


March 2022

# 2022 State of External Data



# Table of Contents

2	Executive overview
3	Intro and key findings
4	Who we surveyed
5	Types of external data purchased
6	Number of data providers sourced
7	External data procurement strategy
8	External data procurement challenges
9	Inhibitors using external data
10	External data and analytics
11	External data and business initiatives
12	External data and getting the most value
13	Annual spend on external data
14	Job responsibilities
15	About Explorium

---

## Executive overview

The external data collection process can be extremely resource-intensive with no guarantee of the data matching the requirements of the use case. Yet, the right data may be crucial to ensuring the accuracy of analytics and predictive models.

How are organizations acquiring and using the right external data?

In our most recent survey, we asked respondents to report:

- What type of external (3rd-party) data does their organization acquire, and from where?
- What is challenging about sourcing external data?
- What business and analytics initiatives do they use external data for?
- What do they need in order to get the most value from the 3rd party/external data they acquire?

---

# Introduction

According to McKinsey, “A well-structured plan for using external data can provide a competitive edge.” (McKinsey, “Harnessing the power of external data” <https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/harnessing-the-power-of-external-data> )

Despite understanding the value, many organizations still struggle to effectively use external data in their analytics and ML programs. This survey paints a clear picture. To get the promised business value and remain competitive, organizations need an external data procurement strategy. Organizations also need to look to new technology solutions that simplify the process of accessing and using external data within their organizations.

At Explorium, we always look to stay on top of the latest trends and insights in the data community. To further that effort, we commissioned Gatepoint Research to speak to data stakeholders inside US companies of more than 100 employees, in industries including Professional Services, Retail, Finance, Healthcare, Utilities and more. We aimed to discover how organizations use external data today. We also wanted to understand the challenges organizations face when building and maintaining their data strategy, and what processes and skills they have in place to find, ingest, onboard, and integrate external data.

---

# Key findings



## Organizations source data from multiple providers

Almost half of respondents said they were purchasing data from 5 or more paid or public sources. Depending on the source this may represent not only significant budget costs (22% of respondents said they were spending over \$500K on external data) but also associated overhead costs. Procurement process, legal negotiation of terms and conditions, and IT security teams maintaining and monitoring usage of the data to ensure compliance - all add time and costs.



## A formal strategy is needed

Only 40% of respondents indicated they had a formal strategy in place to find and integrate external data. The rest outsource the strategy, follow an ad-hoc process, or have general guidelines but no formal guidance. A successful data and analytics strategy for organizations means turning data into actionable insights that drive decision-making within the organization. Forward-thinking organizations understand that having the right data to fit their analytics is vital for success and this includes both internal and external data.

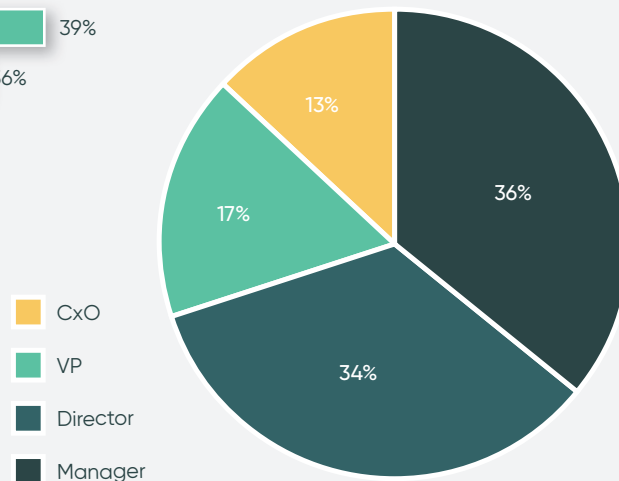
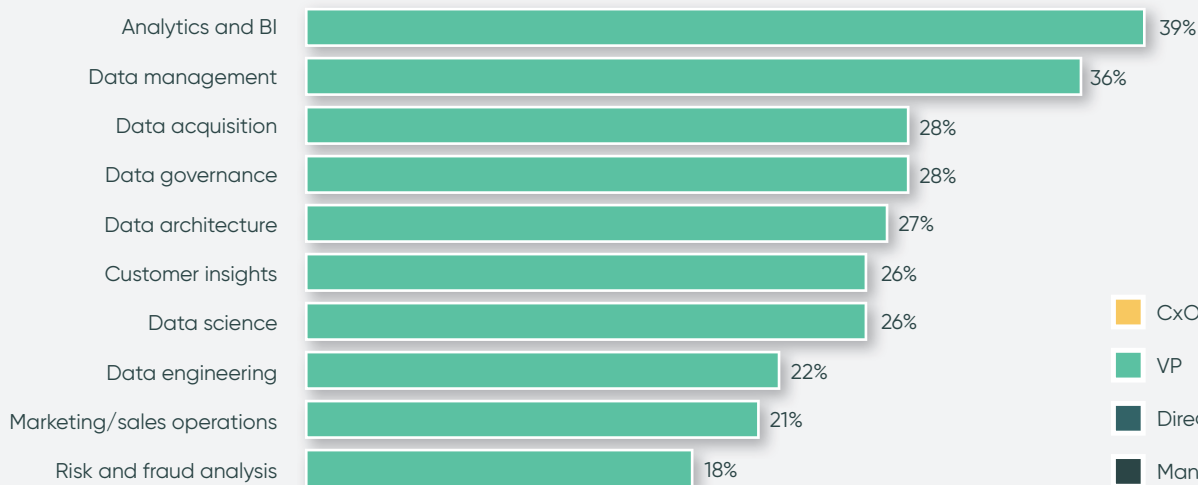


