Vanguard Report

February 2023

Navigating the Path to Successful AI Scaling

Commissioned by

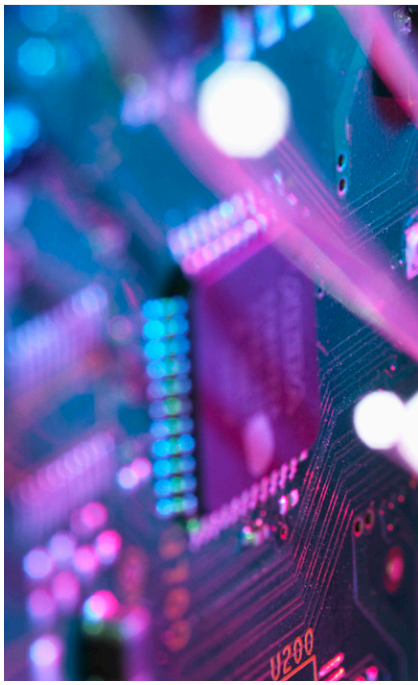
kyndryl

451 Research

S&P Global Market Intelligence

©Copyright 2023 S&P Global. All Rights Reserved.

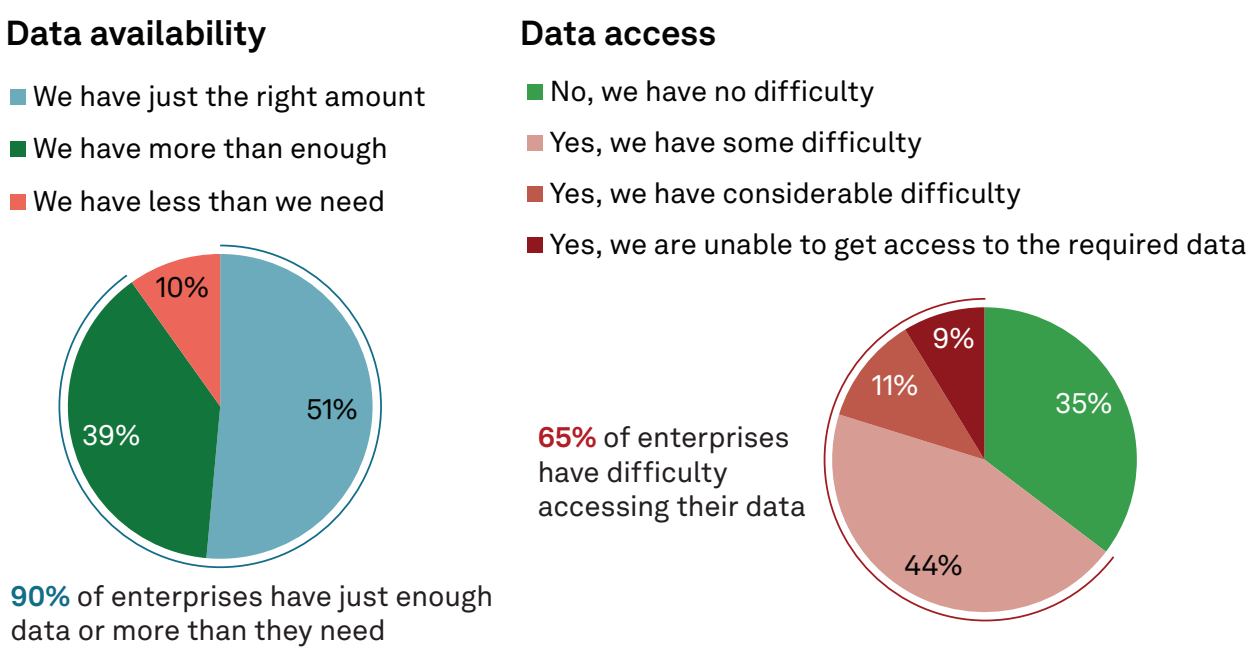
Introduction



Artificial intelligence is now moving into its implementation phase and is beginning to transform industries. However, getting AI applications to scale is proving challenging for various reasons. According to 451 Research's Voice of the Enterprise: AI & Machine Learning, Use Cases 2022 survey, more than one-third (35%) of AI projects don't make it out of the proof-of-concept stage and into production. The reasons for this are numerous, including internal skills shortages, budget limitations and compliance constraints — but access to data can also be a critical barrier to success. This is not due to a lack of data; it's due to lack of access.

