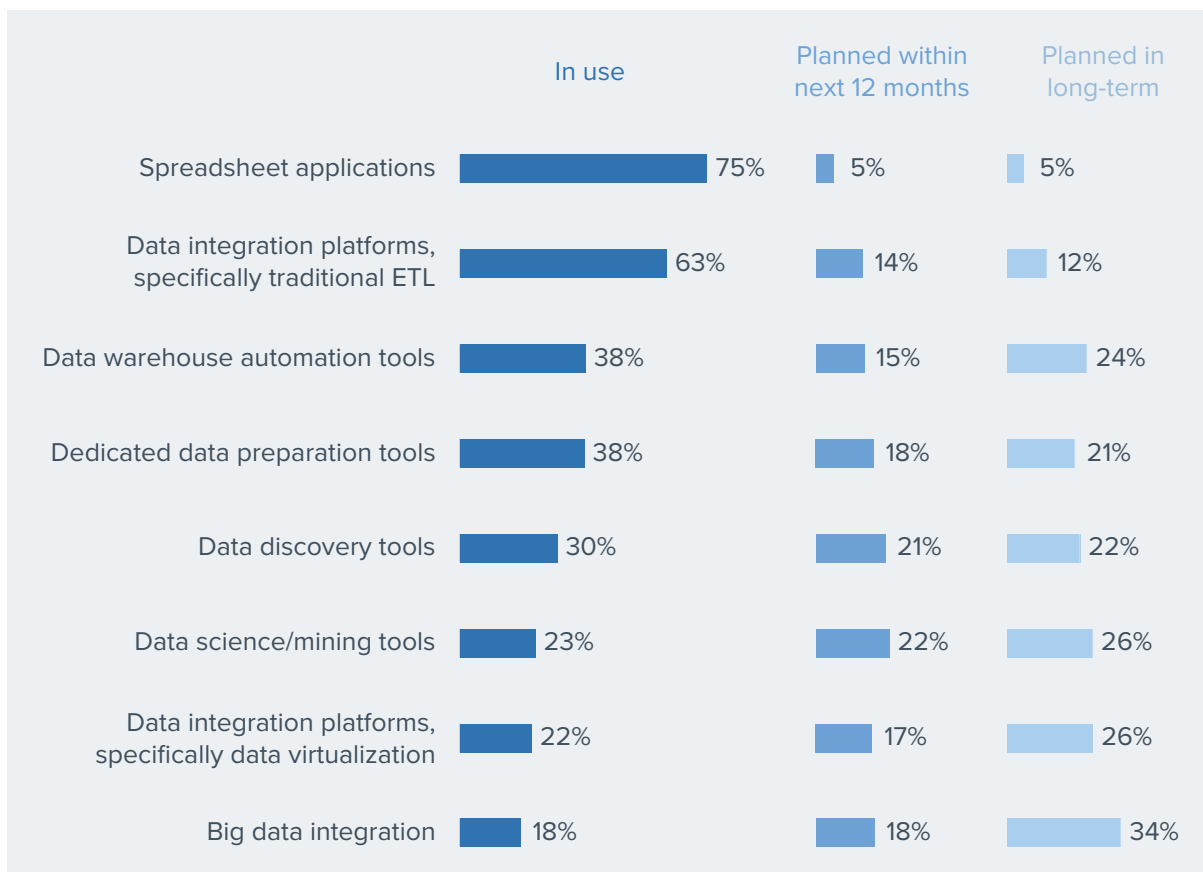


Figure 13: Which tools do you use or plan to use in your company for data preparation? (n=512)

Spreadsheets were yesterday – now it’s time for the right technology

It comes as no great surprise that the use of basic functions is high when it comes to using simple tools based on structured data sources. Calculations for derived key figures, for example, are carried out in 70 percent of companies in the context of data preparation (Figure 14). In the course of data preparation, efforts are also being made in two thirds of companies to carry out cleansing with the aim of improving data quality. With 64 percent, the need to improve the intelligibility and therefore the usability of data by enriching it with semantic information is similarly well represented (modeling). Key figures, hierarchies and attributes, which are intended to facilitate targeted data analysis, are thus also routinely defined during data preparation.

The growing number of available data sources and the distribution of data preparation means that it is essential to catalog the available data in order to use it intensively. However, despite this, not enough companies are taking concrete steps and using suitable functions. Only 42 percent of the companies surveyed use appropriate functions. Cataloging seems to be a matter for BI organizations. While their representatives attest to using it in 53 percent of companies, this figure is below 40 percent for business users and not significantly

higher within IT departments. Spreadsheets barely support this function. In order to better understand and interpret data that has not yet been prepared, profiling has become increasingly established. The fact that 39 percent of companies already use profiling to efficiently analyze the quality, distribution and form of data shows that the bottom-up approach to data preparation is becoming increasingly popular in comparison to the top-down approach frequently used for traditional data integration.

Spreadsheets only support a minimum of functions needed for data preparation.

Data lineage, which is also relevant for transparency, and the use of artificial intelligence are left in the shadows in the area of data preparation for analytics. Artificial intelligence provides the opportunity to assist users in the preparation of data to enable efficient, high-quality preparation. Users are notified of problems with data quality and steps for improvement are proposed. Artificial intelligence can also be used to find suitable candidates for data set joins. This support allows the user to concentrate on the essential aspects of the preparation work.

It is generally shown that data preparation and thus also the tools for data preparation need to deliver more than just data. This data must be made intelligible through the use of semantics. It must be findable through catalogs and enriched by additional data, its security must be ensured, and there must be transparency, in the sense of a lineage, as well as possibilities for joint work on the data preparation. The functions required for this are not available in spreadsheets so it is therefore time to replace them. A wide range of specialized tools is already available to perform these tasks.

Data preparation is not only about integrating data – it's about understanding data and transforming it into an analyzable format.

