mlinsider 2022 Survey

Full Report

The state of Machine Learning at the end of 2022

by cnvrg.io
an Intel Company

Table of Contents

What is ML Insider?

The ML Insider is an annual analysis by cnvrg.io, an Intel company, of the Machine Learning industry highlighting key trends, points of interest, and challenges that AI developers experience every day. This report offers insights into how over 430 AI professionals are building, training, deploying and adapting their machine learning stack to better suit today's modern, complex ML workflow.

Who Took This Survey?

The survey covered 430 participants, representing organizations in size from a few employees to over 5,000. 53% of respondents came from companies with 600 employees or less. The insights from the survey covered dozens of industries; among the most common were: Information Technology Services and Computer Software. Financial Services/Banking, Education, Healthcare, Consumer Goods, Telecommunications, Automotive, Insurance Media/Entertainment, and Defense represent the other half of respondents. 42% of all respondents are data scientists, 16.9% engineering/DevOps and 10.6% in software development roles.

About cnvrg.io

cnvrg.io is a full stack machine learning operating system with everything an AI developer needs to build and deploy AI on any infrastructure. cnvrg.io was built by data scientists, to help data scientists and developers automate training and deployment of ML pipelines at scale. cnvrg.io helps organizations accelerate value from data science.

Key Takeaways and Trends

- AI is recession resilient
  Despite the rough economic circumstances, investment in AI is expected to grow in 2023.
- AI maturity remains low
  57% of respondents reported a low AI maturity with less than 4 models running in production.
- Lack of knowledge and expertise remains a top AI challenge
  Consistent with 2021 data, lack of knowledge and expertise, and hiring AI/data science talent remain the top 2 challenges executing ML programs.
- A majority of organizations plan to address AI explainability in 2023
  43.5% of respondents are planning to introduce explainable AI techniques in the next 12 months, while only 37% already have AI explainability techniques in place.
- Industries with more consumer regulatory pressure tend to have lower AI adoption

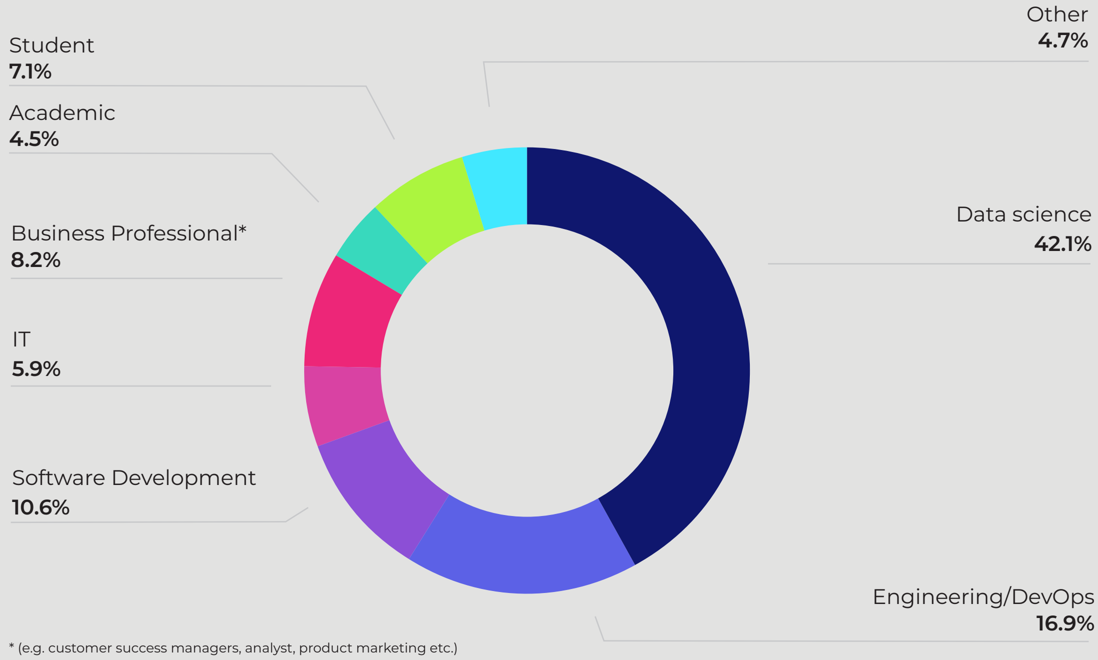
  Defense, Automotive, and Computer Software have the highest AI adoption, while Education/E-learning, Hospital & Healthcare, as well as Media/Entertainment have seen the lowest AI adoption.
- Data scientists are not the only ones with AI in their job description
  The role of an AI developer is evolving with equal responsibility distributed between data scientists, engineers, and software developers.
- Operationalization of AI is still heavily dependent on Developers/DevOps/Engineering
  64% of respondents that found it difficult to successfully execute ML rely on Developers/DevOps/Engineering to operationalize their AI models
- More organizations are adopting hybrid infrastructures

  Compared to 2021, there has been a 13% increase in hybrid compute adoption.
- 89% of organizations are seeing the benefits of their AI solutions
  The majority of organizations investing in AI are benefiting from their AI solutions
- Reducing technical complexity is the key to universal AI adoption and acceptance
  Technical complexity is holding AI back from achieving universal AI adoption and acceptance as a technology. The majority of respondents believe technical complexity of AI development is the biggest challenge to universal AI adoption and acceptance.

ML Insider 2022
Results and Analysis

Figure 1: What is your primary role?

The majority of the respondents described their primary role as data scientist. Though, as we will learn in later survey questions, building AI involves more than just data scientists these days and often involves engineers and DevOps as well.



* (e.g. customer success managers, analyst, product marketing etc.)

Figure 2: What is the size of your team?

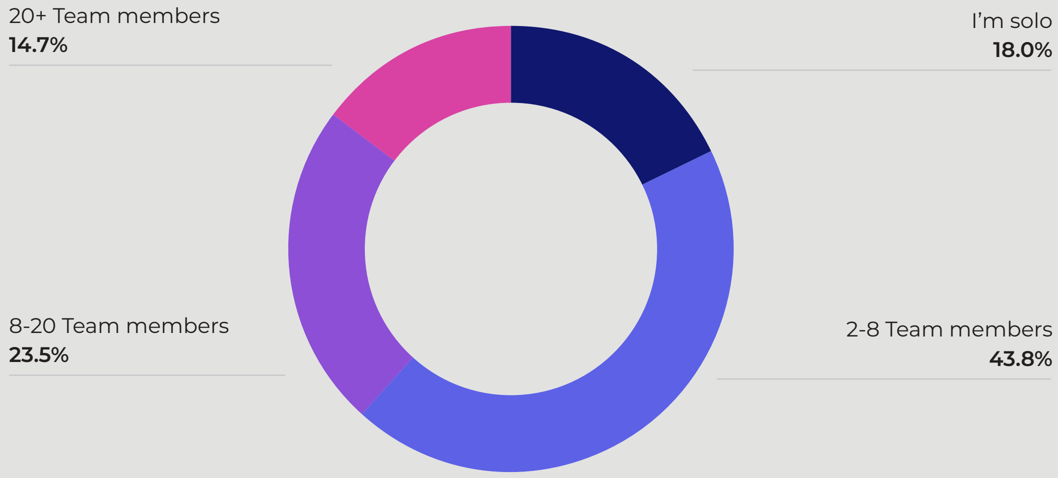


Figure 3: In which industry do you work?

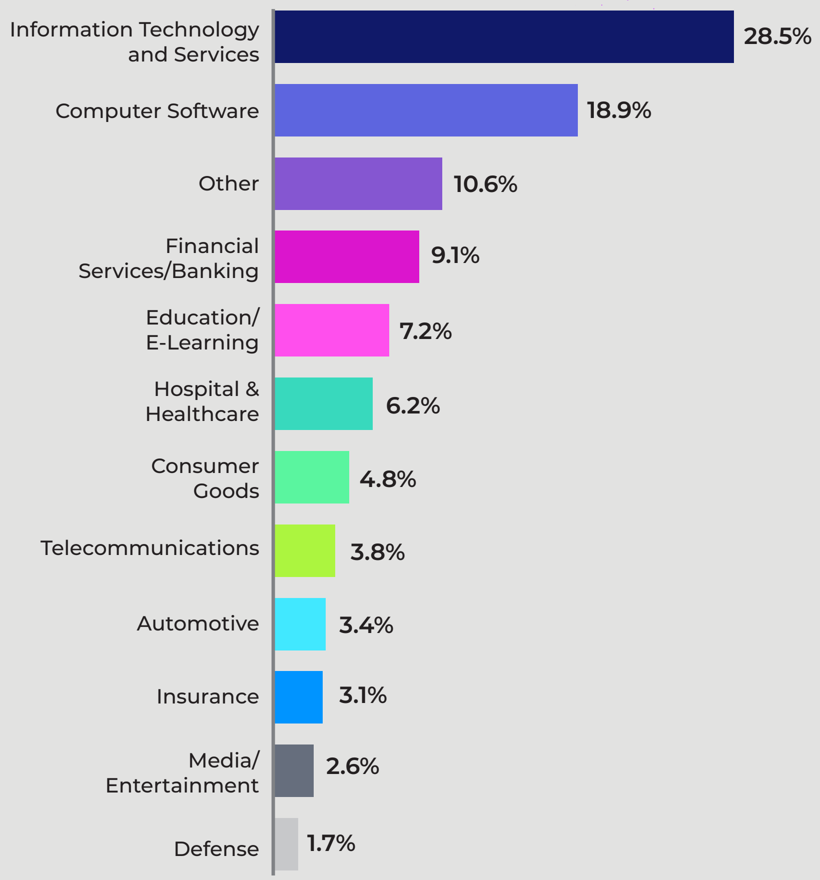
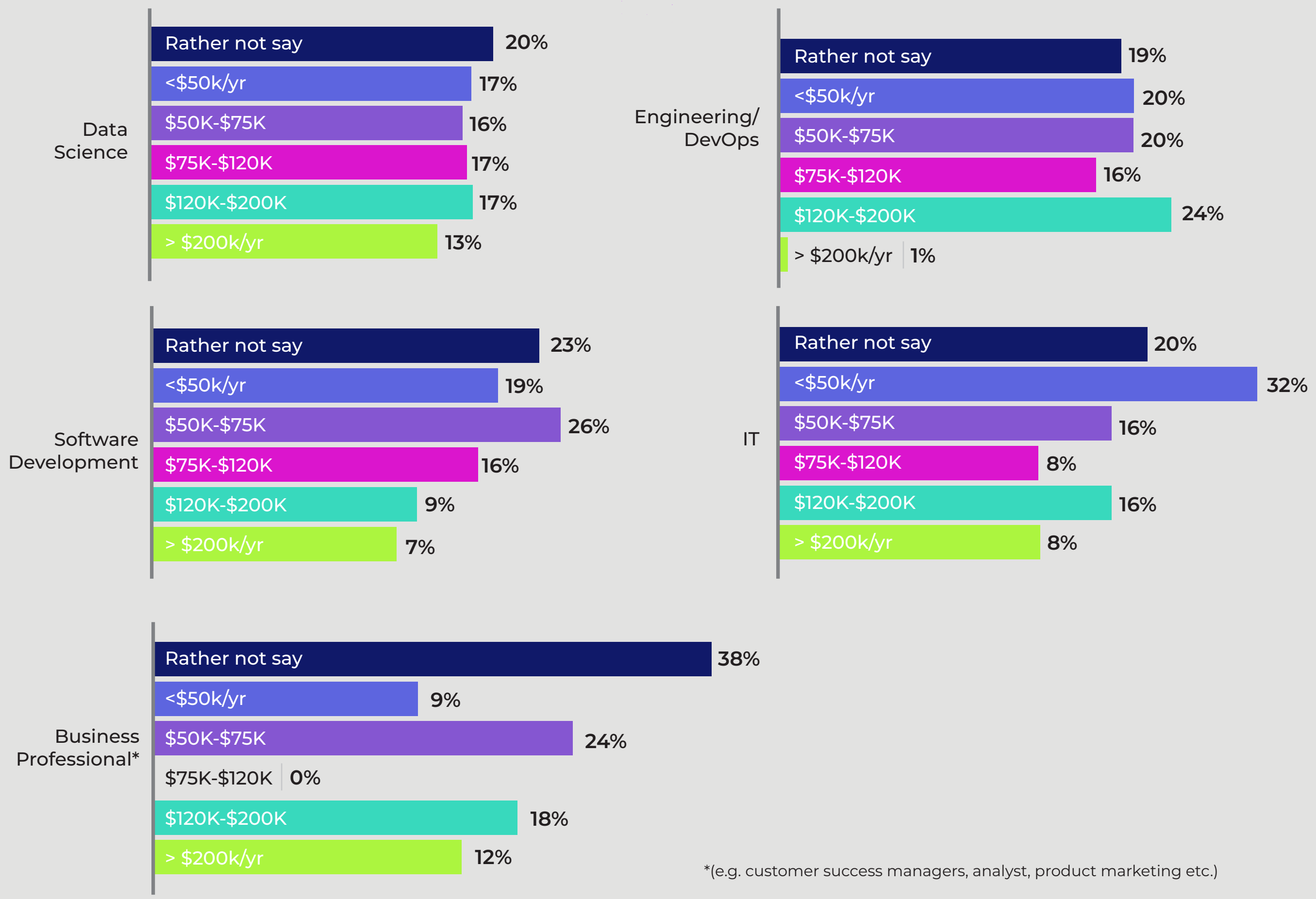


Figure 4: What is the size of the company you currently work in?

