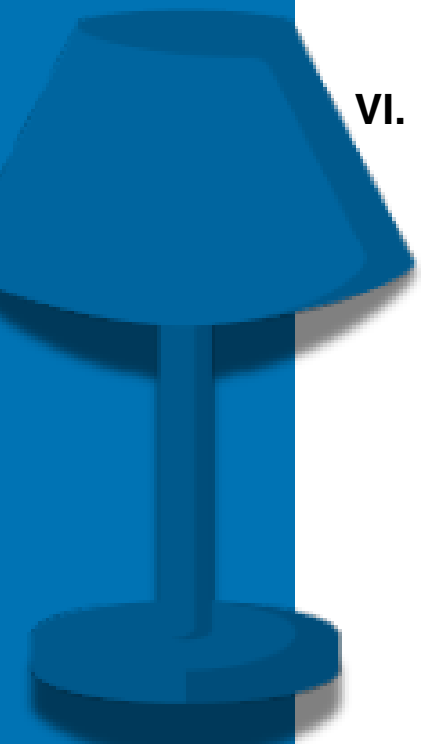
Wakefield Survey

MONTE CARLO

The State of Reliable AI Survey

2024 Edition

Table of Contents



Executive Summary

Data quality has long been a challenge for data teams.

In fact, two-thirds of data teams experienced a data incident that cost their organization $100,000 or more in the past 6 months—and 70% of data leaders say data incidents take longer than 4 hours to detect.

However, the introduction of GenAI has only served to exacerbate both the scope and severity of this problem. Just last year, United Healthcare was purported to have denied millions of dollars in necessary medical coverage to the elderly based on an AI with a known 90% error rate.

How can this be happening? It's because while the data estate has been evolving rapidly over the last 24 months, the general approach to data quality has not.

A staggering 91% of data leaders say they're actively building AI applications, but two-thirds admit to not completely trusting the data they're built on.

What's worse, less than half are implementing sufficient data quality practices to fill those gaps.

This anemic approach to data quality will have a demonstrable impact on enterprise AI applications and data products in the coming months—allowing more data incidents to slip through the cracks, multiplying hallucinations, diminishing the safety of outputs, and eroding confidence in both the AI and the companies that build them.

The best data teams will take a proactive approach to managing data quality at scale across RAG pipelines and the outputs themselves. If AI is to succeed, data quality must be front and center.

Barr Moses
CEO & co-founder
Monte Carlo

Methodology

The Monte Carlo Survey was conducted by Wakefield Research (www.wakefieldresearch.com) among 200 data engineers working in the US, between April 24th, 2024 and May 7th, 2024, using an email invitation and an online survey.

Results of any sample are subject to sampling variation. The magnitude of the variation is measurable and is affected by the number of interviews and the level of the percentages expressing the results.

For the interviews conducted in this particular study, the chances are 95 in 100 that a survey result does not vary, plus or minus, by more than 6.9 percentage points from the result that would be obtained if interviews had been conducted with all persons in the universe represented by the sample.

± Data under “QuickFacts” were derived from the responses, not included as response options that were read during fielding. We include QuickFacts in instances where we feel they will be helpful.

Section 1.

The Truth About Enterprise AI

It's Not "If" But "How"

It's not uncommon for data teams to feel pressure from upstream stakeholders, but the advent of GenAI has transformed the conversation entirely.