## Look for technology solutions to overcome challenges

The number one challenge to using external data by far was difficult data preparation and integration, which has been a traditional barrier to using external data. Matching and integrating external data with internal data is difficult, especially when combining multiple sources together. But new technologies, powered by AI and ML capabilities, are increasingly tackling this challenge. Today, external data platforms do much of the heavy lifting when it comes to entity resolution and integration of external with internal data.

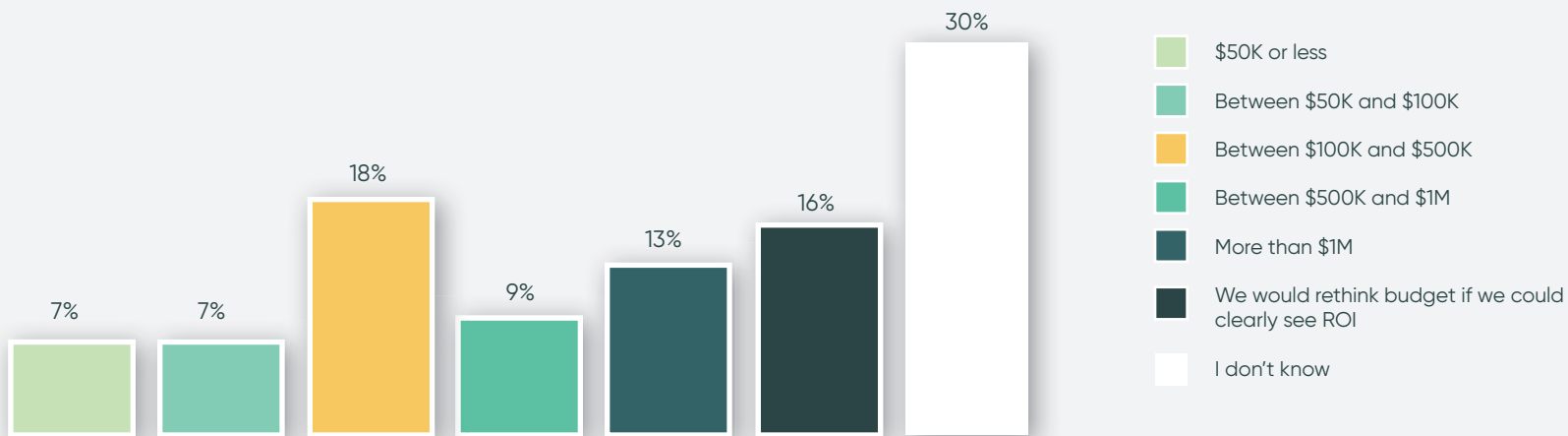
## Who we surveyed

We surveyed data leaders from companies across the United States in various industries how they're using external data and some of the challenges they face. 64% of the respondents hold director or executive level positions in their organizations. Over a third of those surveyed work in analytics and BI (39%) and/or data management (36%).