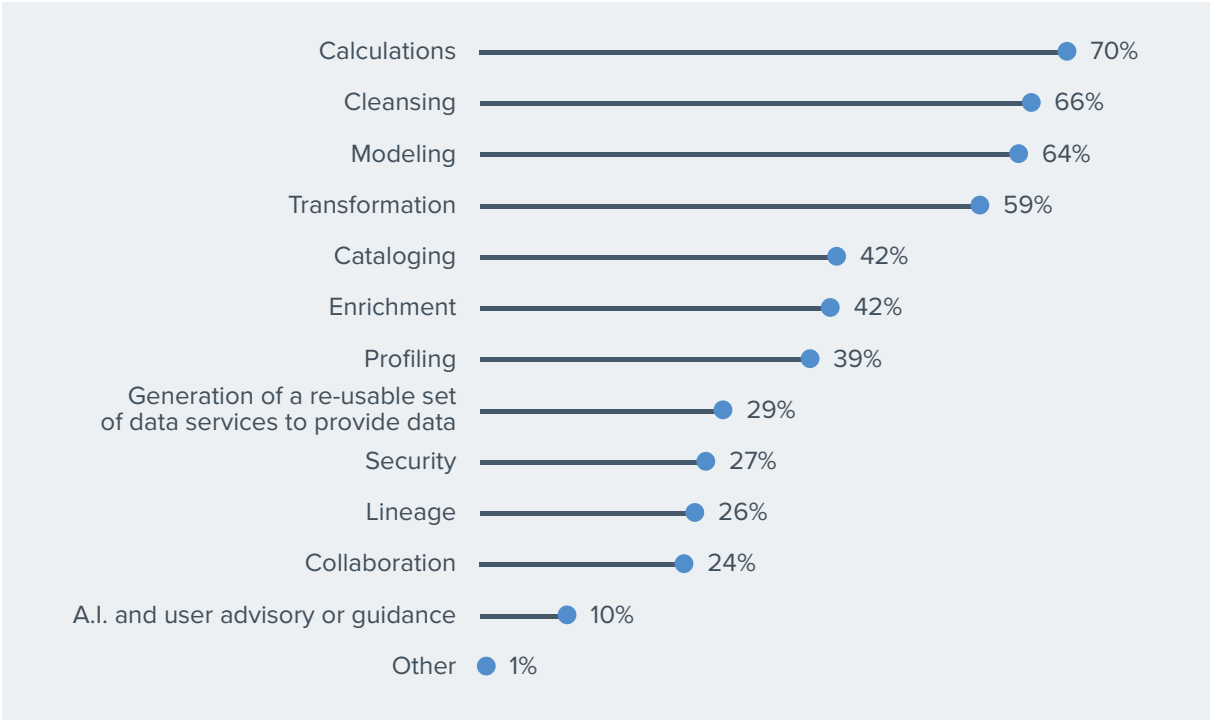


Figure 14: Which of the following data preparation capabilities and functions are used or planned for use in your company? (n=540)

New data sources are only slowly being developed

Another indication that data preparation is mainly used to cover simple use cases is the fact that the predominantly used data sources have always been among the standard sources for reporting and analysis. Data from ERP systems is most frequently prepared for analytics (78 percent, Figure 15).

Due to the ever-increasing exchange of information along value chains, there is a great need for data exchange between companies. This exchange is often carried out via defined interface files – a reason for the intensive use of files as data sources for data preparation in three quarters of companies. The large number of interfaces and the frequently changing data structures mean that the files cannot be integrated in an enterprise data warehouse, or not quickly enough. In order to analyze this data in detail, the preparation has to be organized decentrally and is therefore frequently implemented directly in the business department by means of data preparation.

External data is becoming increasingly important for analyses. Its integration to enrich analyses is a traditional application for data preparation. The fact that 36 percent of the companies surveyed already use external data shows that context-related data preparation can offer high added value because it allows this data to identify drivers for cost or revenue developments, and thus also possible measures, which cannot be detected on the basis of internal company data.

Big data from data lakes, Hadoop, IoT or streaming is currently only used on a manageable scale. With the exception of IoT, data from these sources is used much more frequently for preparing data sets for analytics in North America than in Europe. Due to the high usage, there is great potential in the connection of cloud data sources. At present, these are used in data preparation for analytics in only 18 percent of companies, with usage in North America being more than three times higher than in Europe. Special connectors with simple data access or predefined content can help users quickly and easily tap these sources.

That unstructured data sources or data sources with large amounts of data are less frequently used for data preparation in companies is also due to the fact that efficient access to this data is simply not feasible with the tools predominantly in use (especially spreadsheets). If a company also wants to make the most use of data away from traditional sources, then tools must be used that support this.

There is an increasing need to deal efficiently with data within and between companies. Data is the backbone to enabling value chains.

Working with unsuitable tools prevents the use of potentially valuable data.