As we see in Figure 1 below, 65% of enterprise respondents to 451's Voice of the Enterprise: AI & Machine Learning, Infrastructure 2022 survey said they have difficulty accessing the data they need to build and train their AI/ML models. Removing this barrier is therefore a key element in improving the overall success rate of AI projects. And beyond access to data, any organization on the path toward success in AI needs to understand how building AI-based applications differs from earlier approaches, and must carefully consider how they can get there as efficiently as possible.

Figure 1: Data access is a major barrier to successful AI projects



Q. Which of the following statements best aligns with your current data requirements as it relates to building/training AI/ML models?
Q. Do you currently have difficulty getting access to the data you [require/will require] for your AI/ML workloads?
Base: All respondents (Q1: n=690; Q2: n=688).
Source: 451 Research's Voice of the Enterprise: AI and Machine Learning, Infrastructure 2022

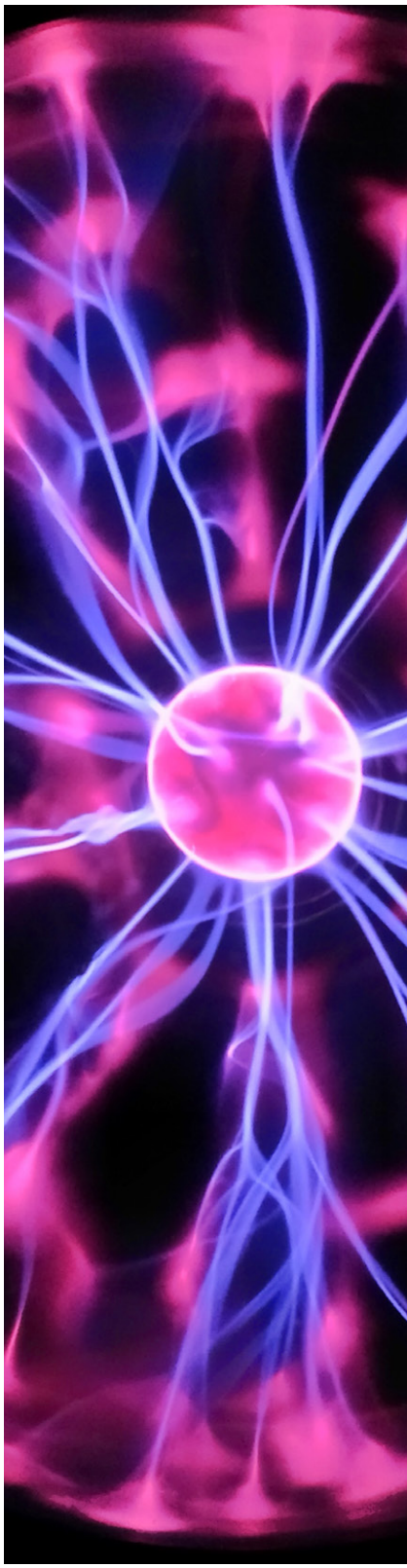
The Take

Getting AI to scale can be challenging, which is why it is important to make sure that the business reasons for engaging in AI projects are well understood up front, and that all relevant stakeholders are involved. After that, organizations must address the technical challenges involved in getting access to the relevant data required to train models, and then make sure they are successfully deployed.

This is why many organizations are increasingly focused on DataOps and MLOps, in much the same way that it was important to focus on DevOps when application development began shifting from waterfall to agile methodologies. By using effective data management and automation practices, organizations can improve the performance and reliability of their AI systems, which can in turn help them scale more effectively.