Figure 5: Salary by profession

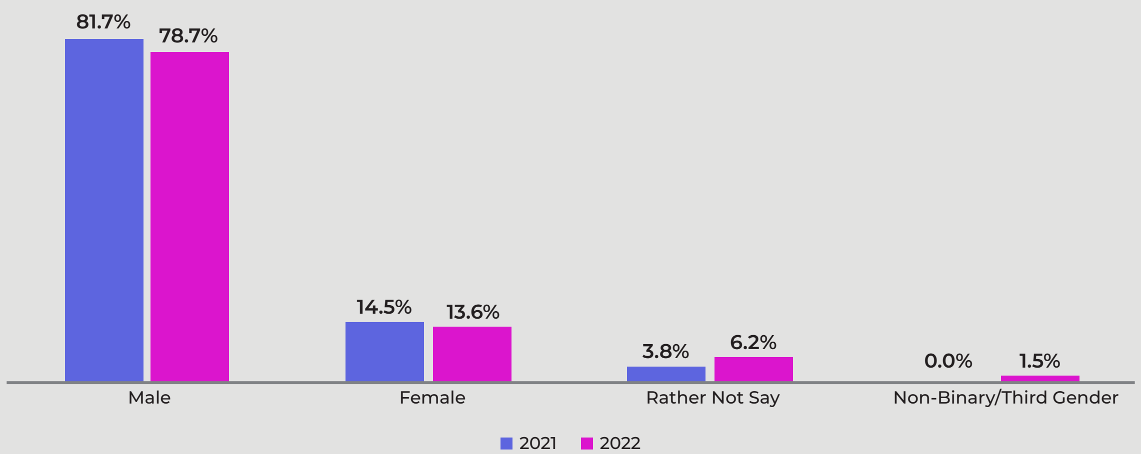


*(e.g. customer success managers, analyst, product marketing etc.)

Figure 6: How do you identify?

Gender 2021 vs 2022 Comparison

A comparison of 2021 and 2022 data on the gender of respondents shows that there is still a major gender gap and lack of diversity in the AI field.



AI Maturity

Overall AI maturity remains quite low. 57% of respondents reported a low AI maturity (0-4 AI models running in their product). In this section we measure AI maturity by the number of models an organization has running in their product. It should be considered that organizations may in fact have AI that is not in production or applied to a product solution.

Figure 7: At what level is AI being used in your organizations applications?

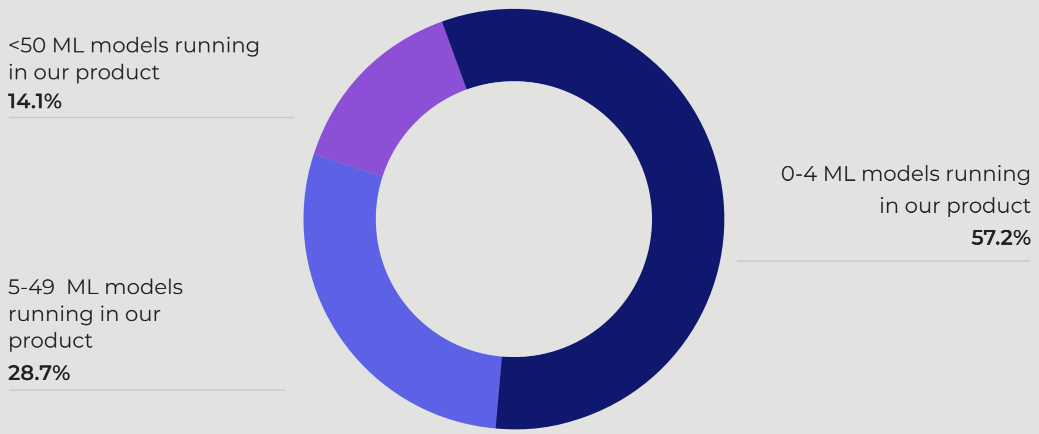


Figure 8: Does your organization run deep-learning tasks?

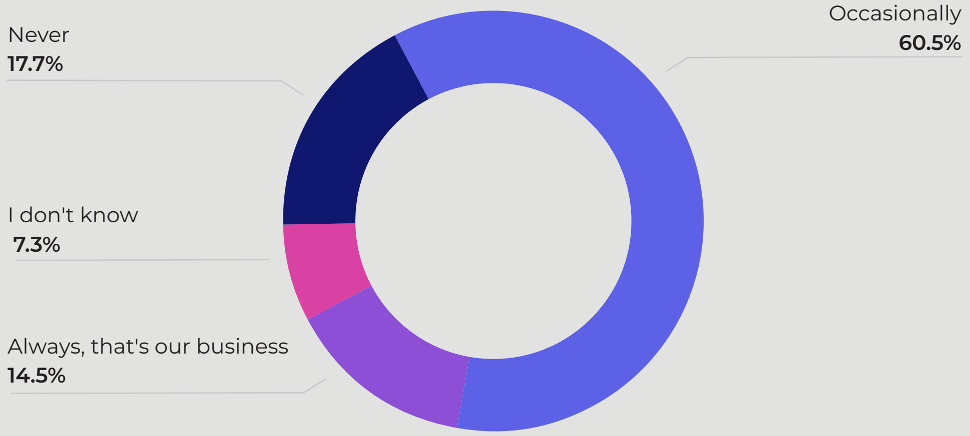


Figure 9: Common Challenges by Level of Maturity

Organizations face different challenges based on their level of AI maturity.

The top 2 challenges for organizations with low AI maturity were the lack of knowledge or expertise and the high cost of AI.

Organizations that fell in the middle, with a medium AI maturity, suffered from an incomplete or insufficient technology stack as well as hiring AI/data science talent. These companies are likely competing for similar talent, and may not have the scalable technology stack to get them to the next level of AI maturity.

For companies with high AI implementation (over 50 models running) hiring AI talent becomes the top challenge. Once organizations reach higher levels of AI maturity, they may have already addressed challenges such as having an effective technology stack, or budget limitations but hiring talent to manage their growing AI operation becomes a more significant challenge.

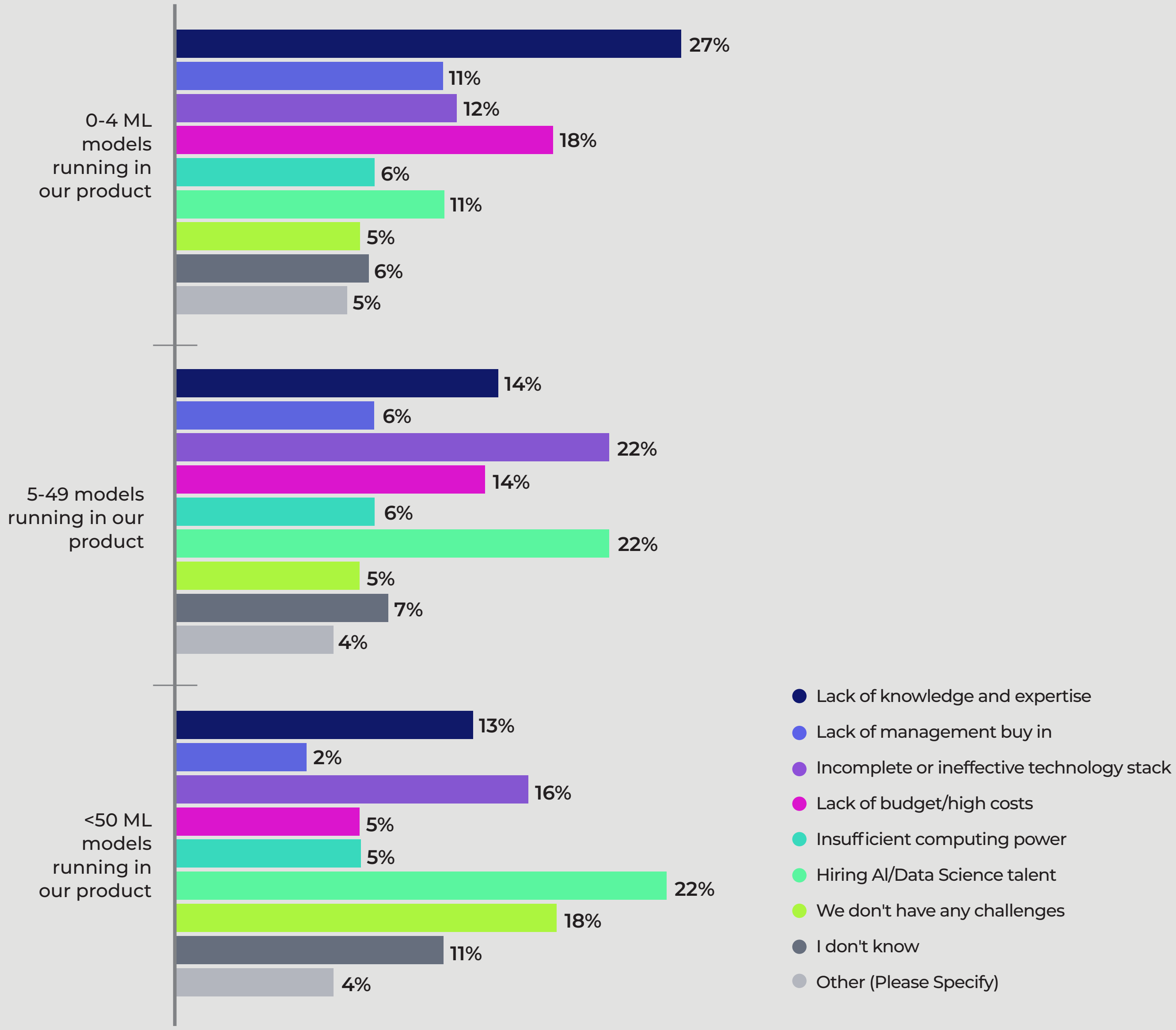


Figure 10: AI Maturity by Industry

Which industries have the highest/lowest AI maturity? Education/E-learning, Hospital & Healthcare, and Media/Entertainment have seen the lowest AI adoption, with the largest number of respondents indicating that they have 0-5 models running. The industries with the highest adoption include Defense, Telecommunications, and Insurance.

