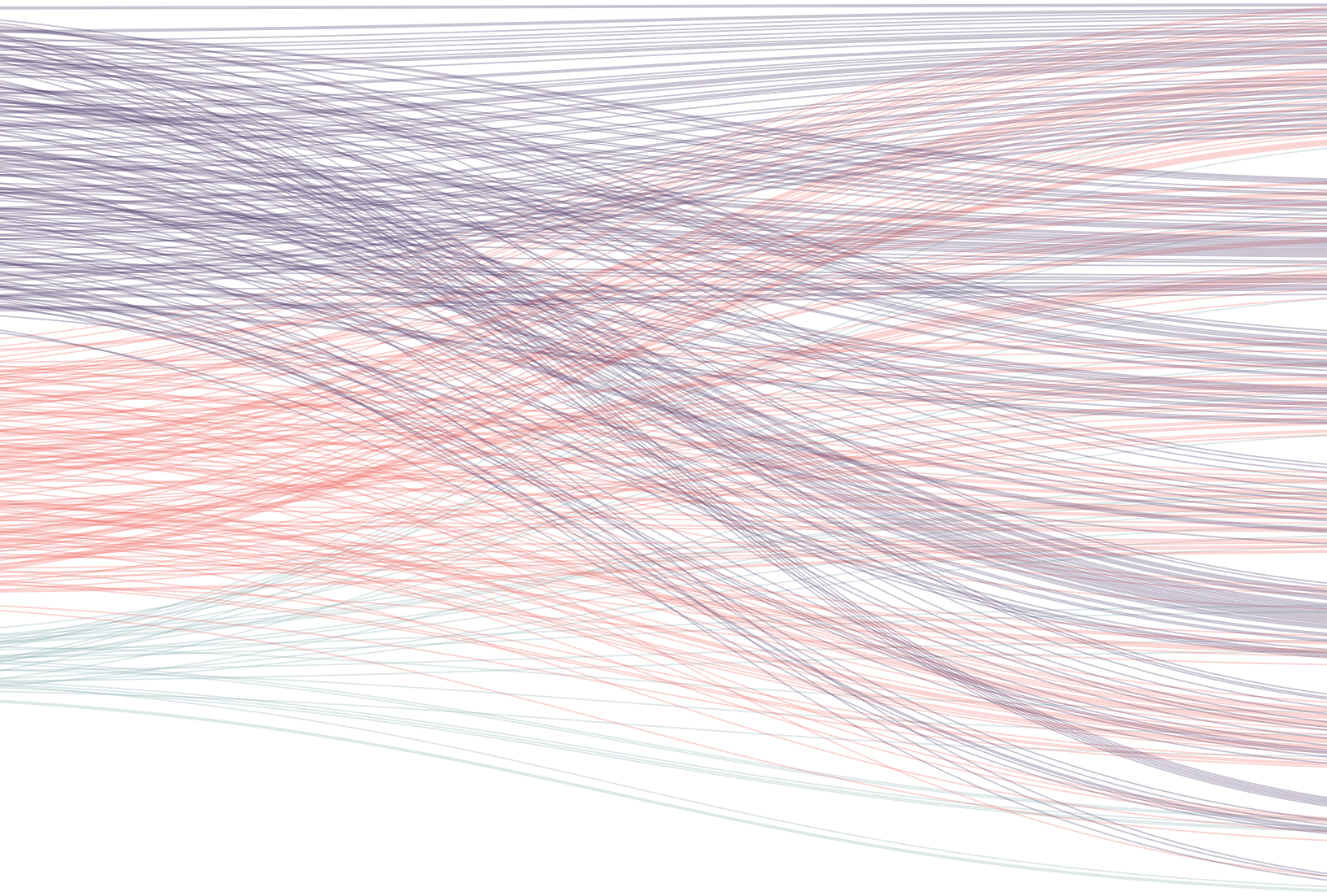


ALGORITHMIA

2021 enterprise trends in machine learning



Contents

Introduction 01

Survey at a glance 02

Report theme 1: Organizations are increasing AI/ML budgets, staff, and number of use cases—but the market’s still dominated by early adopters 04

Report theme 2: Organizations continue to face challenges across the ML lifecycle, and struggle with cross-functional alignment 13

Report theme 3: Despite increased budgets and hiring, organizations are spending more time and resources—not less—on model deployment 23

Report theme 4: With increasingly complex infrastructure needs, organizations report improved outcomes with third-party MLOps solutions 27

Conclusion 32

Methodology 34

About Algorithmia 37

Introduction

2020 was a year of belt-tightening for many organizations due largely to the macroeconomic impacts of the COVID-19 pandemic. In May 2020, Gartner [predicted](#) that global IT spending would decline 8% over the course of 2020 as business and technology leaders refocused their budgets on their most important initiatives.

One thing is readily apparent in the 2021 edition of our *enterprise trends in machine learning* report: AI and ML initiatives are clearly on the priority list in many organizations. Not only has the upheaval of 2020 not impeded AI/ML efforts that were already underway, but it appears to have accelerated those projects as well as new initiatives.

This is our third year conducting a survey about the current state of enterprise machine learning and reporting on the results, and this is our most comprehensive report yet. This year we expanded our questions about organizational maturity, a critical indicator of long-term success that includes organizational alignment, data, training, deployment, management, and governance. To improve the quality of our data and allow more in-depth insights, we also increased the specificity of several questions and the options given to respondents for answering them.

This report is the result of that effort, and includes our detailed analysis of the findings. It is clearly an exciting—and exceedingly important—point in the AI/ML journeys of many organizations. Algorithmia is committed to being good stewards of machine learning technology and critical partners in our customers' success.

“

Machine learning operations is a nascent discipline; it isn't well recognized in the industry. The survey results in this report underscore the importance of MLOps and its impact on optimizing infrastructure consumption, improving application performance, and increasing productivity for data scientists. As ML has come to the mainstream, it's become apparent that MLOps is the key to driving efficiency and scale for organizations of any size. ML is both an important and emerging category that C-level leaders will prioritize if they want to stay competitive.”



Tim Tully
CTO and SVP, Splunk



Survey at a glance

This year's survey revealed 10 key trends that organizations should be paying attention to if they want to succeed with AI/ML in 2021. The trends fall into a few main themes, and the overarching takeaway is that organizations are moving AI/ML initiatives up their strategic priority lists—and accelerating their spending and hiring in the process.

But despite increasing budgets and staff, organizations continue to face significant barriers to reaping AI/ML's full benefits. Specifically, the market is still dominated by early adopters, and organizations continue to struggle with basic deployment and organizational challenges. The bottom line is, organizations simply haven't learned how to translate increasing investments into efficiency and scale.

However, we remain optimistic about the growing potential of AI/ML in 2021 and beyond. As the space continues to mature, the barrier to entry has continued to get lower. The time to invest in AI/ML is now—no matter your organization's size, industry, or unique infrastructure needs.

2021 top 10 trends in enterprise ML

01

Priority and budgets for AI/ML are increasing significantly year-on-year

02

Organizations are expanding into a wider range of AI/ML use cases, with particular focus on process automation and customer experience

03

Most organizations have more than 25 models in production, but there's a gap between AI/ML "haves" and "have-nots"

04

Governance is by far the top challenge for AI/ML deployment, with more than half of all organizations ranking it as a concern

05

The second greatest AI/ML challenge is technology integration and compatibility, with 49% of organizations ranking it as a concern

06

Successful AI/ML initiatives require organizational alignment across multiple decision-makers and business functions

07

Organizational alignment is the biggest gap in achieving AI/ML maturity

08

The time required to deploy a model is increasing, with 64% of all organizations taking a month or longer

09

38% of organizations spend more than 50% of their data scientists' time on deployment—and that only gets worse with scale

10

Organizations that use a third-party machine learning operations solution save money and spend less time on model deployment than those that build their own solution

Continue reading to explore each of these 10 trends in detail, and learn how your organization can make the most of your AI/ML investments in 2021.

Report theme 1: Organizations are increasing AI/ML budgets, staff, and number of use cases—but the market's still dominated by early adopters

The time to invest in AI/ML is now, no matter your organization's size, industry, or infrastructure needs.

43%

Respondents who told us that AI/ML matters “way more than we thought” in a survey this summer

It's clear from this year's data that AI/ML projects have become one of the top strategic priorities in many enterprises. As of last year, organizations had already begun to boost their AI/ML investments; 71% of respondents in our 2020 report said their AI/ML budgets had increased compared with the previous year.

They're not dialing back that spending this year. In fact, companies appear to be doubling down on their AI/ML investments. We ran a [survey](#) this summer to see how organizations were adapting to the pandemic and its impacts, and it showed a new sense of urgency around AI/ML projects.

When we asked respondents why, 43% said their AI/ML initiatives “matter way more than we thought.” Nearly one in four said that their AI/ML initiatives should have been their top priority sooner.

It's not just talk: They're now ramping up their budgets and staffing accordingly. In our summer survey, 50% of respondents indicated that they plan to spend more on AI/ML this year. Roughly one in five said they “plan to spend a lot more”.

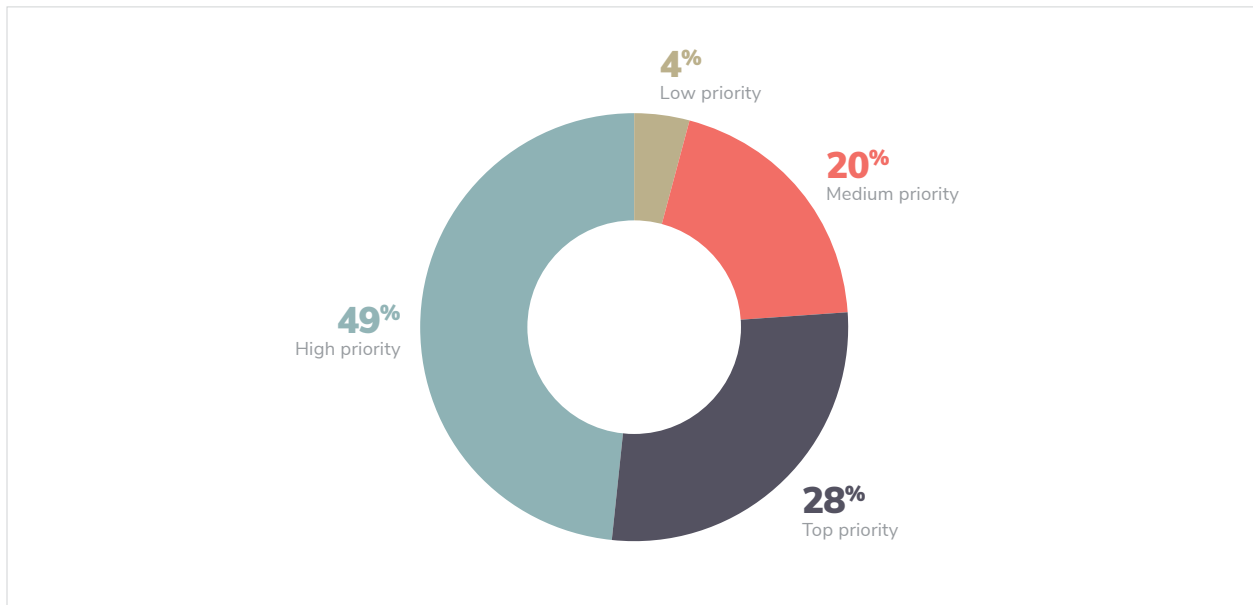
That's reflected in AI/ML-related hiring, too. At a time when many organizations are making difficult decisions about staffing, 76% of respondents in our summer survey said they had not reduced the size of their AI/ML teams—with a full 27% reporting they had increased it.

We see clear evidence in our 2021 survey data that these trends are expanding and extending in the year ahead in spite of continued economic concerns and other uncertainty. In fact, those economic concerns themselves have had a catalytic effect: Organizations are developing a laser focus on AI/ML initiatives as a means of driving top-line revenue while keeping bottom-line costs under control to ensure they are as competitive as possible.