## Annual spend on external data

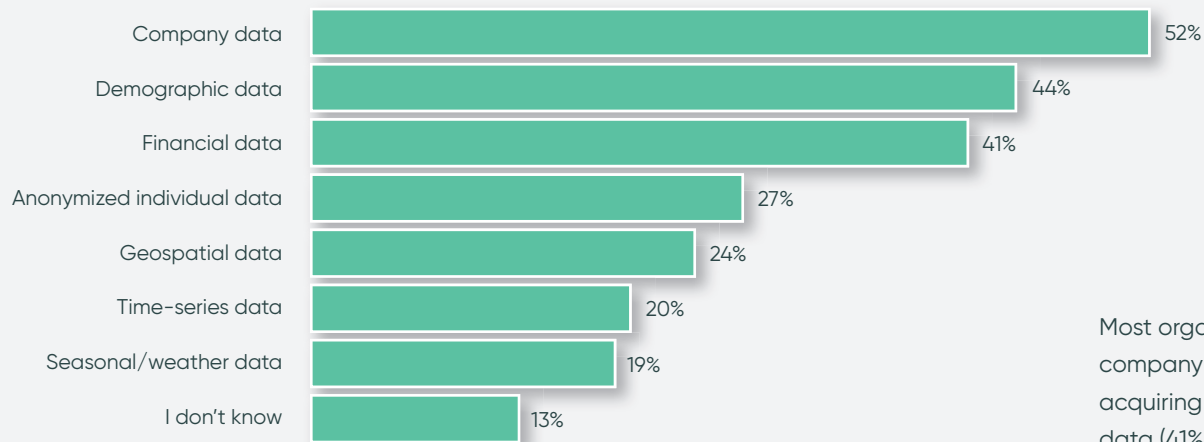
One thing was clear in the survey, organizations are buying external data, in some cases at considerable expense. 22% of respondents said they were spending over \$500K on external data, with 13% saying they spent over \$1M.



Respondents' external data annual spend budgets spike between \$100k and \$500k (18%) and again at more than \$1M (16%).

## Types of external data purchased

When we looked at what types of external data businesses were purchasing, it was a wide variety. Data on companies was the number one response (52%) followed by demographic (44%) and financial data (41%). The high number for the categories suggests that most companies are buying multiple types of data and this correlates with the number of providers organizations are purchasing data from.

