O'REILLY®

# AI Adoption in the Enterprise

## How Companies Are Planning and Prioritizing AI Projects in Practice

Ben Lorica & Paco Nathan

# AI Adoption in the Enterprise

## *How Companies Are Planning and Prioritizing AI Projects in Practice*

*Ben Lorica and Paco Nathan*

Beijing • Boston • Farnham • Sebastopol • Tokyo

**AI Adoption in the Enterprise**
by Ben Lorica and Paco Nathan



Printed in the United States of America.

Published by O'Reilly Media, Inc., 1005 Gravenstein Highway North, Sebastopol, CA 95472.

O'Reilly books may be purchased for educational, business, or sales promotional use. Online editions are also available for most titles (*http://oreilly.com/safari*). For more information, contact our corporate/institutional sales department: 800-998-9938 or *corporate@oreilly.com*.

**Editors:** Mac Slocum and Jenn Webb
**Production Editor:** Kristen Brown
**Copyeditor:** Octal Publishing Services, Inc.
**Interior Designer:** David Futato
**Cover Designer:** Karen Montgomery
**Illustrator:** Rebecca Demarest

February 2019: First Edition

**Revision History for the First Edition**
2019-01-08: First Release



978-1-492-05179-4

[LSI]

# Table of Contents

# Artificial Intelligence Adoption in the Enterprise

## Introduction

In two recent surveys, we identified trends for "The State of Machine Learning Adoption in the Enterprise" and for "Evolving Data Infrastructure", with the latter looking especially at use of public clouds.

We know companies are taking advantage of artificial intelligence (AI), but we wanted to drill down on the details of how they are planning and prioritizing this work. For example, what's the outlook for how AI adoption patterns might change over the course of the next year?

In this survey, we asked respondents to identify the verticals for their organizations and also to indicate the stage of maturity. In other words, to what extent do organizations have revenue-bearing AI projects in production? We use those two variables for segmenting responses.

This survey includes nine additional questions. At a high level, we asked about budgets for AI projects, what kinds of AI technologies and data are being used, which functional parts of the company benefit from these projects, and what main bottlenecks are preventing further AI adoption. Looking into more detailed questions, we asked about the biggest skills gaps related to AI, which risks they check in machine learning models, and what tools are being used.

Notable findings from the survey include the following:

- Eighty-one percent of respondents work for organizations that already use AI.
- More than 60% of respondents work for organizations planning to spend at least 5% of their IT budget (over the next 12 months on AI). One-fifth (19%) work for organizations planning to spend a significant portion—at least 20%—of their IT budget on AI.
- The level of spending depends on the maturity of an organization. Those with a *mature* practice plan to spend on AI at a much higher rate than less-mature companies. Looking at the lack of investment in AI anticipated in the laggards, we expect the gap between leaders and laggards will widen.
- "Lack of data" and a "lack of skilled people" remain key factors that slow down AI adoption within many organizations. Two other common obstacles pertain to organizational challenges: 23% cited "company culture" and 15% cited "difficulties identifying use cases."
- More than half of all respondents signaled that their organizations were in need of machine learning experts and data scientists. Close to half (47%) cited the need for people who can identify use cases that lend themselves to AI solutions.
- Half of all respondents belonged to organizations that used AI for R&D projects, and one-third used it for customer service or IT. We also found that companies are applying AI in functional areas in which they likely have existing analytic applications. For example, about half (45%) of respondents from the technology industry report AI projects in IT. Respondents from finance report higher rates for customer service and finance/accounting, whereas 70% of all respondents from the health sector signaled that they were using AI for R&D projects.
- Respondents who already use reinforcement learning are beginning to build AI systems in some of the application areas for reinforcement learning that we listed in 2017: customer service; operations, facilities, and fleet management; finance; and marketing, advertising, and PR.
- More than half reported that they were already using deep learning, more than one-quarter (28%) use knowledge graphs/

base, and more than one-fifth are using reinforcement learning. We found that more than one-third of all respondents (35%) are already using images and video in their AI systems.

- Although we found that more than half (53%) of all respondents who are already using deep learning use it for computer vision applications (images, video), a lot more are using it for "enterprise data"—86% use it for structured data, and 69% use it for text.
- Our survey results confirm a strong interest in several important issues that go beyond optimizing business metrics: close to half (45%) check for model transparency and interpretability, 41% check for fairness and bias, and one-third are checking that their AI systems are reliable and safe.
- In our 2018 survey, which focused on deep learning, we found the top three deep learning tools to be TensorFlow (used by 61% of all respondents), Keras (25%), and PyTorch (20%). This year, we report a higher rate of usage for Keras (34%) and PyTorch (29%).
- Modern data science platforms include features that can improve productivity, enable experimentation, and enhance collaboration. We found close to half of all respondents want to incorporate tools for model visualization (particularly useful for deep learning) and "AutoML" (model and hyperparameter search), and about one-third want better tools for tracking data lineage.

**NOTE**

For this report, we aren't making a strong distinction between AI and machine learning, so we use the two interchangeably. Except for knowledge graphs, most of the technologies we cover are forms of machine learning.

## Survey Respondents

The survey ran for a few weeks, closing in mid-November 2018; we received more than 1,300 responses.

Throughout the survey, we compared responses across three major industry groups for which we had a sizable number of respondents (Figure 1-1):

- *Technology* = computers, electronics, technology
- *Finance* = financial services
- *Health* = healthcare, life sciences

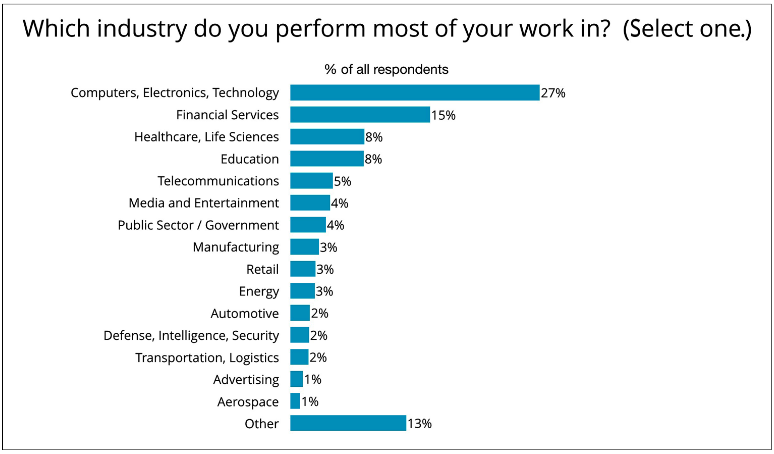


*Figure 1-1. Industry for respondents*

For the remainder of this report, we also adopt the following terminology to describe cohorts from our survey:

*Not yet using AI*
: Respondents who work for organizations that are not using any AI projects yet; however, they answered based on how they would aspire to approach AI projects (for example, projected budget levels).