Trend 1: Priority and budgets for AI/ML are increasing significantly year-on-year

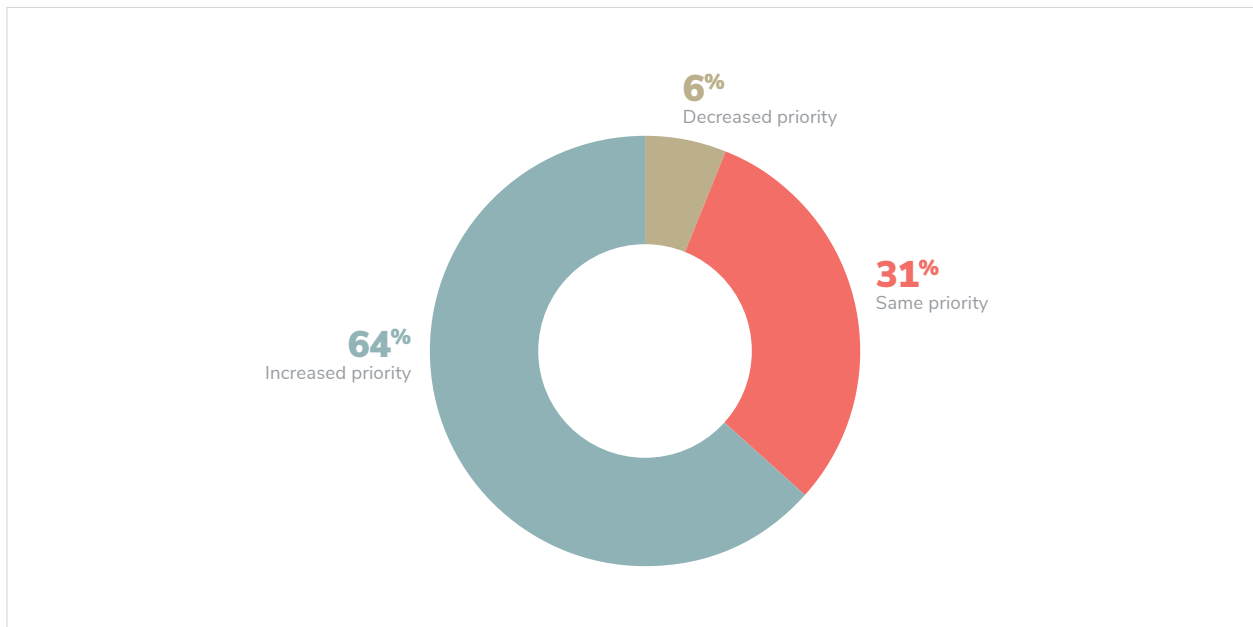
Continuing the trends we saw in our summer survey, our 2021 survey shows an increase in prioritization, spending, and hiring for AI/ML. First off, 76% of organizations say they prioritize AI/ML over other IT initiatives, and 64% say the priority of AI/ML has increased relative to other IT initiatives in the last 12 months.

76% of organizations prioritize AI/ML over other IT initiatives



Respondents were asked to indicate the priority given to AI/ML initiatives relative to other IT initiatives at their organizations. Those who said they were prioritizing AI/ML as either top or high priority relative to other IT initiatives were considered to be prioritizing AI/ML over other IT initiatives. The total percentage of these respondents was calculated with the underlying data before being rounded to the nearest percentage point. Categories do not add up to 100% because they have been rounded to the nearest percentage point.

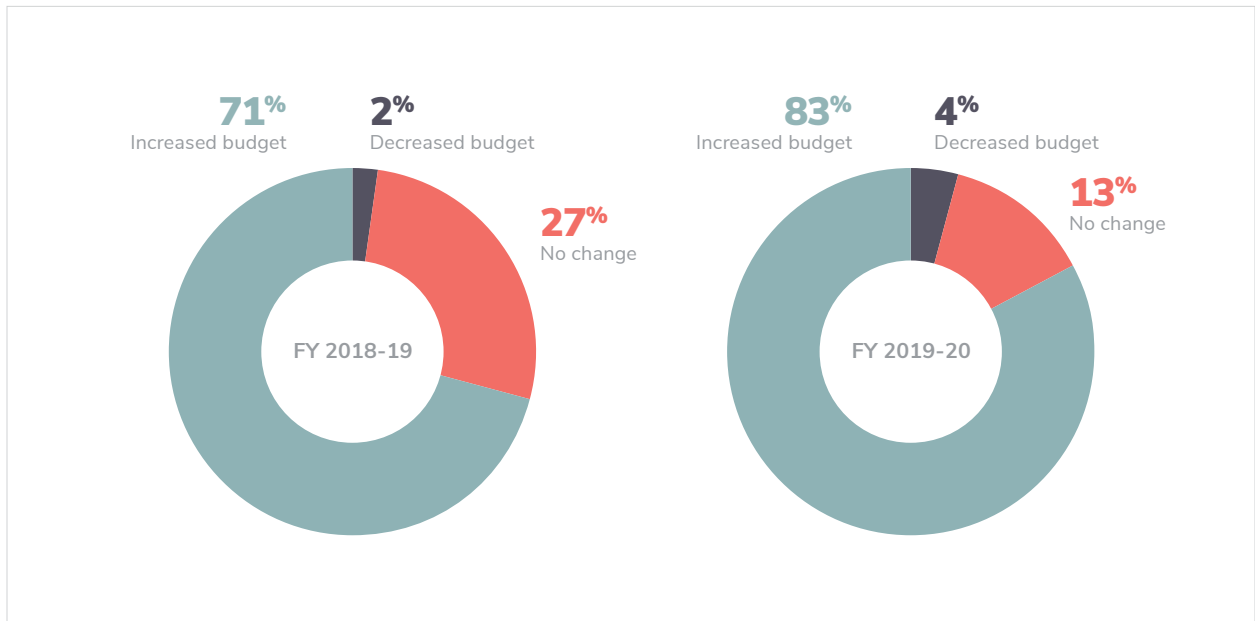
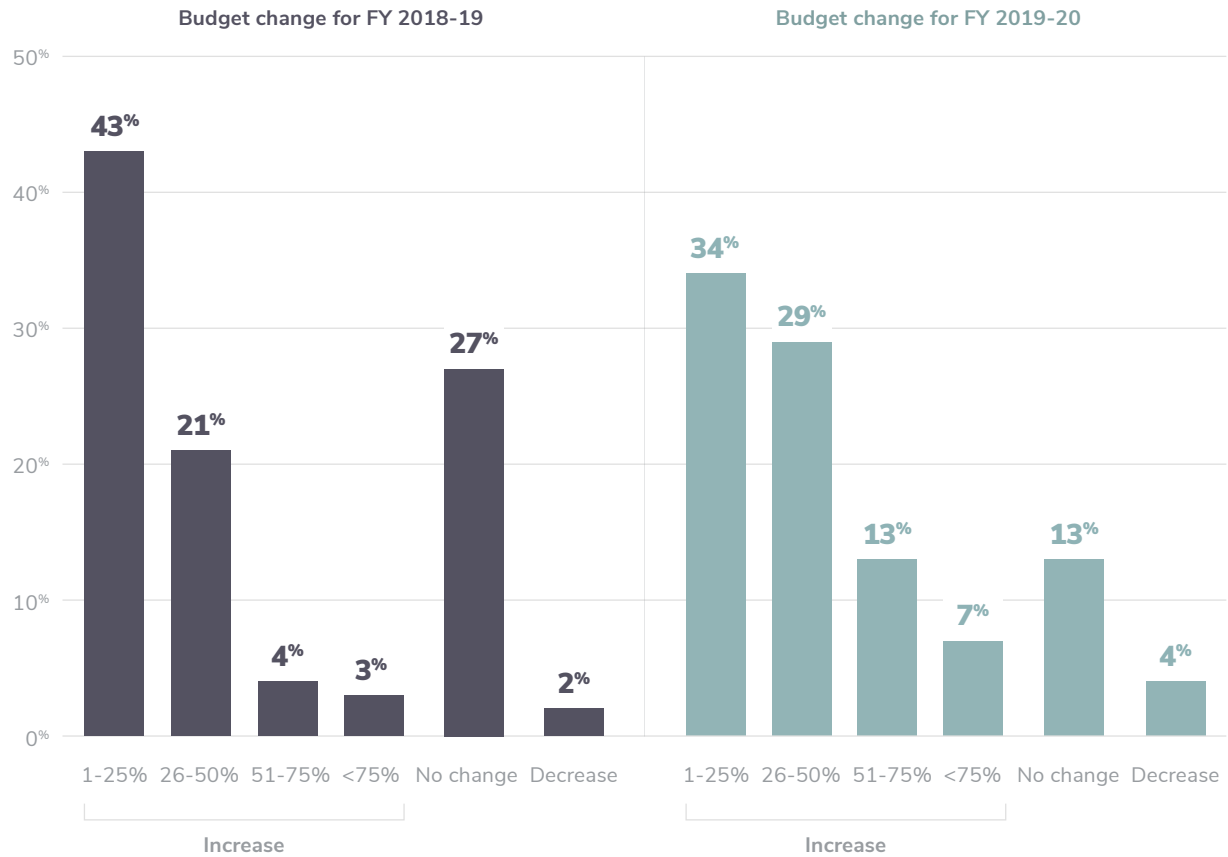
64% of organizations have increased AI/ML priority in the past year



Respondents were asked to indicate how the priority of their AI/ML initiatives has changed relative to other IT initiatives in the past 12 months. Categories do not add up to 100% because they have been rounded to the nearest percentage point.

When it comes to spending, a full 83% of organizations have increased their budgets year-on-year.

83% of organizations have increased AI/ML budgets year-on-year

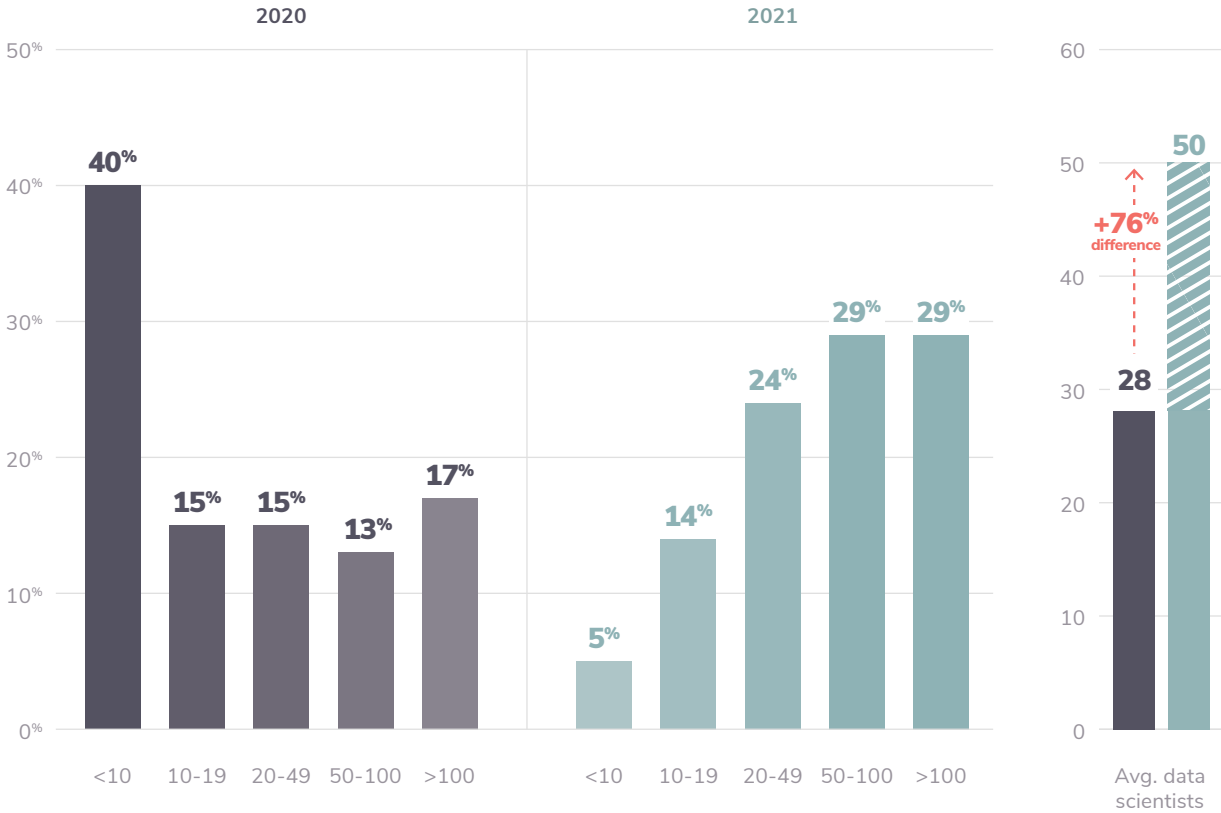


You might think that this is ordinary growth attributable to the general growth of AI/ML. There’s partial truth to this. However, it doesn’t account for the reality that 2020 has been a year of extraordinary change and upset in the industry, a year in which many companies are looking for opportunities to scale back all but their most mission-critical priorities. It also doesn’t tell the full story in terms of how much organizations are increasing their AI/ML budgets.

In last year’s report, just 7% of respondents said their budgets were increasing by more than 50%. This year, 20% reported AI/ML budget increases of more than 50% from FY2019 to FY2020.

Here’s an even more telling indicator of the accelerating pace of AI/ML strategies. Respondents were asked how many data scientists their organizations employ, from which we estimated the average number of data scientists employed by organizations in both the 2020 and 2021 data. Year-on-year, the average number of data scientists employed has increased by 76%. In fact 29% of respondents in our 2021 report now have more than 100 data scientists on their team, a significant increase from the 17% reported last year.