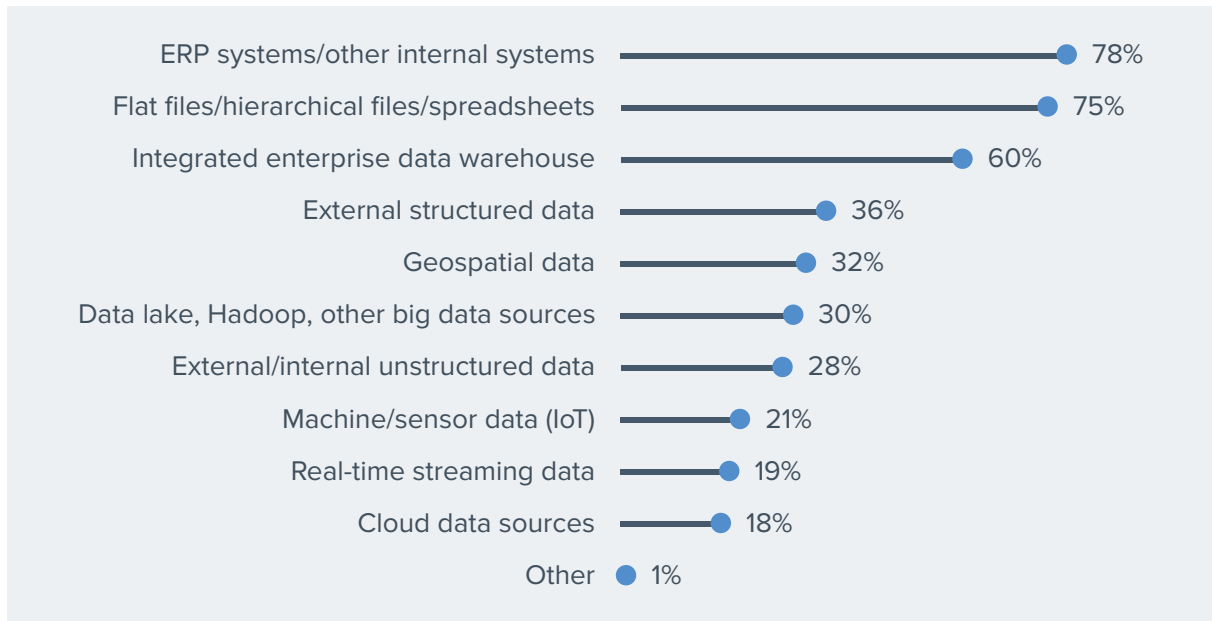


Figure 15: Which of the following sources are/will be used to prepare data for data discovery and advanced analytics? (n=537)

Data preparation does not happen by itself – the right conditions must be created

As the lack of satisfaction with easy access to data for analyses already shows, the situation in companies is not at all bad currently, but there are still significant challenges in accessing different data sources. Here it is evident that improvements in tools and architecture cannot make up for the increasing requirements and challenges caused by the strong rise in data sources (e.g. cloud) and data volumes. It is not enough for a universal democratization of data usage if only 23 percent of companies (Figure 16) see easy access as the norm, while problems and challenges still prevail for 33 percent.

The ability to access data does not just happen. Providing the right data for analysis at the right time is one of the biggest challenges in companies today.

Data preparation is an important piece of the jigsaw on this road to the democratization of data usage in companies. Companies that carry out data preparation systematically have significantly fewer problems accessing relevant data sources. The challenges in accessing data are similarly weighty for companies of all sizes. Here, unlike in other areas, the higher requirements of large companies are balanced by a greater availability of resources.

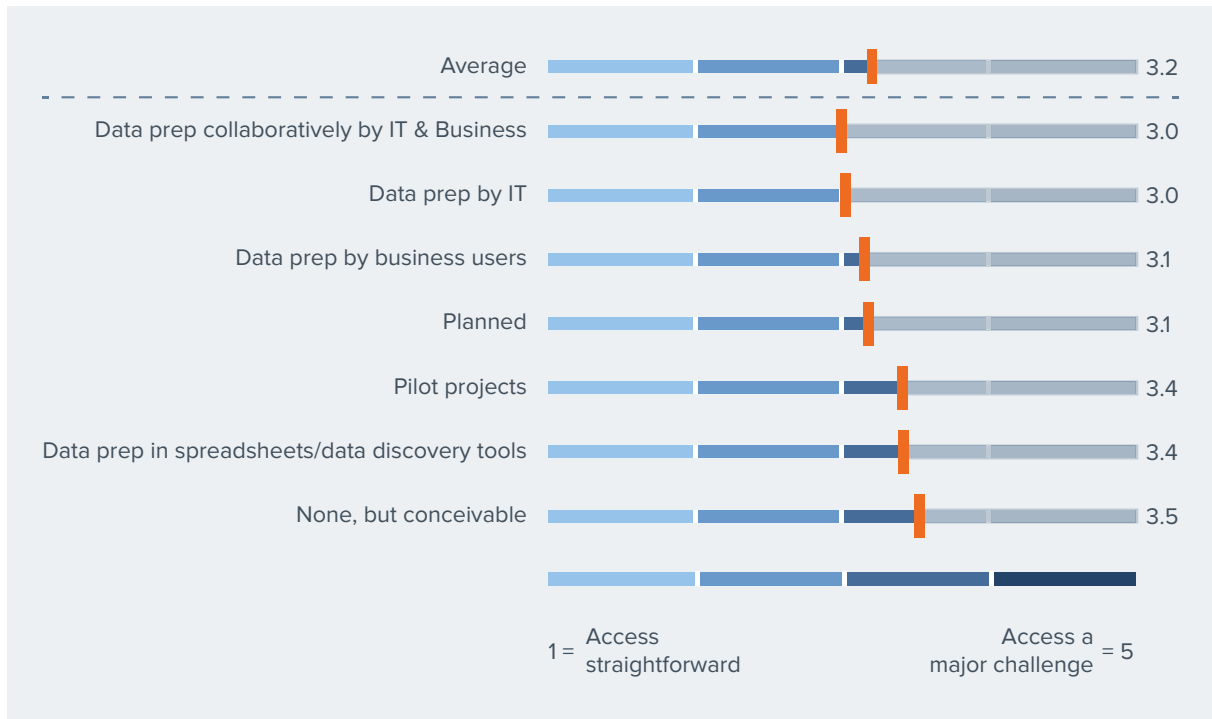


Figure 16: To what degree does your company lack easy access to data in order to perform analysis or explore new areas of analytics? (n=559)

The handling of poor data quality and inconsistencies takes up a lot of time, not only in data warehouse development but also in data preparation. Together with the data preparation (for example, data transformation) itself (51 percent, Figure 17), the handling of inconsistencies and poor data quality (55 percent) clearly heads up the list of time robbers in the analytical process. In addition to these more technical challenges, there are also organizational challenges such as a lack of resources for implementation (38 percent). The participants of this survey attach comparatively little importance to the technical challenge of „translating“ the requirements of business users for the implementers – a strong driver for all self-service initiatives (22 percent). As already shown, a large part of data preparation takes place jointly in IT and business departments, or directly in business departments. The self-service portion of data preparation is thus already high, which minimizes the time required for the necessary translation accordingly. These values are correspondingly lower than average when it comes to implementation in business departments (11 percent).

Poor data quality and inconsistency lead to unnecessary extra effort before data can be analyzed.