In addition to this focus on DataOps and MLOps, it is important that business stakeholders can see the value that AI brings to the organization. This can help ensure that they support the project and provide the resources necessary to help it succeed. By focusing on these critical areas — **establishing clear business reasons, aligning all relevant stakeholders and addressing technical challenges** — organizations can improve their chances of successfully implementing and scaling AI projects.

DataOps and MLOps: Two Keys to Successful AI Deployments



At its heart, getting AI to work means using data to train models and then deploying those models successfully to make predictions. Two techniques, DataOps and MLOps, have come to the fore to help organizations get AI to scale successfully. DataOps refers to the application of more agile and automated approaches to data management to support data-driven business outcomes. MLOps is generally focused on the deployment and post-deployment monitoring and management of models. Both apply DevOps principles to their respective tasks.

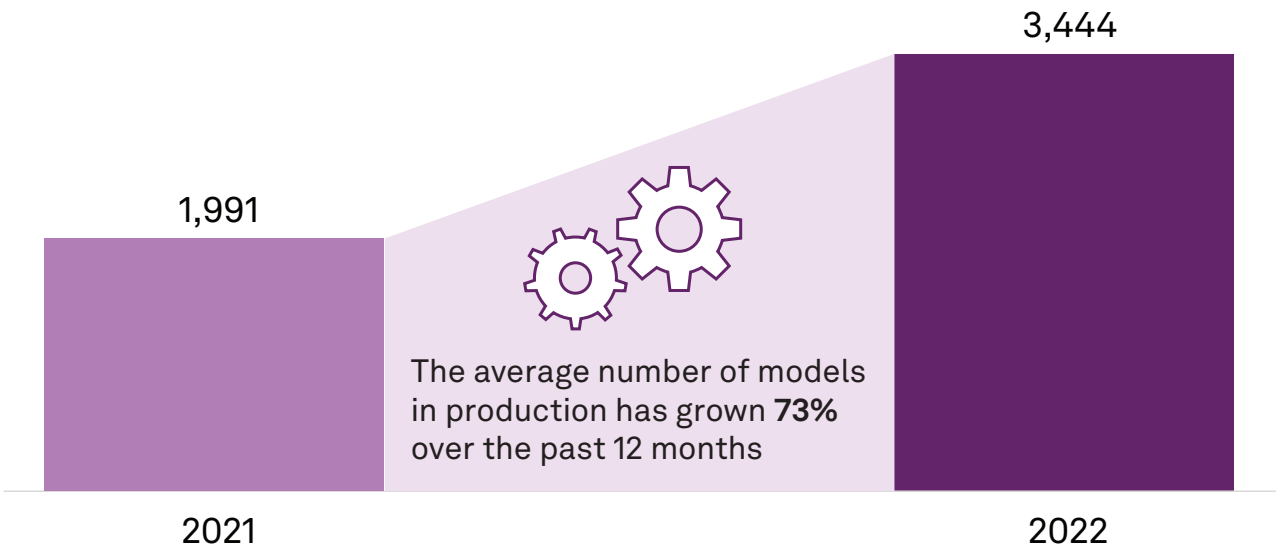
DataOps, which comprises agile and automated approaches to data management, aims to help organizations improve the speed, quality and reliability of their data pipelines and analytics systems by bringing together all those involved in the data lifecycle — including data engineers, data scientists, data analysts and business users.

Data access and integration challenges are most prevalent in the early training stages of AI/ML model building. Once the model has been trained and tested, and is ready to be deployed as part of a wider application, different challenges to scaling AI emerge. But there are techniques that organizations can deploy to help ensure the success of their AI implementations, which fall under the umbrella term of MLOps.

Unlike rules-based models, which only change their behavior when a human updates the rules, ML models can evolve and adapt over time simply through their use. As new data is processed, an ML model may make predictions that are less accurate than when it was first put into production. This can happen if the new data is significantly different from the data the model was trained on. Therefore, the behavior of an ML model can change over time without any explicit intervention from a human. All ML models decay; it's just a matter of time. This needs to be monitored and spotted, and the model retrained accordingly.