The average number of data scientists employed has increased 76% year-on-year



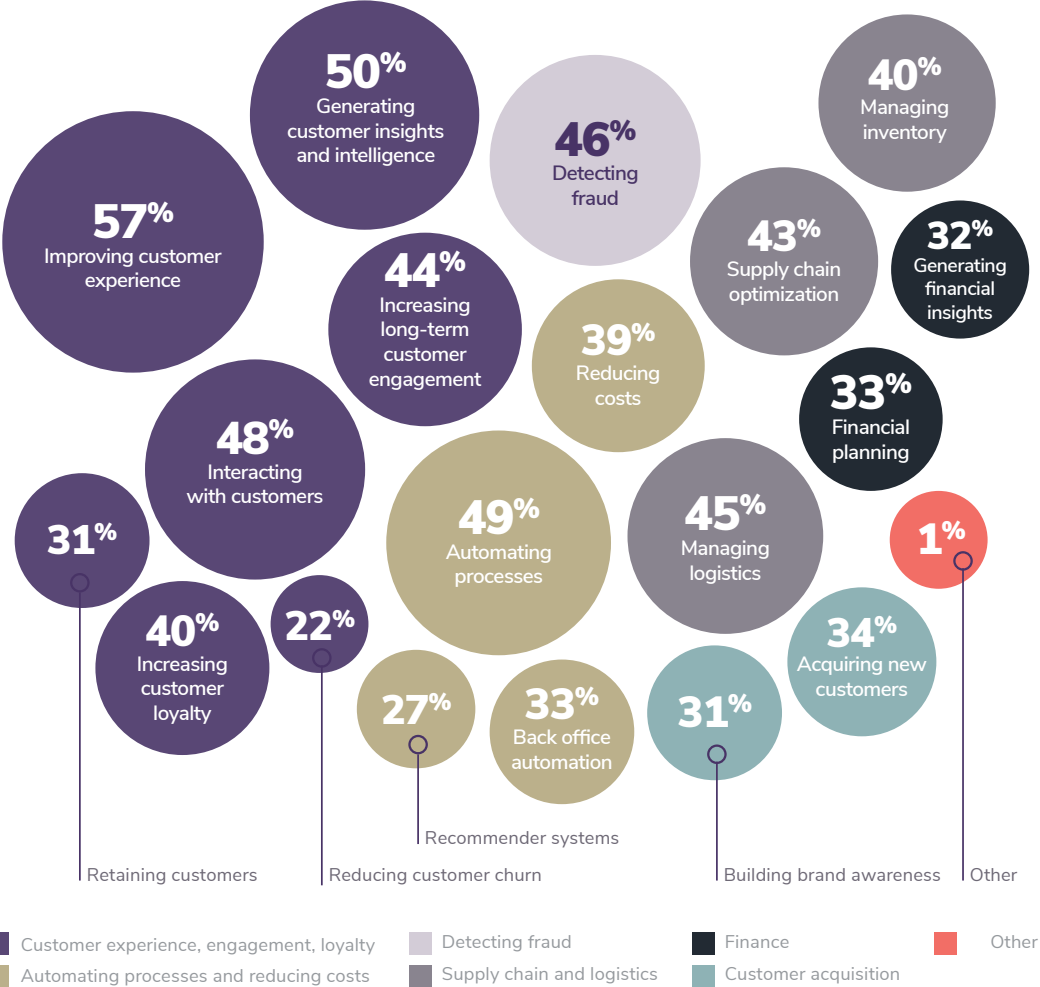
In 2020, respondents were asked to indicate the number of data scientists employed at their organizations through a free-form response field. In 2021, respondents were instead asked to indicate the number of data scientists employed based on predefined ranges, such as “20-49”. To create an accurate year-on-year comparison, 2020 data was first aggregated into the same predefined ranges as 2021. Then, the average number of data scientists employed was estimated by using the lower bound of data scientists in each group (for example, 20 for “20-49”) and averaging across respondents. The percent difference was calculated with the underlying data before rounding to the nearest percentage point. Categories do not add up to 100% because they have been rounded to the nearest percentage point.

It's clear that the organizations making these moves in 2020—prioritizing AI/ML initiatives, increasing spending, and expanding data science teams—expect their strategies will begin producing results sooner rather than later. 2021 is shaping up to be extremely competitive in this regard, as the organizations making serious investments in AI/ML will quickly outpace those still sitting on the sidelines. And these forward-thinking organizations are putting particular emphasis on projects that help them meet the expectations of the modern customer while still keeping their costs under control.

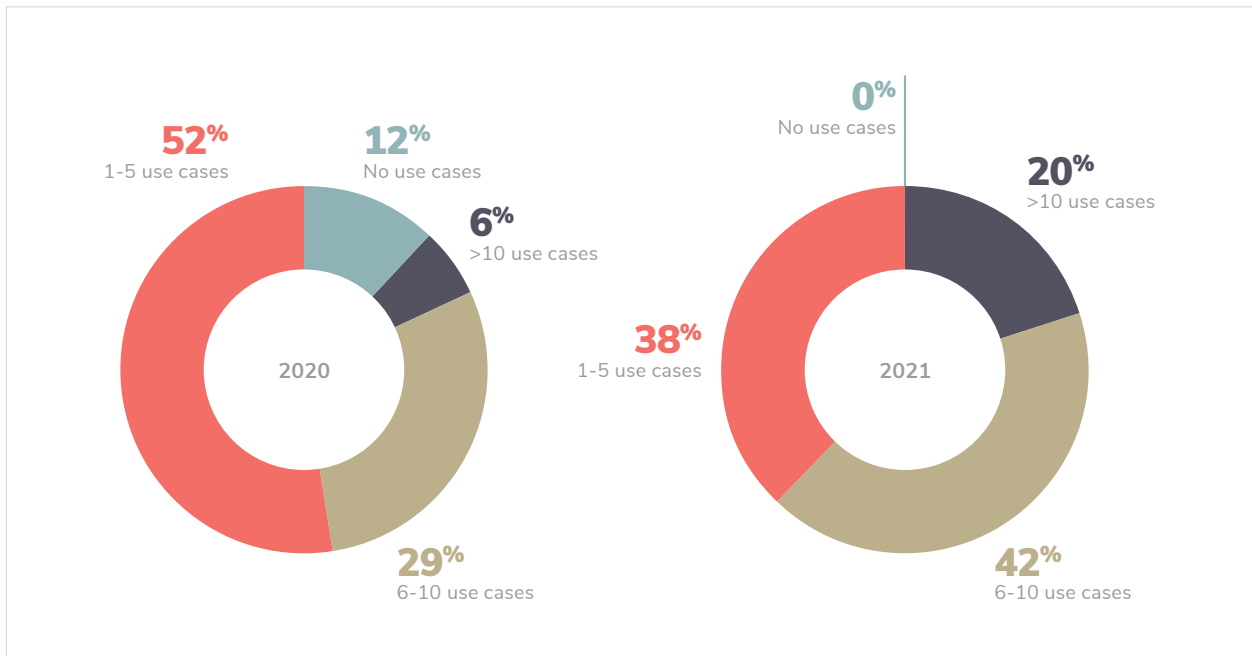
Trend 2: Organizations are expanding into a wider range of AI/ML use cases, with particular focus on process automation and customer experience

Another trend that surfaced in our summer survey and became more pronounced in our 2021 survey data is that organizations are focusing on AI/ML use cases that will reduce costs while improving the customer experience. When respondents were asked about the different ways they're applying AI/ML in their organizations, customer experience and process automation rose to the top as some of the most common use cases respondents selected. We also saw a dramatic (74%) year-on-year increase in organizations that selected more than five use cases from the list of options in the survey.

Customer experience and process automation represent the top AI/ML use cases



The percentage of respondents who indicated they have more than five AI/ML use cases has increased 74% year-on-year



Respondents were asked to select their use cases for AI/ML from a list of options, from which we calculated the total number of indicated use cases per respondent. Respondents were given slightly different options for use cases between 2020 and 2021. See methodology section for a list of use case options per year. The year-on-year percent difference was calculated with the underlying data before being rounded to the nearest percentage point. Categories do not add up to 100% because they have been rounded to the nearest percentage point.

What's more, most organizations are overwhelmingly increasing usage for all of these applications of AI/ML, but especially for customer experience improvement and process automation.

“

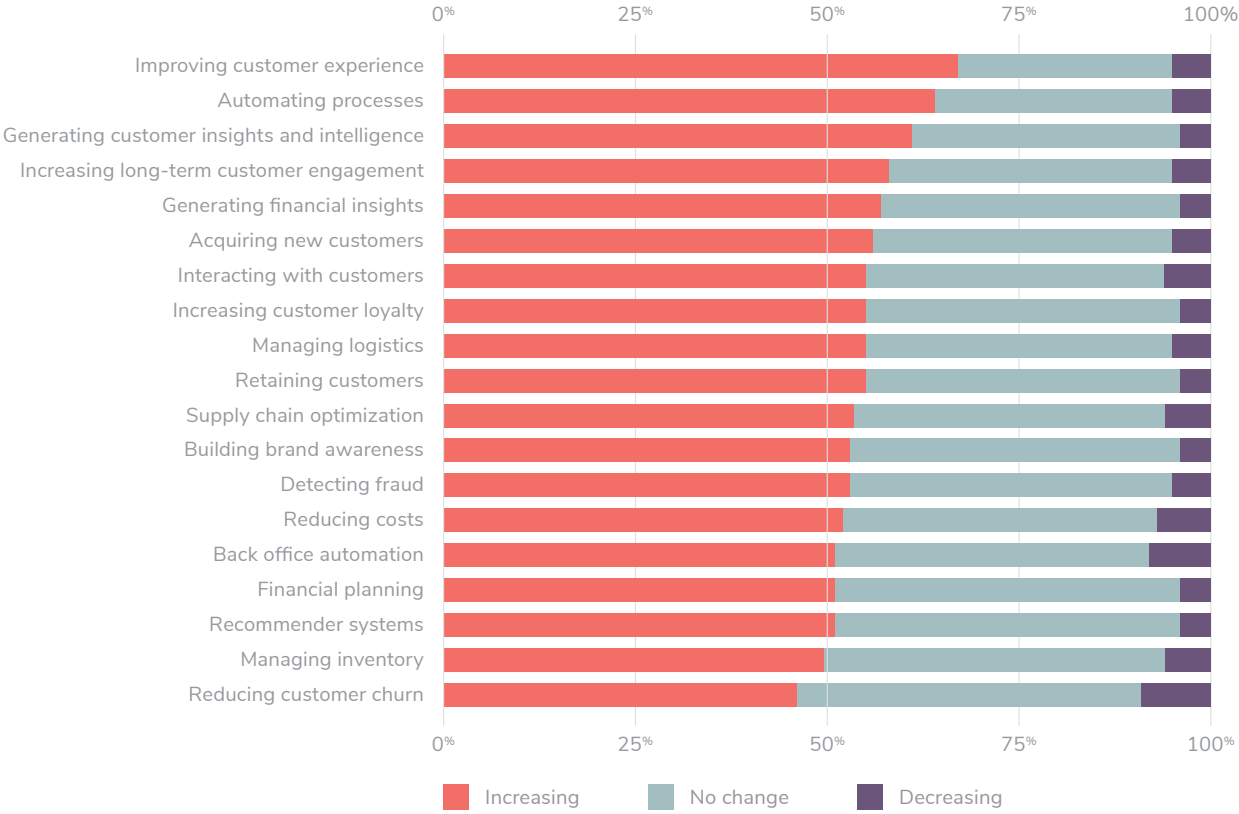
It's illuminating to see what people are doing today with machine learning and the impact it's having on their organizations. When applied to automating processes, it's saving companies money by allowing them to execute at a larger scale with cheaper costs. Machine learning is also having an impact at the top line, focusing on serving existing customers better and driving customer acquisition. That's where you get the real return on investment. And that's why we targeted Algorithmia where we did, to help these organizations unlock value by moving models to production faster, at scale and cost efficiently.”



Kenny Daniel
Founder and CTO, Algorithmia

ALGORITHMIA

For nearly all use cases, 50% or more of organizations are increasing their usage of AI/ML



Respondents were asked to select how usage is changing for different use cases of AI/ML. Per each use case, this chart only includes respondents who have that particular use case.

What does it all mean? The upheaval of 2020 has forced companies to be laser-focused on their most important priorities, including AI/ML initiatives. According to the [VUCA framework](#), we’re simultaneously experiencing volatility, uncertainty, complexity, and ambiguity. This underlines the urgency of pursuing AI/ML use cases that have a clear business ROI. The luxury of doing an AI/ML project for the sake of it no longer exists.

This is why we’re seeing such emphasis on use cases that focus on customer experience and process automation—they both have direct lines to tangible ROI. AI/ML use cases that enhance customer experience drive top-line growth by capturing new customers and retaining existing customers. Process automation generates short-term savings that boost the bottom line.