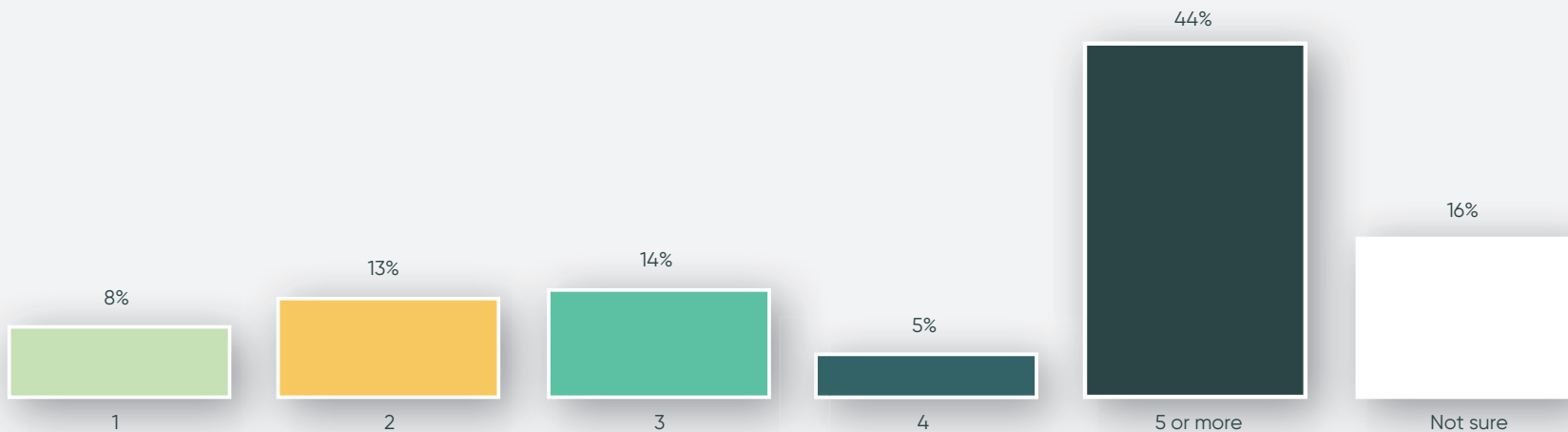


Most organizations surveyed acquire company data (52%). Also common is acquiring demographic (44%) and financial data (41%).



## Number of data providers sourced

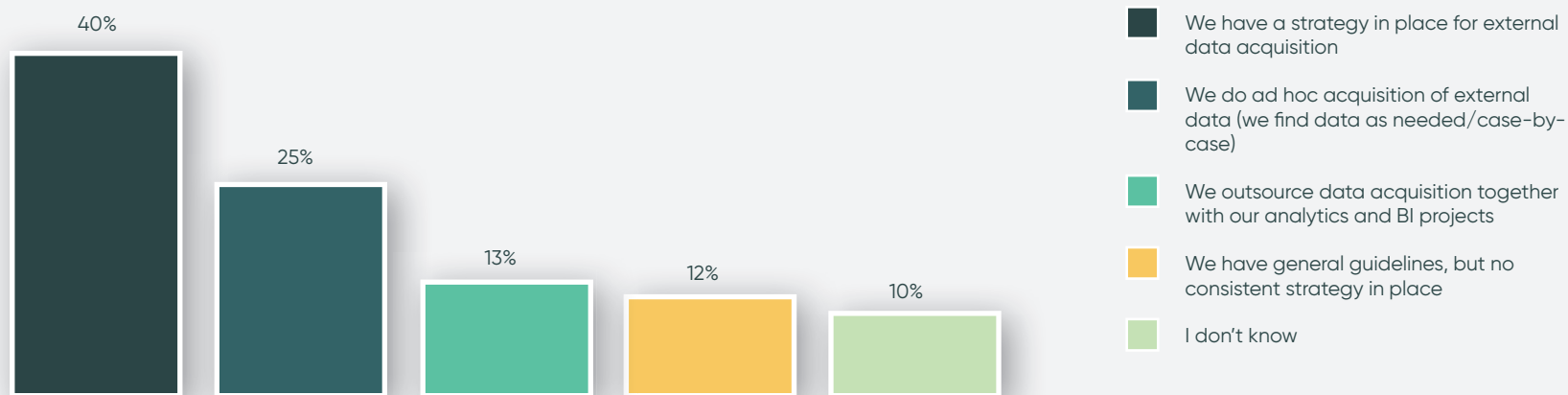
We asked how many providers organizations were purchasing data from and 44% of respondents said 5 or more. Similar to the types of data being purchased, this indicates that companies are buying a variety of external data from a variety of specialized providers or brokers to get the data they need.



Respondents reveal that almost half (44%) acquire external data from 5 or more paid or public sources.