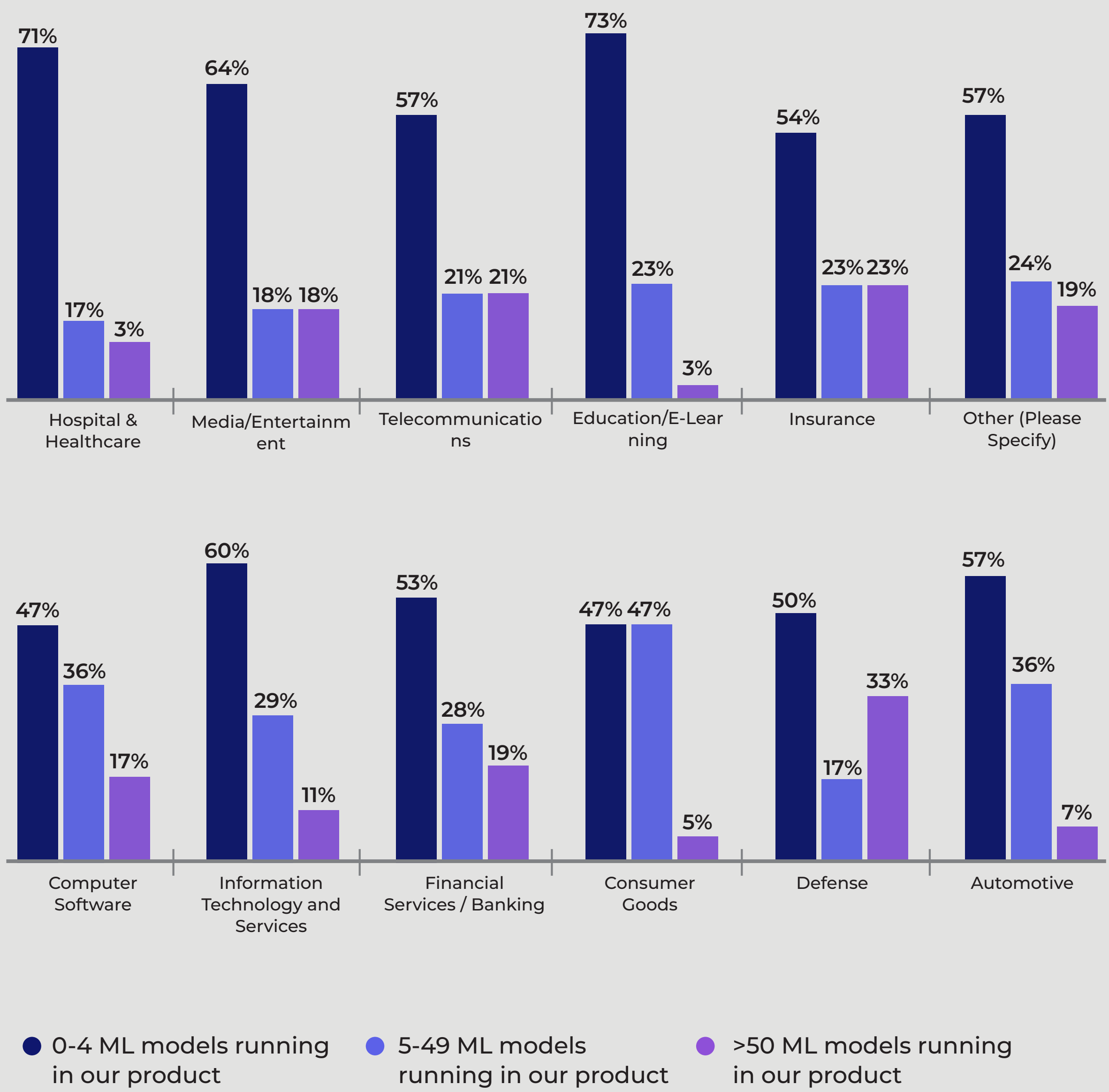
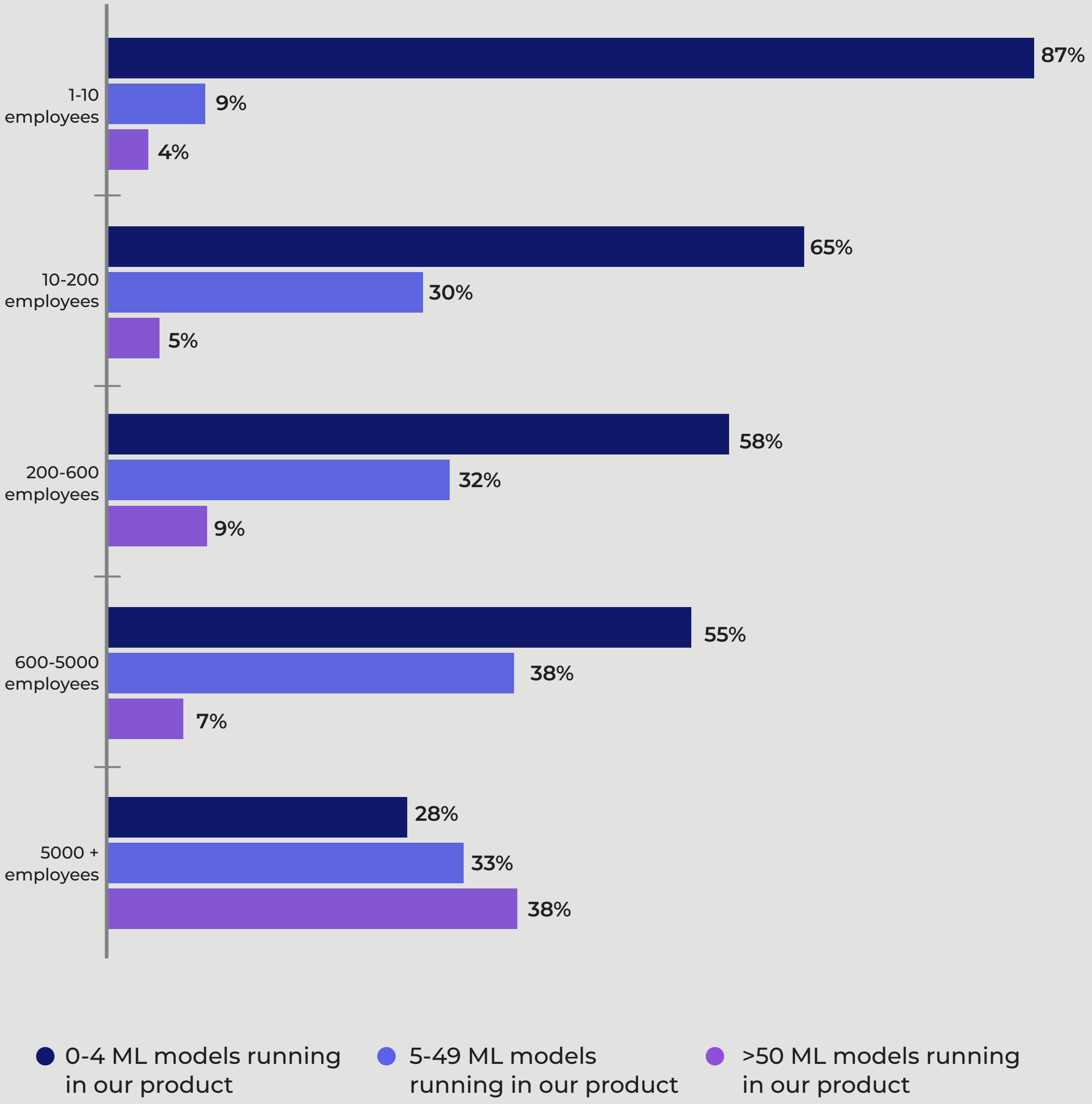


Figure 11: AI Maturity by Company Size

Not all large companies have high AI maturity. While 38% of large companies have high AI maturity, 62% of large companies have under 50 models running.



Applying AI

Who, what, where, and why are organizations applying AI? This section of ML Insider examines all aspects of applied AI across the industry to tell a story. See how your organization stacks up against others by examining team sizes, roles responsible for AI, and use cases that have helped improve different lines of business in the organization.

Figure 12: Who is building AI solutions in your organization? (Select all that apply)

Historically, data scientists have been solely responsible for building AI solutions and had to be highly skilled in research, engineering, and programming. The survey shows that the responsibility of AI development is evolving. As respondents indicate here, there is a changing definition of 'AI developer'. Data scientists, engineers, and developers are equally responsible for building AI solutions in an organization.

AI is now more commonly used in applications and in more and more cases this relies heavily on the skills of engineers and developers to deploy and maintain those models in production. In addition, AI development is becoming more accessible, allowing different roles to build AI. In order to stay competitive, companies will want to cater to the challenges of those specific roles and simplify the AI process. By enabling more professionals to build AI solutions, organizations will also relieve hiring challenges.

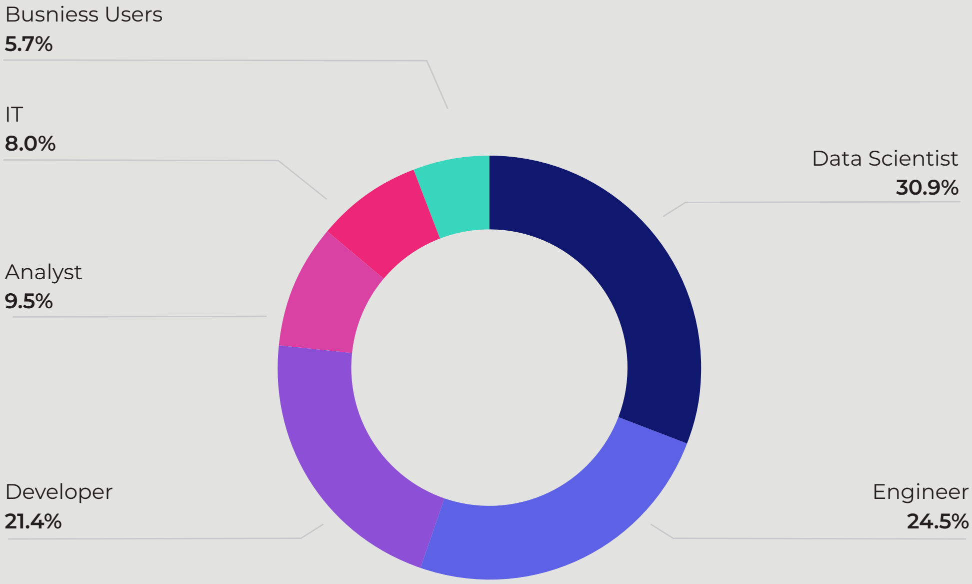


Figure 13: How many of each role are there in your company?

As one might expect, the size of a company correlates with the amount of AI developers they have. Though, compared to small and medium-sized companies, large companies overwhelmingly indicated having over 50 of each role in their company. Small to medium-sized companies with 600 employees or less are able to work with leaner data science teams.

Like data scientists, ML Engineers remain scarce in small to medium-sized companies. Compared to data scientists, ML engineers tend to be a luxury that only large companies can afford. In Figure 5, ML Engineers seem to earn the highest salary.

The number of data engineers also correlates with the size of a company. This is likely due to data size. Larger companies tend to have access to much larger datasets to manage which requires more data engineers.

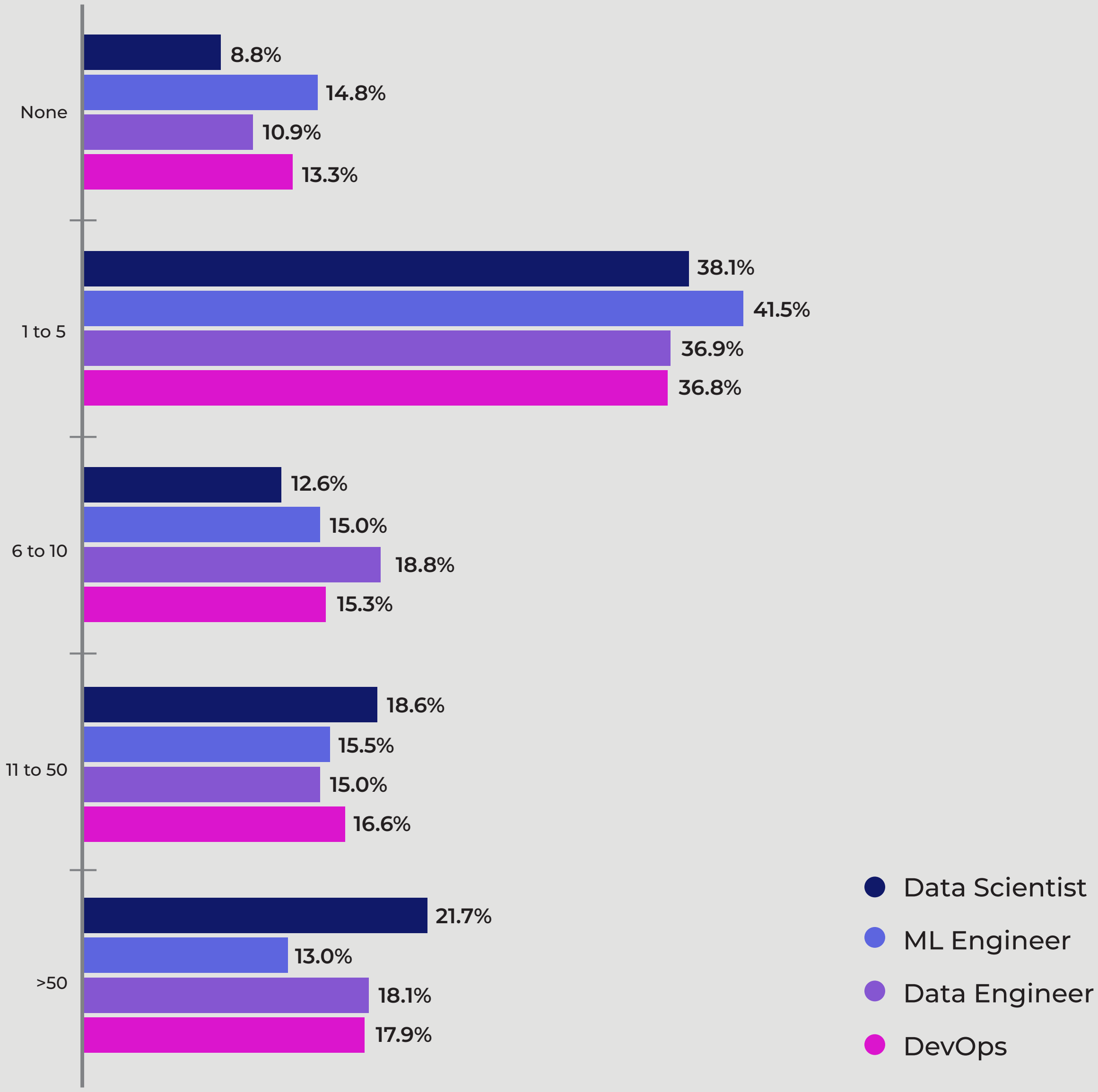


Figure 14: What lines of business are using ML in your company?