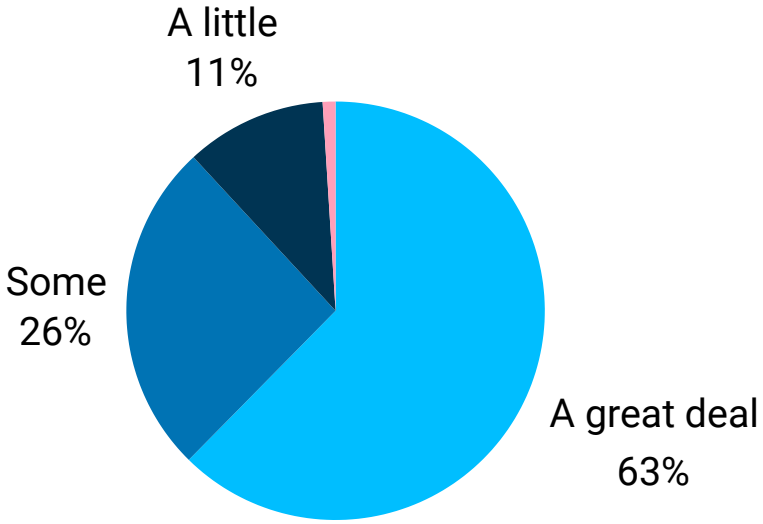
Whether or not data leaders feel that AI is the most valuable initiative on their roadmap has been rendered largely irrelevant by executives fighting to land on top in the AI race.

For most modern data teams, the discussion around AI is no longer if or when, but how they will be building for GenAI in the next 12 months.

But many data professionals believe their executives have unrealistic expectations.

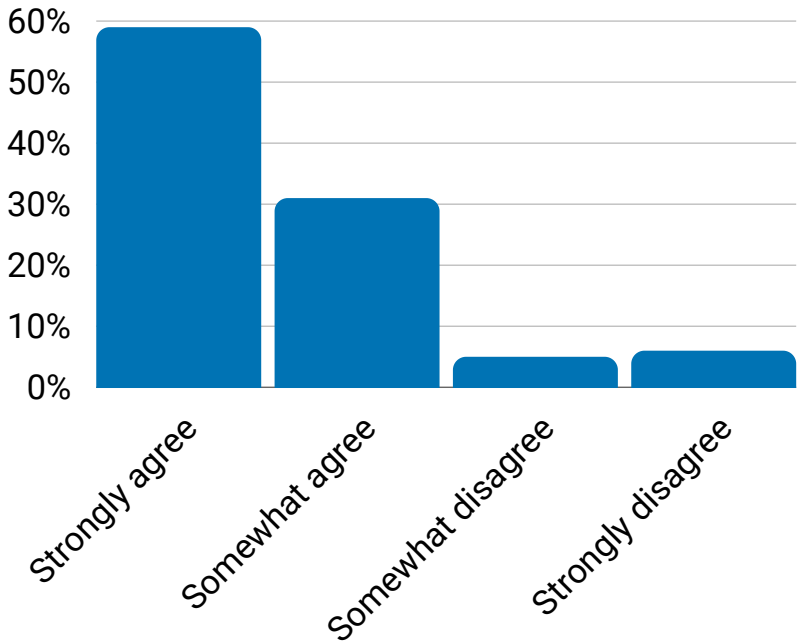
Leadership Pressure

Q How much pressure do you feel from your leadership to implement a Generative AI strategy or build Generative AI products?



Leadership Expectations

Q Leadership at my organization (VP, CEO, board, etc.) does not have a realistic expectation of what Generative AI can do, both in its technical feasibility and its ability to drive business value.



Data Teams Say It's Not All Hype

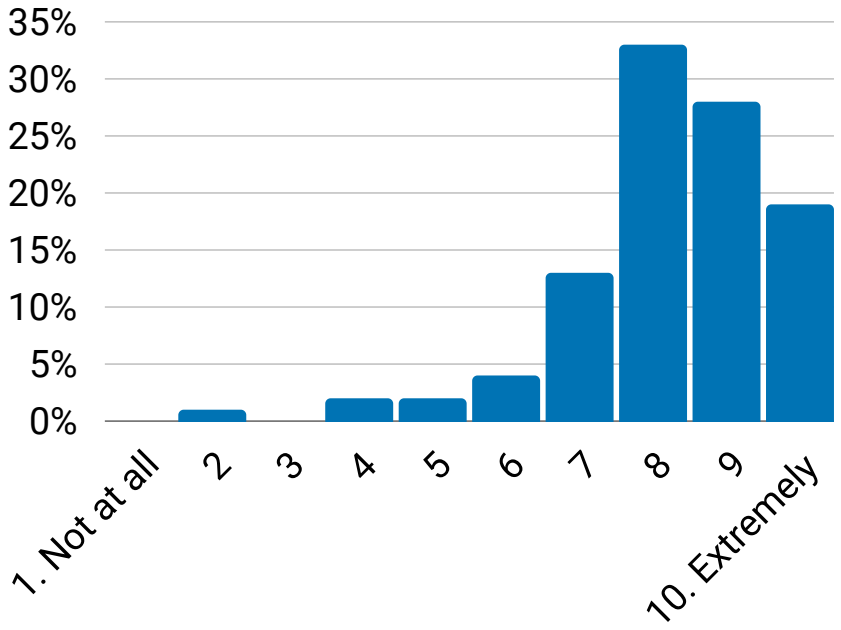
Just because executives may over-inflate the value of AI doesn't mean data leaders think it's all hype.