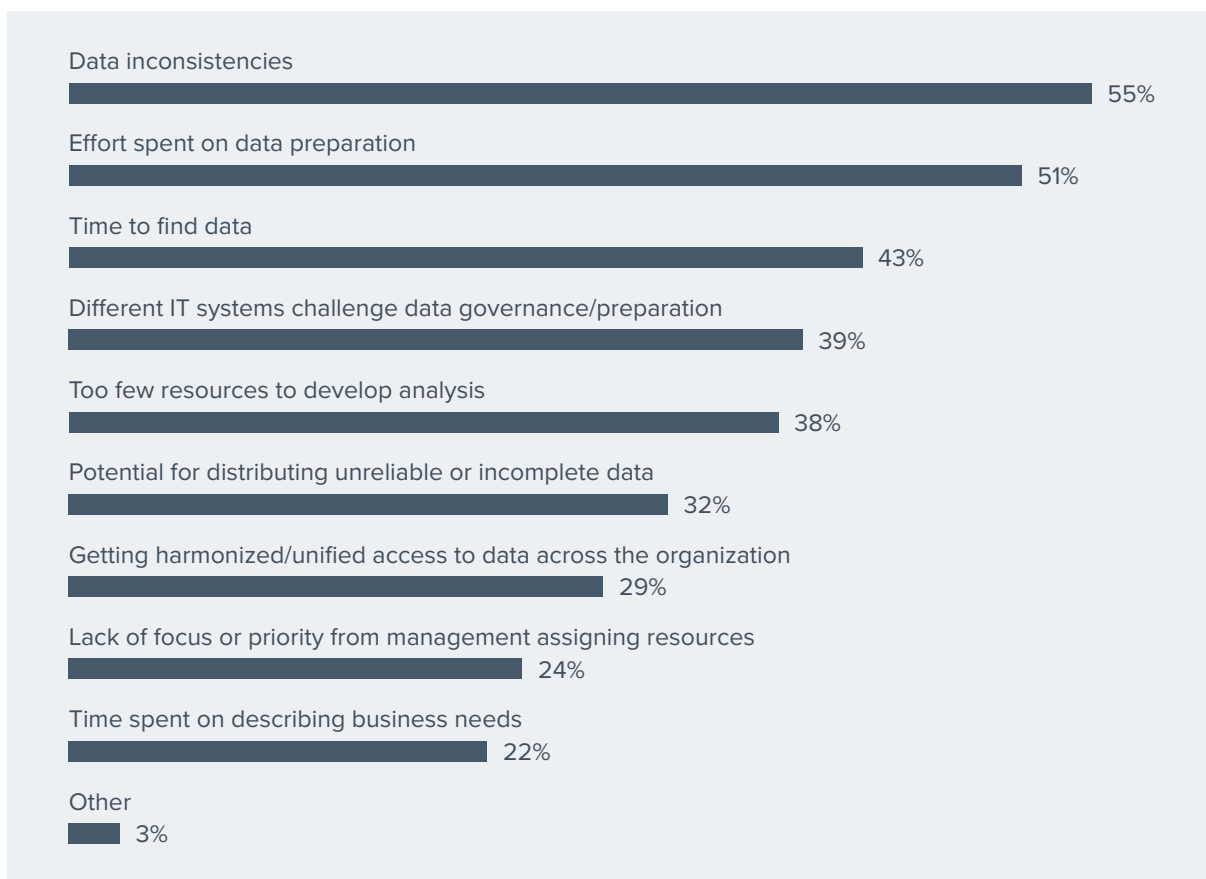


Figure 17: What are your company's key challenges in terms of time spent from idea to final analysis? (n=689)

No one masters data preparation without problems

In the course of this report we have revealed user expectations of data preparation, as well as who is driving initiatives and has the skills to implement them. What challenges do companies face today when it comes to data preparation? One thing is clear: almost all companies are experiencing difficulties, since only four percent indicate that they have a firm grip on data preparation and are not having any problems (Figure 18). This result also illustrates why satisfaction with the current state of implementation is rarely high.

One of the biggest obstacles to the introduction or professionalization of data preparation is the availability of know-how. Over half of the participants (53 percent) in this BARC survey believe that business users lack the necessary knowledge and skills. On the IT side, a third of participants still see significant gaps. Of note here is the large spread of results between Europe and North America. West of the Atlantic, business users are much better acquainted with data preparation, while IT has better conditions for preparing data east of the Atlantic. These differences are also noteworthy as they do not appear to cause any major differences in the organization of data preparation or the drivers, even though the presence (or lack) of know-how should be an important indicator of the organizational anchoring.

A lack of know-how is the top challenge facing companies in their data preparation today. This can seriously hinder their ability to explore the full potential of data.

Lack of budget or resources (a problem for 39 percent) is a direct result of a lack of sponsors (31 percent) and the low level of attention given by decision-makers (see Figure 9). Scarce resources contribute to predominantly simple applications being implemented, while more complex applications are not implemented due to costs, even though they may have potentially greater benefits.

Companies today do not consider the tools available to be among the biggest obstacles to implementing data preparation initiatives. As already mentioned, limitations in terms of know-how and budgets prevent wider use. While there are concerns that data preparation may affect the enforcement of data governance and thus present a risk, on average this figure is low (17 percent) and lower than in North America or in IT.