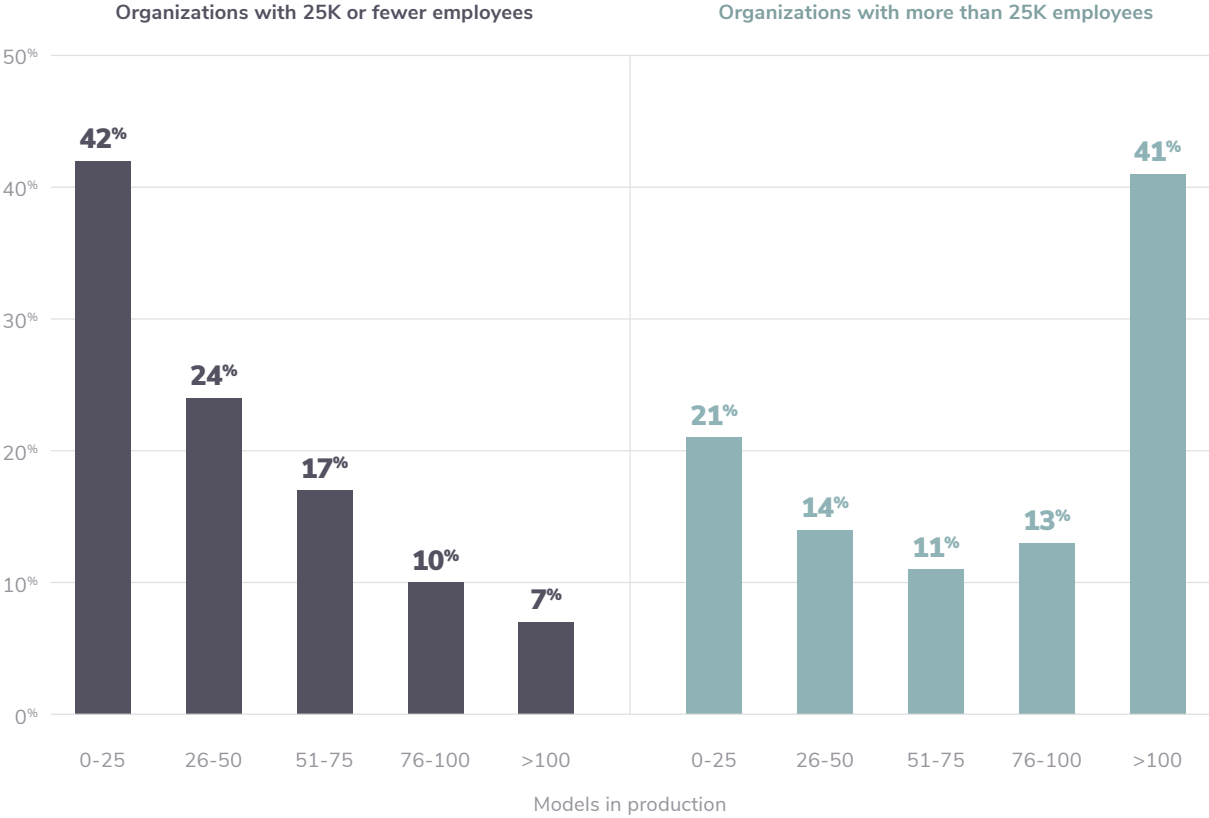
As organizations achieve tangible ROI wins in areas like these, they also build momentum for additional innovation. We believe this is a key reason why we’re seeing organizations expand into a wider range of use cases. These organizations are clearly moving past initial experimentation and scaling up their efforts to achieve significant ROI.

Trend 3: Most organizations have more than 25 models in production, but there’s a gap between AI/ML “haves” and “have-nots”

First, some good news: The average organization has already deployed a significant number of ML models to production. More than half of all respondents have more than 25 models in production, and 40% have more than 50.

However, when we compare the largest enterprises in our sample (organizations with more than 25K employees) with everyone else, we see a jarring divide. Only 7% of organizations with 25K or fewer employees have more than 100 models in production, but among the very largest enterprises, 41% have more than 100 models.

The world’s largest enterprises are dominating the high end of model scale



”

The tsunami of data generated by digital business can no longer be addressed with human scale solutions. Organizations must move faster, and adopt new, ML and cloud-scale services to deliver the real-time continuous intelligence they need to build and secure great customer experiences. Don't wait, start applying these new technologies today... experiment, learn, improve and move forward.”



Dione Hedgepeth
Chief Customer Officer, Sumo Logic

sumo logic

There's a massive opportunity to capitalize on the next wave of AI/ML adoption in 2021.

As a previous [McKinsey report](#) found, the world's largest technology firms have benefited from both vast resources and early adoption. These companies have collectively spent billions of dollars on R&D and other investment areas, and have been doing so for years. That gives them a significant head-start.

That's not to say that these companies are the only ones with the right approach. Rather, it illustrates the importance of adopting AI/ML as early as possible and making meaningful investments in its long-term success.

As McKinsey wrote in its report, “evidence suggests that AI can deliver real value to serious adopters and can be a powerful force for disruption.” Moreover, “early AI adopters that combine strong digital capability with proactive strategies have higher profit margins and expect the performance gap with other firms to widen in the future.”

Some more good news: The barrier to entry is much lower today than it was when the earliest adopters of AI/ML began investing in the technology. We're on the cusp of a second wave of adoption that is far more accessible to organizations of all sizes and in all industries. AI/ML is no longer the sole purview of technology titans with vast resources.

But a new gap may now be forming, and not between the largest early adopters and everyone else. Rather, the next divide is occurring (and growing) between enterprises that are making tangible AI/ML investments now and those that continue to take little or no action.

There's a massive opportunity to capitalize on the new next wave of adoption, but not everyone is paying attention. Given that a growing number of organizations of all types are prioritizing AI/ML—with increasing budgets and staff to match—the message is clear: The time to act is now.

Report theme 2: Organizations continue to face challenges across the ML lifecycle, and struggle with cross-functional alignment

When organizations struggle to adopt and scale AI/ML, the reasons why have become increasingly clear. Success across the entire ML lifecycle requires a high level of [machine learning operations](#) (MLOps) maturity. We've outlined the path to that maturity in a [whitepaper on the topic](#).

”

Algorithmia's 2021 report clearly demonstrates the rising importance of AI for companies across industries. Google was an early mover in this space and solved many operational challenges associated with deploying ML models at scale. These challenges still exist, and all organizations looking to adopt AI will have to solve them. Today, however, companies can access pre-built solutions without having to grow this expertise in-house. Organizations that can cleverly take advantage of these solutions and leverage AI will find competitive advantages.



Anna Patterson
Founder and Managing Partner,
Gradient Ventures at Google



Trend 4: Governance is by far the top challenge for AI/ML deployment, with more than half of all organizations ranking it as a concern

When we asked survey respondents to indicate the challenges they face when deploying models, we saw a wide range of difficulties that occur across the ML lifecycle. However, the most common challenge by far was with [AI/ML governance](#), an issue that is primarily a concern at the end of the ML lifecycle when models have already been developed and organizations need to minimize their risk.

“

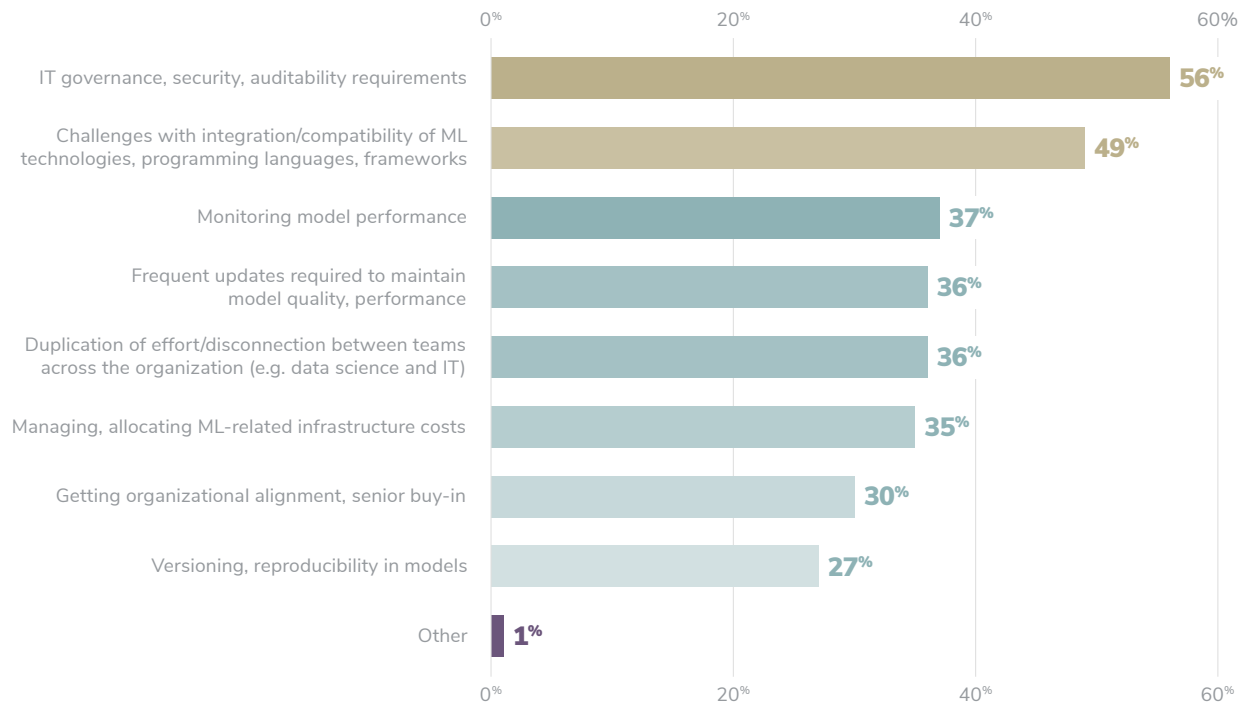
The opportunity for AI is to allow banks to provide services in much more personalized, highly scalable, and customized ways. The challenges include the ability to explain your AI and to ensure confidentiality, since a lot of the data in finance is personal information or highly confidential. User trust is key. The person on the other side wants to trust you with their most valuable assets, and with their most valuable information. The challenge for MLOps is making governance and security highly scalable in a time when investment in AI and the number of production models we are seeing is accelerating.”



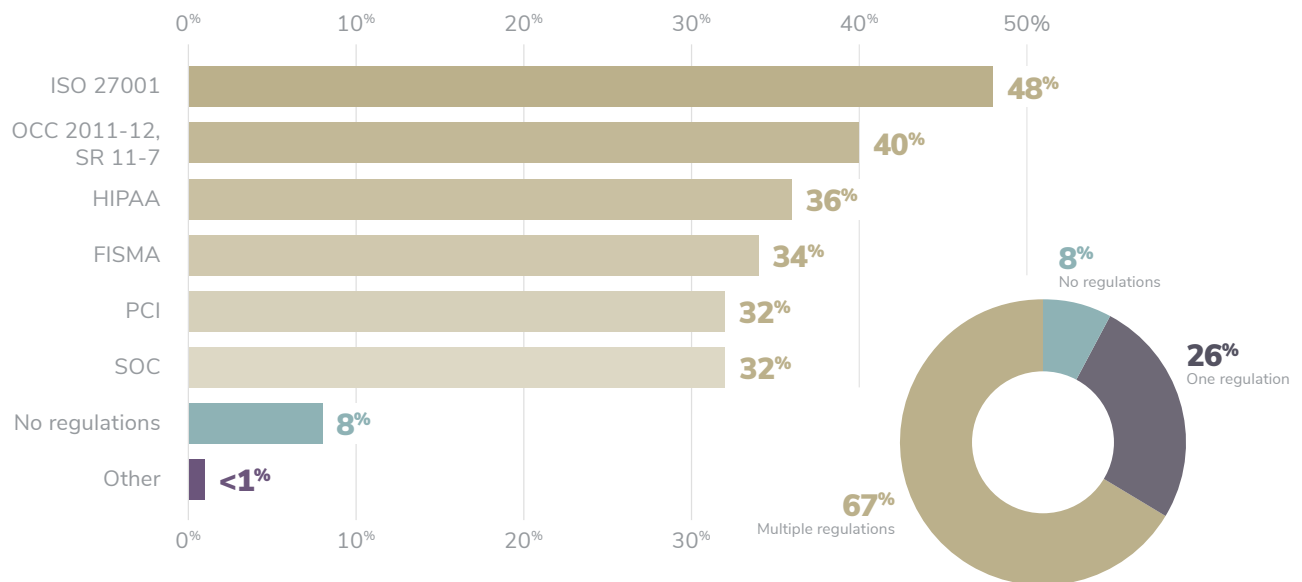
Apoorv Saxena

Global Head of AI Technology, JP Morgan Chase & Co

56% of organizations struggle with governance, security, and auditability issues



67% of all organizations must comply with multiple regulations



Respondents were asked to identify regulations that their organizations are required to comply with. This chart shows the percentage of respondents who selected each regulation. Respondents who didn't select any regulations are shown in the "None" category. For respondents who selected "Other", "Other" was counted as a regulation. In the donut chart, categories do not add up to 100% because they have been rounded to the nearest percentage point.

Respondents also indicated that they face a wide range of regulatory requirements, from SOC to HIPAA to PCI. Nearly every organization is dealing with some level of compliance burden; when asked which regulations they needed to comply with, 67% of respondents selected multiple regulations and only 8% didn't select a single regulation.

Since governance is a broad, longstanding term in IT, it's important to give it clear meaning in this specific context. We define AI/ML governance as the overall process for how an organization controls access, implements policy, and tracks activity for models. From a regulatory compliance standpoint, governance is a must-have to minimize organizational risk in the event of an audit.

