

A Forrester Consulting
Thought Leadership Paper
Commissioned By HPE

March 2021

AI Plus HPC: The Future Of Advanced Analytics

AI And HPC Will Support Each Other And
Merge Over Time

Table Of Contents

- 3** Executive Summary
- 4** AI Is The Focus Of Most Firms' Data Innovation Strategies
- 7** AI + HPC: The Future Of Advanced Analytics
- 12** AI And HPC Collaboration Is Critical To Shared Success
- 14** Key Recommendations
- 15** Appendix

Project Director:

Andrew Magarie,
Principal Market Impact Consultant

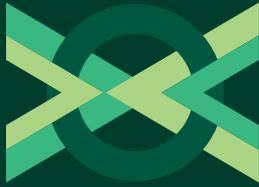
Contributing Research:

Forrester's Infrastructure &
Operations research group

ABOUT FORRESTER CONSULTING

Forrester Consulting provides independent and objective research-based consulting to help leaders succeed in their organizations. Ranging in scope from a short strategy session to custom projects, Forrester's Consulting services connect you directly with research analysts who apply expert insight to your specific business challenges. For more information, visit forrester.com/consulting.

© 2021, Forrester Research, Inc. All rights reserved. Unauthorized reproduction is strictly prohibited. Information is based on best available resources. Opinions reflect judgment at the time and are subject to change. Forrester®, Technographics®, Forrester Wave, RoleView, TechRadar, and Total Economic Impact are trademarks of Forrester Research, Inc. All other trademarks are the property of their respective companies. For additional information, go to forrester.com. [E-49428]



Synergies between AI and HPC can enhance the value of both, with shared infrastructure being a critical component of unifying those disciplines.

Executive Summary

Combining artificial intelligence (AI) and high-performance computing (HPC) will unlock the potential of each of these powerful analytics disciplines, leading to increased business agility, innovation, and competitive differentiation. To do this successfully, organizations must 1) integrate AI and HPC infrastructure to create synergies through shared resources and improved flexibility and 2) improve collaboration between AI and HPC — traditionally siloed organizations.

HPE commissioned Forrester Consulting to evaluate the current state of HPC and AI infrastructure. Forrester conducted an online survey with 464 global decision-makers and practitioners, divided into business decision-makers relying on AI and/or HPC, AI experts, and HPC experts, as well as three interviews, to explore this topic. We found that synergies between AI and HPC can enhance the value of both, with shared infrastructure being a critical component of unifying those disciplines.

KEY FINDINGS

- › **AI is an emerging discipline with high expectations.** While many firms are currently investing in their AI disciplines, early returns show that AI capabilities are limited, and most are still in early stages of maturity. Sixty-one percent of AI experts say their firms are investing in infrastructure improvements currently, but just 26% of AI projects are in full deployment, and fewer than half of proofs of concept (POCs) deliver expected business value.
- › **Combining AI and HPC is the future.** Most HPC experts say workflows incorporating machine learning algorithms to improve HPC speed and reduce costs will happen in the next year. Half of AI experts say they are using HPC infrastructure to improve unsupervised learning and machine learning (ML) model training by expanding processing flow with higher-performance data sourcing and calculating capabilities. As a result, experts expect unifying AI and ML to deliver critical benefits to innovation, competitive differentiation, business agility, and cost savings.
- › **Firms need integrated infrastructure to capitalize on the promise of AI and HPC.** More than eight in 10 AI experts say that they will need to improve their infrastructure to meet future plans for AI. Meanwhile, over half of HPC experts say they need infrastructure upgrades to meet even current needs, with an additional 35% saying while their current infrastructure meets needs, they will require future improvements. With many of these future infrastructure improvements providing benefits to both AI and HPC, Forrester predicts growing support for integrating AI and HPC infrastructure.

Experts expect unifying AI and ML to deliver critical benefits to innovation, competitive differentiation, business agility, and cost savings.

AI Is The Focus Of Most Firms' Data Innovation Strategies

An explosion of enterprise data has brought interest in AI to a fever pitch. Organizations are piloting AI projects across the business, ramping up capabilities to serve many use cases and stakeholders. However, while AI capabilities rapidly proliferate across the business, data challenges are surfacing that hamper the promise of AI. High-performance computing capabilities can be the key to overcoming these challenges and unlocking the potential of AI.