In fact, most data teams believe that AI will be fairly useful to their businesses.

The challenge is identifying those use cases which could actually provide value—and of those, the use-cases your team has the first-party data to implement in a valuable way.

AI Usefulness

Q How useful do you think Generative AI can be for your business?



Quick Facts 82% of respondents rated the usefulness of GenAI at least and 8

"Generative AI for the sake of Generative AI won't solve any problems. Just like any engineering puzzle, you have to know the problem you are trying to solve before asking what tools can help us get there.

Generative AI is a tool in the tool belt. It might be a Swiss army knife but it's still a tool and should be considered as an option to solve a problem, not a whole solution." [Read more]

Anonymous Contributor
Data Engineering

Everyone is Building

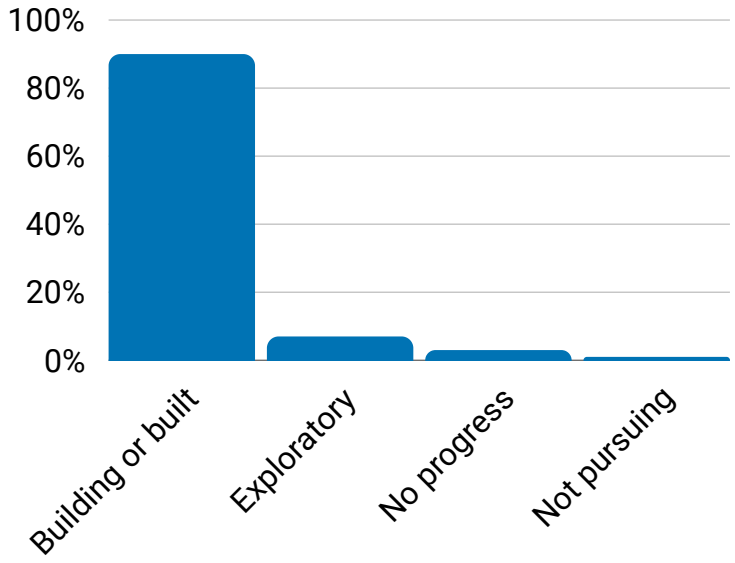
Executives are clamoring for more AI and data teams are answering.

The number and types of AI use cases are increasing by the day as data teams find new ways to leverage the nascent and legacy data for their stakeholders.

In terms of those teams building GenAI today, the distribution of customers is fairly evenly distributed with 75% of respondents claiming to be building or have built internal-only GenAI products (like transcription tools), and 72% having built or building customer-facing products.

Current AI Initiatives

Q How would you describe your organization's Generative AI initiatives today?



Quick Facts

90% are building or have built some form of AI product

"I've seen examples of industries with huge amounts of documentation that want to enable their internal teams to retrieve answers out of tens of thousands of pages of records...That's the right approach, because the risk is low — it allows you to get your hands dirty, provides a lot of value, and doesn't create a lot of risk. At Databricks, we have an internal chatbot that helps employees figure things out and look at their data. And we see a lot of value there."

Adam Conway
SVP of Products
Databricks

Section 2.