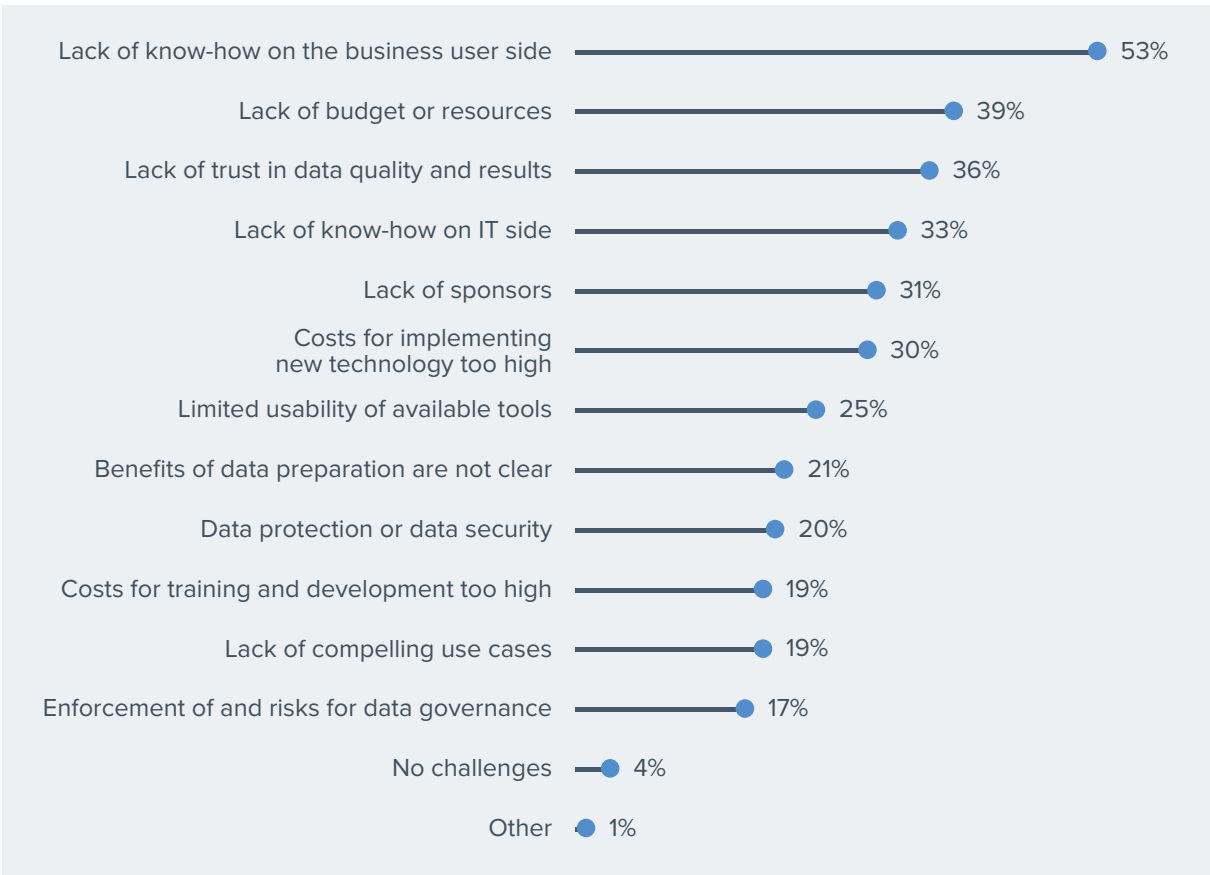


Figure 18: In your company, what problems do you face when using data preparation? (n=647)

Conclusion

The survey confirms that data preparation is not only a relevant discipline for creating value from data, but also an effective means for flexible information management as well as a basis for more comprehensive use of BI and analytics. The wide range of possible applications and the extent of the benefits already achieved are clear to see. But its potential is still not being fully exploited today. We see the implementation of mainly simple applications as the focus of respondents, who have far from exhausted the available functional range of modern solutions. Much is changing and companies are using opportunities identified in discussions about digitalization and big data to align their data management to new requirements. But orientation is missing in many places and the process is therefore rather slow.

Especially in the context of data preparation, the survey shows the direction of change in organization, governance and technology clearly. For example, companies are already aware that it is especially the joint cooperation of business departments and IT that creates the most benefit – the symbiosis of technological standards and decentralized flexibility in an agile environment that makes work with data and analyses possible. In particular, the use of „all relevant“ company data places direct requirements on data and thus on a set of rules to create structures and ensure data quality, definitions and data security. A major challenge for half of the companies surveyed is to define the right data governance principles, achieve the goals mentioned above, and combine them with the desired flexibility. From a technological standpoint, the survey shows that a revolution in the use of tools has begun. Spreadsheet applications do not meet current needs when it comes to data preparation. Users‘ needs are pushing them out of their comfort zone, thus revealing weaknesses. Development is moving away from spreadsheets for data preparation towards modern technologies with a significantly greater scope of services in order to improve the value added from data through data preparation and, in many cases, to enable it in the first place. The cause is recognized, but it is still difficult for many companies to develop further. The lack of know-how in business departments, lack of budgets and resources, as well as inadequate data quality are cited as particular hurdles.

Anyone who wants to create value from data has to do something to make this happen! The combination of bottom-up and top-down methodology has proven to be effective.



More than ever before, companies‘ data management objectives need to be specified and goals, structures, responsibilities and actions need to be derived therefrom. A strategy for analytics helps to translate business objectives into goals for analytics, to determine guidelines for orientation, to understand and meet business requirements, and to

create the prerequisites for a successful operation. Define a strategy to bring together different interests and to use data effectively for the benefit of the entire company.



Working together with data places higher requirements on data quality, common definitions and standards. Therefore, it is essential to create a framework for the use and analysis of data across business departments. The sphere of governance must be defined and balanced with requirements in terms of flexibility. This should start by institutionalizing responsibility for data. An authority is needed which, for example, is responsible for the customer domain and enhances this data asset for the benefit of the company, irrespective of the specific context in which it is used.



Strengthen the collaboration between business and technical experts for the work with data. Drive collaboration and create a contact point, for example through a competence center or communities of practice for analytics. The task of the competence center should also be to develop analytical awareness, acceptance and analytical skills within the company.



Encourage business departments, collect experiences directly and build knowledge and skills through data preparation directly in business departments.



Parallel to the strategic top-down approach, bottom-up experiences must be gathered. This helps to produce benefits and to gain attention and resources from management.



Test modern technologies for data preparation. Beyond spreadsheets, there are powerful and functionally more far-reaching tools for data preparation, and not only in one tool category. Based on your business, functional, technical and organizational requirements, first determine the tool category. Then evaluate tools from the appropriate category in relation to specific tasks.



Data discovery and advanced analytics (including data preparation) require new skills and methods compared to traditional BI tasks. Use available know-how on the market efficiently to learn internally from best practices and experiences and build up your own skills.

Attachment A: Methodology and Demographics

The online user survey was conducted worldwide in March and April 2017. BARC promoted this survey through Web sites, at events and in email newsletters. A total of 695 people participated in the survey. Most participants came from Europe (71 percent), with a further 18 percent coming from North America and 7 percent from Asia and Pacific (see Figure 19).