## External data procurement strategy

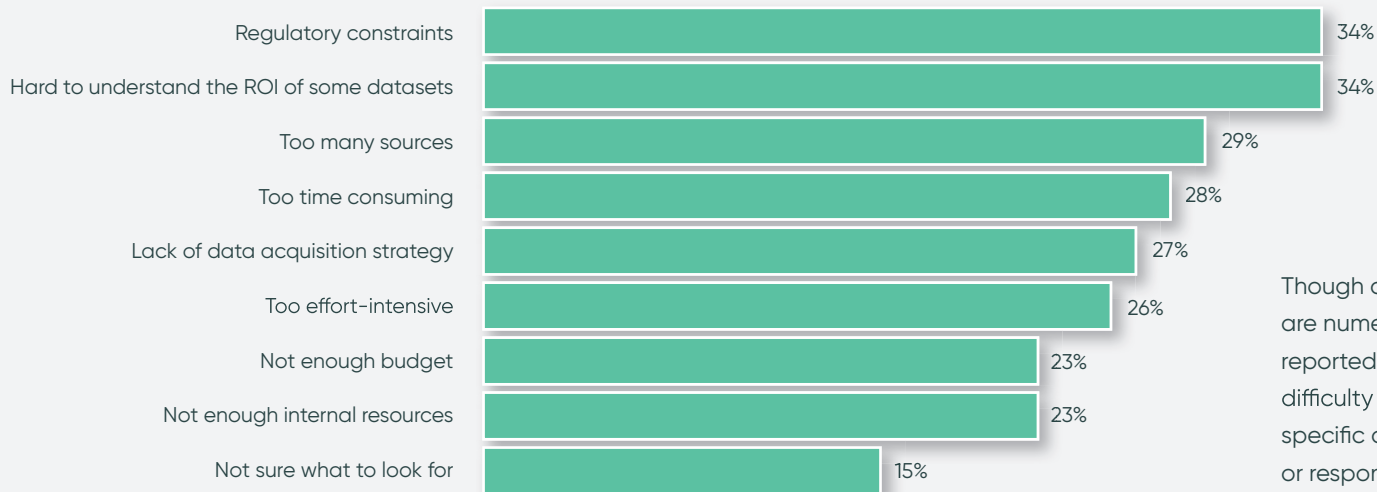
While most respondents are using a variety of data types and providers, when asked about their external data procurement strategy, only 40% of companies have a strategy in place. This is an improvement in our survey from last year when only 28% responded they had a formal strategy in place. But 37% of respondents said they either followed general guidelines or an ad-hoc process to get external data.



Across all respondents, data acquisition strategies vary, with 40% adhering to a predetermined strategy versus 25% acquiring external data on an ad hoc basis as needed. 13% report outsourcing their data acquisition project by project.

# External data procurement challenges

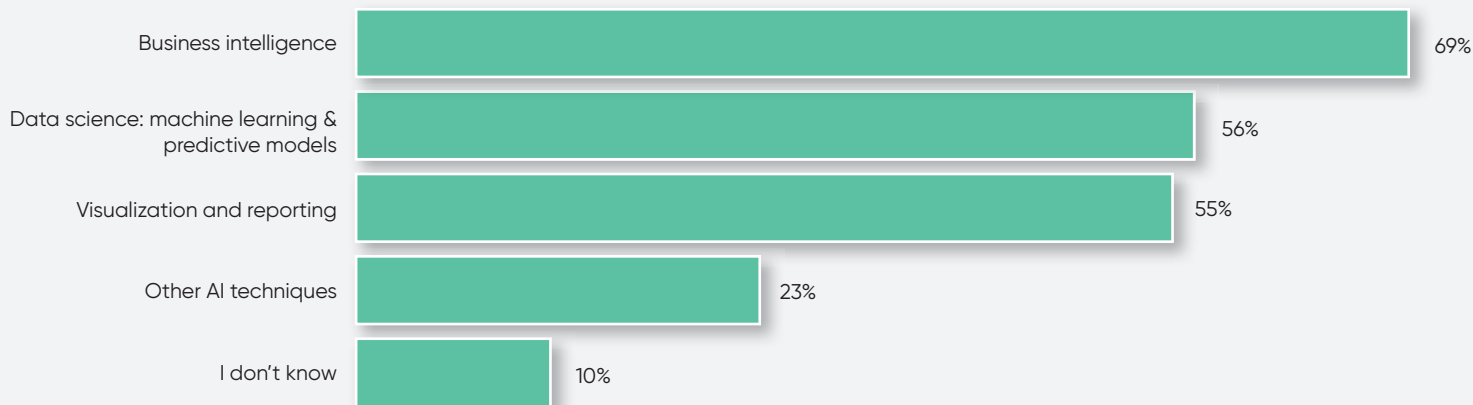
There are many challenges organizations face when buying external data. The leading problems reported are regulatory hurdles and difficulty understanding the ROI of specific datasets acquired, cited by 34% of respondents each.



Though challenges in acquiring data are numerous, the leading problems reported are regulatory hurdles and difficulty understanding the ROI of specific datasets acquired, cited by 34% of respondents each.