AI Infrastructure in 2024

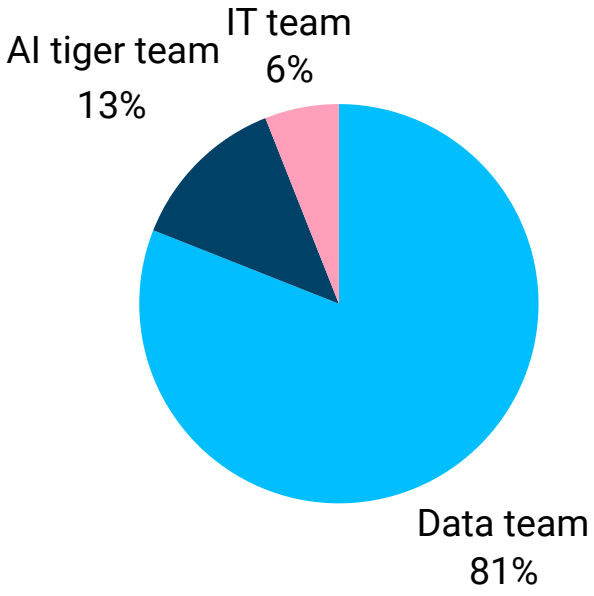
Data Teams Need to Sharpen their AI Skills

While data professionals may disagree on the priority or their organization's preparedness for GenAI at the given moment, they feel strongly that the data team should own still own it—and their leadership should be experienced enough to guide them through it.

This underscores the necessity for data teams to be become proficient in strategies like RAG and fine-tuning, as well as the data quality requirements, to make these products both valuable to stakeholders and successful over the long-term.

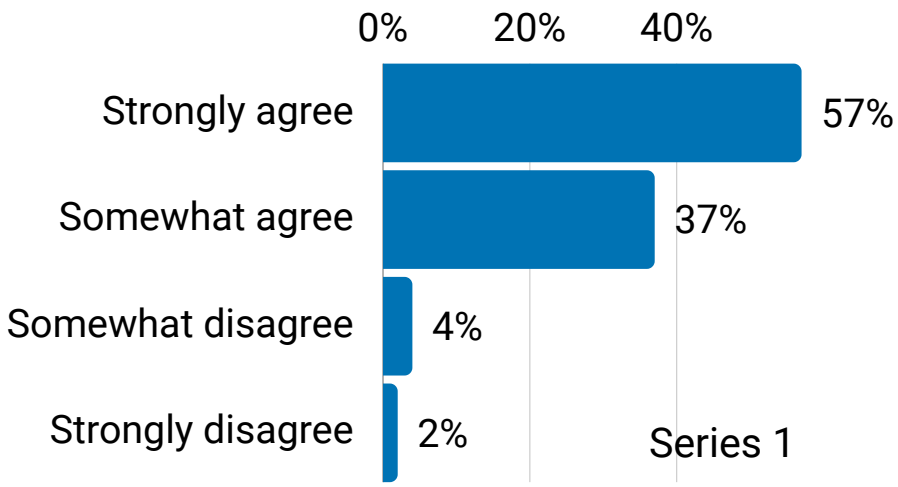
AI Ownership

Q Who is primarily responsible for Generative AI in your organization?



Leadership Requirements

Q In 2024, the leader of a tech company needs to have a background in AI for the company to be successful.



Architecture Is A Mixed Bag

In the early days of AI, there's no one-size-fits-most approach to building.