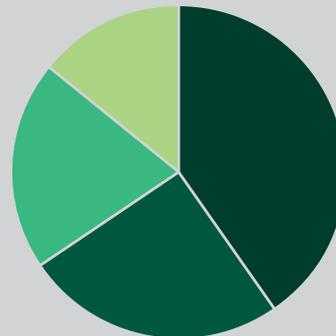
Our survey of business decision-makers, AI experts, and HPC experts shows:

- > Businesses are banking on the AI revolution.** Many businesses are investing heavily in AI and creating centers of excellence to support AI initiatives (see Figure 1). AI is currently a work in progress for most, with just 26% of AI projects in full deployment today and 57% of respondents reporting that AI POCs are not delivering expected business value. In addition, less than half of all respondents (43%) rate their AI discipline as mature. It is still early days for most AI programs, so investing intelligently in AI infrastructure and strategy will be crucial to proving the value of AI and moving beyond POCs.

Figure 1

Firms Are Centralizing Their AI Organizations

- 40%** We have a center of excellence that supports the work of many business unit-focused teams.
- 25%** We have a central team of experts who do most of the work.
- 20%** We are too early on our journey to say we have a real operating model.
- 14%** We have teams at the business unit level that work somewhat independently.



And Investing To Meet Infrastructure Needs

63% Currently investing in AI infrastructure

37% AI infrastructure needs to be upgraded to meet current needs

45% AI infrastructure needs to be upgraded to meet future needs

AI Is Still Evolving For Most Organizations

26%
Average percent of AI models in full deployment


Fewer than half of all respondents (43%) rate their AI disciplines as mature.*

Base: 209 global decision-makers or practitioners responsible for AI/ML

*Base: 464 global decision-makers or practitioners responsible for high-performance computing or AI/ML

Source: A commissioned study conducted by Forrester Consulting on behalf of HPE, October 2020

- › **AI infrastructure investment is key to future success.** Thirty-seven percent of AI experts in our study report needing infrastructure updates to meet current AI needs while 45% say they will need to invest in the future. Firms are investing to meet this demand: 61% of AI experts say their companies are currently investing in infrastructure improvements.
- › **Synergies between AI and HPC drive value for both.** Firms can combine AI and HPC capabilities to improve both disciplines. For example, training AI models, especially complex models, can be very compute-intensive, often bottlenecking model training capabilities. HPC can help solve these bottlenecks by powering deep learning using very large data sets. Half of AI experts we surveyed say they are expanding their AI processing flow with higher-performance data sourcing and calculating capabilities. Our survey also shows that firms with mature HPC disciplines are more likely to be further along in their AI development, with more full-scale model deployment and fewer proof-of-concept projects.

TO UNLOCK AI, SOLVE NEXT-LEVEL DATA CHALLENGES

Data fuels AI models, and advancements such as deep learning have only increased firms' hunger for it. However, firms are running into next-level data challenges that current AI infrastructure is not well equipped to solve. We asked 464 AI and HPC experts and business decision-makers about the most important data challenges they must solve to meet today's business goals (see Figure 2). Top challenges include the following:

- › **Challenge #1: Source more data from the edge of networks.** The data downpour keeps coming. Eighty percent of all respondents say sourcing and using unstructured and edge data from sensors and physical processes represents the next frontier of untapped insight. Two-thirds of respondents say, however, that increasing use of this data is an important challenge to overcome.
- › **Challenge #2: Improve AI processing capabilities.** Two-thirds of survey respondents say scaling data management technology and becoming more agile are important challenges to solve. The underlying difficulty is clear: As more raw data is generated every day, especially at the edge, even "modern" data architectures such as data lakes and big data can't store it all. Because of this, making the right data available for AI models is also a next-level challenge to solve, with close to two-thirds of respondents indicating its importance.
- › **Challenge #3: Reduce analytics silos.** While AI may be emerging as a centralized resource for the business, other analytics disciplines operate in silos. For example, 55% of respondents say that HPC is a "black box" to senior leadership, and fewer than half say the value of HPC is well understood by most. Part of why HPC is not well understood is that it is more likely to be a separate team of experts. These silos can limit firms' ability to source, manage, analyze, and act on data. HPC capabilities may be able to improve AI model training, but only if the two work in concert.

"Supervised vs. unsupervised training [deep learning] determines HPC's role in training. When data is more granular, we can implement supervised learning and keep costs down. When it's unsupervised, you need a large, powerful cluster to sort through."

Director of AppDev and operations at an enterprise pharmaceutical firm



"In the imaging side of the business, we need to get into the 10,000s of images in order to train a model."