But while regulatory compliance is important, it's only one part of the governance equation. Effective governance is the bedrock for minimizing risk to both an organization's bottom line and to its brand. Organizations with effective AI/ML governance not only have a fine-grained level of control and visibility into how models operate in production, but they unlock operational efficiencies by integrating AI/ML governance policies with the rest of their IT policies. They can document and version models, tracking both the inputs and outputs of those models to understand all the variables that might affect model results. This enables organizations to quickly identify and mitigate issues such as model drift that degrade the accuracy of results and the performance of applications—issues that can directly impact the business' bottom line and erode customer trust in the brand over time.

Many organizations (more than half of those surveyed) are struggling with this important aspect of AI/ML strategy. It's also reasonable to assume that some organizations don't yet realize they're struggling with it; governance is a late-lifecycle concern, so it doesn't always become obvious to organizations until they're in the later stages of the ML lifecycle. As organizations head into 2021, they should consider whether they need to focus more on AI/ML governance, even if it's not currently a top concern.

”

Data scientists have traditionally been forced to enlist developers to help create custom solutions for operationalizing and monitoring ML model performance. As the number of models grows, so do the number of people involved and the complexity of operating and maintaining these custom solutions with cross-functional teams. This forces a reexamination of the entire approach. To address this challenge, organizations should look to products that integrate operational tooling with off-the-shelf monitoring infrastructure so that data scientists rapidly assemble ML monitoring solutions on their own. Simplifying the monitoring process has the potential to eliminate model performance as a widespread challenge.”



Tim Hall
VP Products, InfluxData



Trend 5: The second greatest AI/ML challenge is technology integration and compatibility, with 49% of organizations ranking it as a concern

While governance topped the list of AI/ML issues, there was the clear runner-up: 49% of respondents said they experience challenges with the integration or compatibility of their ML technologies, programming languages, or frameworks. This means organizations are still stumbling at the beginning of the AI/ML lifecycle.

It's not without reason: ML technology is unique and always developing. It's not always ready for plug-and-play integrations with other systems. Specialized hardware, such as GPUs, is critical to modern machine learning, but presents a number of compatibility and software challenges.

The bottom line? This is still a relatively young industry and the technology is rapidly evolving. Staying current with the AI/ML technology landscape requires regular monitoring and updates. That creates a burden on organizations and their data science teams, especially those that take a do-it-yourself approach.

“

I think this survey shines a light on what I would call the 'unsexy part of MLOps' in that, more teams cited major blockers to deploying models being security, governance, and audibility.”



Demetrios Brinkmann
Community Coordinator MLOps Community

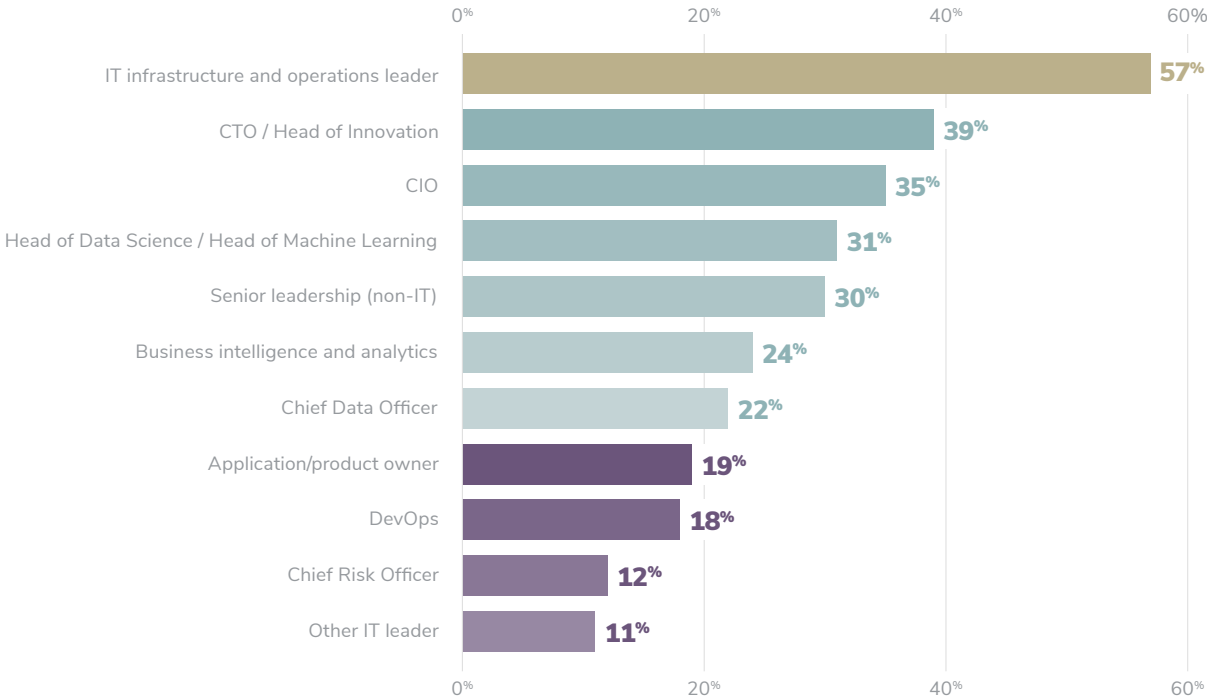


Trend 6: Successful AI/ML initiatives require organizational alignment across multiple decision-makers and business functions

Indeed, as the stakes and complexity around AI/ML initiatives grow, so does the need for collaborative, cross-functional AI/ML teams that span departments. Strong, sustained results absolutely depend on this team-building approach.

We asked survey participants to indicate who sets the priorities for AI/ML in their organizations, and got widely diverse responses—though infrastructure and operations (I&O) leaders clearly rose to the top, with CTOs, CIOs, and heads of data science following closely behind.

Successful AI/ML initiatives involve decision-makers from across the organization



“

I have always found that when C-level executives are behind a technology initiative then it is more likely to succeed. Any deep technology initiative is likely to have change rippling across the organization, and C-level executives can enable the change management. So interesting to see CTO and CIO in second and third place: I believe those organizations are more likely to succeed with AI than others.”



Michael Azoff
Chief Analyst, Kisaco Research



One of the key takeaways that we see here is that AI/ML is no longer the sole or primary purview of data scientists. Rather, a much broader cross-section of roles are leading or becoming involved with AI/ML strategy and operations. Consider the different domains and responsibilities of these roles:

Infrastructure and operations

I&O is a well-established role in IT. As the name suggests, I&O is essentially the backbone that makes fast, secure, and scalable systems possible. I&O pros typically take the lead on integrating new technologies and initiatives into existing systems and processes in an organization, such as the SDLC or CI/CD pipeline.

This means that I&O leaders are also increasingly responsible for scaling AI/ML initiatives in production. They need to be convinced that AI/ML initiatives don't introduce new risks or operational burdens. We believe governance is becoming an increasing issue in AI/ML partially because I&O leaders (rightly) take it seriously and are playing higher-profile roles in their organizations' AI/ML strategies.

CIO

This is the top technical leadership role at most companies; this person and their direct reports are often responsible for driving digital transformation within the company. CIOs are also increasingly strategic and collaborative, meaning they're working closely with their C-suite counterparts on the organization's most pressing goals. They're also focused on AI/ML—and how to measure its success—from a leadership level.

Head of data science

This is an emerging role with a fluid definition, but the head of data science is generally responsible for unlocking the value in an organization's data for the business. The head of data science is responsible for leading the team that owns the organization's AI/ML capabilities and initiatives, and is responsible for actually developing them and moving them forward. And the value proposition of this role is just beginning to come to fruition, as cross-discipline teams work together to pursue quantifiable business value from their data.

We believe one reason why tool and technology integration is a common concern is because this issue falls within the purview of data science teams, who continue to be key players in AI/ML decision-making.

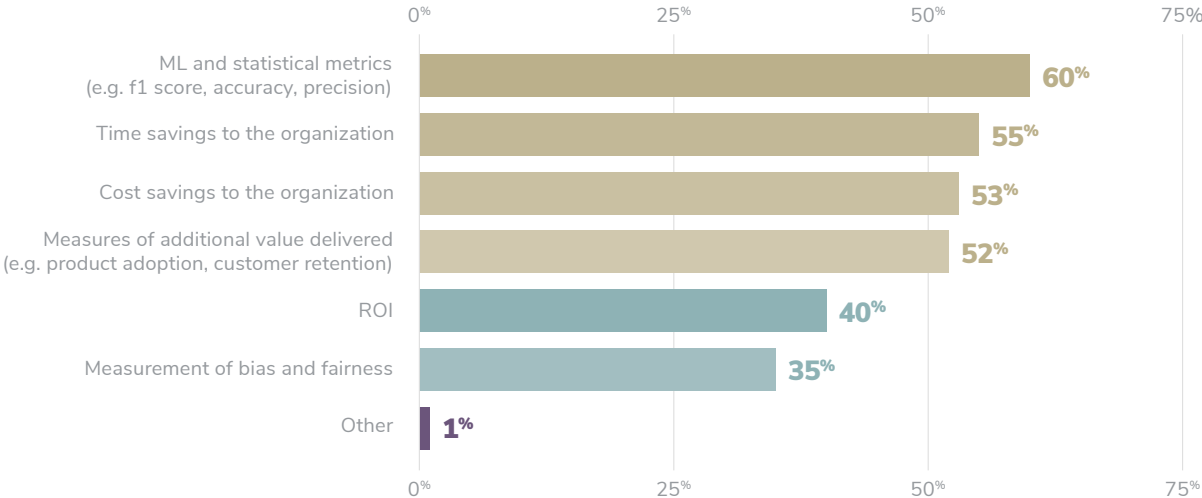
Business team

Business teams are those who create the products that actually use machine learning on the back end. These professionals are responsible for driving market growth and helping the organization maintain competitive edge, so they are commonly focused on the customer-centric use cases mentioned previously.

As you can see, there are now a wide range of roles and even organizational levels—from the C suite on down—involved in AI/ML strategy and implementation, though I&O has become the major leader in setting priorities.

The cross-functionality of AI/ML is also evident in the metrics organizations are using for success. We asked organizations to indicate which metrics they use to evaluate the success of their AI/ML initiatives, and got a wide range of responses.

Organizations are using a variety of success metrics for AI/ML initiatives



“

In our conversations with F1000 CIOs and their teams, we find that they explore ML technologies and ML use cases to gain a competitive advantage, and justify their investments in the current term by communicating the efficiencies gained and the abilities to drive better customer intimacy. Ultimately, as they crack the code and understand how best to capitalize on their efforts, they begin to shift the justification of their investments through business outcomes. [This report] is an outstanding visualization of the shift and recognizes that organizations who lean forward, and use disruptive technologies to thrive through disruptive times, will outlast organizations that remain risk- and change-averse.

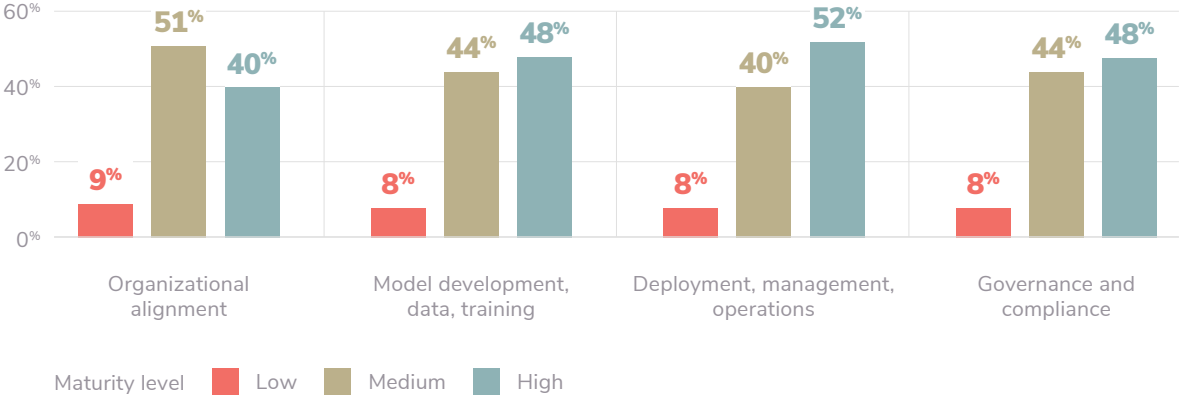


Todd Tucker
VP, Standards and Education, Technology Business Management Council

Trend 7: Organizational alignment is the biggest gap in achieving AI/ML maturity

Despite the cross-functionality of AI/ML, when we asked organizations to rank their current level of maturity against the maturity framework we outline in [our whitepaper](#), organizational alignment was the weakest point. While in most areas, a greater number of organizations ranked themselves as high maturity than medium maturity, this trend was reversed when it comes to organizational alignment.

Organizational alignment is the biggest gap in achieving AI/ML maturity



It's interesting that this was the lowest-ranked area of maturity in respondents' self-assessments, because in fact it's the most foundational area for success. Not only does maturity in this area enable maturity in more complex and technical areas (such as governance), but it is simply crucial to the success of any project due to the cross-functional nature of AI/ML discussed above. Organizations that ignore this critical need will diminish their potential ROI.

“

I believe that the competitive advantage of AI exists at the intersection of data, technology, and people.