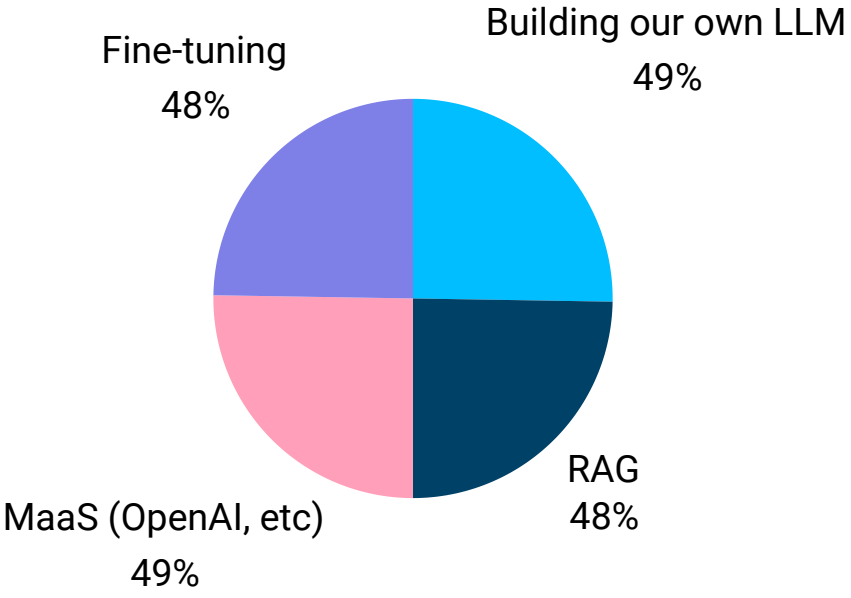
As data teams look to develop AI applications for their organizations, they'll need to return to tried and tested scoping questions to define the right approach:

- What are we looking to achieve?
- How will this approach add value?
- What do we have the expertise to build and maintain right now?
- Where is our time best spent?

And at the end of the day, the value of any AI application, regardless of approach, will come down to your own first-party data.

AI Strategies

Q What strategies is your team implementing for Generative AI?



Quick Facts

Fine-tuning was most popular amongst data engineers, be least favored amongst data leaders.

"Regardless of which one you use...your AI application is going to require pipelines that feed these models with company data. At the end of the day, if generative AI is used in internal processes to extract analysis and insight from unstructured data – it will be used in... drumroll... a data pipeline."

Lior Gavish
CTO & Co-Founder
Monte Carlo

Data sources are increasing to power new use-cases

Even before the advent of GenAI, organizations were dealing with far greater data than they had in decades past.

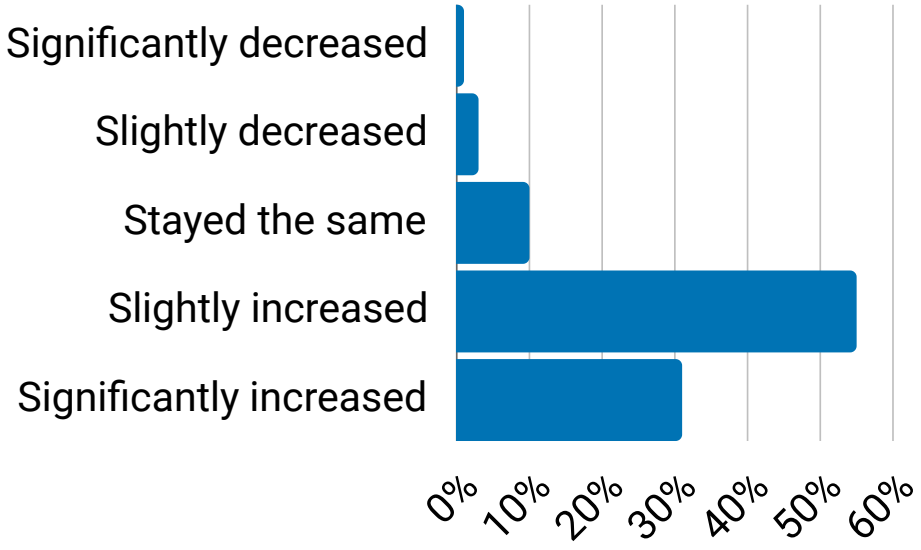
And that increased complexity has proven a worthy tradeoff for data teams and their stakeholders.

However, since adopting GenAI programs, both applications and the number of critical data sources are increasing yet again—deepening the complexity and scale of data estates to new levels.

And as complexity grows, data governance takes center-stage.

Data Sources

Q How has the amount of data sources changed in the past year in your adoption of Generative AI?



Quick Facts

Fine-tuning was most popular amongst data engineers, be least favored amongst data leaders.