2021 vs 2022 comparison

Compared to 2021, organizations have increased the use of ML to address operations and supply chain lines of business in 2022. Organizations appear to be addressing the world's current supply chain issues with AI.

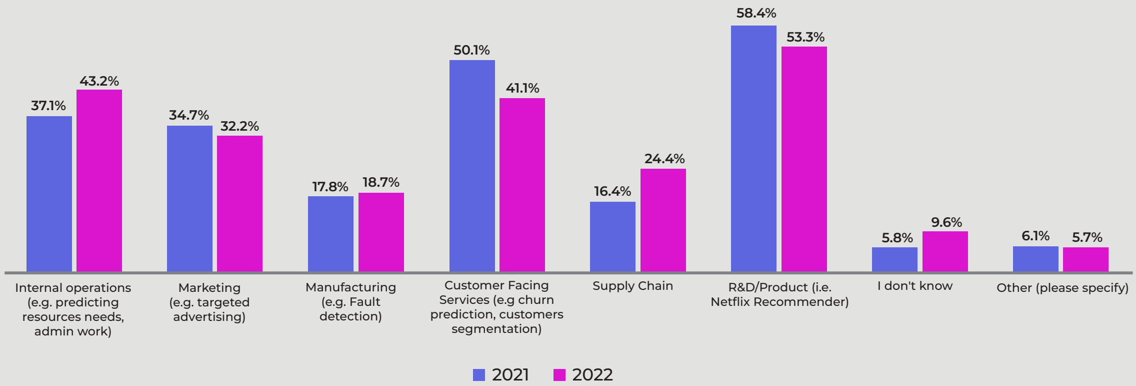


Figure 15: What lines of business are using ML in your company? (2022) (Select all that apply)

The top lines of business that are using ML within a company in 2022 are R&D/ Product (i.e. recommender system, digital voice assistant), Internal operations (e.g. predicting resource needs, admin work), and customer-facing services (e.g churn prediction, customer segmentation).

Leveraging the benefits of AI requires a good understanding of the capabilities of AI, as well as a good understanding of the problem it is trying to solve. The fact that R&D and technical use cases are more prevalent may indicate a need for more accessible solutions in those lines of business to push AI adoption further.

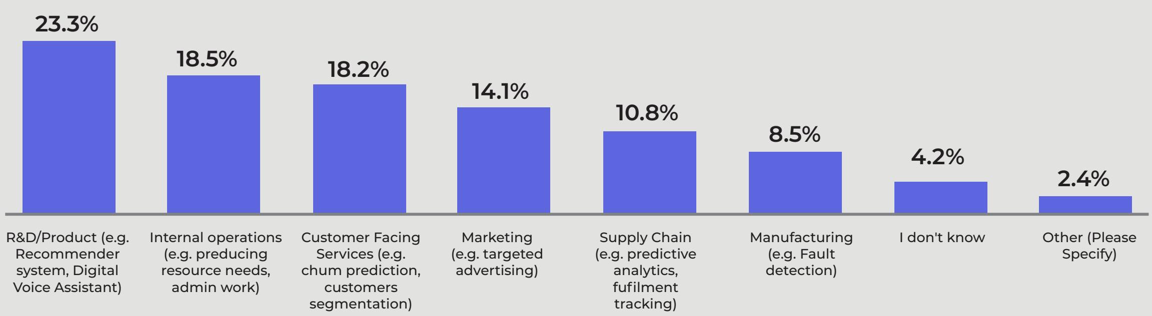


Figure 16: What is your main goal for implementing AI? (Select all that apply)

Organizations are focused on using AI to improve their product and R&D workflows and in some companies are looking to leverage AI to optimize their internal processes, their sales and customer support.

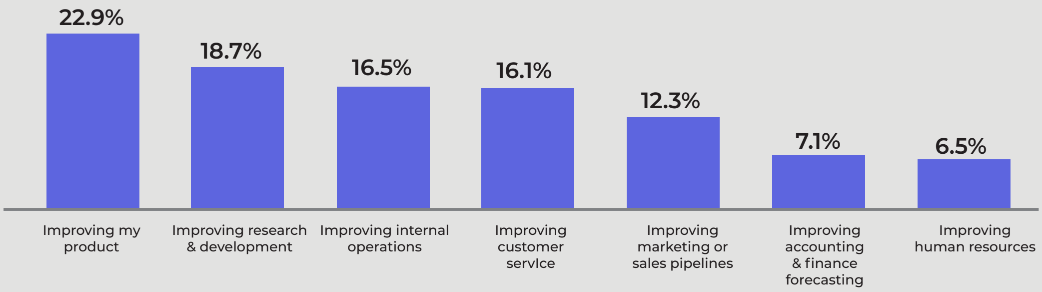


Figure 17: Which use case do you use, or are you looking to use in your organization? (Select all that apply)

Predictive analytics/Forecasting is the most popular use case for AI.

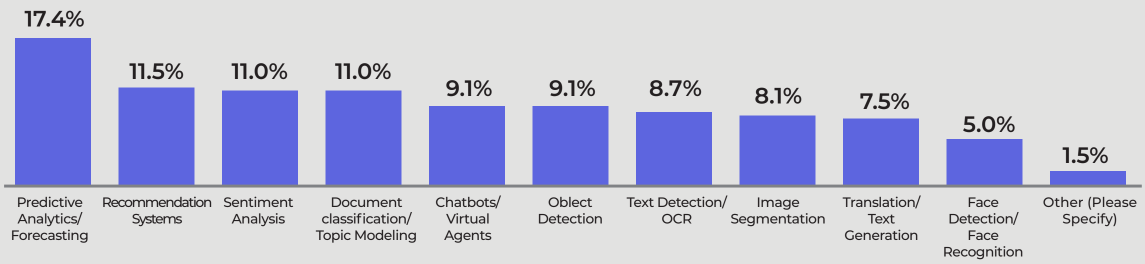
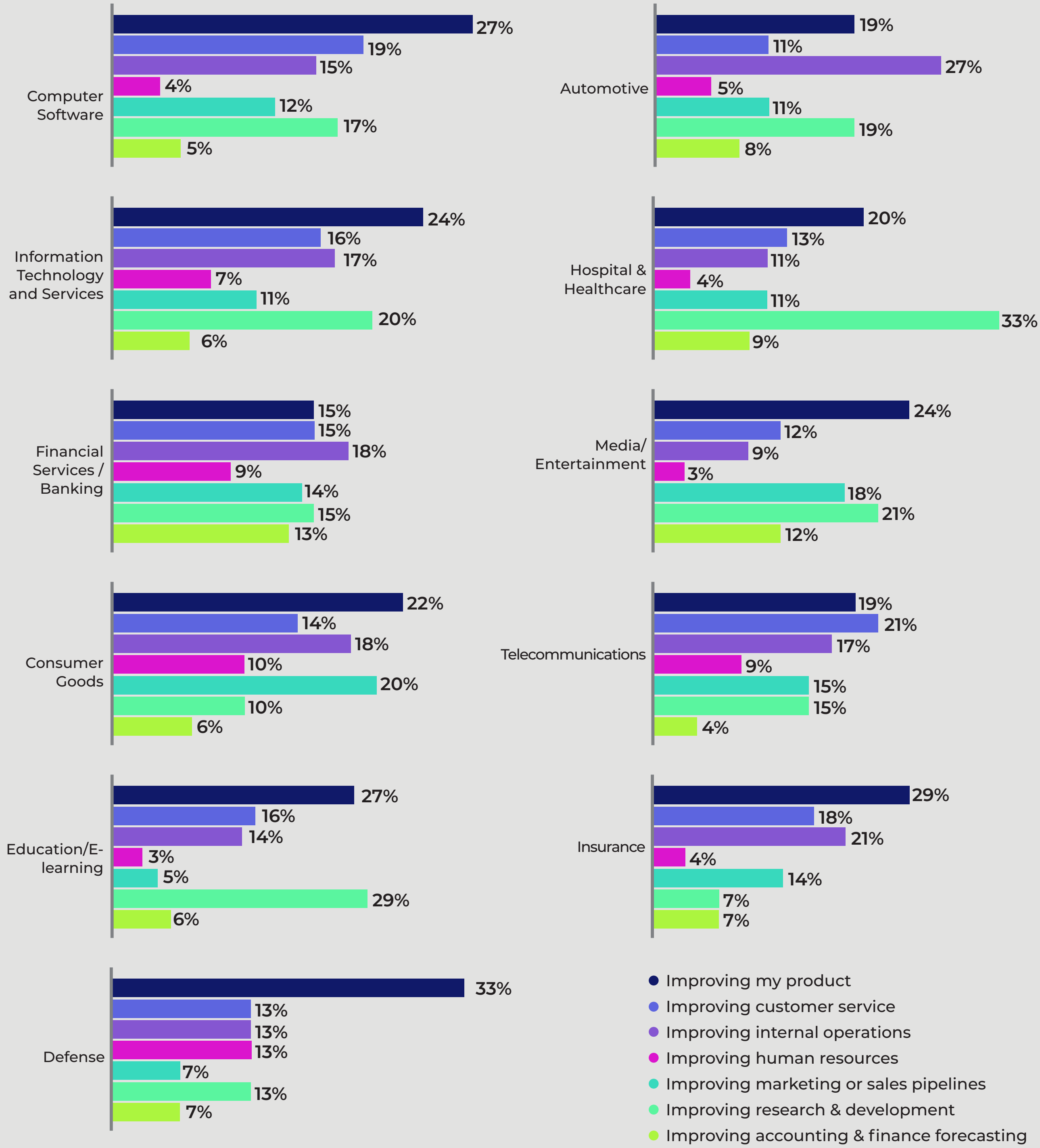


Figure 18: AI Goals by Industry

The goals for AI differ by industry. Most industries are heavily product-oriented. Healthcare emphasizes research & development as the primary goal for implementing AI. Customer service is a key focus in the Telecommunications industry, while the Automotive and Financial Services industries stated their main goal is to improve internal processes and operations.



AI Challenges

It's not news that AI is difficult to execute successfully. This section dives into the main challenges organizations are seeing most often, whether by industry, company size, or tools. Compared to 2021, lack of knowledge and expertise and hiring AI talent remain the top challenges for most organizations. When combined with the fact that teams tend to be small anyway, the difficulty in obtaining talent points to the need to get the most out of already existing talent.

Figure 19: In your experience, how difficult is it to successfully execute AI projects?

65% of respondents found it difficult, extremely difficult and even impossible to execute. Only 7% of respondents found it simple.

Figure 20: What is your company's main challenge in executing ML programs?

Lack of knowledge and expertise remains the most common AI challenge across organizations.

Other challenges mentioned include lack of clean or high-quality data, framing correctly, poor senior tech leadership, privacy issues in data management, cybersecurity issues of ML products, GDPR compliance, accessibility by stakeholders, trusting AI results, and more.

The respondents stated a long list of challenges with AI execution, although over time more solutions are becoming available to solve many of these challenges.