This process might be relatively straightforward with only a handful of machine-learning models in production, but when the number of models reaches thousands or more (see Figure 2), it can become a major challenge. This is where MLOps comes in, which describes the tools and processes of deploying, managing and monitoring ML models. MLOps aims to improve the quality and performance of ML models, and to make the process of building and deploying ML applications more efficient and scalable.

Figure 2: Number of machine learning models in production



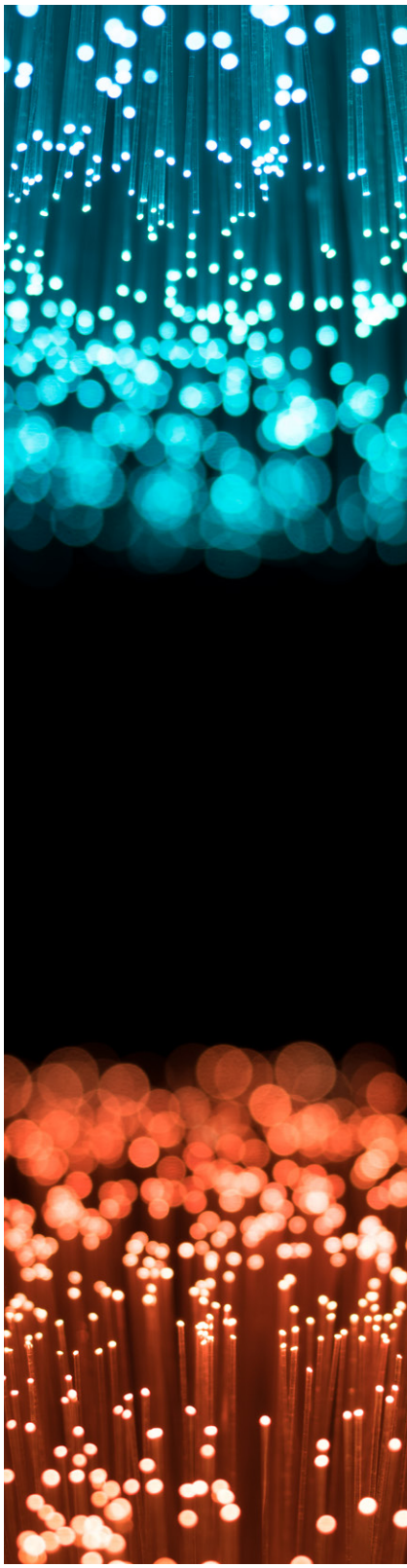
Q. Approximately how many distinct machine learning models are currently in production across your organization?
Base: ML is in production (2022: n=315; 2021 n=281).
Source: Voice of the Enterprise: AI and Machine Learning, Infrastructure 2022.

Maximizing the Success of AI/ML Projects Through Partners

Developing AI models for their own sake is pointless. Instead, organizations should start by clearly defining the business problem they are trying to solve with AI/ML. This means figuring out what success looks like in terms of key performance indicators (KPIs) and setting specific, measurable goals. One key to success is ensuring that data science teams are closely integrated with the business, rather than working in isolation. This enables them to better understand the problem at hand and develop models that are directly aligned with business needs.

Many organizations face a shortage of in-house skills and expertise around AI/ML. In such cases, service providers and consultants can be valuable partners, helping to bridge the gap and bring AI/ML applications to fruition. Organizations that use such partners to build their AI applications can have greater success than those that attempt to develop systems wholly in-house.