You need to make sure you have the right customer data to craft the right prompt so the answer is personalized to the end user asking the question...You need to make sure the LLM has something to ground its answers on. You don't want the LLM to use its general knowledge, which might not be accurate for your company." [Read more]

Nga Phan
SVP of Product Management
Salesforce AI

Section 3.

The Data Quality Problem

Bad Data Is Still Expensive

Every year data becomes more valuable—and bad data becomes more costly.

In the age of AI, that's more true than ever.

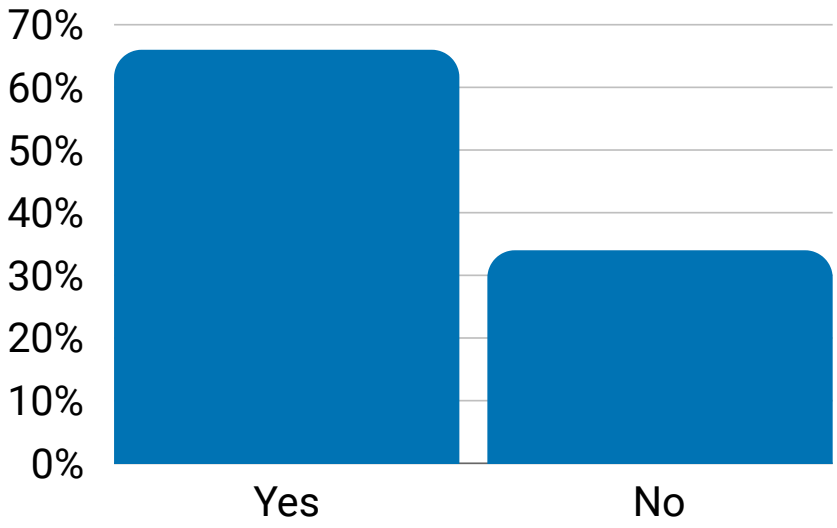
So, it's no surprise that more than half of respondents reported experiencing a data incident that cost their organization more than $100K. And we didn't even ask how MANY they experienced.

What's worse, 68% of respondents weren't fully confident that their data was reliable enough for AI use-cases. And "not fully" leaves a lot of room for error when it comes to AI.

Fortunately, if we understand the cost of bad data, it becomes easier for teams to build a business case to improve it.

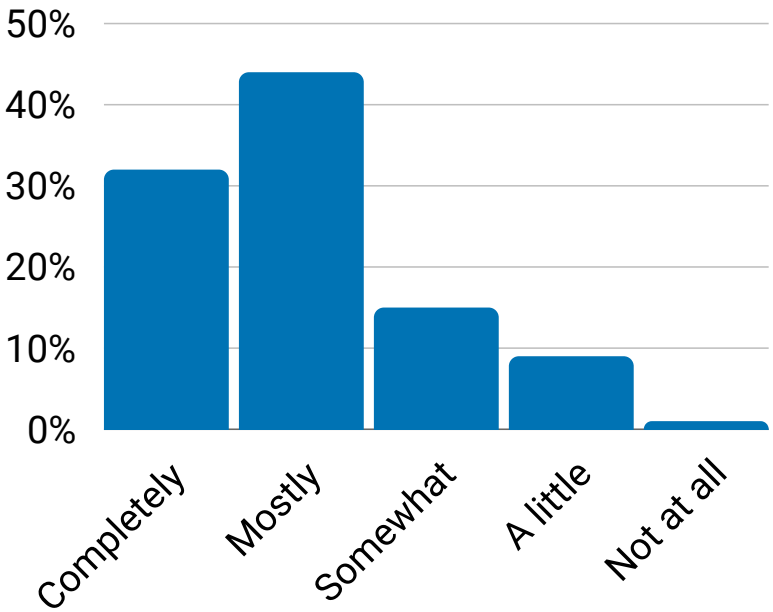
The Cost of Bad Data

Q Have you experienced a data incident that has cost your organization $100K or more in the past 6 months?



The Cost of Bad Data

Q How confident are you that your data is currently of high enough quality to reliably feed a large language model (LLM)?



Faster Detection Reduces Cost

The equation for data downtime is [#incidents x (detection + resolution).