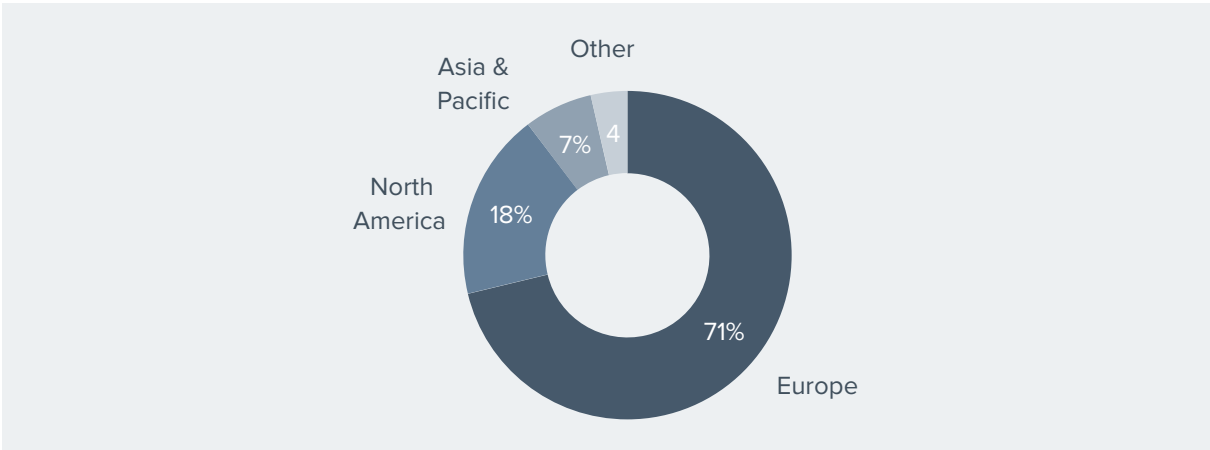


Figure 19: Respondents by region (n=695)

A closer look at the professional backgrounds of the respondents reveals an almost even split between IT (41 percent) and business departments (46 percent) (see Figure 20). 13 percent of respondents also belonged to a multi-departmental BI organization.

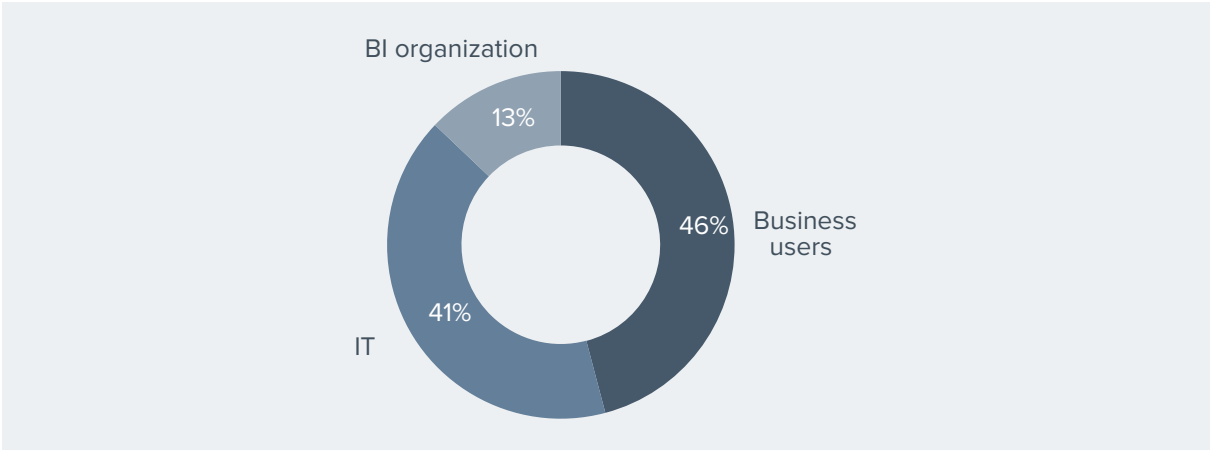


Figure 20: Respondents by department (n=695)

Respondents came from a wide range of industries (see Figure 21), most notably manufacturing (18 percent), IT (15 percent), financial services (12 percent) and retail (12 percent).

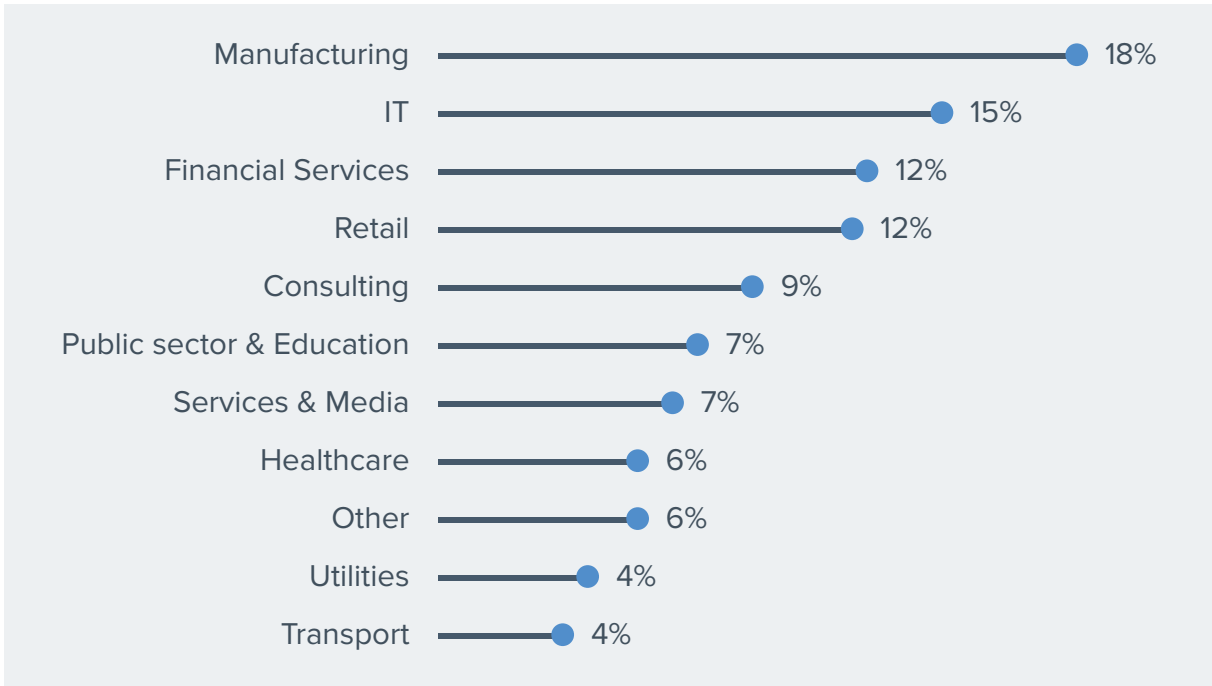


Figure 21: Respondents by industry (n=695)