## External data and analytics

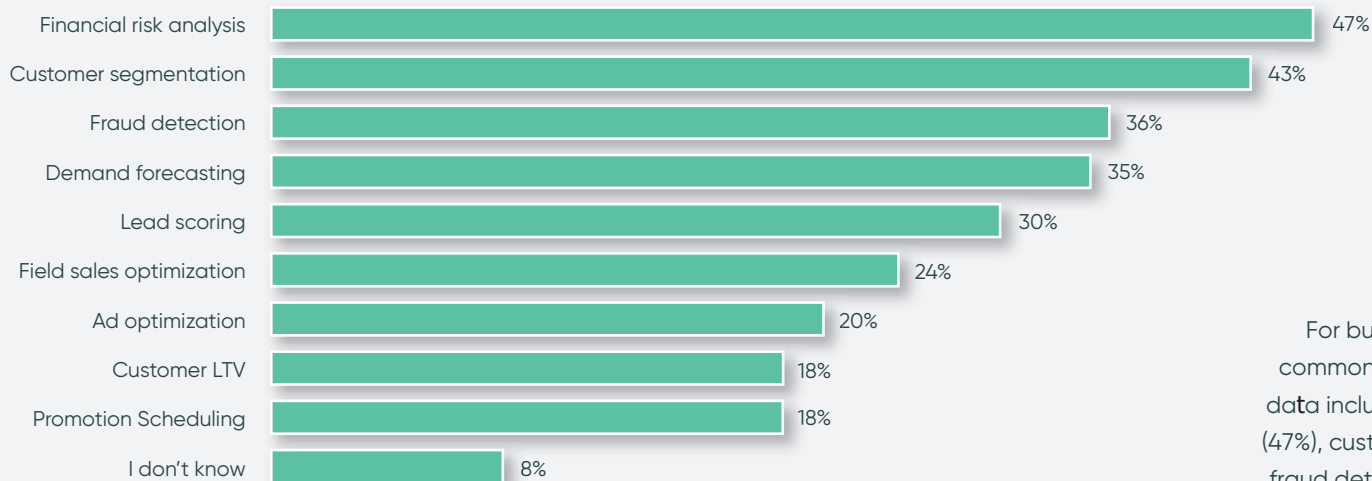
The leading usages of external data within analytics are business intelligence (69%), data science (56%) and visualization and reporting (55%). The relatively low number for data science and other AI techniques only receiving 23% suggests that there is still ample room for improvement in how organizations leverage external data in their advanced analytics initiatives.



Leading usages of external data in analytics initiatives include business intelligence (used by 69%), data science (56%) and visualization and reporting (55%).

## External data and business initiatives

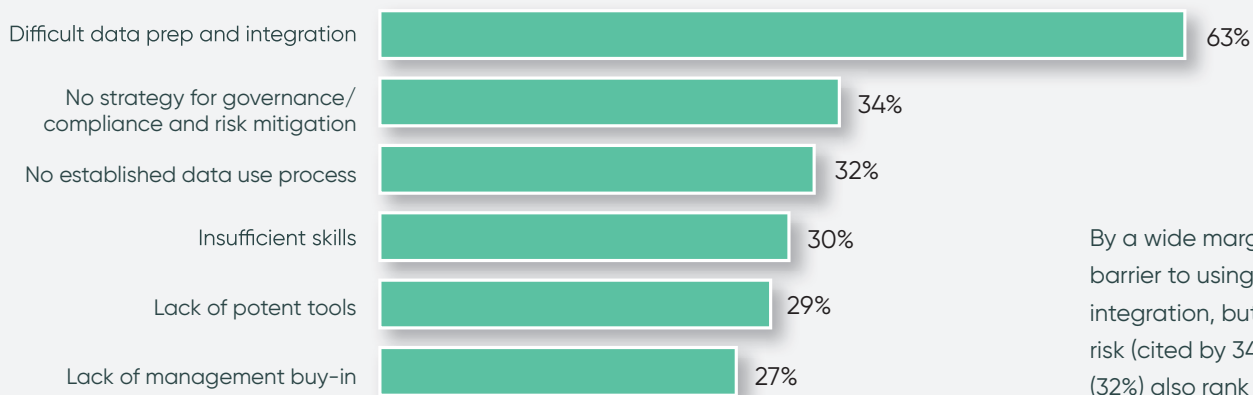
The most common use cases that external data is used for include financial risk analysis (47%), customer segmentation (43%), fraud detection (36%) and demand forecasting (35%). As the value of external data becomes more apparent and is used to create competitive advantages, we expect the number of use cases to increase significantly as organizations seek out more ways to leverage it across their key business initiatives.