*Evaluation stage*
: Respondents who work for organizations that so far have been limited to trial evaluations and "proof of concept" (PoC).

*Mature practice*
: Respondents who work for organizations that have revenue-bearing AI projects in production.

As Figure 1-2 shows, more than half of respondents work for organizations that are evaluating AI technologies, more than a quarter (27%) have revenue-bearing AI projects in production, and one-fifth (20%) are not yet using AI.

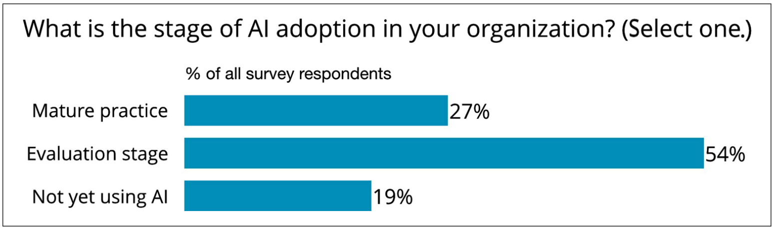


*Figure 1-2. Stage of maturity for AI adoption*

Maturity varies by industry: only one-tenth (11%) of those who work in finance belong to organizations that have not yet used AI (Figure 1-3), compared to 29% in the public sector. In contrast, one-third (30%) of those who work in finance describe having a mature AI practice, compared to 16% from the public sector who describe their organizations as having a mature practice.

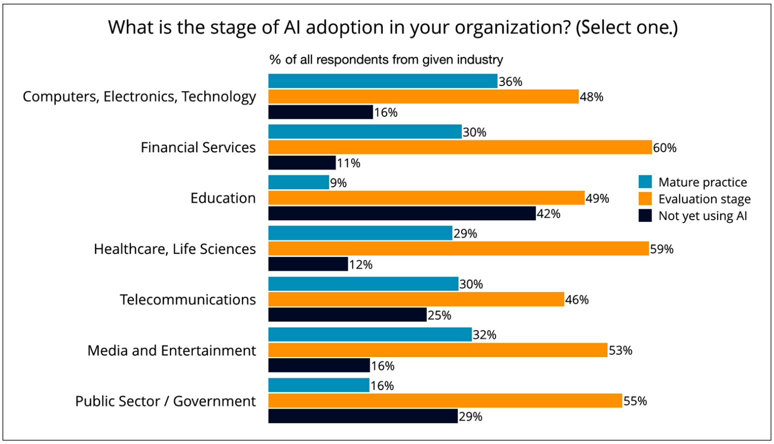


*Figure 1-3. Industry, by stage of maturity*

# Investing in AI

We wanted to gauge the level of investment in AI over the near-term. As shown in Figure 1-4, we found that more than 60% planned to spend at least 5% of their IT budget toward AI, including a fifth (19%) who plan to spend more than 20% of their budget.

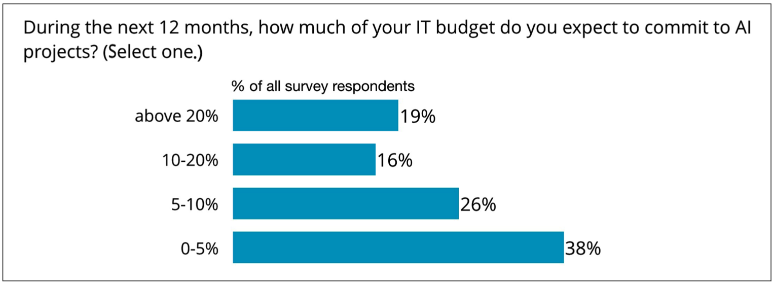


*Figure 1-4. Budget allocation over the next year*

The level of spending depends on the maturity of an organization. To the extent that AI solutions are deployed wisely, the gap between those with mature practices and those who have yet to adopt AI will continue to widen over the next 12 months, as demonstrated in Figure 1-5.

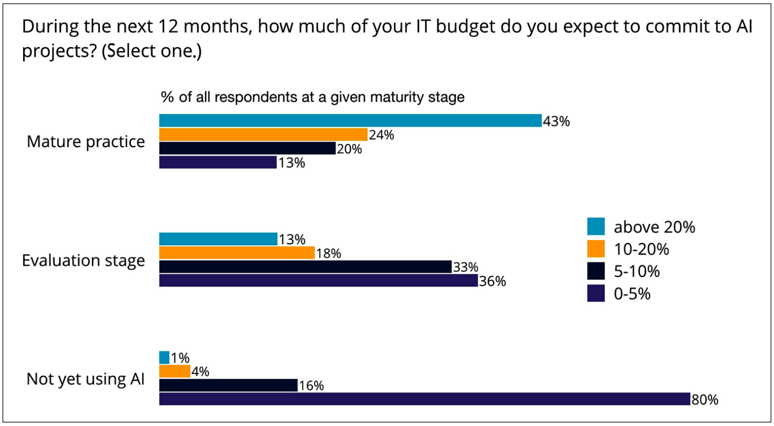


*Figure 1-5. Budget, by stage of maturity*

Looking at investment plans across several industries (Figure 1-6), the technology, health, and retail sectors are planning to invest most aggressively in AI, whereas the public sector is planning to invest the least.

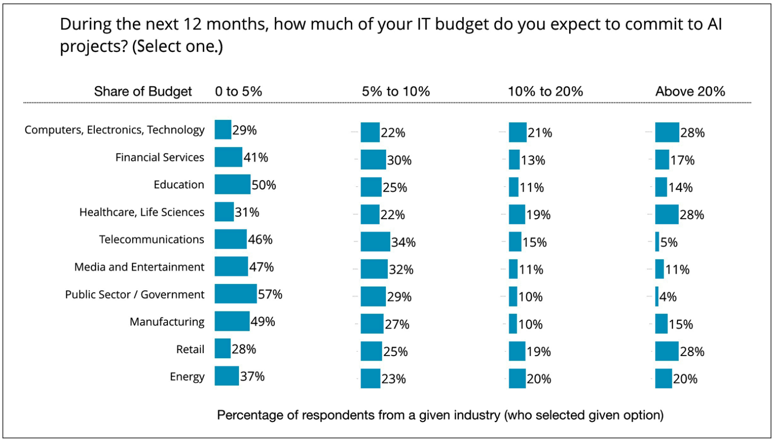


*Figure 1-6. Budget, by industry*

# What Is Holding Back Adoption of AI

Conversations with members of the community as well as previous surveys point to "lack of data" and a "lack of skilled people" as key factors that slow AI adoption within many organizations. Two other common obstacles pertain to organizational challenges—Figure 1-7 shows that respondents cited "company culture" and "difficulties identifying use cases" as serious challenges.