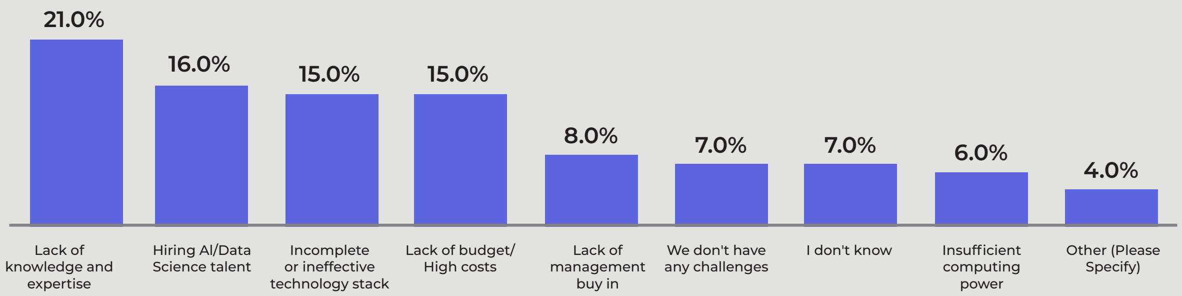


Figure 21: Those that find AI execution difficult by level of MLOps adoption

In most situations, MLOps can be a critical component to the success of an organization's AI. Though, of the respondents that found it difficult to execute AI successfully, 40% are using in-house developed MLOps tools. For some organizations, maintaining an MLOps solution in-house can add complexity to your AI development cycle, making it more difficult to successfully deliver the final AI solution.

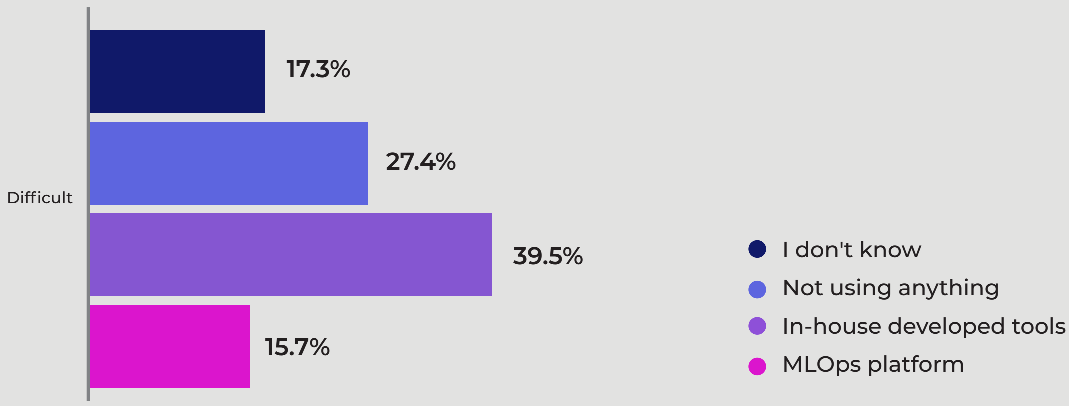


Figure 22: Which phase of the AI pipeline would you consider to be most challenging in your company?

Real-world data is extremely messy. Data operations such as data acquisition and data cleansing are typically the most challenging phases of the AI pipeline. Feature stores are often used to simplify data operations and centralize common data resources and make the process more repeatable. The second biggest challenge is explainability which is becoming a growing concern within organizations.

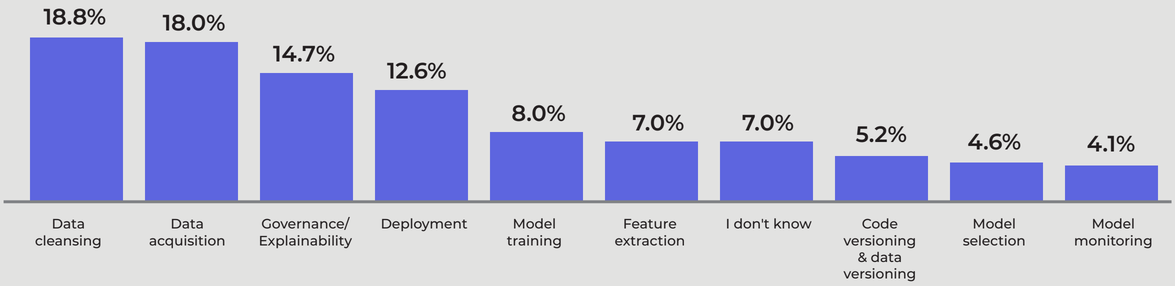


Figure 23: How long does it typically take to get from experimentation (model staging, training, testing) to production?

Figure 24: Difficulty Level by Industry

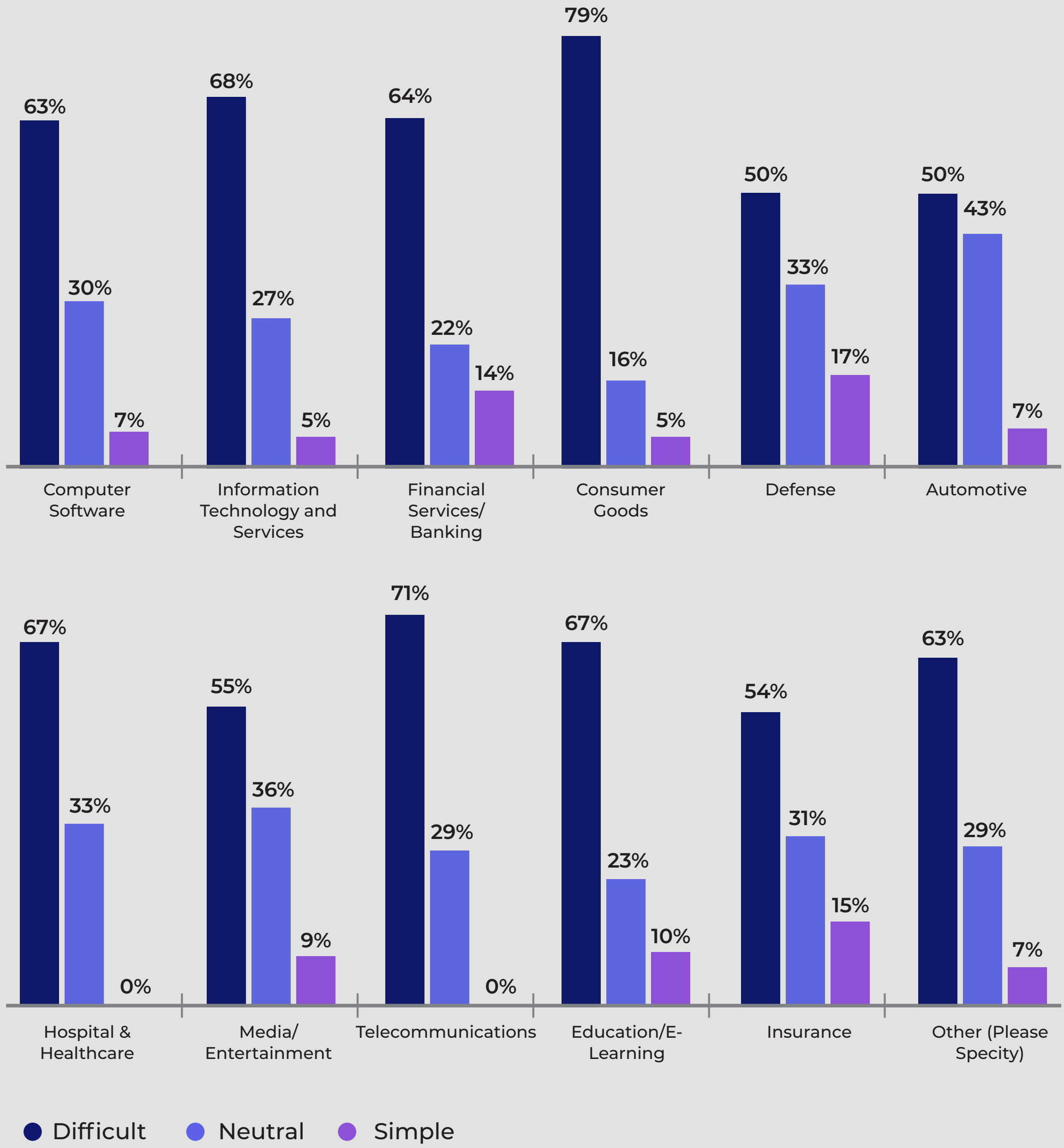


Figure 25: Is your organization seeing the benefits of your AI solutions?

89% of organizations are seeing the benefits of their AI solutions.

Figure 26: My organization is struggling to move from AI experimentation to production

Operationalization of AI used to be a main challenge for organizations. 46% of respondents do not find deployment from experimentation to production to be a challenge. Still, over half of respondents find it challenging to move from AI experimentation to production.

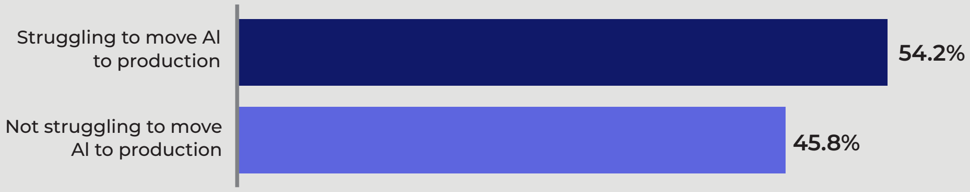


Figure 27: AI Operationalization by MLOps Adoption Level

MLOps is an effective solution for operationalizing AI.

Who is succeeding in operationalizing their models? Of the 45.5% of respondents who reported no struggle to move from AI experimentation to production, 66% have implemented some form of MLOps. 47% of those succeeding in operationalizing their models have reported using in-house MLOps tools, while 19% use a commercial MLOps solution.

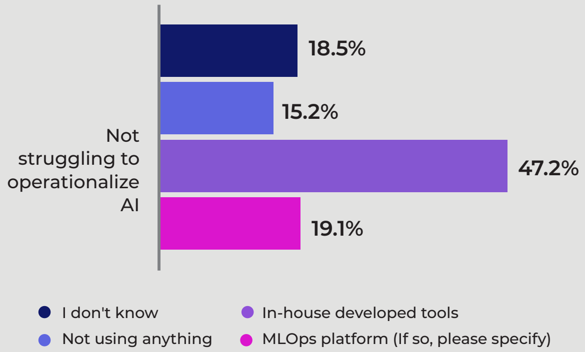
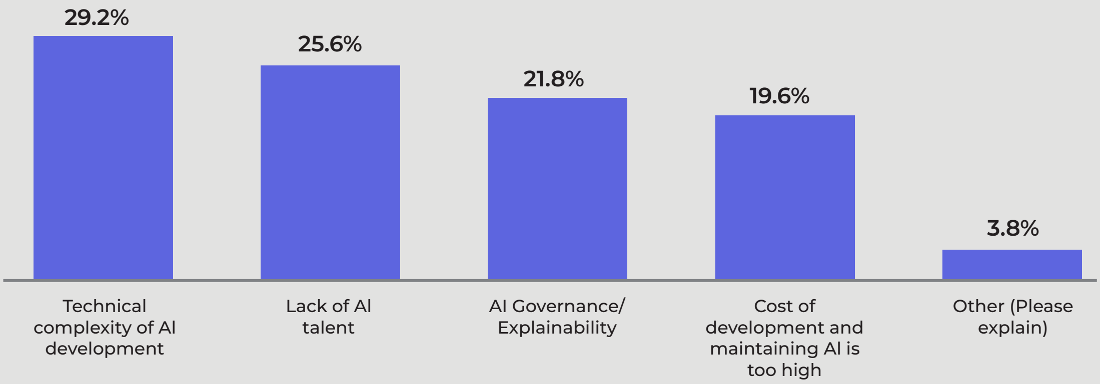


Figure 28: I rely on Developers/DevOps/Engineering to operationalize my model

66% of respondents rely on Developers, DevOps or Engineering to operationalize their models. The technical complexity of operationalizing AI models requires the technical skill of developers, engineers, and DevOps which can add bottlenecks and increase time to production.

Figure 29: What do you believe is the biggest challenge to universal AI adoption and acceptance?

The majority of respondents believe the technical complexity of AI development will be the biggest challenge to universal AI adoption and acceptance. Today, it is quite difficult to build AI, and there is little transparency and understanding about how these applications are being made. In order to achieve universal AI adoption and acceptance, there will need to be simpler ways to both develop and interact with AI applications.



Tools & Technology

What tools are organizations using to combat AI challenges? This section reviews infrastructures, tool stacks, and levels of automation across organizations. How does your organizations technology stack up against others?