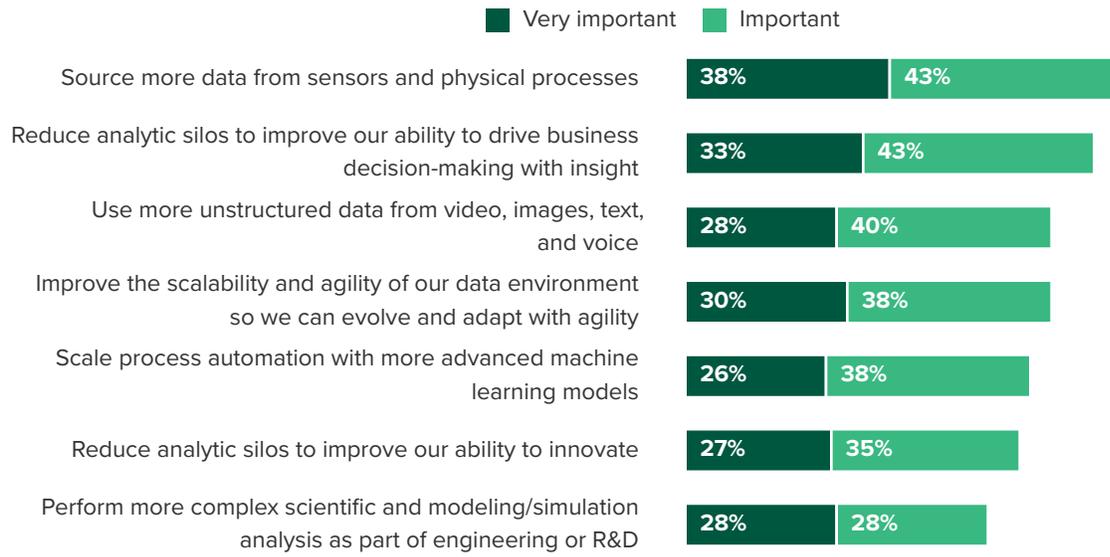
Director of AppDev and operations at an enterprise pharmaceutical firm



Three in four respondents say it is important to reduce analytics silos to drive better decision-making; 62% say it is important to drive innovation improvement.

Figure 2

“How important is it for your organization to solve the following data and analytics challenges to meet your top business goals?”



Base: 464 global decision-makers or practitioners responsible for high-performance computing or AI/ML
Source: A commissioned study conducted by Forrester Consulting on behalf of HPE, October 2020

AI Plus HPC: The Future Of Advanced Analytics

Our study found synergies between HPC, AI, and big data that lead to top- and bottom-line business benefits. For example, big data reinvigorated the science of neural networks that led to deep learning breakthroughs. Now deep learning is improving the benefits from HPC — for example, in processing of medical imagery. In fact, experts in our survey see many potential linkages between the use cases of AI, HPC, and big data (see Figure 3). Top strategic business activities in which both AI and HPC can play a critical role include business modeling and simulation, process automation, and risk analysis.

Experts in our survey saw many potential linkages between the use cases of AI, HPC, and big data.

Figure 3

AI And HPC Have Significant Overlap In Business Use Cases, Value To Business Strategy

	Top 5 business strategy activities	AI	HPC	Big data
#1	Marketing offers	✓		
#2	Fraud detection			✓
#3	Business scenario modeling and simulation	✓	✓	✓
#4	Process automation	✓	✓	
#5	Risk analysis	✓	✓	✓

✓ Ranked as top 5 in how important experts say each tech category (AI, HPC, big data) is to the business activity

Base: 464 global decision-makers or practitioners responsible for high-performance computing or AI/ML
 Source: A commissioned study conducted by Forrester Consulting on behalf of HPE, October 2020

THE BENEFITS OF UNIFYING AI AND HPC

Experts we surveyed see benefits to unifying the AI and HPC, including (see Figure 4):

- › **Cost reduction benefits.** Using AI to optimize HPC workflows and increase model targeting accuracy enables firms to run fewer simulations to get the desired results, saving money on compute costs. Sixty-three percent of experts say they will optimize HPC workflows using ML in the next year. In addition, optimized job scheduling for AI and HPC resources on the same infrastructure can create cost savings. Two-thirds of experts see this as an expected benefit of unification.
- › **Increased competitive advantage through faster time-to-insights.** Specifically, we found two benefits: 1) adding HPC to accelerate ML model training and 2) using AI to optimize HPC workflows. Together, these enable firms to discover insights more quickly, allowing them to adapt and respond to changing business conditions. Sixty-five percent of experts see this as an expected benefit of unification.
- › **Innovation possibilities from new insights gained.** HPC and AI have different strengths. HPC is good at solving computationally complex problems while AI is improving how we process imagery, video, and sensor data into the ability for robots to automate processes and engage with humans more meaningfully. Experts see the innovation possibilities in fields such as medical robots, which has never been possible before. Sixty-one percent of experts see increased innovation as an expected benefit of unification.