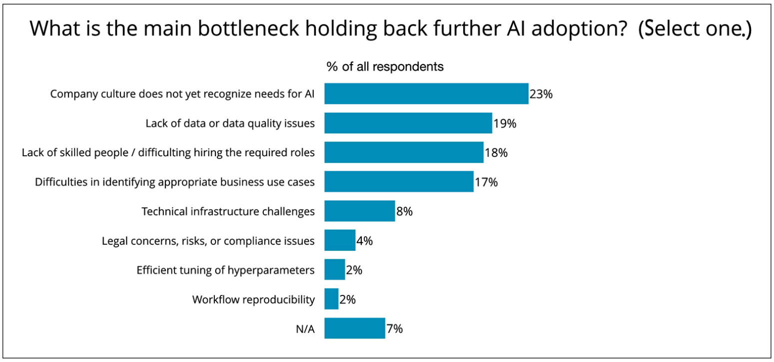


*Figure 1-7. Challenges for AI adoption*

Given that 81% of respondents work for companies that already use AI, it's not surprising that 8% of respondents cite technical infrastructure challenges as their main bottleneck. Across a broader audience (fewer AI adopters), this would likely be seen as a more common obstacle.

Also, given the difficulty of workflow reproducibility, it's surprising to see only 2% cite that as a bottleneck. Considering how the top four challenges can create "non-starter" scenarios, it's possible that workflow reproducibility doesn't surface until later stages of maturity for an organization.

Challenges posed by company culture, lack of data, and a skills gap held across technology, finance, and the health industry. However, respondents from the health industry had less trouble identifying appropriate use cases for AI, as illustrated in Figure 1-8.

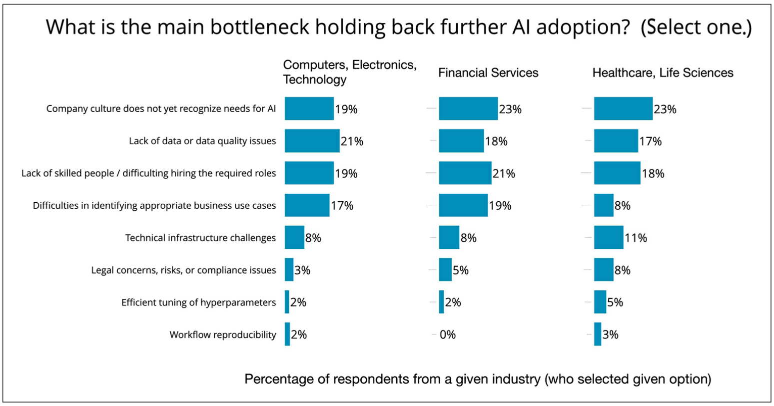


*Figure 1-8. Challenges, by industry*

Figure 1-9 offers a look at the challenges segmented by stage of maturity. It's interesting that mature practices seem to have overcome the challenges related to company culture or not recognizing use cases. Even so, mature practices adhere to the top two cited bottlenecks we've seen in previous surveys: data and talent.

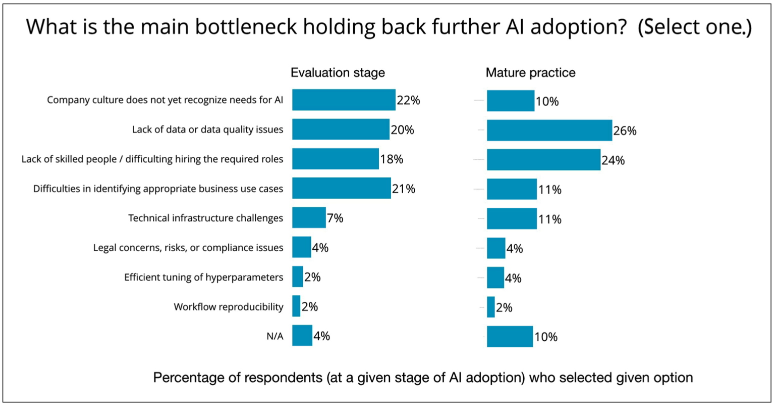


*Figure 1-9. Challenges, by stage of maturity*

# Understanding the Skills Gap

In a recent study, LinkedIn found that within the United States, demand for data scientists is "off the charts." Our survey confirms this strong demand (Figure 1-10): more than half of all respondents signaled their organizations were in need of machine learning experts and data scientists. But just as with any (new) technology, companies are also in need of people who can identify use cases that lend themselves to AI solutions.

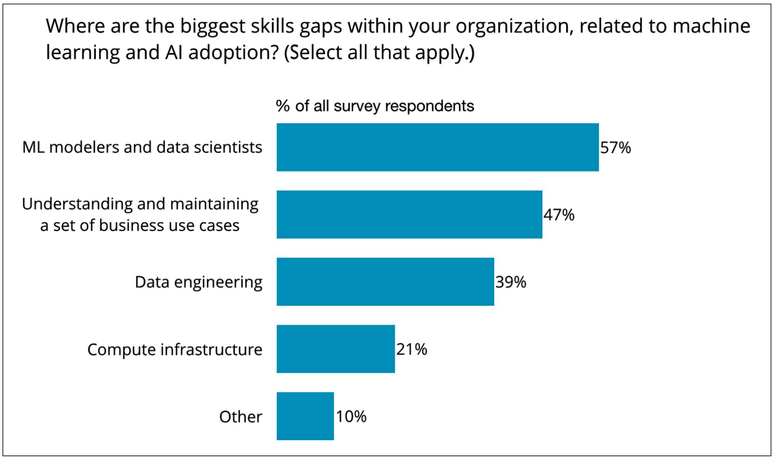


*Figure 1-10. Skills gap*

Figure 1-11 looks at hiring needs across three major industries. Compared to the technology sector, respondents from health and finance have a larger need for data and infrastructure engineers.

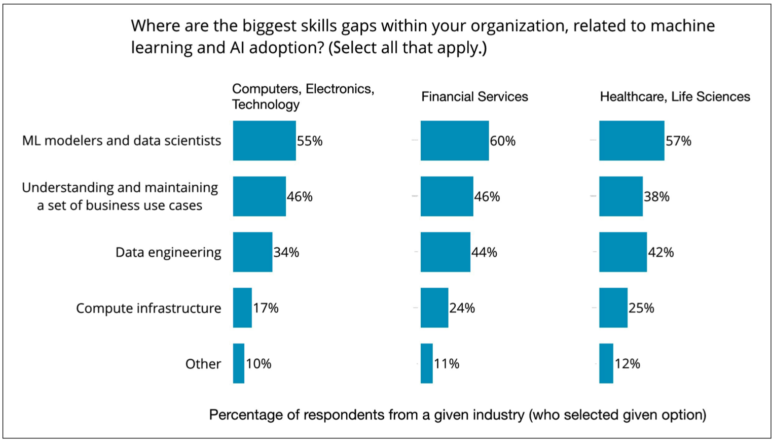


*Figure 1-11. Skills gap, by industry*

Slicing the skills gap by stage of maturity, the mature practices have need for more machine learning experts and data scientists, although less so than companies at the evaluation stage. Those roles are likely to be staffed already when companies have a mature practice. However, priorities for data engineering and compute infrastructure remain the same, as illustrated in Figure 1-12.