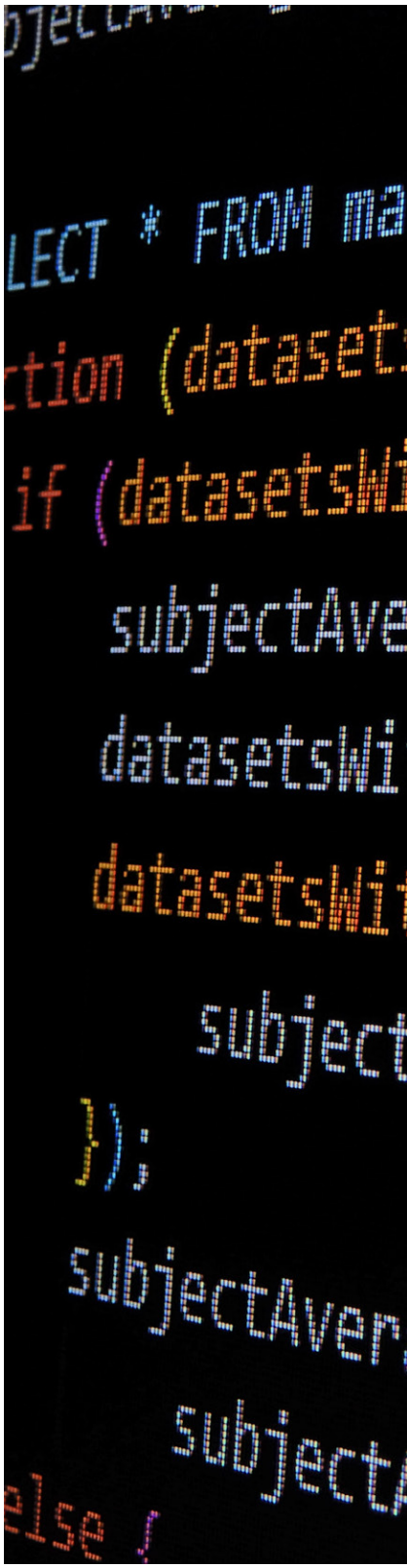
According to 451 Research's Voice of the Enterprise: AI & Machine Learning, Use Cases 2022 survey, the project abandonment rate among organizations that use service providers for their primary AI/ML development strategy is below the overall rate of 35% of projects being abandoned. And this rate is 10 percentage points lower than those organizations that use cloud-based AI/ML services as their primary strategy (29% vs. 39%, respectively). This is likely due to service providers using specialist expertise and providing support throughout the planning and development process, helping to ensure that the final AI application is well-designed, scalable and aligned with business needs.



Conclusions

Organizations embarking on AI projects must understand some of the things that make such projects different from earlier ways of building and deploying applications. They should ensure that the business reasons for engaging in AI projects are well understood and that all relevant stakeholders are involved.

A focus on DataOps and MLOps to manage and automate data for improved performance and reliability of AI systems will help improve the chances of successful implementation and scaling of AI projects. And organizations should consider all their development options, whether in-house or using partners such as consultants and service providers. The right partner can ensure that organizations have access to the specialized expertise required and understand all the potential pitfalls and challenges, as well as helping ensure that projects come in on budget and are deployed and maintained successfully.



kyndryl™

At Kyndryl, we are committed to the continuous improvement of the digital economy. We design, manage and modernize the mission-critical technology systems that the world depends on every day. In collaboration with our partners and customers in 100+ countries, we co-create solutions that help enterprises reach peak digital performance. Because we know that a healthy digital ecosystem not only benefits our customers but also fuels innovation and transforms the world, we stand up to further human progress day after day.

Kyndryl. The Heart of Progress™

Learn more

About the author

Nick Patience

Research Director, AI Applications & Platforms

Nick Patience is lead analyst for AI and machine learning for 451 Research, a part of S&P Global Market Intelligence. He is part of the company's Data, AI & Analytics research channel but also works across the entire research team to uncover and understand use cases for machine learning, an area he has been researching since 2001. Nick also oversees 451's Workforce Productivity & Collaboration research and is a member of 451 Research's Center of Excellence for Quantum Technologies.

Nick is a frequent speaker on the industry use cases for AI and the infrastructure that underpins its development and deployment.

Nick has a long background in research into how applications can take advantage of data – in particular unstructured data – using AI and machine learning. He is a cofounder of 451 Research and rejoined the team in 2015 after almost three years running product marketing at machine learning-driven eDiscovery software company Recommind (now part of OpenText). He has held various senior management roles at 451 Research in both in New York and London since 1999.