For business initiatives, most common applications of external data includes: financial risk analysis (47%), customer segmentation (43%), fraud detection (36%) and demand forecasting (35%).

## Inhibitors using external data

An overwhelming majority of respondents (63%) said the top inhibitor to using external data is difficult data preparation and integration. Lack of strategies to address compliance and risk (cited by 34%) and lack of process for usage (32%) rounded out the top 3. Interestingly, lack of management buy-in (27%) was the least chosen inhibitor.

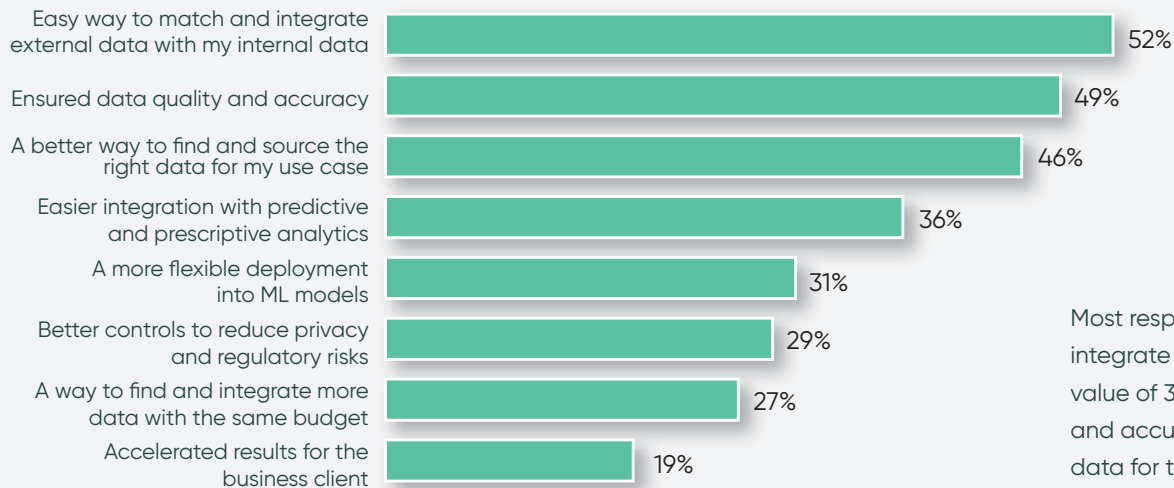


By a wide margin respondents report the most common barrier to using external data is data prep and integration, but lack of strategies to address usage and risk (cited by 34%) and lack of process for data usage (32%) also rank in the top three usage inhibitors.

## External data and getting the most value

When we asked how you can get more value out of external data, most respondents primarily need to be able to easily integrate internal and external data. This comes as no surprise as it's also the biggest inhibitor to using external data. Respondents also wanted assurances on data quality and accuracy (49%) and the ability to locate the right data for the job (46%).

The key takeaway here is that organizations need to seek out new technologies to address the challenges with data preparation and integration and expand their use of external data.



Most respondents primarily need to be able to easily integrate internal and external data to maximize the value of 3rd party data, but also demand data quality and accuracy (49%) and the ability to locate the right data for the job (46%).

Request a Demo

For more information, please visit us



---

## About Explorium

Explorium's belief is that lack of data shouldn't hold back your creativity or ability to solve business challenges. Relevant external data has the power to unlock new discoveries and insights that power competitive advantages and drive smarter business decisions.

Explorium's External Data Cloud combines the world's best data sources in a single platform for easy access, discovery, and integration of relevant enrichments.

With Explorium, data professionals can build better models, discover relevant audiences, and create disruptive solutions. Explorium is the world's first External Data Cloud founded on our strengths in connecting data with data science. As the most complete external data offering, Explorium has won the trust of many enterprise customers.