Figure 4

“Which of the following business benefits do you think unifying AI, HPC, and big data will have?”

67% Cost reduction through higher utilization of shared infrastructure (vs. operating different silos)

65% Increased competitive advantage through faster results/insights

61% The ability to extract new insights that were not previously possible from your data estate

57% Greater agility by being able to apply methods from all disciplines in a single developer workflow to address new business challenges/opportunities

51% Improved top talent retention through ability to cross-train/learn from a broader set of subject matter experts in a more diverse team

48% Potential to evaluate using supercomputing technology architecture more broadly across your data center (increased use of low-latency interconnects, high-speed storage, high-performance processors and accelerators)

Base: 464 global decision-makers or practitioners responsible for high-performance computing or AI/ML
Source: A commissioned study conducted by Forrester Consulting on behalf of HPE, October 2020

- › **Greater agility through optimized workflows.** Our study found that many of the data pipeline and compute resource load balancing operations required of AI and HPC work are similar. Most respondents in our survey expect to apply methods from all disciplines in a single, optimized workflow. This will create better efficiency and greater agility. Fifty-seven percent of experts see this as an expected benefit of unification.

FIRMS ARE ALREADY INTEGRATING AI AND HPC — AND PLAN FOR MORE

While help is on the way for AI infrastructure in the form of much needed funding, most experts say that their HPC environments need infrastructure updates to meet current demand, but less than 40% of firms are investing today. Many of those needed updates synergize with AI or complement investments that would also benefit AI. Specifically, we found that experts in HPC plan to:

- › **Improve the flexibility of their HPC infrastructure to handle more AI.** Seventy-two percent of HPC experts are planning to use more flexible infrastructure to handle both HPC and AI jobs in the next five years (see Figure 5). One way this can be done is by employing parallel processing capabilities for HPC and AI: 49% of HPC experts do this today, and an additional 45% plan to in the future.

Seventy-two percent of HPC experts are planning to use more flexible infrastructure to handle both HPC and AI jobs in the next five years.

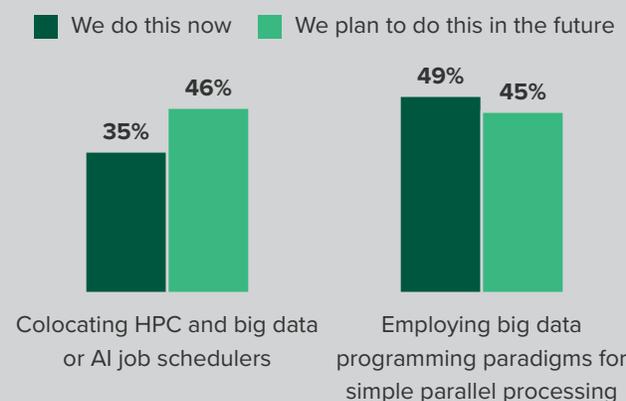
“I’m anticipating synergy between AI and HPC in the next three to five years in the form of infrastructure convergence. That would enable us to expand the platforms to new users, democratize the data more. In theory, there should be speed as well.”

Sr. strategic sourcing manager, multinational enterprise manufacturer



Figure 5

“Which of the following are you doing or planning to do to improve/expand your HPC infrastructure?”



“Do you think that any of the following AI and HPC convergences will happen in the following timeframes?***

- ➔ AI and HPC team will work off data in the same unified file system.
- ✓ 19% report this is happening today.
- 🕒 65% of respondents will need this in the next one to five years.

“On what timeframe do you see your AI operations needing the following?†

- ➔ More sophisticated job scheduling to deconflict resources among compute resources for model training and tuning
- 🕒 64% of respondents will need this in the next one to five years.