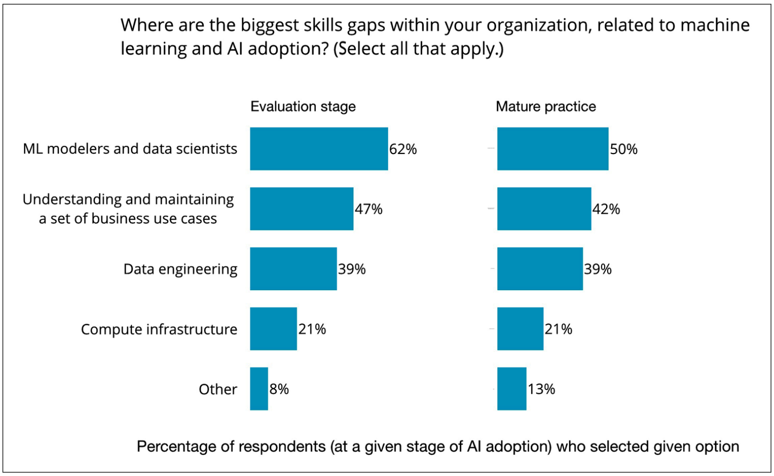


*Figure 1-12. Skills gap, by stage of maturity*

## How Organizations Are Using AI

Half of all respondents belong to organizations that use AI for R&D projects (Figure 1-13), and one-third use it for customer service or IT. As we noted in an earlier post, IT is an area that lends itself to (partial) automation; thus, many AI solutions already target IT systems.

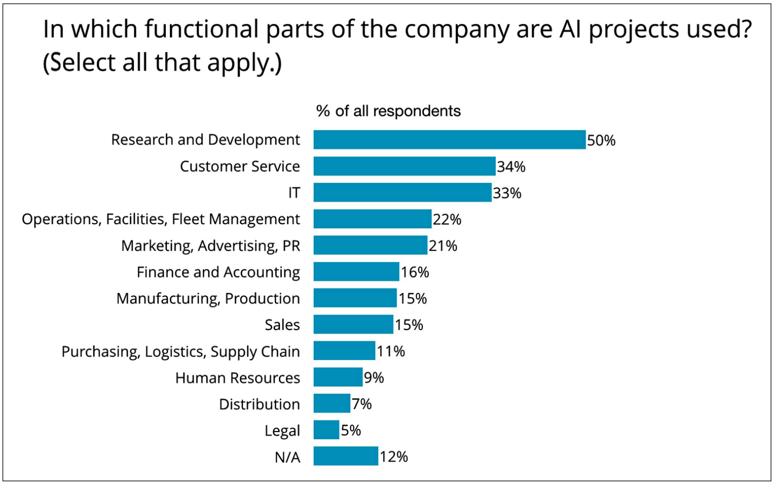


*Figure 1-13. AI projects*

We also found that companies are applying AI in functional areas in which they likely have existing analytic applications. For example, Figure 1-14 demonstrates that about half (45%) of respondents from the technology industry report AI projects in IT. Respondents from finance report higher rates for customer service and finance/accounting, whereas 70% of all respondents from the health sector signaled they were using AI for R&D projects.

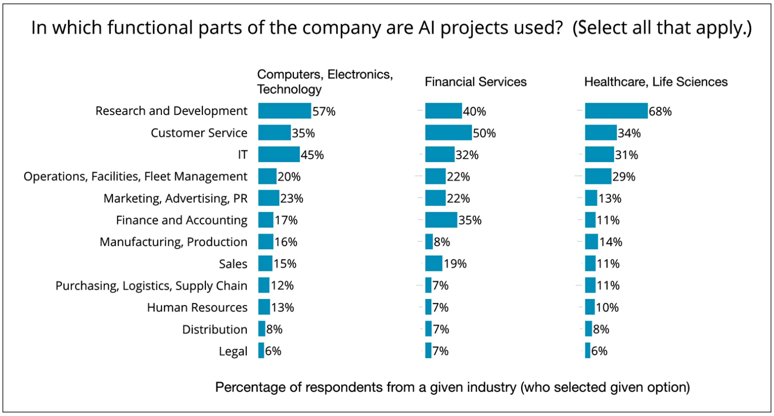


*Figure 1-14. AI projects, by top three industries*

# Building-Block Technologies

In Figure 1-15, we can see that more than half of respondents reported they were already using deep learning, nearly one-third use active learning ("human-in-the-loop"), more than a quarter (28%) use knowledge graphs/base, and more than a fifth are using reinforcement learning.

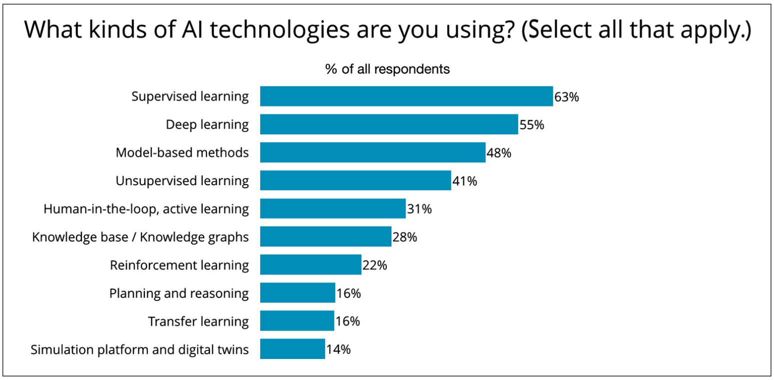


*Figure 1-15. AI technologies used*

Looking at technologies used across key industries (Figure 1-16), we found that respondents from technology and health industries use reinforcement learning and knowledge graphs/base at a higher rate than those from the finance sector.