You can only **get value from your AI** when it's woven with deep business accountability and expertise. In order for enterprises to unlock the real power of AI and realize its value, they must first **build a strong partnership between data science and the business.**”



Nir Kaldero

Chief Data, Analytics, and AI, NEORIS
Adjunct Executive, Global Data Science, CEMEX

Report theme 3: Despite increased budgets and hiring, organizations are spending more time and resources—not less—on model deployment

Clearly, the economic disruption of the pandemic has brought a renewed sense of importance to AI/ML initiatives in many organizations. As explained previously, organizations are moving AI/ML up their priority lists and increasing budgets and staffing, too.

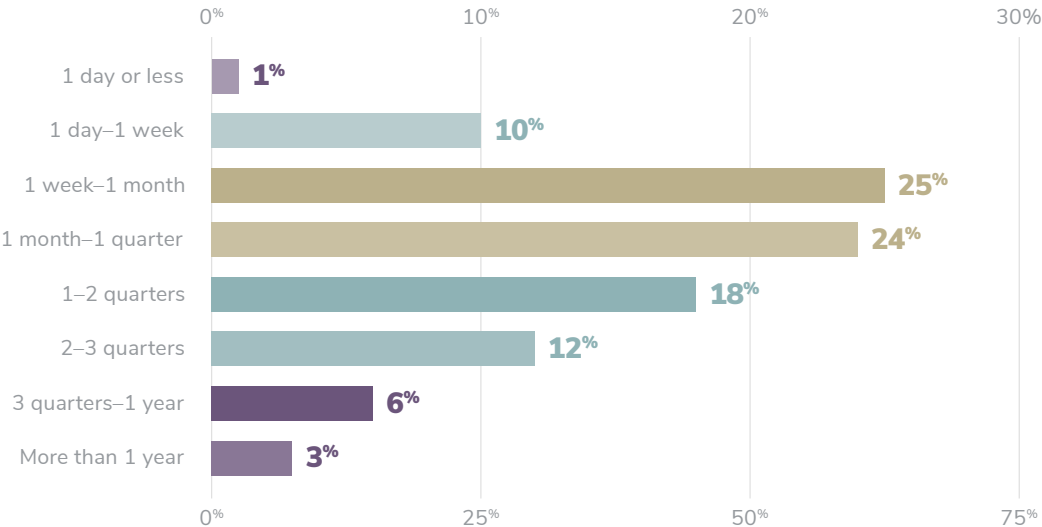
However, this doesn't mean organizations can solve all the challenges they encounter in their AI/ML work overnight. In fact, our survey revealed that as organizations increase their AI/ML investments, they are actually spending more time and resources on model deployment, not less. We believe that organizations are using their increased headcounts to manually scale AI/ML efforts rather than addressing underlying issues with operational efficiency. But a growing AI/ML staff will have a much bigger impact if it's able to focus on data science instead of constantly paying down operational overhead.

Trend 8: The time required to deploy a model is increasing, with 64% of all organizations taking a month or longer

Our data shows that the total timeline required to develop and deploy a model is significant. First off, we polled respondents about the average amount of time it takes them to develop a trained model once they've actually defined a use case (a new question in this year's survey). A majority (66%) reported that this process takes them a month or longer to complete.

Then, once a trained model has been developed, it must still be put into scaled production. When we asked respondents about the average amount of time it takes them to do that, only 11% indicated they can deploy a trained model to production within a week. The majority of organizations (64%) are taking a month or longer.

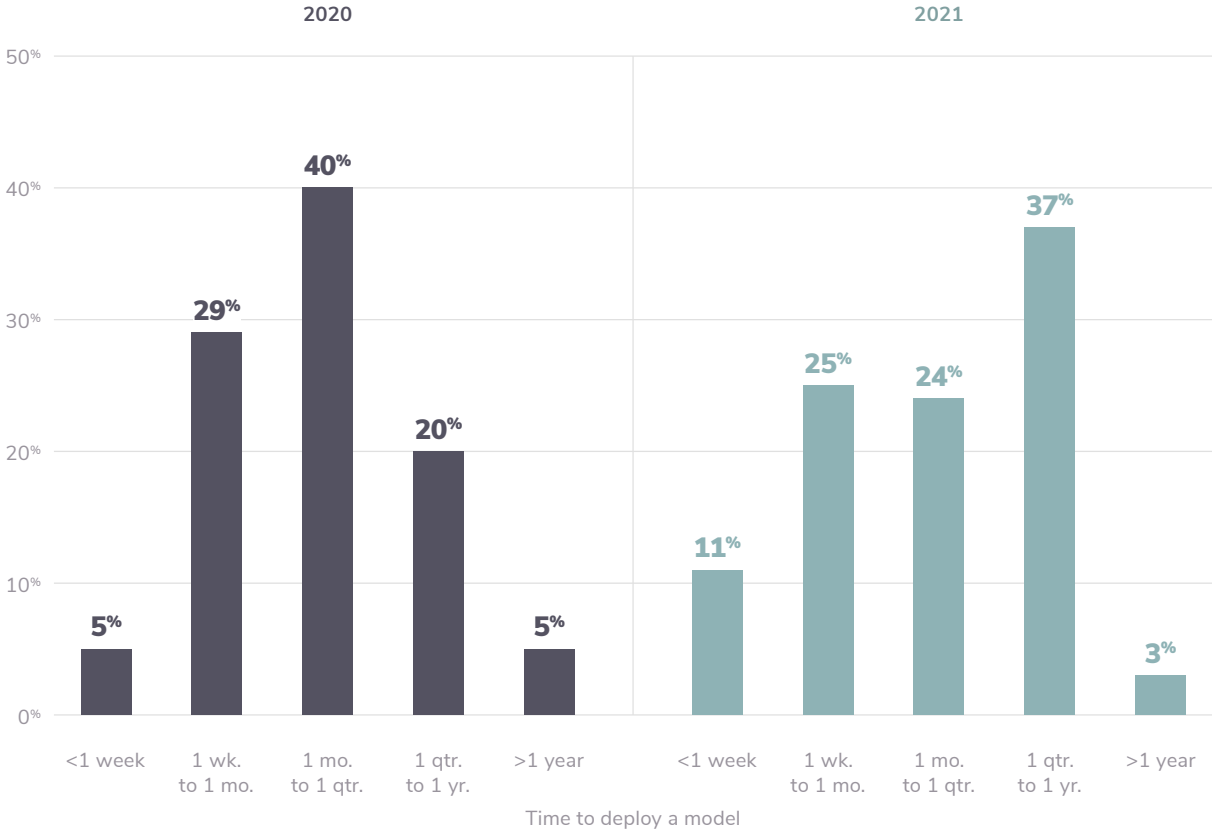
Only 11% of organizations can put a model into production within a week, and 64% take a month or longer



The total percentage of respondents who selected a month or longer was calculated with the underlying data before being rounded to the nearest percentage point. Categories do not add up to 100% because they have been rounded to the nearest percentage point.

This means that once organizations define a use case for AI/ML, they still face a significant timeline—months to even years—to get to the point where they have a trained model developed for it and scaled to production. What’s more, the time required to deploy a model (once it’s been developed) is actually getting longer. Since we asked respondents this question in last year’s survey too, we were able to compare the results, and found that it is indeed increasing year-on-year.

The time required to deploy a model is increasing year-on-year

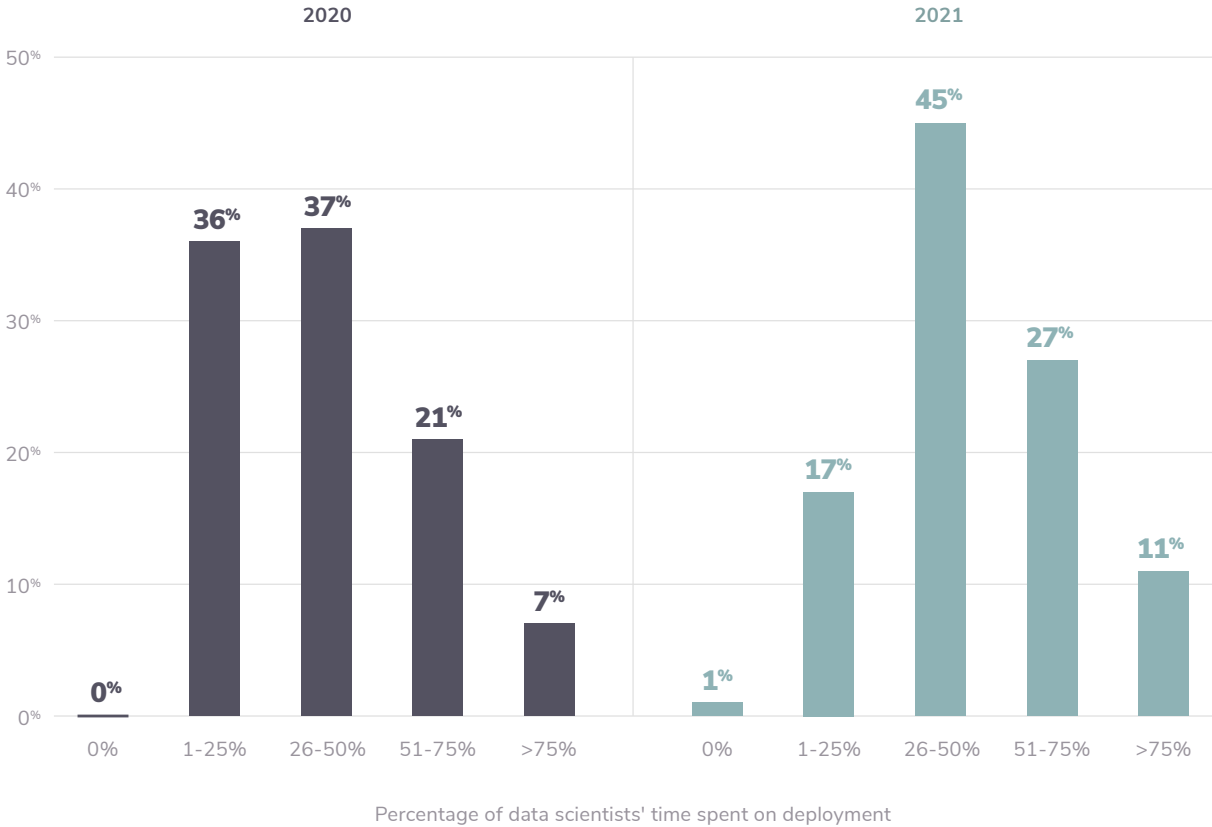


Data from 2020 doesn’t include respondents who selected “I do not know or I am unsure”. Respondents were given slightly different options for time ranges between 2020 and 2021. See the methodology section for a list of time range options per year, and how they were aggregated for accurate year-on-year comparison. Categories do not add up to 100% because they have been rounded to the nearest percentage point.

Trend 9: 38% of organizations spend more than 50% of their data scientists' time on deployment—and that only gets worse with scale

Our 2021 data also shows clearly that data scientists are spending too much of their valuable time on model deployment. We asked respondents how much of their data scientists' time is being spent on deploying models, which we defined in the question as “prepping trained models and deploying them where they can be consumed by apps or used with other models”. Our data showed that a full 38% of organizations are spending more than 50% of their data scientists' time on these tasks.

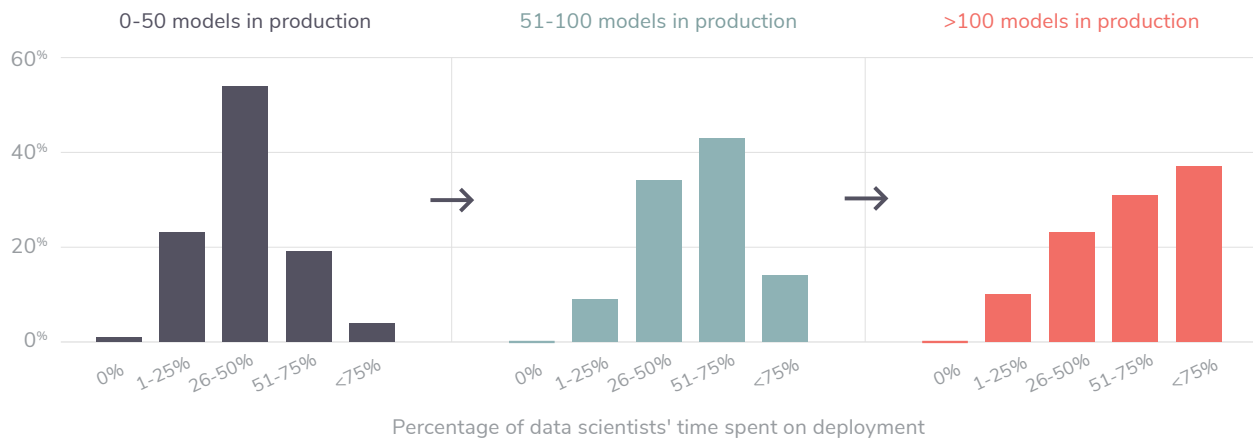
38% of organizations spend more than 50% of their data scientists' time on deployment



Data from 2020 doesn't include respondents who selected “I do not know or I am unsure” or “My team does not/has not deployed models”, or who didn't answer the question. Categories do not add up to 100% because they have been rounded to the nearest percentage point.