Figure 30: What tools are you using for preprocessing/ETL?

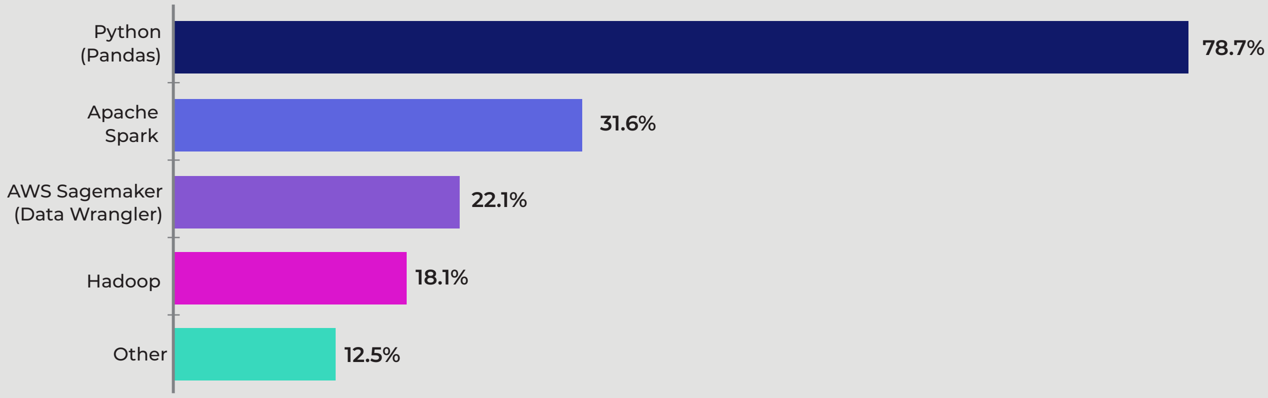


Figure 31: Where is your data stored?

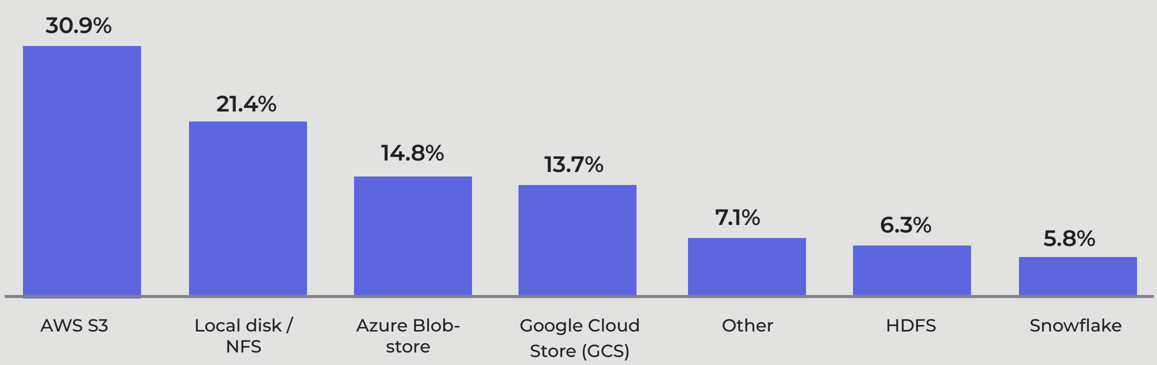


Figure 32: Where is your data stored?

Data storage 2021 vs 2022 Comparison

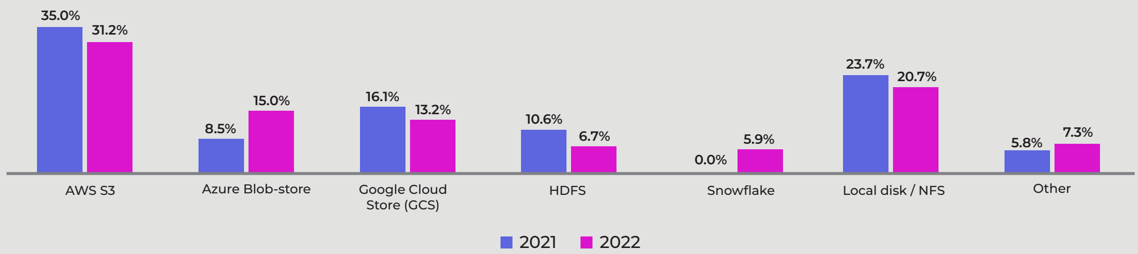


Figure 33: What cloud providers are you using? (Select all that apply)

Many respondents selected more than one cloud provider, illustrating the popularity of multi-cloud infrastructures. The majority of cloud users selected AWS, Azure, and GCP as the most commonly used cloud providers.

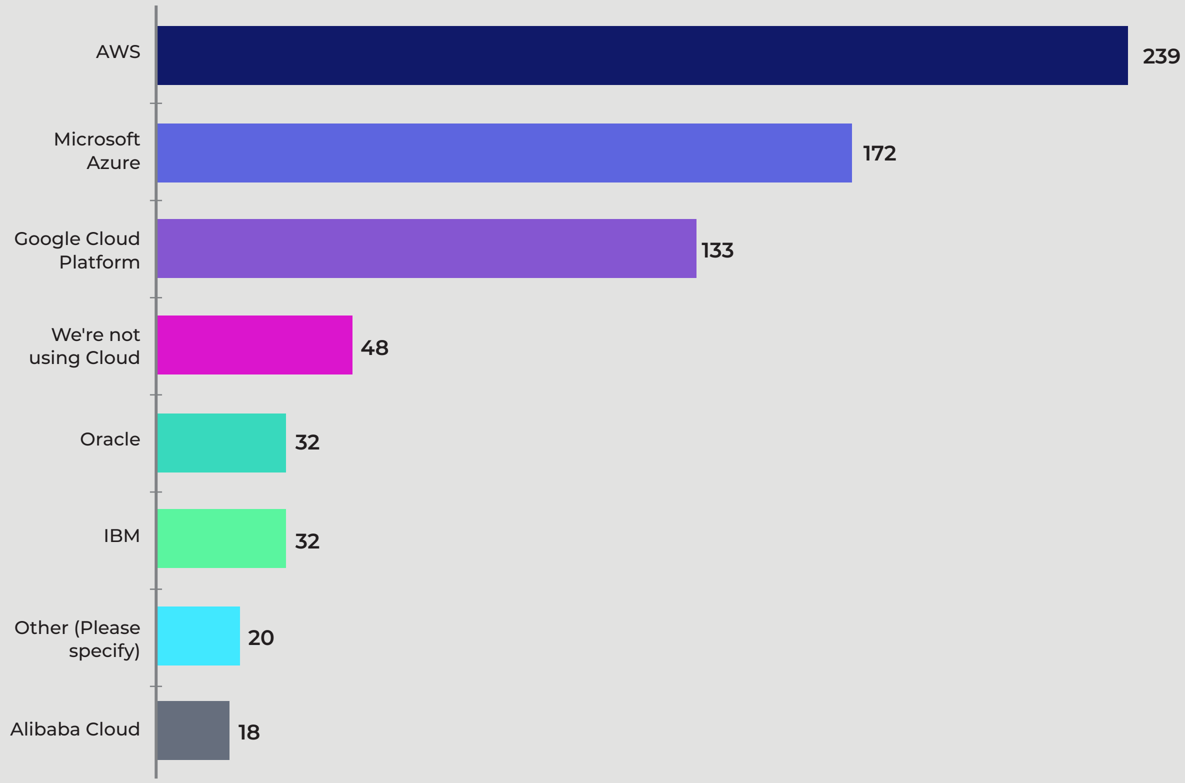


Figure 34: Where do you (mostly) run your ML workloads?

Compared to 2021, there has been a 13% increase in hybrid cloud adoption. Organizations are shifting away from on premises-only and cloud-only infrastructures. The trend in 2022 is that hybrid cloud is becoming more popular. This is likely due to the flexibility to utilize different types of computing for different types of tasks, as well as the high cost of cloud. Hybrid infrastructures allow organizations to reduce costs and improve AI workload performance. While hybrid infrastructures have their benefits, it can still be a challenge for organizations to manage hybrid infrastructures. Organizations will need to adapt to these challenges and enable AI developers to easily shift workloads between clouds and on premises resources to fully take advantage of hybrid infrastructure benefits.

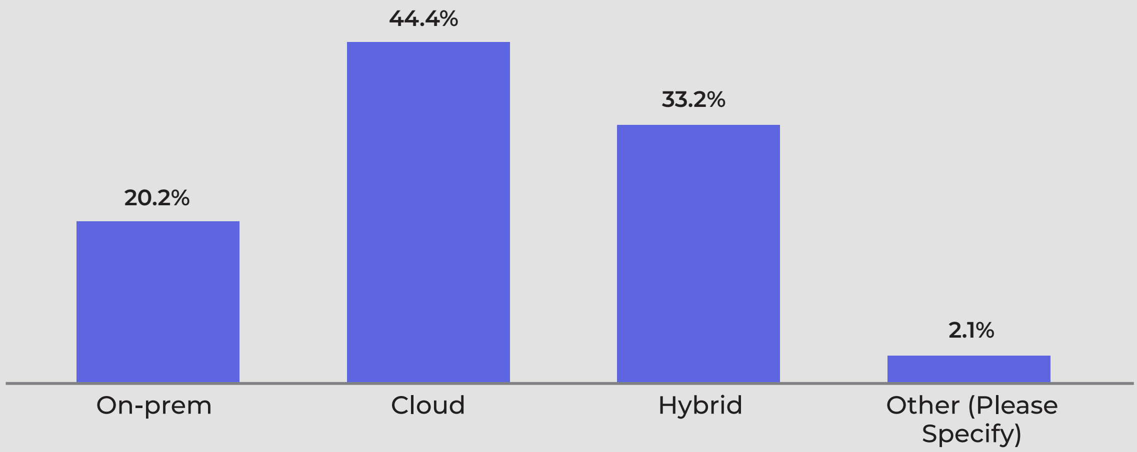


Figure 35: Where do you (mostly) run your ML workloads?

Compute 2021 vs 2022 Comparison

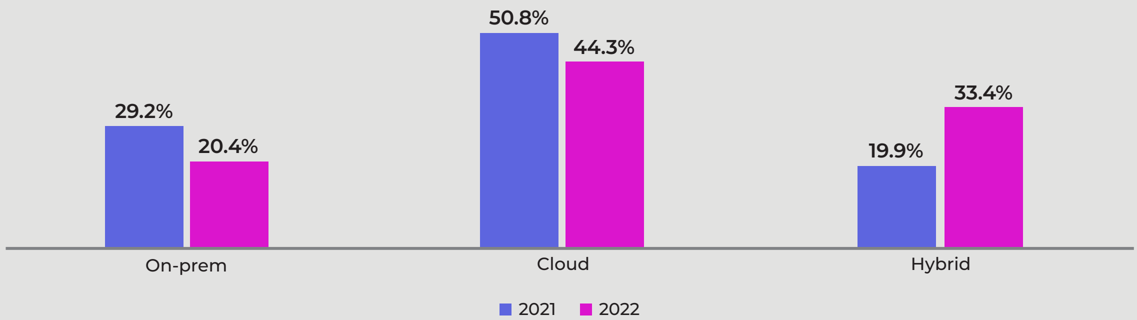


Figure 36: How important is the ability to shift AI workloads between clouds?

As organizations begin to adopt a hybrid infrastructure, the ability to shift AI workloads between clouds becomes more important. Here, we can see that 73% of respondents with hybrid infrustructures said the ability to shift AI workloads between clouds was important or even critical.

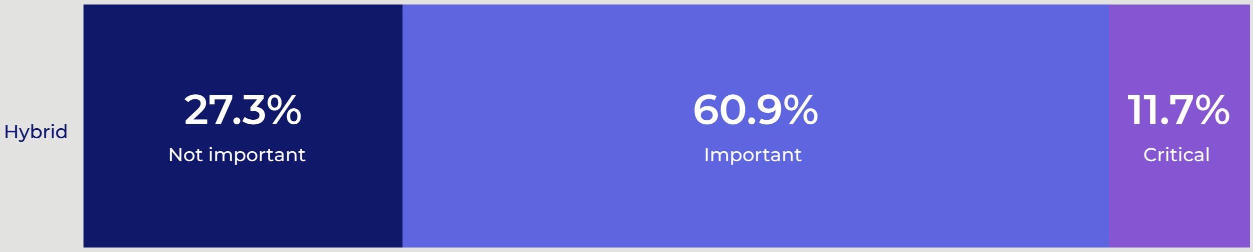


Figure 37: On average how much are you spending per month on cloud costs (USD)?