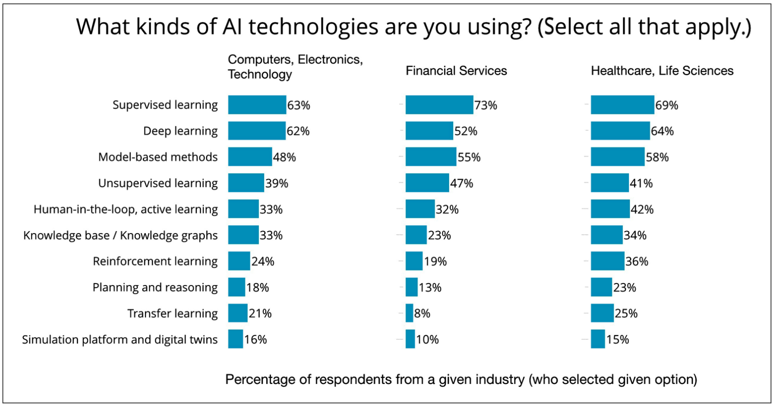


*Figure 1-16. AI technologies, by industry*

Looking at technologies by stage of maturity, the mature practices use more technologies across the board, as would be expected. The largest contrast is in transfer learning, as depicted in Figure 1-17, where mature practices report using nearly three times the rate compared with evaluation-stage practices. There are a wider range of use cases for transfer learning that can be identified after an organization reaches a higher level of maturity.

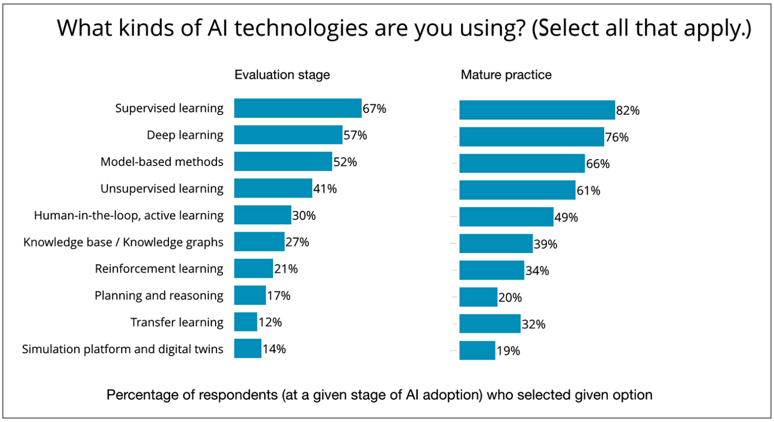


*Figure 1-17. AI technologies, by stage of maturity*

## Data Types

Although the resurgence of deep learning can be traced back to important breakthroughs in computer vision and speech technologies, as we noted earlier, enterprises are using AI in areas for which they already have some data and analytics in place (refer back to Figure 1-14). Thus, it's no surprise that organizations are using structured data and text to train their AI systems. One of the areas in which AI (specifically, deep learning) has made enough progress to be productized is computer vision. In fact, as Figure 1-18 illustrates, we found that more than one-third of all respondents (35%) are also already using images and video in their AI systems.

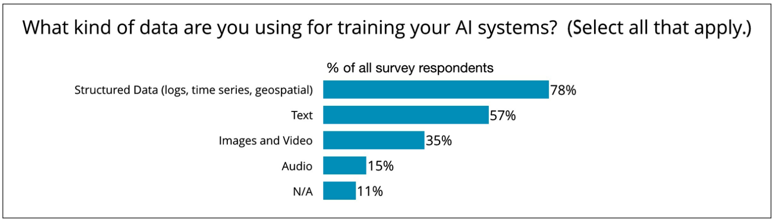


*Figure 1-18. Data types*

Looking at data types in some key industries, we found that respondents from the health sector used computer vision (images and video; Figure 1-19) at a higher rate than respondents from finance and technology. This is in line with anecdotal information that computer vision is beginning to make inroads in medical imaging and radiology.

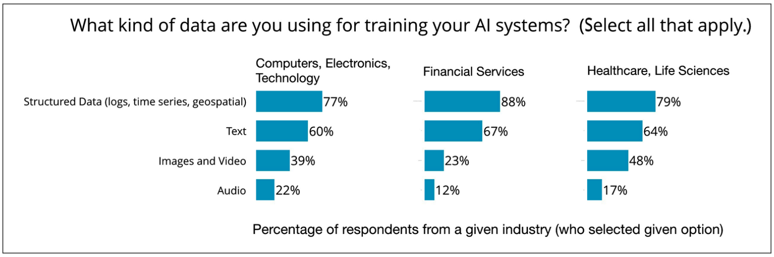


*Figure 1-19. Data types, by industry*

Segmenting the data types by stage of maturity (Figure 1-20), we found that the use of structured data stays relatively the same, whereas the other types of data are used substantially more.

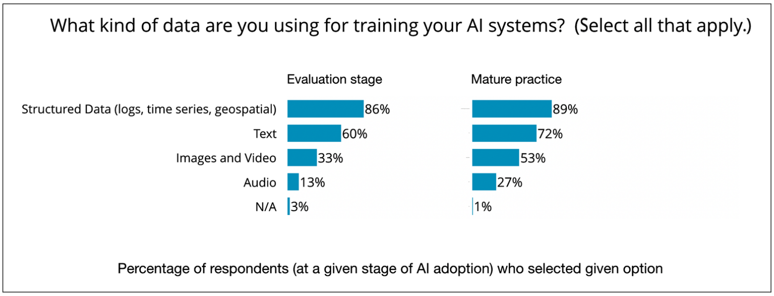


*Figure 1-20. Data types, by stage of maturity*

# Deep Learning and Reinforcement Learning

The AI applications of tomorrow will be hybrid systems, composed of several components and reliant on many different methodologies. With that said, much of the excitement around AI is due to progress in deep learning and reinforcement learning. In this section, we briefly examine respondents who are already using these methodologies and determine which AI projects their organizations have been working on. To do so, we isolate respondents who signaled that they are already using deep learning and reinforcement learning.

First, even though we found that more than half (53%) of all respondents who are already using deep learning use it for computer vision applications (images, video), Figure 1-21 shows that a lot more are using it for structured data (86%) and text (69%).

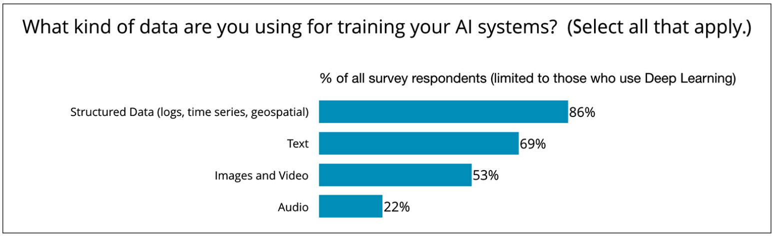


*Figure 1-21. Data types, limited to deep learning*

Again, respondents from the health industry report higher rates of use of deep learning for computer vision (images and video) and speech technologies (audio), as depicted in Figure 1-22.