The amount of time that data scientists spend on deployment actually increases with the number of models in production too. Our data shows that organizations with the most models in production spend the greatest percentage of their data scientists' time on model deployment.

Organizations with more models spend more of their data scientists' time on deployment, not less



Once again, this implies that organizations are using increased headcount to manually scale AI/ML. Organizations that have scaled their models in this way are significantly limiting their potential ROI and the long-term sustainability of their initiatives. They would be far better served by improving operational efficiency and scale so their data scientists can focus on building innovative models—not performing manual operational tasks.

“

Data scientists are expensive and increasingly hard to find, hire, and retain. Yet companies are letting an increasingly large portion of that scarce capacity go to model deployment, maintenance, and management. Data scientists are neither good at nor do they like doing this stuff. This is just crazy. The companies that get this right will leave the ones that continue on this course in their dust.”



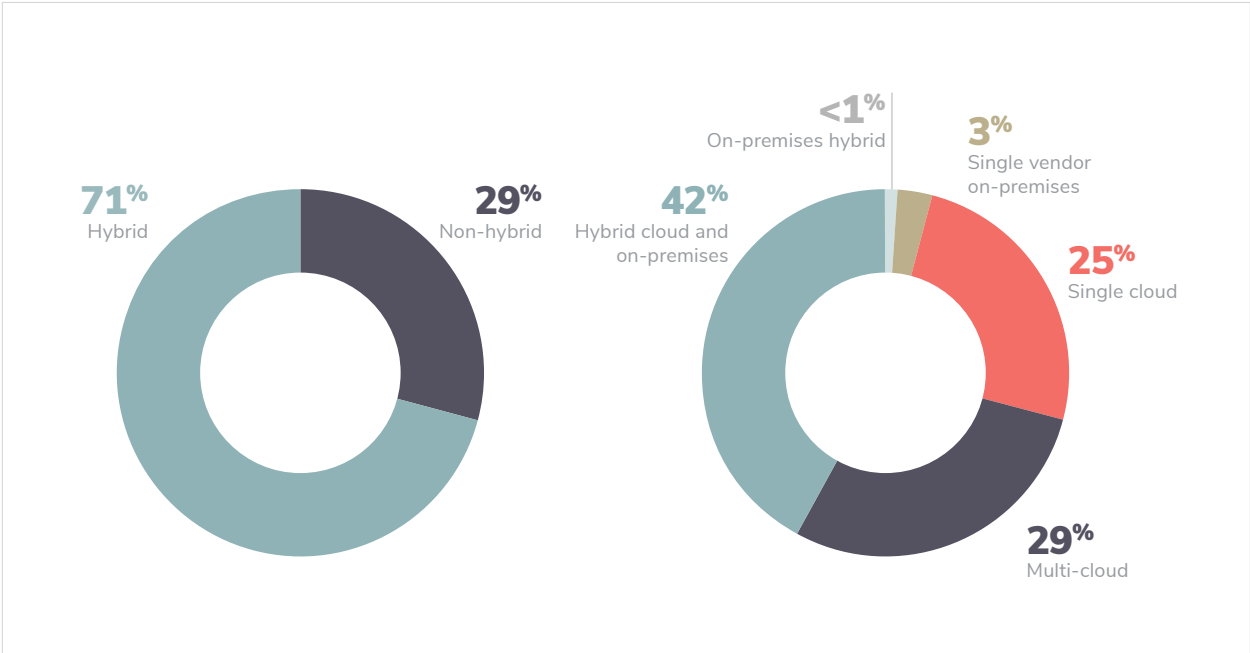
H. P. Bunaes
Founder AI Powered Banking



Report theme 4: With increasingly complex infrastructure needs, organizations report improved outcomes with third-party MLOps solutions

As the AI/ML market matures, we see interesting trends in how organizations are approaching infrastructure. First off, organizations have increasingly complex environments for deploying models. In our 2021 survey, 71% of all respondents indicated that they use a hybrid environment (consisting of more than one cloud or on-premises infrastructure provider) to deploy models, and 42% of all respondents have a hybrid environment consisting of both cloud and on-premises solutions.

71% of all organizations have hybrid environments, and 42% have a combination of cloud and on-premises infrastructure



Excludes respondents who selected "Other" (less than 1% of respondents). Categories do not add up to 100% because they have been rounded to the nearest percentage point.

We also know that hybrid environments that include both cloud and on-premises solutions are becoming more common. After excluding respondents who selected “I do not know or I am unsure” or “Other”, 16% of respondents in the 2020 survey indicated they have a hybrid environment consisting of both cloud and on-premises solutions. This is significantly lower than the 42% of respondents who indicated this in the 2021 survey.

And with increasingly complex environments come increasingly complex infrastructure needs. So how are organizations handling those infrastructure needs? We asked respondents how their organizations approach model deployment and management infrastructure, with four* different options:

1. We build and maintain our own system from scratch
2. We integrate open-source components into a system that is maintained in-house
3. We integrate commercial point solutions into a system that is maintained in-house
4. We use a third-party platform supported by a vendor

“

We see a recurring theme around tooling consolidation: two out of three teams we talk to, say they are building an internal ML platform, which is really their attempt to standardize tooling and process around ML development and deployment. Their effort is typically motivated by a goal to cut model operating costs and the need for transparency over production ML applications.”



Alessya Visnjic
CEO Why Labs
ex-Amazon Machine Learning

WHYLABS

In the early days of AI/ML, any organization that wanted to deploy models at scale was essentially required to build and maintain their own system from scratch. We see many organizations still gravitating toward this approach, but this group is disproportionately skewed towards organizations with a large number of models. Among organizations with more than 100 models, 60% chose to build and maintain their own systems from scratch, but only 35% made this choice among other organizations. Our hypothesis is that these organizations represent early adopters in the space, who have gravitated toward building their own solutions because it may have been their only choice.

In contrast to the organizations that are building and maintaining their own systems from scratch are those that are using a third-party solution—either integrating commercial point solutions into their systems or using a third-party platform supported by a vendor. For the rest of this report, we will be comparing respondents who selected option #1 above (which we’ll call “build from scratch”) with respondents who selected either option #3 or option #4 (which we’ll call “buy a third-party solution”). This corresponds to the [build vs. buy decision](#) for MLOps, which we’ve written about before.

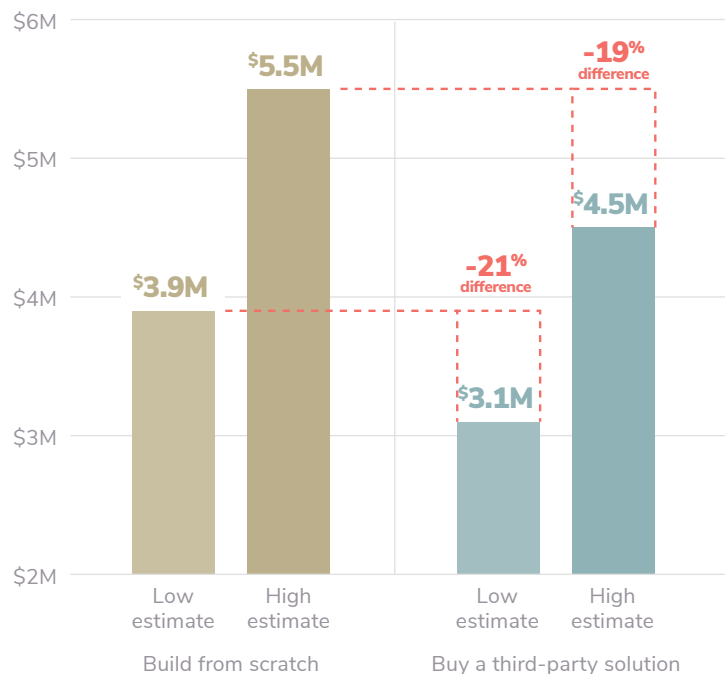
*We also provided a fifth option for “Other,” but it received less than 1% of all responses.

The market has matured significantly since the early-adopter enterprises first began investing in AI/ML. As we've seen, organizations now have increasingly complex infrastructure environments, including hybrid cloud and on-premises environments. Regardless of the specifics, IT infrastructure and applications tend to be far more distributed than they were in the past. A third-party MLOps platform can be better equipped to handle these complex environments, and shave off some of the total infrastructure costs as your models scale. We see this reflected in our 2021 survey, as organizations that buy a third-party solution tend to spend less money on infrastructure and less time on model deployment than organizations that build from scratch.

Trend 10: Organizations that use a third-party MLOps solution save money and spend less time on model deployment than those that build their own solution

Our 2021 data found that buying a third-party solution yields better outcomes in multiple areas than building from scratch. First off, we asked respondents to report their estimated annual infrastructure costs for all models in production. Based on these responses, we created both a low estimate and a high estimate for infrastructure costs an organization might expect. For both the low and high estimates, we saw cost savings for organizations that buy a third-party solution. Indeed, according to our data, organizations that buy a third-party solution spend an average of 19-21% less on infrastructure costs annually—and we have reason to believe the actual cost savings seen in production could be much higher. This is because we used a conservative method to calculate the average infrastructure cost*, which discounted the higher end of the cost spectrum. Since a greater portion of organizations at that high end of the cost spectrum are building their own solutions from scratch, their average costs were likely underestimated.

Buying a third-party solution costs 19-21% less than building your own

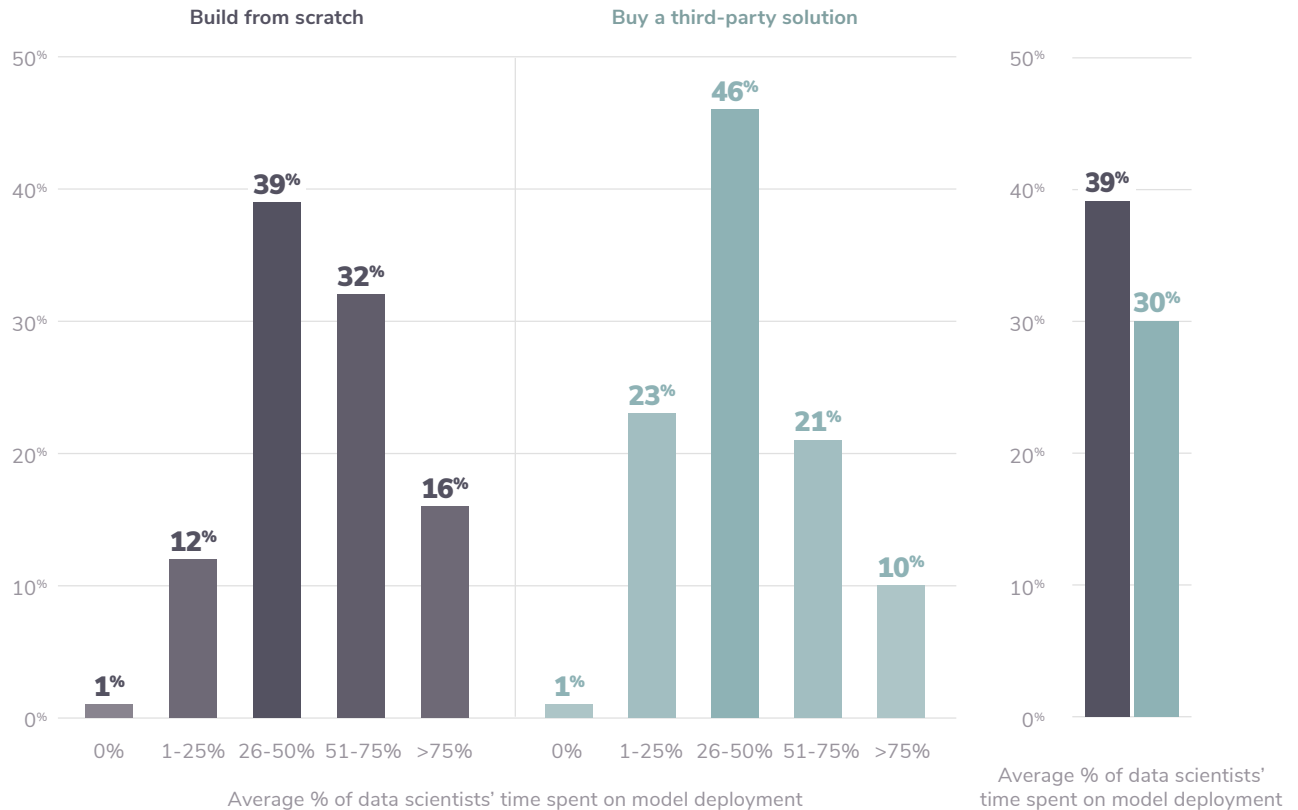


Respondents were asked to indicate their average annual infrastructure costs based on predefined ranges, such as "\$51-\$100K". The total average annual infrastructure cost was then estimated as a range. The low estimate is based on the lower bound for each predefined range (for example, \$51K for "\$51-\$100K"). The high estimate is based on the upper bound for each predefined range (for example, \$100K for "\$51-\$100K"). For the pre-defined range that represented the greatest cost ("more than \$10M"), the lower bound of the range was used for both the high and low estimate. The percent difference was calculated with the underlying data before rounding to the nearest percentage point.