Figure 22 shows respondents' company sizes by the number of employees. Organizations with more than 2,500 employees were the best represented in this study. However, companies of other sizes were also well represented with 25 percent having less than 250 employees, and 34 percent with 250 to 2,500 employees.

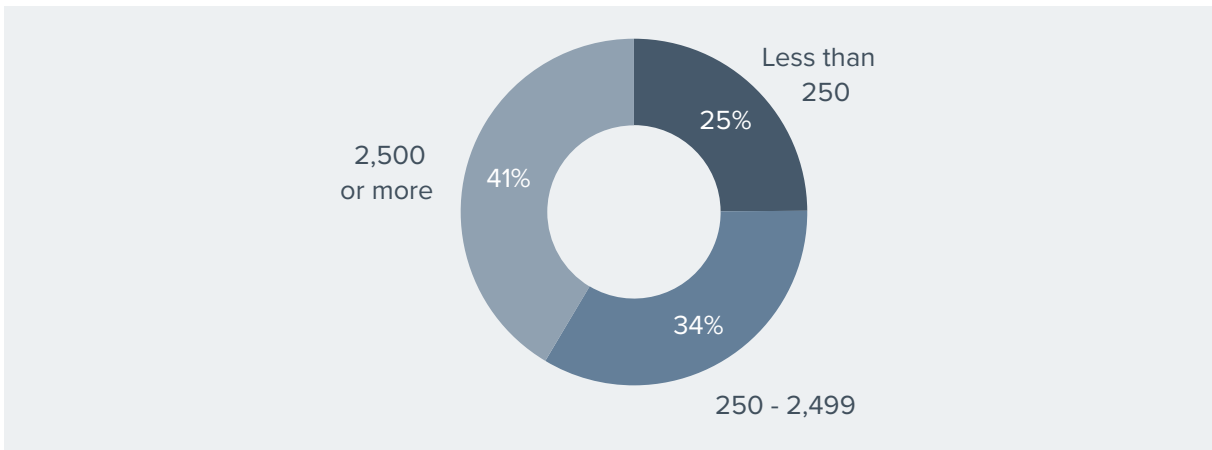


Figure 22: Respondents by company size (n=695)

About the authors

Timm Grosser has been working in the BI and data management area for more than 10 years. He works as a Senior Analyst and Consultant at the Business Application Research Center (BARC) focusing on BI, data management and big data. His core competencies are decision supporting information systems with a special focus on strategic topics in data management and big data. During his time as a consultant, he designed numerous solutions in BI/big data strategy, organization, architecture and tool selection with customers or in the BARC test lab. He is a frequent speaker at conferences and seminars as well as the author of numerous industry articles and market studies.



Contact Details

Timm Grosser

Senior Analyst

Data Management

tgrosser@barc.de

Robert Tischler is Managing Director and Senior Analyst at the Business Application Research Center (BARC) Austria. His specialty areas are Business Intelligence Front Ends and process oriented Business Intelligence applications. He supports companies in the software selection process, during the introductory phase and with reviews to optimize architecture and organization. Robert Tischler has many years of experience as a project manager of BI projects, hands-on expertise with many BI tools on the front end, as well as in the backend, designing, coaching and implementing reporting, analysis and planning solutions with a focus on self-service BI, Information Design and advanced planning.



Contact Details

Robert Tischler

Senior Analyst

Business Intelligence

rtischler@barc.at

BARC – a CXP Group Company

BARC is a leading enterprise software industry analyst and consulting firm. Major companies, government agencies and financial institutions rely on BARC's expertise in software selection, consulting and IT strategy projects. BARC has specialized in core research areas including Data Management (DM), Business Intelligence (BI), Customer Relationship Management (CRM) and Enterprise Content Management (ECM).

Along with CXP and Pierre Audoin Consultants (PAC), BARC forms part of the CXP Group – the #1 independent European research and advisory firm in the field of digital, software and IT services.

CXP Group supports its customers in their decision making thanks to a unique 360° approach encompassing assessment capabilities covering global and local IT markets, products & technologies, services, and their business usage.

Every year the group supports more than 1,500 large and medium-sized business-user organizations in their sourcing decisions and investment projects, as well as their choice and optimization of software solutions, and offers consultancy services in many areas (BI, HR, ECM, IT management, finance, ERP, CRM, etc.).

Capitalizing on 40 years of experience, and with 140 seasoned professionals present in 8 countries, CXP Group supports more than 300 software vendors and IT service providers with market research products, strategic analysis, sales support & marketing services.

CXP Group has 3 subsidiaries: Le CXP, BARC (Business Application Research Center) and PAC (Pierre Audoin Consultants).



BARC · | e CXP · PAC

Germany

BARC GmbH

Berliner Platz 7

D-97080 Würzburg

+49 931 880651-0

www.barc.de

Austria

BARC GmbH

Goldschlagstraße 172 /

Stiege 4 / 2.OG

A-1140 Wien

+43 1890 1203-451

Switzerland

BARC Schweiz GmbH

Täferstraße 22a

CH-5405 Baden-Dättwil

+41 (0) 56 470 94 34

Rest of the World

+44 1536 772 451

www.barc-research.com

Denodo

www.denodo.com



Company profile

Organizations have long used data as a strategic tool to better run their businesses. But accessing business-critical information in a timely fashion is not easy, because relevant data resides in disparate and often siloed systems spread across the enterprise. Business users waste time piecing together the data from these different systems and then correlating them, while IT teams try to automate the process using legacy data integration techniques like extract-transform-load (ETL), which are time-consuming and expensive.