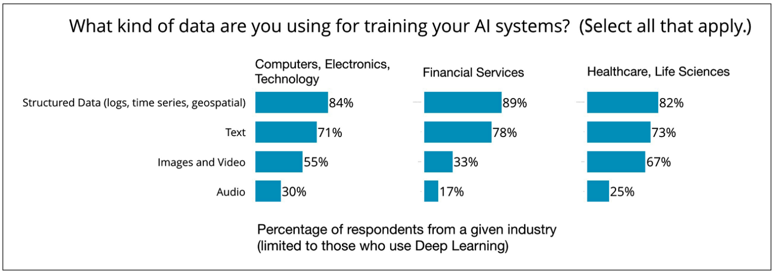


*Figure 1-22. Data types, by industry (percentage of respondents)*

So, what areas are deep learning and reinforcement learning users focusing on? First, let's look at the AI projects with which deep learning users have been involved. Aside from R&D, we found that more than one-third of all respondents (Figure 1-23) who are already using deep learning are applying AI in IT and customer service.

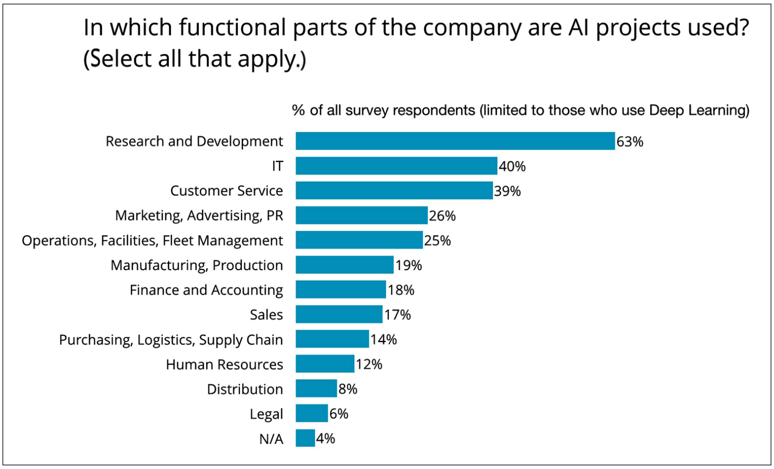


*Figure 1-23. AI projects, limited to deep learning*

Looking at technologies used across key industries, we found results to be in line with other AI technologies (see Figure 1-14). Figure 1-24 illustrates that respondents from finance report higher rates for customer service and finance/accounting, whereas 76% of all respondents from the health sector are using AI for R&D projects.

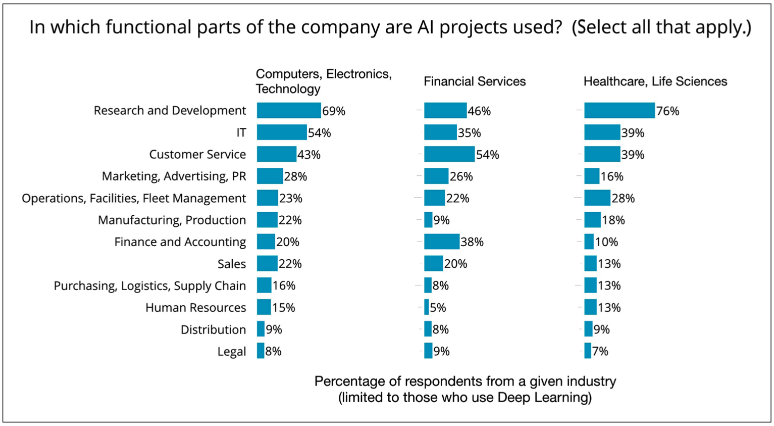


*Figure 1-24. Projects, by industry, limited to deep learning (percentage of respondents)*

At the end of 2017, we listed areas where we found real-world applications that used reinforcement learning. At the time, we highlighted robotics and industrial automation, text, speech and dialogue systems, media and advertising, and other areas. To get a sense of potential applications of reinforcement learning, we isolated respondents who already use reinforcement learning (Figure 1-25). We found that reinforcement learning users are beginning to build AI systems in some of the application areas we listed in 2017: customer service; operations, facilities, and fleet management; finance; and marketing, advertising, and PR.

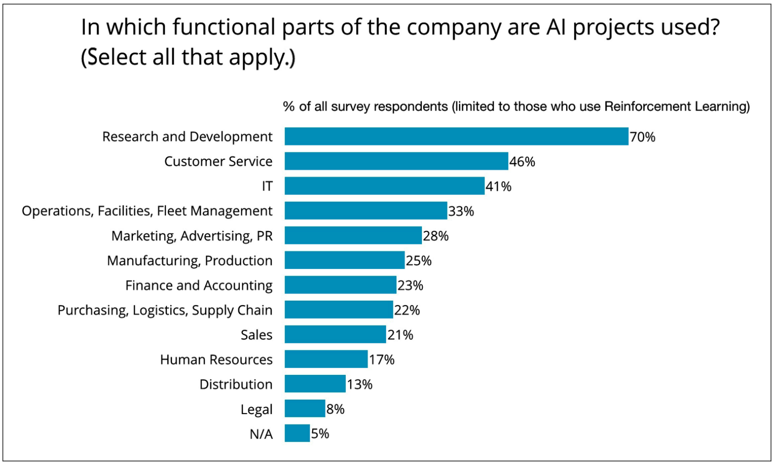


*Figure 1-25. Projects, by industry, limited to reinforcement learning*

> **NOTE**
>
> There aren't enough respondents who use reinforcement learning to do a version of the previous chart "by industry."

## Risks

In a recent post, we observed that when it comes to AI and machine learning, there are many important considerations that go beyond optimizing a statistical metric. Our survey results, presented in Figure 1-26, confirm strong interest in several important issues: close to half (45%) check for model transparency and interpretability and one-third are checking that their AI systems are reliable and safe.

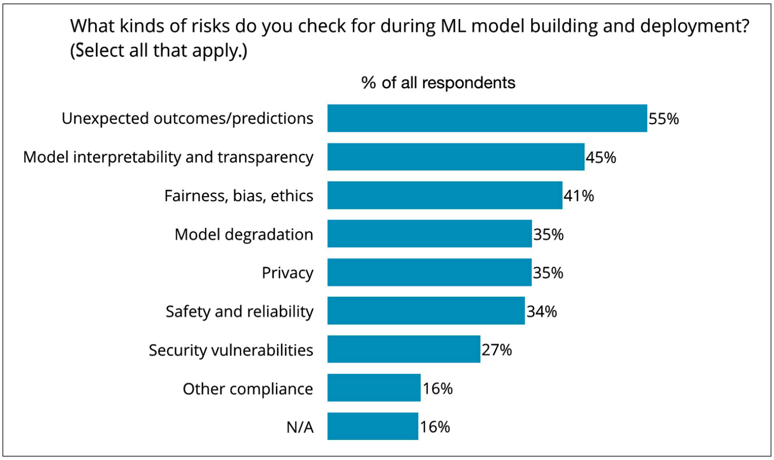


*Figure 1-26. Risks checked (percentage of respondents)*

Examining results in key industries, more than half of all respondents from finance and health check for model transparency and interpretability, and more than half of all respondents from the health sector already check for reliability and safety, as shown in Figure 1-27.