Base: 169 global decision-makers or practitioners responsible for high-performance computing

*Base: 355 global decision-makers or practitioners responsible for high-performance computing or AI/ML

†Base: 209 global decision-makers or practitioners responsible for high-performance computing or AI/ML

Source: A commissioned study conducted by Forrester Consulting on behalf of HPE, October 2020

- › **Migrate to more shared infrastructure.** Almost half of HPC experts plan to migrate to more shared infrastructure in the future. Nearly two-thirds of respondents say they need more sophisticated job scheduling, however, to make the efficient use of shared infrastructure resources. Furthermore, while only one in five HPC experts says AI and HPC employ the same data systems today, two-thirds see it happening in the next five years.

FIRMS ARE CHANGING WHERE THEY RUN AI AND HPC JOBS

Our study found that both AI and HPC experts are changing where they plan to run their jobs, which could affect the impact of planned AI and HPC investments (see Figure 6). Specifically:

- › **AI experts plan to train more models on-premises.** Two-thirds of AI and HPC experts say the cost of training models at scale in the public cloud is a challenge. As a result, AI experts plan to decrease model training in the public cloud in the next three years and increase model training on owned infrastructure.
- › **But they plan to run more models in the cloud and at the edge.** Fifty-five percent of AI experts plan to run AI models in the cloud in the next three years (up from 43% today), and 43% plan to run them at the edge, up from just 28% today. This makes sense as the cloud gives firms the agility to run instances of AI models in different physical regions, while edge computing lets firms run models close to where data, processes, and people actually are.
- › **HPC experts are trying out the cloud, but some aren't happy.** Seventy-four percent of experts say they plan to run HPC jobs in the cloud in the next five years. However, HPC services offered by public cloud vendors are still immature and do not offer the full range of capabilities that experts need. Further, we found that firms do not always realize the cost savings of cloud for HPC workloads.
- › **Firms could benefit from combining on-premises infrastructure for AI and HPC.** As firms run more HPC in the cloud and discover issues, some will shift resources to improve on-premises HPC infrastructure instead. We predict a few forward-thinking firms will further explore integrating on-premises AI training and HPC infrastructures. We conclude they will find significant benefits by doing so.

“Everything [HPC infrastructure] was on-prem two years ago. We had a large migration effort to move everything to the cloud. The issue is that costs have actually gone up when we did that.”

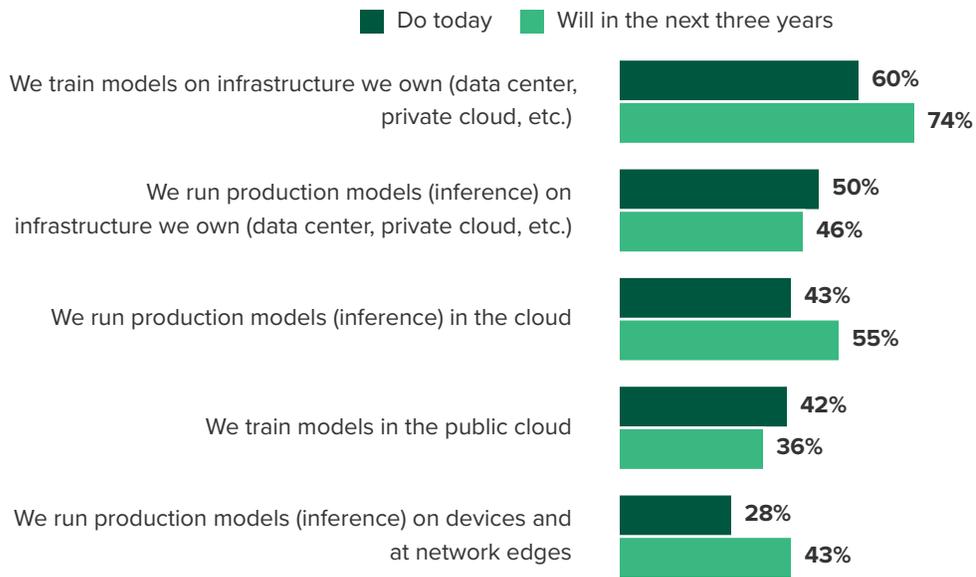
Director of AppDev and operations at an enterprise pharmaceutical firm



According to Forrester Analytics' Global Business Technographics® Telecommunications And Networking Survey, 2020, 40% of decision-makers believe edge computing gives them the flexibility to handle both present and future AI demands, making AI support the top benefit of edge computing.¹

Figure 6

“To what extent do you rely on public cloud vendor services for advanced machine learning model training today?”