*In both the high and low estimates, all organizations reporting more than \$10M in annual infrastructure spend were assumed to only be spending \$10.000001M.

Organizations that buy a third-party solution also tend to spend less time on model deployment. When we calculated the average percentage of data scientist time spent on model deployment, it was lower for organizations that buy a third-party solution rather than building from scratch.

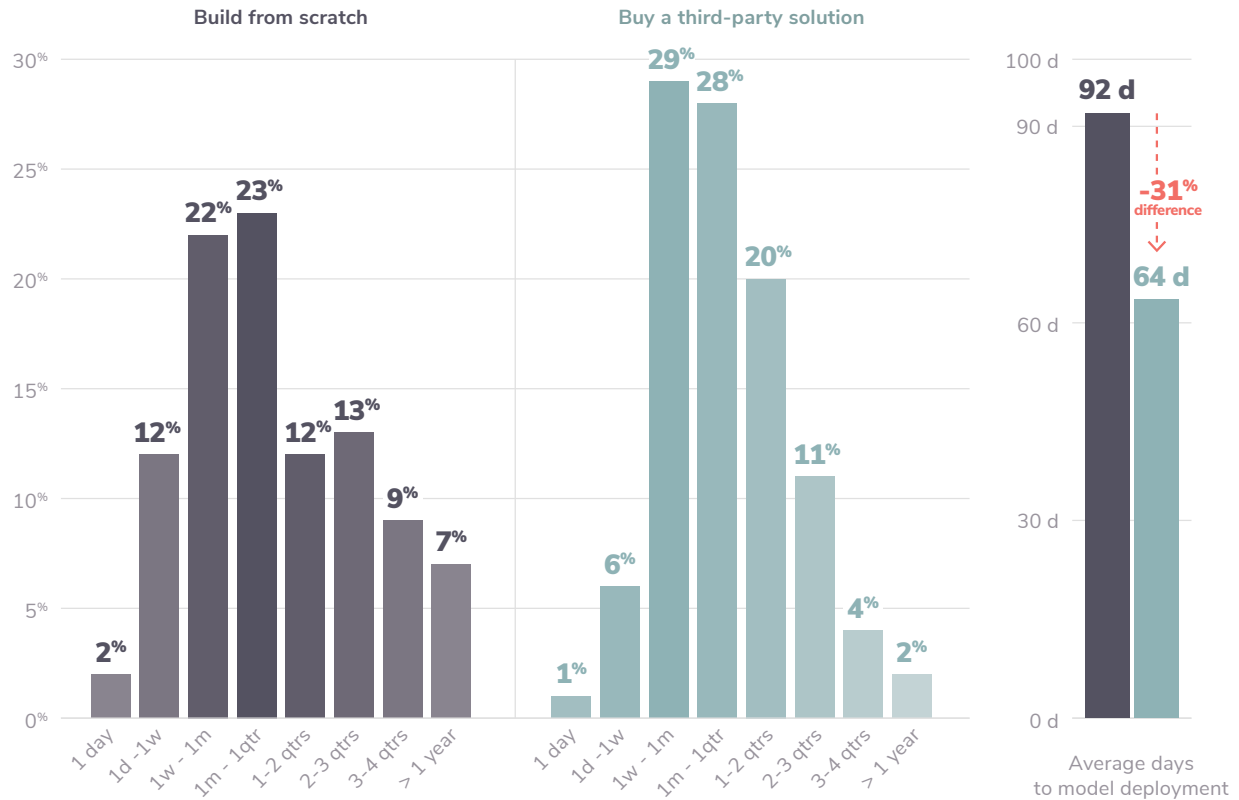
Organizations that buy a third-party solution spend less of their data scientists' time on model deployment



Respondents were asked to indicate the average percentage of their data scientists' time spent on model deployment based on predefined ranges, such as "26-50%". The total average percentage of data scientist time spent on model deployment was then estimated by using the lower bound of the data in each range (for example, 26% for "26-50%") and averaging across respondents. Categories do not add up to 100% because they have been rounded to the nearest percentage point.

Lastly, these organizations tend to get their models into production more rapidly. On average, the number of days it takes them to put a trained model into scaled production is 31% lower than for organizations that build from scratch.

The time required to deploy a model is 31% lower for organizations that buy a third-party solution



Respondents were asked to indicate the average time it takes to deploy a model based on predefined ranges, such as “1 month-1 quarter”. The total average time to deploy a model was then estimated by using the lower bound of the data in each range (for example, 1 month for “1 month-1 quarter”) and averaging across respondents. For comparability between ranges, all data was converted to days before averaging it, with 1 month being equal to 30 days and 1 quarter being equal to 90 days. The percent difference was calculated with the underlying data before rounding to the nearest percentage point. Categories do not add up to 100% because they have been rounded to the nearest percentage point.

As organizations face increasingly complex infrastructure needs, third-party MLOps solutions can help organizations save on infrastructure costs while also helping to solve one of the most prominent problems organizations currently face: Speeding up model deployment while reducing the operational burden on data science teams. Third-party solutions can also lower the barrier to entry for AI/ML. Truly, this field is no longer limited to enterprises that can invest in building and maintaining their own infrastructure entirely from scratch.

”

Today MLOps doesn't need any introduction—it's central to the AI strategy of any organization. Organizations are now moving away from custom open-source tooling to more standardized platforms for their model management, which will make this space even more competitive. Looking forward, what's even more interesting is how organizations are handling complex issues of ethics, explainability, and bias as they roll out more AI. These issues have far-reaching impacts on the way models are built and governed, and should be important considerations for any organization using AI.”

Sacin Porwal

General Manager – AI Solution Engineering, Wipro

Conclusion

2021 will be a crucial year for AI/ML initiatives. There are plenty of reasons for optimism, including a new sense of urgency and importance about AI/ML within many organizations, plus growing investments in terms of prioritization, spending, and staffing. And AI/ML is far more accessible than ever before. You no longer need to build and maintain your own infrastructure from scratch just to get started with AI/ML. Organizations of all sizes, industries, and infrastructure needs can now get started with AI/ML more quickly, or scale their existing AI/ML efforts with greater ease.

What's also clear is that the organizations that will reap the greatest benefits from AI/ML in 2021 are those that invest in operational efficiency and scale. Those organizations will be able to more effectively direct their AI/ML investments to the efforts that drive the most significant top- and bottom-line impacts for their businesses.

2021 will certainly be a year when the gap grows between those organizations that take bold steps to scale their AI/ML initiatives, and those that get mired in operational and organizational issues. It's time to act—your AI/ML future depends on it.

“

COVID-19 has caused rapid change which has challenged our assumptions in many areas. In this rapidly changing environment, organizations are rethinking their investments and seeing **the importance of AI/ML to drive revenue and efficiency during uncertain times.**

Before the pandemic, the top concern for organizations pursuing AI/ML initiatives was a lack of skilled in-house talent. Today, organizations are worrying more about how to get ML models into production faster and how to ensure their performance over time.

While we don't want to marginalize these issues, I am encouraged by the fact that the type of challenges have more to do with how to **maximize the value of AI/ML investments** as opposed to whether or not a company can pursue them at all.”



Diego Oppenheimer

CEO and Founder, Algorithmia

Methodology

The purpose of Algorithmia's 2021 *enterprise trends in machine learning* report is to report on the latest developments and trends in enterprise machine learning and how they have evolved over the past year. The report is based on data that Algorithmia collected in November 2020 in a survey effort that returned 403 responses.

The survey asked 29 questions about AI/ML initiatives, challenges, infrastructure, company demographics, and more. The survey questions were developed by Algorithmia, and an independent third-party company conducted the survey on Algorithmia's behalf to ensure survey attribution anonymity and remove bias for or against Algorithmia on the part of the respondents.

Respondents voluntarily participated in the survey in exchange for access to content or a service, such as free Wi-Fi. Respondents received no monetary payment for their participation.

The third party screened participants using the following questions:

- What is your company size? (Only respondents at companies with \$100M+ in revenue were included)
- Which best describes your role? (Respondents with roles of Consultant or Student were excluded)
- Are you involved with artificial intelligence (AI) and/or machine learning (ML) projects at your company? (Only respondents who answered "Yes" were included)

In this way, Algorithmia amassed a group of 403 individuals with a level of insight into the machine learning efforts of their companies across a random sampling of industries and machine learning maturity levels.

In all charts and analysis, percentages have been rounded to the nearest percentage point.

Changes from previous years

Algorithmia's annual report about enterprise machine learning is an evolving project and we seek to make improvements every year. We made multiple changes to this year's survey and report to improve both the reliability and relevance of our insights to AI/ML leaders.

This year, we limited the survey to respondents at companies with \$100M or more in revenue. This resulted in a slight shift towards organizations with more than 1,000 employees, with respondents being distributed accordingly:

- Fewer than 500 employees: 3% (down from 18% in 2020)
- 500-999 employees: 17% (flat year-on-year)
- 1,000-4,999 employees: 33% (up from 29% in 2020)
- 5,000-9,999 employees: 19% (up from 14% in 2020)
- 10,000-25,000 employees: 11% (up from 9% in 2020)
- More than 25,000 employees: 18% (up from 13% in 2020)

We made this change to improve the relevance of results to enterprise IT environments, and plan to continue this focus in subsequent years. As we establish a consistent baseline of company sizes in future years, we expect the reliability of our results to further improve.

In previous years, we have also distributed the survey to individuals who have engaged with Algorithmia through various channels in the past, such as by attending a company webinar, downloading a whitepaper, or meeting with our team at an industry trade show. While we still sent a survey to this audience this year, we ultimately did not include the results in the final report. To remove possible sources of bias, we only included responses from the blind survey in this year's report. When referencing data from last year's report, we also only used data from the blind survey conducted that year. However, while we did not include the data from this version of the survey, we did use its results to analyze trends and further bolster our confidence in our findings.

To gain deeper and more relevant insights from the survey, we also added and modified multiple questions in the survey. Anytime this may have impacted year-on-year comparisons, we have included a footnote in the relevant chart.

Specifically, in both our 2020 and 2021 surveys, respondents were asked to select the ways that AI/ML is being used at their organizations. Respondents were given slightly different options between these two years, as follows:

- Included in both years: Improving customer experience, Retaining customers, Reducing customer churn, Interacting with customers, Increasing long-term customer engagement, Increasing customer loyalty, Reducing costs, Acquiring new customers, Building brand awareness, Recommender systems, Other
- Included in 2021 survey only: Automating processes, Back office automation, Managing inventory, Managing logistics, Supply chain optimization, Financial planning, Generating financial insights
- Included in 2020 survey only: Increasing customer satisfaction, Processing automation for internal organization, Increasing conversion rates, Predicting demand fluctuations, Filtering assets and content

In a question about the infrastructure used to deploy models, respondents were also given slightly different options in 2020 and 2021, as follows:

- Included in both years: AWS, Azure, Google Cloud Platform, Other
- Included in 2021 survey only: VMWare-based, OpenShift-based
- Included in 2020 survey only: On-premises, A mix of cloud providers, A mix of cloud and on-premises solutions, I do not know or I am unsure

Lastly, when asked about the time it takes to put a trained model into scaled production, respondents were given slightly different options in 2020 and 2021. When comparing data year-on-year, data was aggregated accordingly:

- 2021 options: 1 day or less (aggregated to “1 week or less”), 1 day-1 week (aggregated to “1 week or less”), 1 week-1 month, 1 month-1 quarter, 1-2 quarters (aggregated to “1 quarter-1 year”), 2-3 quarters (aggregated to “1 quarter-1 year”), 3 quarters-1 year (aggregated to “1 quarter-1 year”), More than 1 year
- 2020 options: 0-7 days (aggregated to “1 week or less”), 8-30 days (aggregated to “1 week-1 month”), 31-90 days (aggregated to “1 month-1 quarter”), 91-365 days (aggregated to “1 quarter-1 year”), More than 1 year, I do not know or I am unsure

We will continue to conduct the annual survey to increase the breadth of our understanding of machine learning in the enterprise, and share our insights into how the industry is evolving. As we continue to build on this work in subsequent years, we aim to reach ever more relevant and applicable insights to help AI/ML leaders drive innovation in this space.

About Algorithmia

Algorithmia is machine learning operations software that manages all stages of the ML lifecycle within existing operational processes. Put models into production quickly, securely, and cost-effectively.

Unlike inefficient and expensive do-it-yourself solutions that lock users into specific technology stacks, Algorithmia automates ML deployment, optimizes collaboration between operations and development, leverages existing SDLC and CI/CD systems, and provides advanced security and governance.

Over 120,000 engineers and data scientists have used Algorithmia's platform to date, including the United Nations, government intelligence agencies, and Fortune 500 companies.

Learn how Algorithmia can help solve your most pressing ML challenges and put you in position for scalable success. Visit algorithmia.com/product.

About the cover

The cover image is a parallel set chart—similar to a Sankey diagram. Each line-set represents a specific data category. The width of each line-set's path is determined by the proportional amount of the category total.

The line-set on the left depicts survey participants' current state of maturity in the following area: Organizational alignment for AI/ML initiatives. The line-set on the right displays approximately how many ML models their organization has in production.