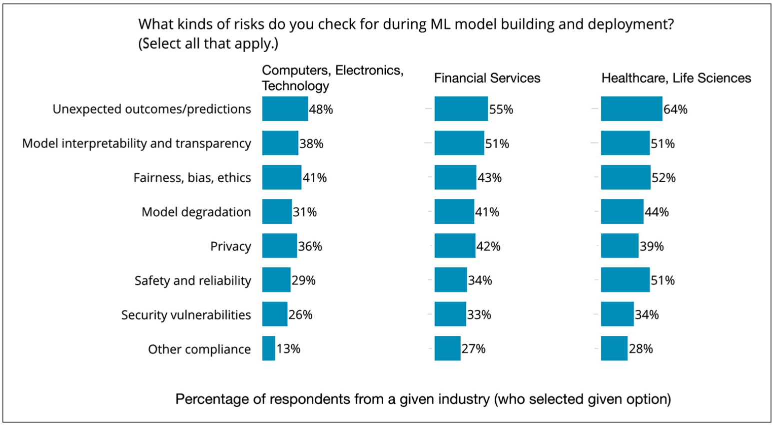


*Figure 1-27. Risks checked, by industry (percentage of respondents)*

Looking at the risks checked by maturity level (Figure 1-28), the mature practices show more consideration in each area. This fits with previous surveys—as organizations deploy machine learning models in production, with experience they tend to learn to consider the associated risks more closely.

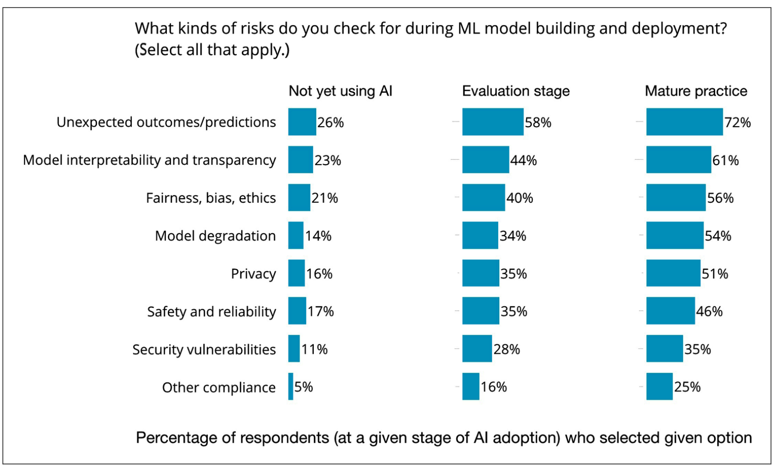


*Figure 1-28. Risks checked, by stage of maturity (percentage of respondents)*

# Tools for Building AI Applications

In our 2018 survey, which focused on deep learning, we found the top three deep learning tools to be TensorFlow (at the time, used by

61% of all respondents), Keras (25%), and PyTorch (20%). This year, we report a higher rate of usage for Keras (34%) and PyTorch (29%), as demonstrated in Figure 1-29.

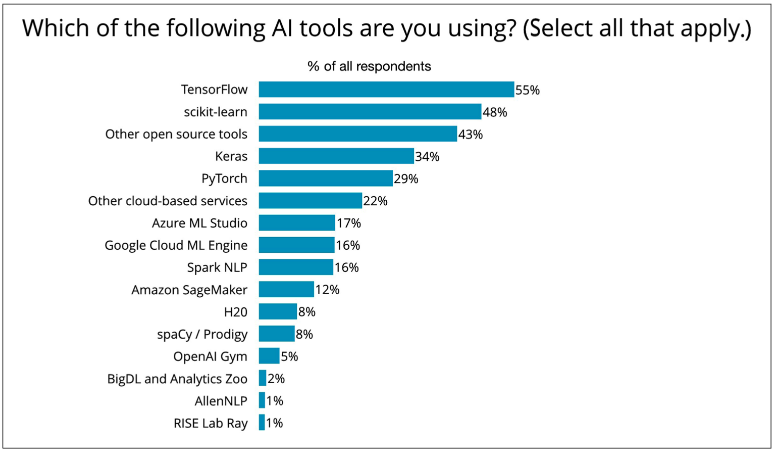


*Figure 1-29. AI tools used*

In a recent post, we reported on some of the first production systems that use Ray (a flexible, high-performance distributed execution framework). Although only 1% of respondents reported *directly* using Ray, it's worth pointing out that Ray is being used in other systems. From its initial libraries focused on reinforcement learning and hyperparameter tuning, Ray is now helping users of the popular Python library Pandas scale to larger datasets via a related project called Modin. Microsoft's acquisition of Bonsai means users of its autonomous AI products are also using Ray (Ray is used in Bonsai's backend). Also, Amazon Web Services (AWS) recently began supporting Ray in its SageMaker service.

Figure 1-30 presents the tools used in three key sectors.

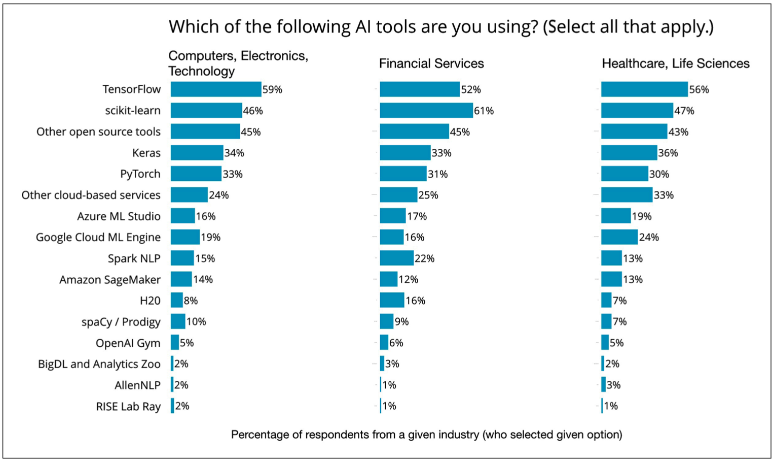


*Figure 1-30. AI tools, by industry (percentage of respondents)*

Beyond individual libraries, companies are looking to build platforms that will increase the productivity of their data scientists and machine learning engineers. Modern data science platforms include features that improve productivity, enable experimentation, and enhance collaboration. As Figure 1-31 illustrates, we found that close to half of all respondents want to incorporate tools for model visualization (particularly useful for deep learning) and "AutoML" (model and hyperparameter search), and about one-third want better tools for tracking data lineage.