Almost 50% of cloud users are spending under $12k per month in cloud costs

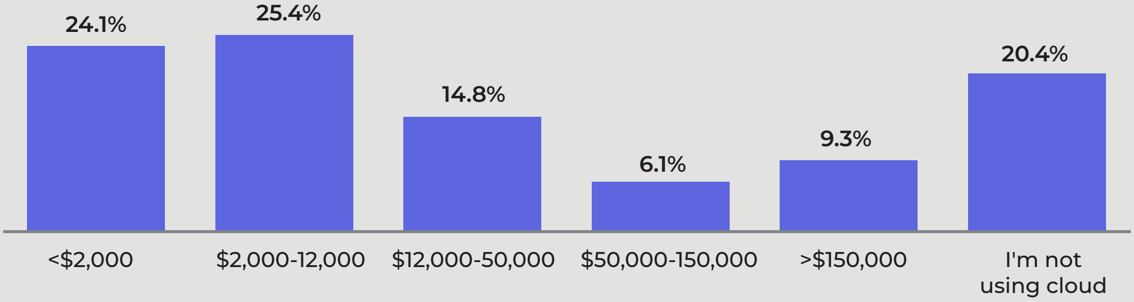


Figure 38: What is the average CPU count per task that you or your team use for running ML tasks?

23% of respondents use over 33 CPUs for running ML. 24% of respondents didn't know the average CPU count per task for ML which indicates a lack of visibility into compute utilization, and can be solved with an infrastructure monitoring.

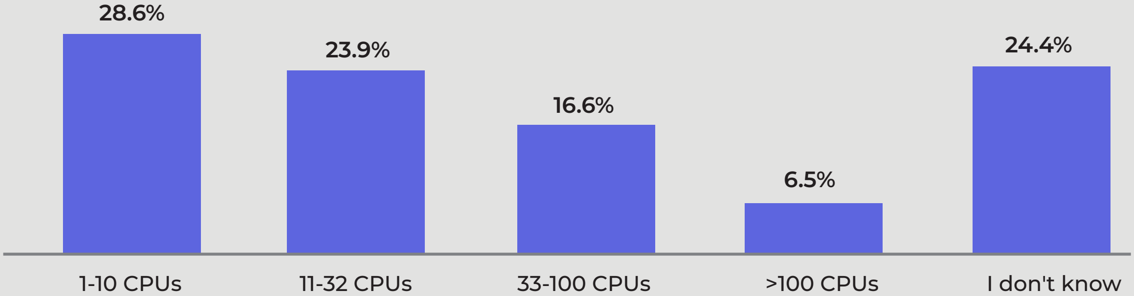


Figure 39: What is the average GPU count per task that you or your team use for running ML tasks?

5.2% of respondents use over 30 GPUs for running ML. 28% of respondents did not know how many average GPU's are utilized per task for running ML. This shows a lack of visibility into GPU utilization.

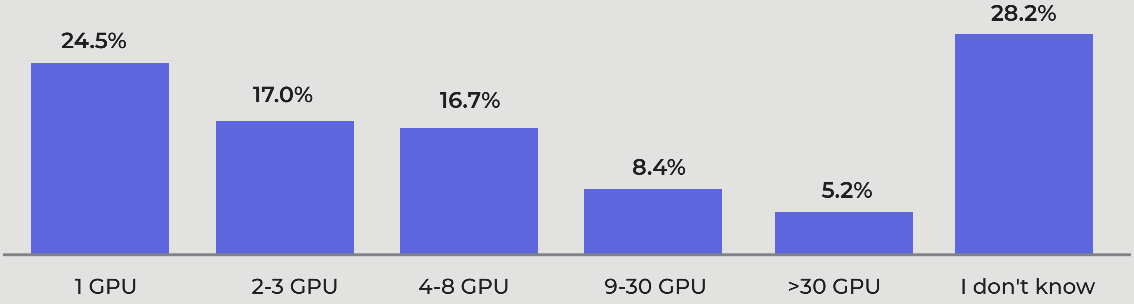


Figure 40: What is the average Memory count per task that you or your team use for running ML tasks?

5.8% of respondents use over 1TB per ML task. 26% of respondents didn't know the average memory count per task for ML, which indicates a lack of visibility into memory usage.

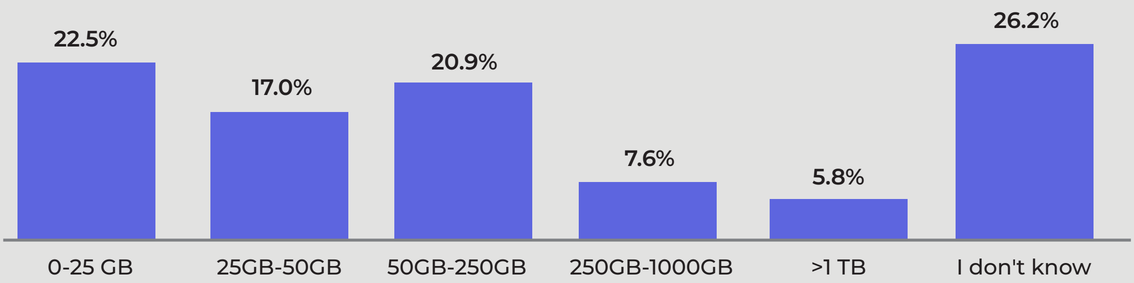


Figure 41: What is the most common type of data your company analyzes?

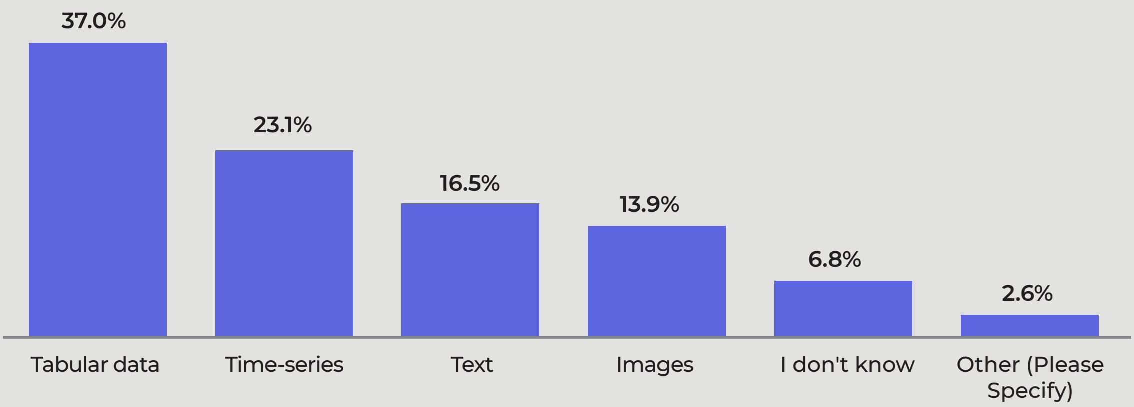


Figure 42: What is the typical size of a dataset you analyze in your company?

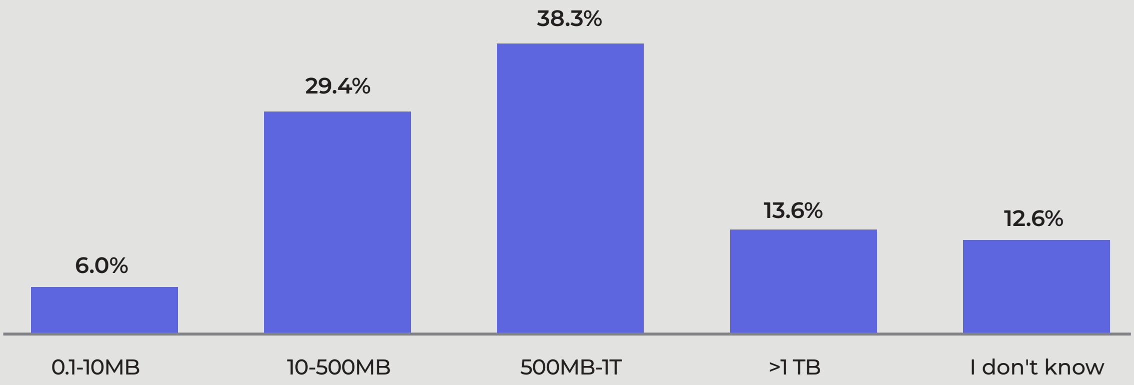


Figure 43: How often are your datasets updated/refreshed?

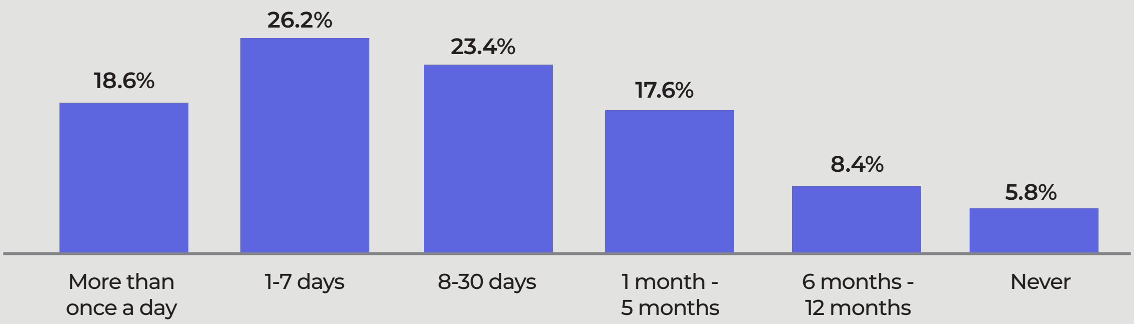


Figure 44: What level of automation do you have for training and deploying AI models?

Figure 45: What does your company use for MLOps?

Only 56.9% of respondents reported having MLOps implemented into their AI operations. 40% of respondents use in-house built tools for MLOps while only 16% of respondents are using a designated platform for MLOps.

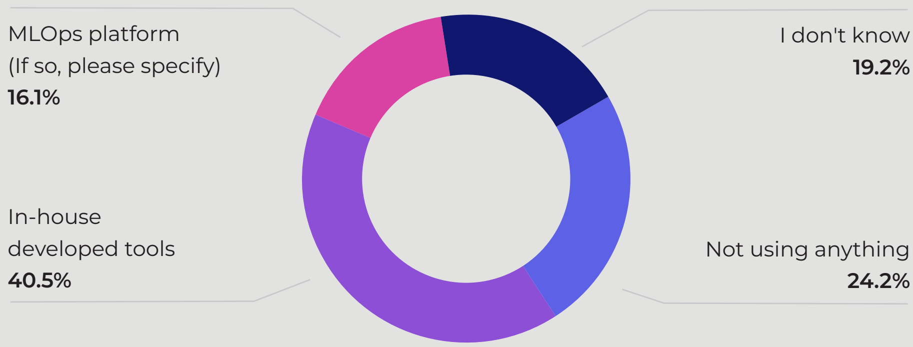


Figure 46: Are you using containers for ML?

76% of respondents are using containers for their ML

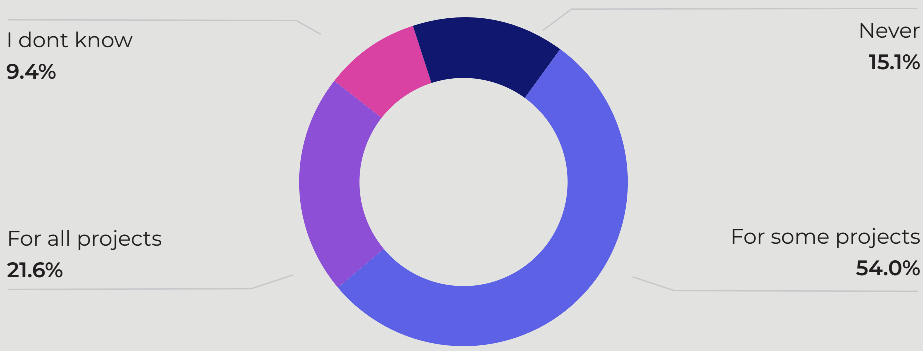
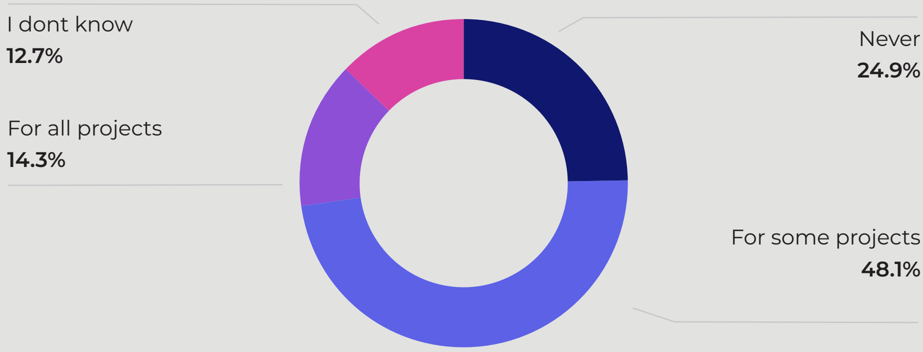


Figure 47: Are you using Kubernetes for ML tasks?