Base: 209 global decision-makers or practitioners responsible for AI/ML

Source: A commissioned study conducted by Forrester Consulting on behalf of HPE, October 2020

AI And HPC Collaboration Is Critical To Shared Success

While our study found that HPC and AI are still separated organizationally, 80% of decision-makers say there is some level of communication between teams, and almost half say there is collaboration with shared goals (see Figure 7). Further, two-thirds believe there will be increased collaboration between AI and HPC teams within the next year (see Figure 8).

That said, we also found that a significant number of experts do not think they will ever completely merge.

Based on the findings of this study, we disagree. As firms shift more AI model training on-premises and upgrade their HPC infrastructure, the benefits of organization and infrastructure unification will become more evident, pushing against perceived data and security obstacles. While firms consolidate model training and HPC on-premises, they will seek to integrate and automate workflows that deploy AI runtime code to cloud and edge environments. This will demand even more sophisticated on-premises infrastructure to address efficiency, data cost, and security needs.

“Unified data architecture for AI and HPC sounds sexy, but I don’t see it happening anytime soon. The data is too complex, and privacy and security are huge hurdles that will only get more challenging in the future.”

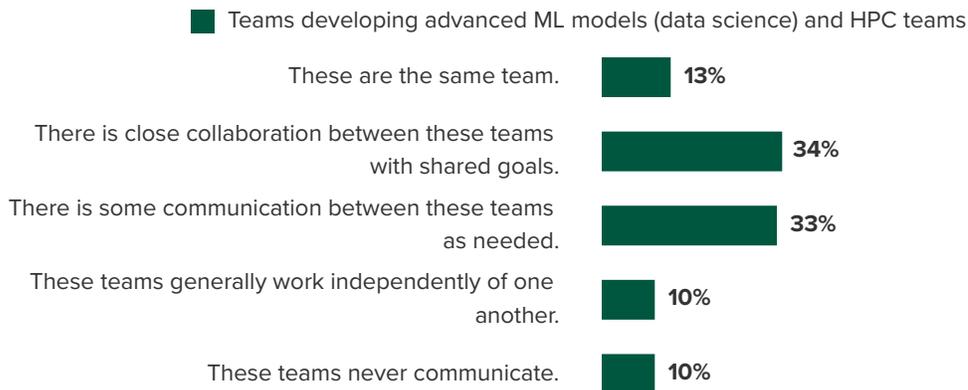
Director of IT, US life sciences firm



More firms will unify on-premises AI training and HPC infrastructure over time, and this will naturally unify the infrastructure operations teams as well.

Figure 7

“Please indicate the level of communication and collaboration between the following teams.”