Prior to starting 451 Research, Nick was a financial and technology journalist with ComputerWire (now part of Datamonitor) in London and New York. Nick has a master's degree in computing science from the University of London, and a BA in Philosophy and Music from Middlesex University.

About S&P Global Market Intelligence

S&P Global Market Intelligence's Technology, Media and Telecommunications (TMT) Research provides essential insight into the pace and extent of digital transformation across the global TMT landscape. Through the 451 Research and Kagan products, TMT Research offers differentiated insight and data on adoption, innovation and disruption across the telecom, media and technology markets, backed by a global team of industry experts, and delivered via a range of syndicated research, advisory and go-to-market services, and live events.

CONTACTS

The Americas
+1 877 863 1306
market.intelligence@spglobal.com

Europe, Middle East & Africa
+44 20 7176 1234
market.intelligence@spglobal.com

Asia-Pacific
+852 2533 3565
market.intelligence@spglobal.com
www.spglobal.com/marketintelligence

Copyright © 2023 by S&P Global Market Intelligence, a division of S&P Global Inc. All rights reserved.

These materials have been prepared solely for information purposes based upon information generally available to the public and from sources believed to be reliable. No content (including index data, ratings, credit-related analyses and data, research, model, software or other application or output therefrom) or any part thereof (Content) may be modified, reverse engineered, reproduced or distributed in any form by any means, or stored in a database or retrieval system, without the prior written permission of S&P Global Market Intelligence or its affiliates (collectively S&P Global). The Content shall not be used for any unlawful or unauthorized purposes. S&P Global and any third-party providers (collectively S&P Global Parties) do not guarantee the accuracy, completeness, timeliness or availability of the Content. S&P Global Parties are not responsible for any errors or omissions, regardless of the cause, for the results obtained from the use of the Content. THE CONTENT IS PROVIDED ON "AS IS" BASIS. S&P GLOBAL PARTIES DISCLAIM ANY AND ALL EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, ANY WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE OR USE, FREEDOM FROM BUGS, SOFTWARE ERRORS OR DEFECTS, THAT THE CONTENT'S FUNCTIONING WILL BE UNINTERRUPTED OR THAT THE CONTENT WILL OPERATE WITH ANY SOFTWARE OR HARDWARE CONFIGURATION. In no event shall S&P Global Parties be liable to any party for any direct, indirect, incidental, exemplary, compensatory, punitive, special or consequential damages, costs, expenses, legal fees, or losses (including, without limitation, lost income or lost profits and opportunity costs or losses caused by negligence) in connection with any use of the Content even if advised of the possibility of such damages.

S&P Global Market Intelligence's opinions, quotes and credit-related and other analyses are statements of opinion as of the date they are expressed and not statements of fact or recommendations to purchase, hold, or sell any securities or to make any investment decisions, and do not address the suitability of any security. S&P Global Market Intelligence may provide index data. Direct investment in an index is not possible. Exposure to an asset class represented by an index is available through investable instruments based on that index. S&P Global Market Intelligence assumes no obligation to update the Content following publication in any form or format. The Content should not be relied on and is not a substitute for the skill, judgment and experience of the user, its management, employees, advisors and/or clients when making investment and other business decisions. S&P Global keeps certain activities of its divisions separate from each other to preserve the independence and objectivity of their respective activities. As a result, certain divisions of S&P Global may have information that is not available to other S&P Global divisions. S&P Global has established policies and procedures to maintain the confidentiality of certain nonpublic information received in connection with each analytical process.

S&P Global may receive compensation for its ratings and certain analyses, normally from issuers or underwriters of securities or from obligors. S&P Global reserves the right to disseminate its opinions and analyses. S&P Global's public ratings and analyses are made available on its websites, www.standardandpoors.com (free of charge) and www.ratingsdirect.com (subscription), and may be distributed through other means, including via S&P Global publications and third-party redistributors. Additional information about our ratings fees is available at www.standardandpoors.com/usratingsfees.