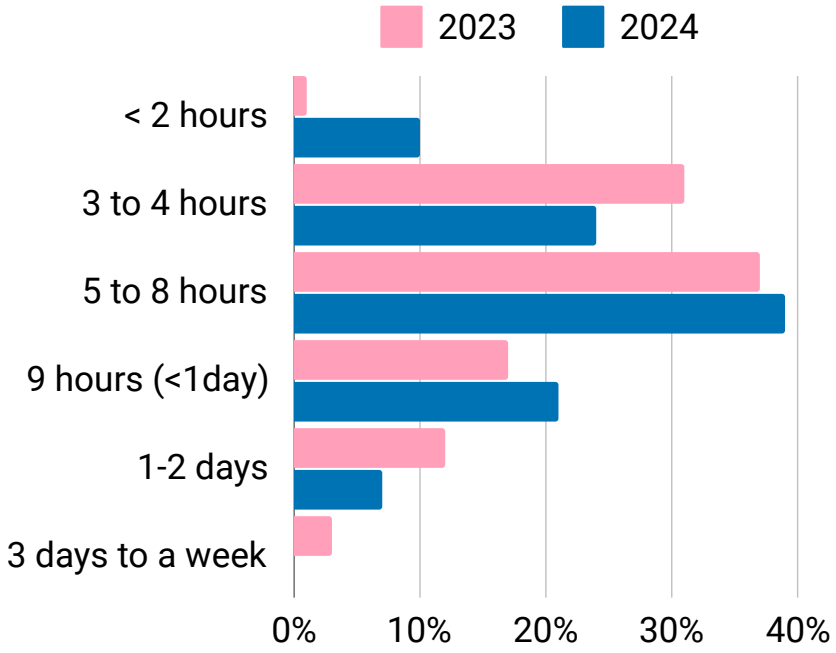
In our experience the longer it takes to detect a data incident the more likely it was to be discovered by someone outside the data team.

Who discovers an incident matters. If it's an internal data consumer, it erodes data trust. If it's an external customer, it creates churn risk. And if it's Wall Street...

Of course, finding an incident is only one half of the equation. Resolving it is a different story entirely.

Time to Detection

Q How long does it typically take to detect a data incident?



Quick Facts

66% of respondents claimed that it took more than 4 hours to detect a data quality incident.

"...we did have some process issues where a job wouldn't run and no one would get notified soon enough.

Any data quality issue that made its way past the data team chipped away at the faith and trust the data-driven Sub-Zero team had in their reports.

It doesn't look good if we receive a complaint from a user on an issue that we were unaware of and it could or should have been caught 6 hours ago." [Full Story]

Justin Swenson
Data & AI Governance Lead
Sub-Zero

Expect 8 Hours of Data Downtime Per Incident

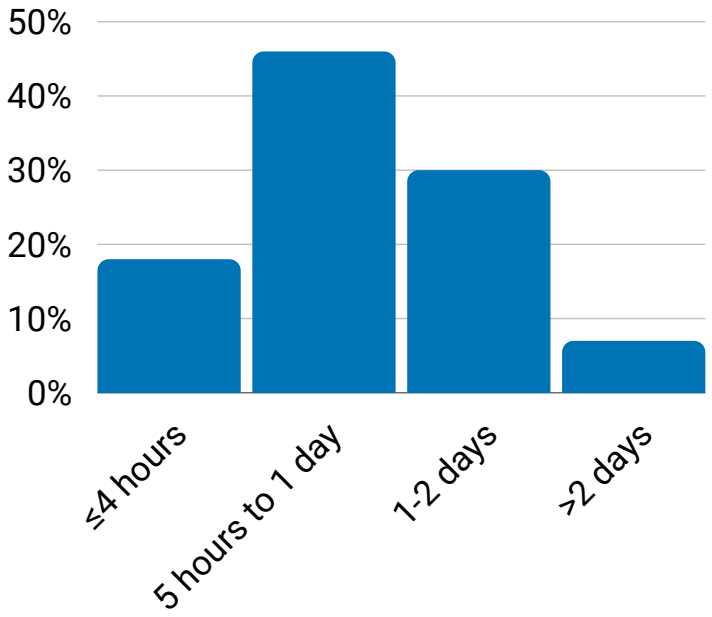
As AI continues to expand the scale and complexity of data systems, data quality issues will continue to expand as well.