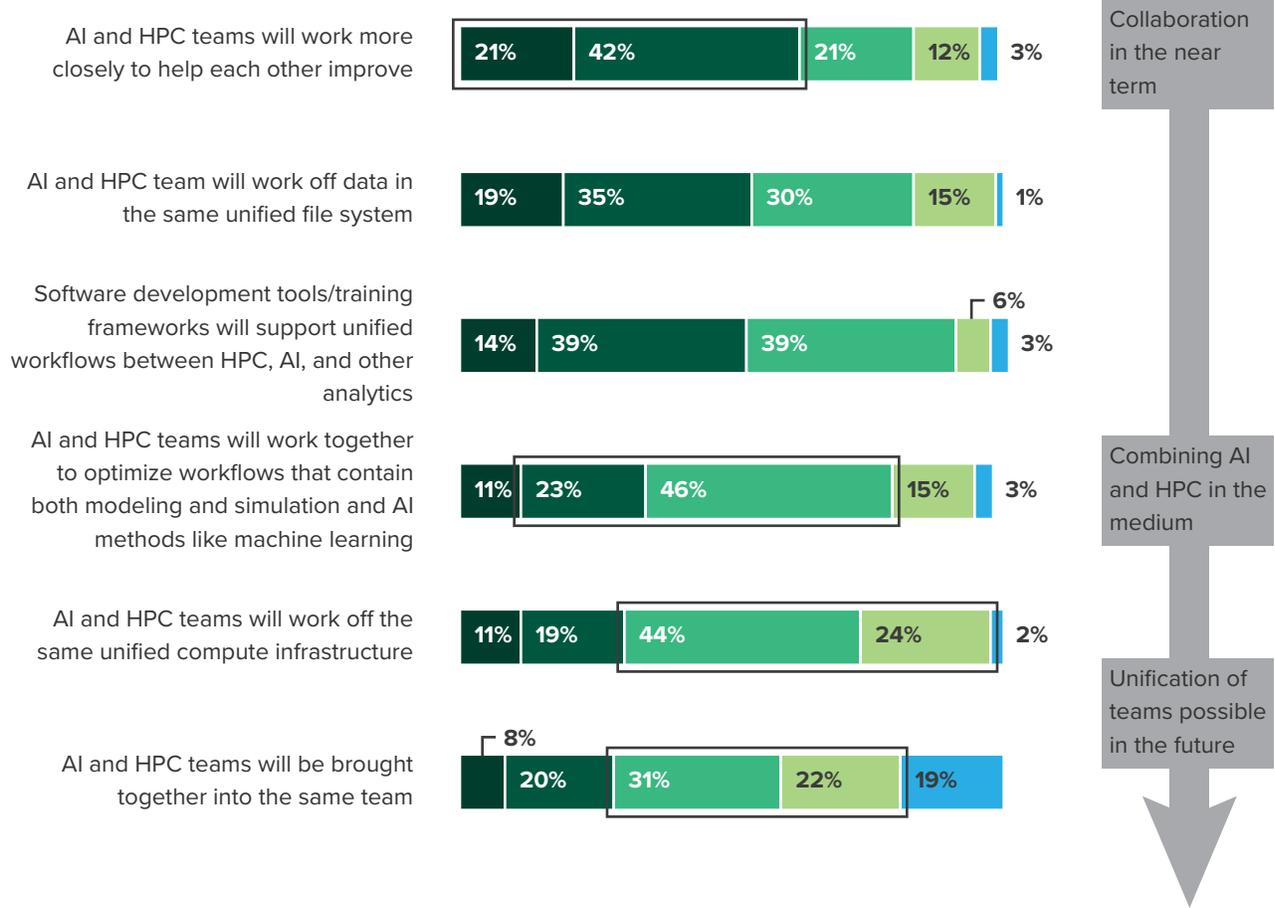
Base: 355 global decision-makers or practitioners responsible for high-performance computing or AI/ML
Source: A commissioned study conducted by Forrester Consulting on behalf of HPE, October 2020

Figure 8

“Do you think that any of the following AI and HPC convergences will happen in the following timeframes?”

■ This is happening today.
 ■ This will happen soon (within the next year).
 ■ This will happen eventually (within the next five years).
 ■ This may happen, but not within the next five years.

■ This will not happen at all.



Base: 355 global decision-makers or practitioners responsible for high-performance computing or AI/ML
 Source: A commissioned study conducted by Forrester Consulting on behalf of HPE, October 2020

Key Recommendations

Forrester's in-depth survey of business decision-makers, HPC experts, and AI experts yielded several important recommendations:



Look for synergies across workflows. Align organizations for better synergies across collaborative HPC and AI efforts to avoid data silos and to improve infrastructure efficiencies, enable better decision-making, and drive innovation improvement.



Investigate a hybrid architecture. Increasing data growth and analysis for AI and HPC will require firms to invest in on-premises compute infrastructure. Choose technologies that support multiple use cases and mixed workflows (for example, simulation, analytics, and AI). Look deeply at hybrid architectures to reduce bottlenecks when training complex AI models on large data sets.