62% of respondents are using Kubernetes for ML tasks.



AI Explainability

These questions aim to gauge how important AI explainability is, and understand who in an organization cares most about explainability. Is it the AI developers themselves? Or is it the key stakeholders and business leaders pushing AI explainability?

Figure 48: As an AI developer, AI explainability is a top concern for me

A comparison between Figure 48 and Figure 49 shows that AI developers and key business stakeholders are mostly aligned on the importance of AI explainability. 75% of AI developers find AI explainability to be a top concern, while slightly less (69% of respondents) feel AI explaibility is a top concern for key stakeholders.

Figure 49: As an AI developer, AI explainability is a top concern for key stakeholders at my organization

Figure 50: How is your team addressing explainable AI?

At what level are organizations addressing AI explainability? 44% of respondents plan to introduce explainable AI techniques in the next 12 months, while only 29% already have AI explainability techniques in place.

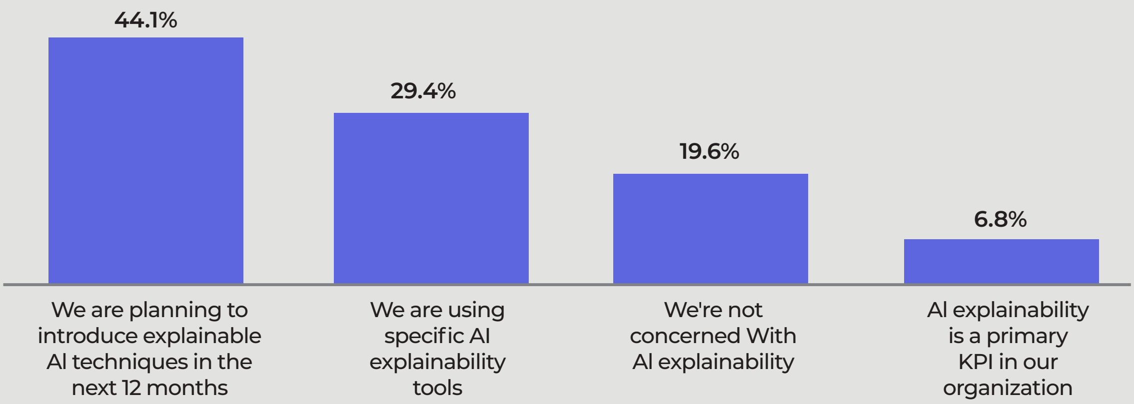
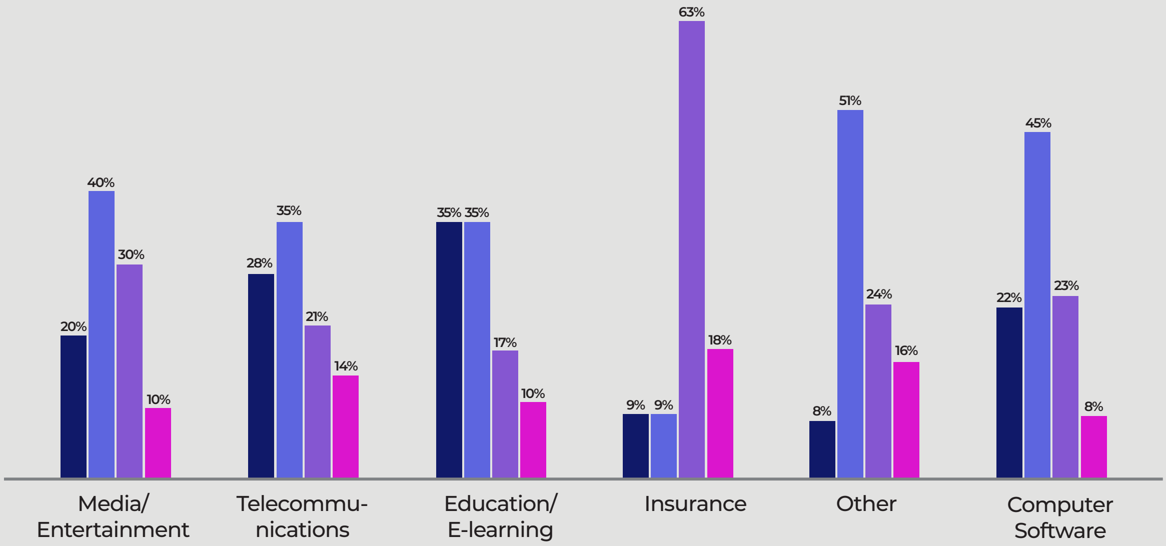


Figure 51: AI Explainability Maturity by Industry

Some industries are more focused on AI explainability than others. Highly regulated industries such as the Insurance industry have already adopted AI explainability tools. However, the respondents from the Automotive and Defense industries suggested that they do not have AI explainability tools in place, but they plan to in the next year.

Industries with lower consumer regulatory pressure such as Education/E-learning, Telecommunications, Computer Software, and IT services are less concerned with AI explainability.



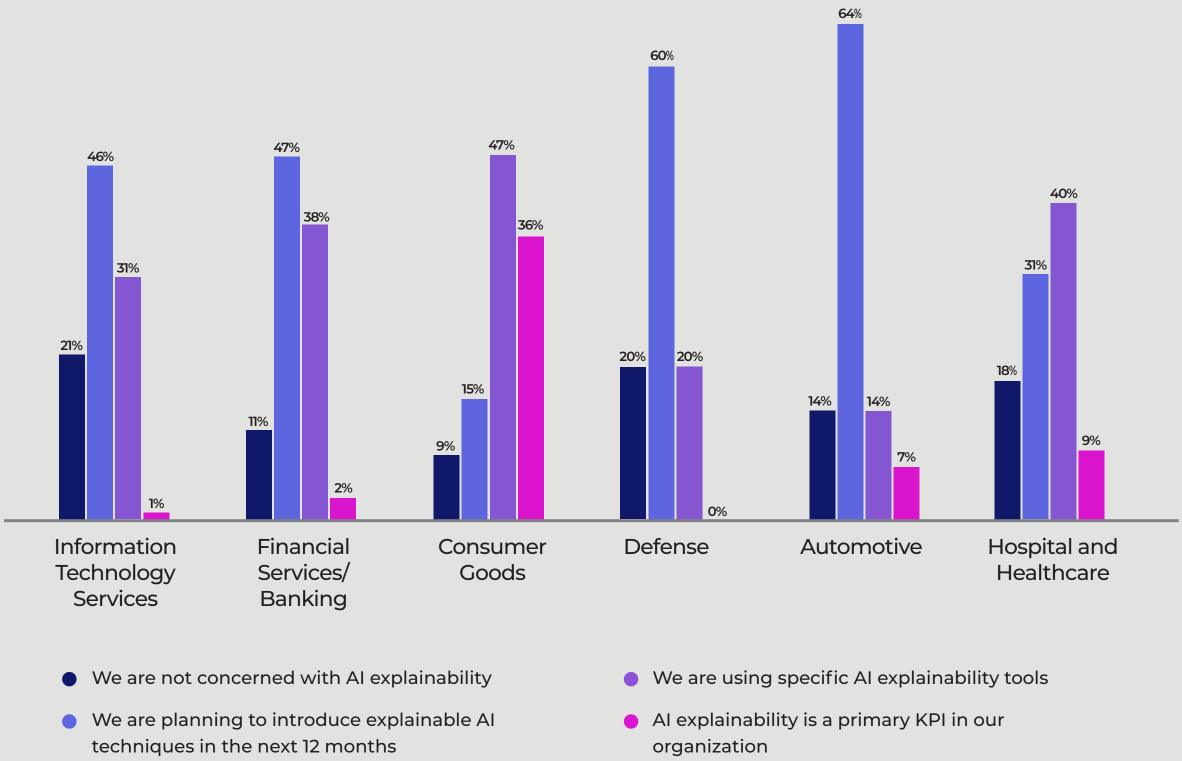
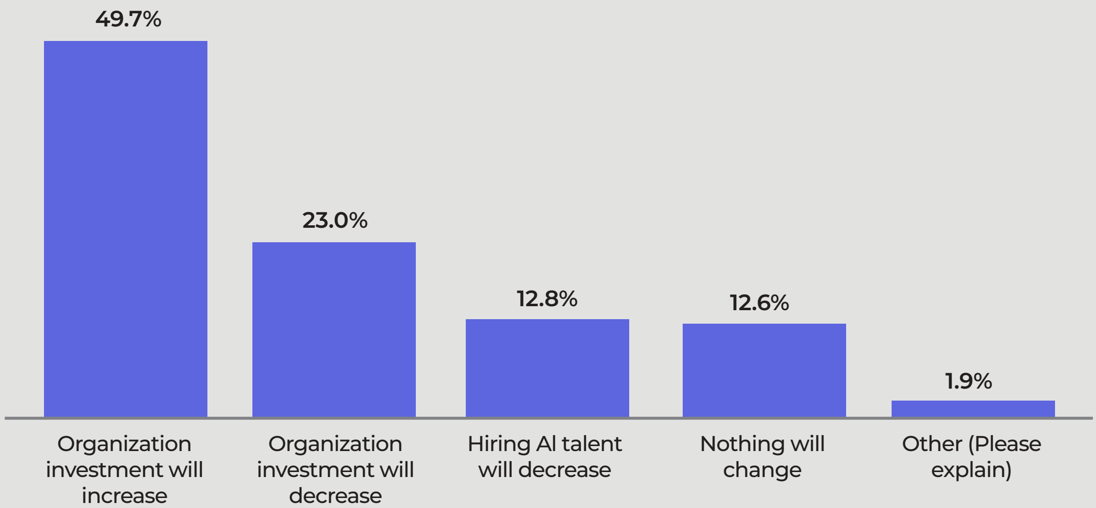


Figure 52: What do you believe the biggest impact will be on AI development from the current economic situation?

Amid high inflation, low market performance, supply chain issues, and uncertainties from both the war in Ukraine and the lingering pandemic, it is safe to say that the global economic state in 2022 has been less than ideal. The tech sector in particular has been hit hard by the economic situation which has led to massive company layoffs, budget cuts, and decreasing earnings from tech leaders. So, what does that mean for the future of AI in 2023?

Nearly 50% of respondents indicated that they believe organization investment in AI development will actually increase.



Summary

With 89% of organizations seeing the benefits of their AI solutions, there is no doubt that artificial intelligence in the industry is growing more valuable to organization's bottom line. However, with 57% of respondents reporting a low maturity, there is still a long way for AI to grow and become more and more adopted and accepted across different industries. AI maturity is still in its infancy and has a lot of potential for further development. Today, many organizations have hurdled the challenge of making AI operational, with 45.8% of respondents reporting that they are not struggling to operationalize their AI. According to ML Insider results, organizations will need to focus on making AI accessible, explainable, and less technically complex to develop.

It will be the responsibility of AI developers to set the tone for AI going forward, with 76.8% of all AI responsibilities leaning on data scientists, engineers, and developers. Leading infrastructure technologies and automation solutions will need to innovate and create tools to support all AI developers and reduce technical complexity at all phases of the AI pipeline. Enabling not only data scientists, but engineers and software developers to successfully build AI will be key to the advancement of AI, as most respondents found AI knowledge and expertise, as well as hiring AI talent to be a top challenge.

Start applying AI today

Browse dozens of AI use cases in the AI Blueprint Marketplace with pre-built ML pipelines that can be quickly integrated with any application.

Learn more about cnvrg.io AI platform to reduce bottlenecks for your AI team.

Thank **you**

by **cnvrg.io**
an **Intel** Company