What's worse, if teams can't resolve issues efficiently, the impact will compound. According to previous surveys, poor time to resolution strongly correlates to more incidents—and consequently—more time spent on data quality.

Based on survey averages, most data teams can expect at least 8 hours of data downtime per incident (detection + resolution), with 37% of respondents reporting a full day or more for resolution alone.

Time to Resolution

Q Once a data incident is discovered, how many hours does it typically take to resolve?



Quick Facts

More than 80% of respondents reported that it takes longer than 4 hours to resolve a data quality issue once detected.

"When we looked into it we realized that much of our capacity issues came from data quality issues popping up.

We discovered the process of finding the root cause of data quality issues took 200 hours per quarter across our entire team." [Full Story]

Kineret Kimhi
Data Engineering & BI Manager
BlaBlaCar

The Data Estate is Evolving. Data Quality Isn't.

Despite the scaling complexity of the data estate, many data teams still rely on tedious and unreliable methods like manual testing to detect data quality issues.

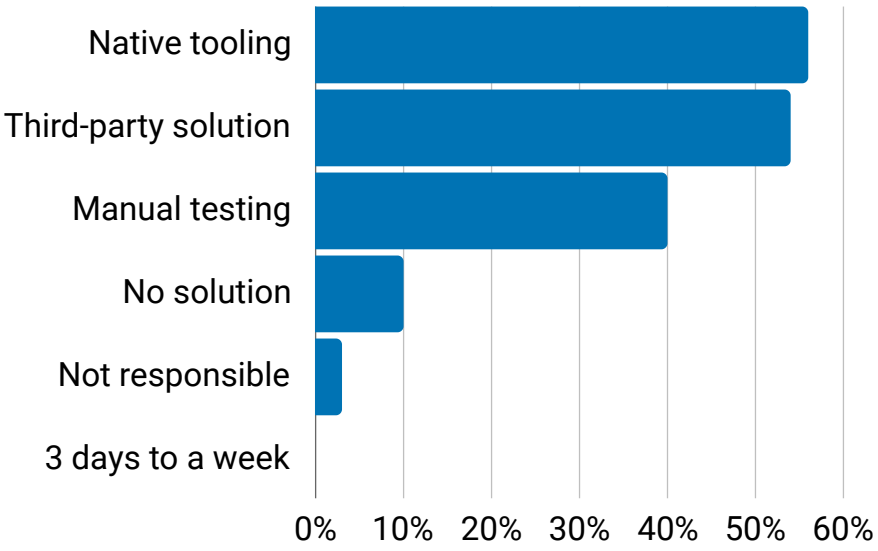
The success of AI depends on the data's quality—and the success of the data team depends on its ability to efficiently detect and resolve the issues that impact it.

At the dizzying scale of AI, data teams need modern scalable solutions like data observability to reduce time to both detection and resolution in one platform.

Data is evolving. Data quality needs to evolve with it.

Data Quality Methods

Q How is your data team addressing the quality of the data for feeding large language models (LLMs)?



Quick Facts

54% of teams surveyed still rely on manual testing or have no initiative in place to address data quality in LLMs.

"We don't have any choice — we need to be able to observe the data. We need to understand what data we're putting into the LLM, and if the LLM is coming up with its own thing, we need to know that — and then know how to deal with that situation.

If you don't have observability of what goes into the LLM and what comes out, you're screwed." [Full Story]

Vishnu Ram
VP of Engineering
Kredit Karma

Additional Resources

Don’t let this be the end of your data quality journey! Check out more helpful resources including:

- Data Downtime Blog: Get fresh tips, how-tos, and expert advice on all things data.
- Data Observability Product Tour: See data observability in action to find out how it’s changing the data quality conversation.
- Request A Demo: Talk to our team to find out how much data downtime is costing your team—and how Monte Carlo can help.