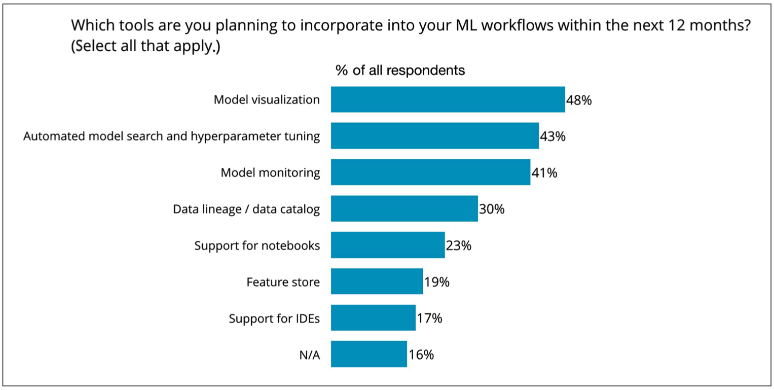


*Figure 1-31. Workflow features planned over the upcoming year*

Looking at the sought-after tooling in three key sectors (Figure 1-32), the push toward incorporating AutoML within the next year is even more pronounced in finance, and there's a large bump in model visualization for healthcare.

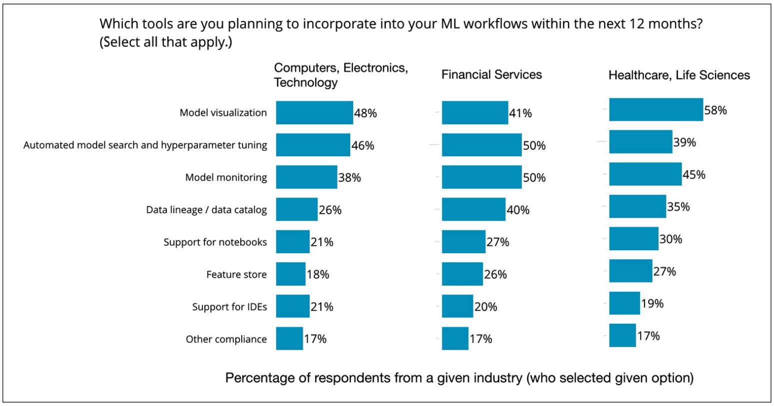


*Figure 1-32. Workflow features, by industry (percentage of respondents)*

Looking at the sought-after tooling by stage of maturity (Figure 1-33), the mature practices show more interest in each category of tooling, but especially a large push toward AutoML—nearly twice that of the evaluation stage.

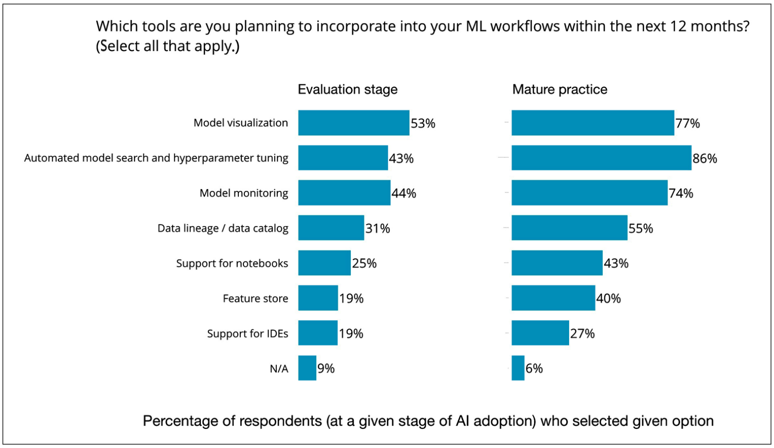


*Figure 1-33. Workflow features, by stage of maturity (percentage of respondents)*

# Overall Analysis

We can draw some high-level conclusions using these results to complement earlier analysis.

There continues to be strong interest in important issues beyond merely optimizing for business metrics. That fits well with an earlier survey this year about AI adoption that indicated how serious work for AI in production regarding accountability, compliance, ethics, and so on is much more than a passing fad. Key risks considered include model transparency and interpretability, checks for fairness and bias, and checks that AI systems are reliable and safe. In stark contrast to Silicon Valley tech firms and "lean startup" methodology, the banks—which are arguably the original "data-driven" organizations—tend to consider key risk indicators (KRIs) more so than key performance indicators (KPIs).

Companies are applying AI in functional areas in which they likely have existing analytic applications, building atop that base. Meanwhile the scope of "AI" is expanding. Half of the organizations in our survey already use deep learning. One-third use human-in-the-loop. One-quarter use knowledge graphs. One-fifth use reinforcement learning. Note that reinforcement learning is probably more widely used in production in industry than has been generally perceived, and watch for reinforcement learning to become much more pervasive among enterprise solutions.

Transfer learning provides an interesting nuance, given how its use in production tends to require more experienced practitioners. We see mature practices making use of transfer learning at nearly three times the rate of evaluation stage companies. There's value in applications of transfer learning, although those are perhaps not as apparent to the uninitiated.

Having said that about the variety of use cases and technologies, realistically there are four categories of data used for AI so far: images and video, audio, text, and structured data. Plan accordingly.

In terms of personnel, most organizations need machine learning experts and data scientists. There's clearly a talent crunch in that direction. Even so, avoid building a lopsided team: mature practices might have hired small armies of data scientists but still lack other crucial roles. In particular, you'll need people who can identify use cases that fit AI solutions.

## About the Authors

**Ben Lorica** is the Chief Data Scientist at O'Reilly Media, Inc. and is the Program Director of both the Strata Data Conference and the Artificial Intelligence Conference. He has applied Business Intelligence, Data Mining, Machine Learning and Statistical Analysis in a variety of settings including Direct Marketing, Consumer and Market Research, Targeted Advertising, Text Mining, and Financial Engineering. His background includes stints with an investment management company, internet startups, and financial services.

**Paco Nathan** is known as a "player/coach," with core expertise in data science, natural language processing, machine learning, and cloud computing. He has over 35 years of tech industry experience, ranging from Bell Labs to early-stage start-ups. He is the co-chair of Rev and advisor for Amplify Partners, Deep Learning Analytics, and Recognai. His recent roles include Director, Learning Group at O'Reilly Media and Director, Community Evangelism at Databricks and Apache Spark. He was cited in 2015 as one of the "Top 30 People in Big Data and Analytics" by Innovation Enterprise.