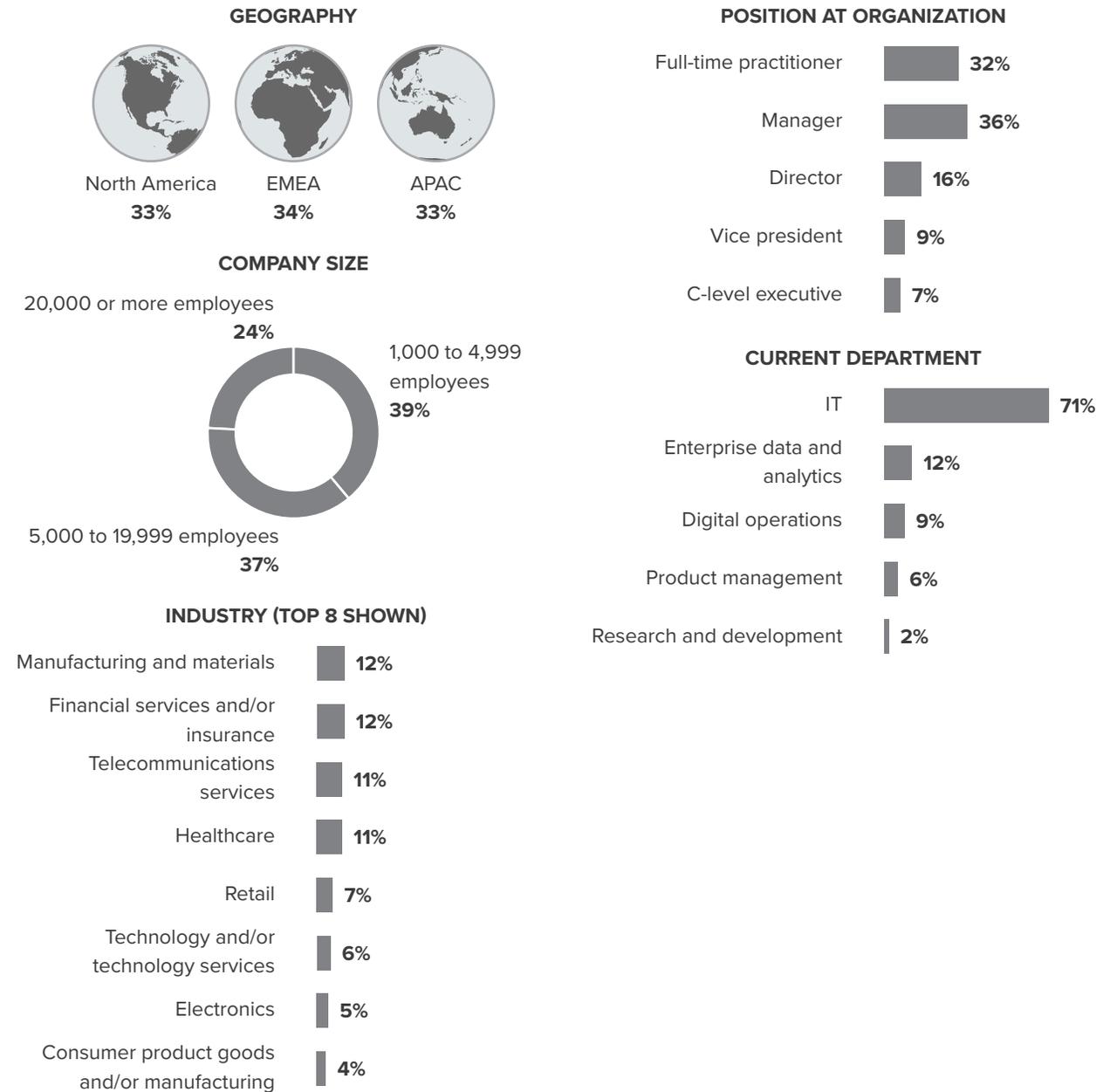


Expand to the edge. Growth of data at the edge will require firms to improve AI and analytics capabilities of the core data center and public cloud to deliver the insights organizations require. Choose infrastructure that can ingest and process data from the edge efficiently. Interconnectivity of edge and core will become increasingly important.

Appendix A: Methodology

In this study, Forrester conducted an online survey of 464 decision-makers and practitioners, divided into business decision-makers, AI experts, and HPC experts, in Canada, China, France, Germany, India, Japan, the UK, and the US, to evaluate the current state of AI and HPC infrastructure capabilities and investments. We also conducted three interviews, one with each type of decision-maker, to explore this topic. Business decision-makers relied on either AI or HPC to meet performance goals, and experts either used their respective technology on a daily basis or were involved the in the strategy and architecture of it. The study was completed in October 2020.

Appendix B: Demographics/Data



Base: 464 global decision-makers or practitioners responsible for high-performance computing or AI/ML
 Source: A commissioned study conducted by Forrester Consulting on behalf of HPE, October 2020

Appendix C: Supplemental Material

RELATED FORRESTER RESEARCH

“Now Tech: AI Infrastructure, Q1 2020,” Forrester Research, Inc., February 7, 2020

“Hardware Remains Critical In A Software-Defined World,” Forrester Research, Inc., January 10, 2020

“AI Deep Learning Workloads Demand A New Approach To Infrastructure,” Forrester Research, Inc., May 4, 2018

Appendix D: Endnotes

¹ Source: Forrester Analytics Global Business Technographics® Telecommunications And Networking Survey, 2020