Data virtualization is the more agile data integration style that presents the relevant, interrelated data, in real-time and in a consistent format, irrespective of the underlying database systems, structures, and storage. Since data virtualization doesn't replicate or store data, it delivers the complete, business-critical information at a fraction of the cost and time. Business users benefit from faster access to data, and IT teams continue to deliver data undisrupted without worrying about changes to the underlying systems.

Denodo is the leader in data virtualization providing agile, high performance data integration, data abstraction, and real-time data services across the broadest range of enterprise, cloud, big data, and unstructured data sources at half the cost of traditional approaches. Denodo's customers across every major industry have gained significant business agility and ROI by enabling faster and easier access to unified business information for agile BI, big data analytics, Web and cloud integration, single-view applications, and enterprise data services. Denodo is well-funded, profitable, and privately held. For more information, visit www.denodo.com or call +49 (0) 89 599 904 50.

Denodo DACH

Karlstraße 10

80333 Munich

Germany

+49 (0) 89 599 904 50

www.denodo.com

SAS

www.sas.com



Company profile

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 83,000 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 more than 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 compliance. 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 managed 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. Heidelberg has been the home of SAS' German headquarters since 1982. Customers in the DACH region include Allianz, Continental, Commerzbank, HUK Coburg, Boehringer Ingelheim, Fraport, Nestlé, DER Touristik, Migros oder ERSTE Bank Österreich.

SAS Institute GmbH

In der Neckarhelle 162

69118 Heidelberg

Germany

+49 (0) 6221 415 - 0

www.sas.com

Tableau

www.tableau.com



Company profile

Tableau (NYSE: DATA) helps people transform data into actionable insights that make an impact. Easily connect to data stored anywhere, in any format. Quickly perform ad hoc analyses that reveal hidden opportunities. Drag and drop to create interactive dashboards with advanced visual analytics. Then share across your organization and empower teammates to explore their perspective on data. From global enterprises to early-stage startups and small businesses, people everywhere use Tableau's analytics platform to see and understand their data.

Tableau

An der Welle 4

60322 Frankfurt am Main

Germany

+49 (0)69 58997 6700

www.tableau.com

Tableau Desktop

Get results fast with intuitive visual analytics from Tableau Desktop. Connect to any data in just a few clicks. Leave chart builders behind—make discoveries with live visualizations and interactive dashboards. Apply powerful analytics, from forecasting to regressions, that answer deeper questions. Quickly spot trends and outliers to reveal everyday opportunities and eureka moments alike. Tableau Desktop gives you results that matter with analytics that work the way you think.

Tableau Server

Give your business the freedom to explore data and discover opportunities in a trusted environment. Share insights and collaborate with data sources, interactive dashboards, and ad hoc analyses. Ensure the security of your data with fine-grained control of user and content-level permissions. Choose how to deploy—on-premises or in the cloud—and scale up as your business grows. Tableau Server is true enterprise-scale analytics your business will love, made easy to deploy, manage, and scale.

Tableau Online

Share insights across your organization with a fully-hosted analytics solution. Start publishing data sources and dashboards in seconds, empowering everyone in your business with access to interactive visual analytics. With Tableau Online, your analytics are hosted in the cloud. Anyone—from coworkers to clients and partners—can collaborate with data from a browser or mobile device. Say goodbye to VPNs, software upgrades, and capacity limits.

Tableau Public

Tableau Public is a free platform that lets anyone create, publish, and share interactive visualizations online. Drag and drop to explore your data and create richly interactive data stories. Easily design custom dashboards for desktops, phones, or tablets. Share your work on social media in just a few clicks or embed dashboards on your site or blog, no programming required. Data in. Brilliance out.

TimeXtender

www.timextender.com

TIMEXTENDER

Company profile

At TimeXtender, we believe that business leaders should never waste precious time waiting to access trustworthy data, or that talented IT professionals should ever waste their time on tedious tasks that can readily be automated. That's why we've developed Discovery Hub®. We believe in leveraging the power of automation so that the human mind can be focused on innovating rather than being wasted on repetitive and redundant tasks.

With a successful ten-year track record and more than 2,700 customers in numerous industries across 60 countries, we are the world's leading provider of Data Warehouse Automation for the Microsoft® SQL Server®.

Our DNA and culture

Time is precious. With Discovery Hub®, you can achieve clarity sooner, make better decisions faster and spend your time more wisely. At TimeXtender, we believe in spending time on mindfulness in order to build a balanced organisation with balanced people. This approach empowers our organisation to make the right decisions at the right time and we hope to inspire others to do the same. Because time matters, we engage fully charged. When we are fully charged, we get more done, have better interactions and enjoy greater levels of success.

Discovery Hub® – governed self-service Business Intelligence

Our Discovery Hub® is an innovative approach to the IT architecture behind Business Intelligence (BI). It brings together the best of both worlds: self-service BI and data governance. Discovery Hub® leverages the power of automation, enabling business leaders to rapidly explore and interpret trustworthy data while at the same time freeing IT and BI specialists from tedious and redundant tasks.

You retain the freedom to select any visualisation or reporting tool, whether it is Power BI, Qlik, Tableau, Excel or any other.

By combining the power of automation with a metadata-driven approach, Discovery Hub® gives you results five times faster than traditional data warehouse methods. It also delivers a 70% reduction in build costs and a 60-80% reduction in maintenance costs.

As it sits on top of the SQL Server stack with full visibility, both on-prem and in the cloud, it is readily customisable.

GDPR compliance and documentation

Discovery Hub® is the ideal solution for supporting GDPR compliance of your BI and Reporting & Analytics architecture, enabling you to automatically document:

- The data you hold
- Where you hold it
- What it is used for
- Who has access to it

TimeXtender

Bredskifte Allé 13

DK-8210 Aarhus V

Denmark

+45 8620 5700

411 108th Avenue NE

Bellevue

WA 98004

+1 877 572 3477

Herriotstraße 1

60528 Frankfurt

Germany

+49 (0) 69 6773 3148

info@timextender.com

www.timextender.com



BARC · Ie CXP · PAC