Expert Insights

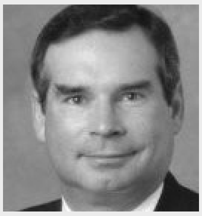
IBM Institute for Business Value

Dealing with the AI data dilemma

The right approach to integration, governance, and tools

IBM

Experts on this topic



Michael Haydock

IBM Fellow, Vice President and Chief Scientist, Global Business Services, IBM
mhaydock@us.ibm.com
linkedin.com/in/drmikehaydock

Michael is a leading artificial intelligence expert using deep learning for IBM Global Business Services (GBS). He specializes in numerical optimization and quantitative machine learning areas, applying these exponential technologies to customer relationship understanding and supply chain optimization for enterprises. Michael holds a Ph.D. in Operations Research from Walden University, Minneapolis, MN.

Steven Eliuk

Vice President, AI and Governance Automation, Global Chief Data Office (GCDO), IBM
Steven.Eliuk@ibm.com
linkedin.com/in/m3the01

Steven leads the development of platform and infrastructure components for AI in the Cognitive Enterprise Data Platform, IBM's data lake. He also applies AI in the enterprise to internal processes while maintaining governance, security, privacy, and trust. His teams also develop, maintain, and operate IBM's global set of Privacy Services, including GDPR, CCPA, and other regulatory requirements. Before joining IBM, Steven led the design of high-performance computing (HPC) infrastructure for artificial intelligence at Samsung Research America. Steven earned a Ph.D. in Computer Science from the University of Alberta.

Susara van den Heever

Director, Cloud and Cognitive Software Expert Labs
svdheever@fr.ibm.com
linkedin.com/in/susaravandenheever

Susara serves as the Director of Expert Labs EMEA, with a focus on creating client value via IBM Cloud and Cognitive Software. Her passion is to use data and AI to make the world a better place. She works to help people and businesses be more efficient, save time, manage resources better—even make better environmental choices. Her career has spanned many industries: from helping a beer producer create a better supply chain, to ensuring fair allocation of bandwidth to telecom providers, to helping energy companies create sustainable production, to making cities better places to live. Susara holds a Ph.D. in Chemical Engineering from Carnegie Mellon University.

Why do about half of AI projects fail? In a word, data.

Key takeaways

To get the most value from AI, always start with the business problem. Then seek out multiple data types—structured and unstructured, internal and external, qualitative and quantitative—to tackle the problem and enrich the solution.

Embed robust, permission-based governance that establishes data provenance to build trust in the data and AI insights.

Plan for the challenges of rigorous data preparation and the complexities of merging disparate data sources. Reuse data, automate processes, and adopt the right tools.

—

AI's unique data challenges

Artificial intelligence (AI), no longer the new kid on the block—reserve that spot for quantum computing—is rapidly being adopted and embedded into all kinds of business and societal uses. Early in the pandemic, 84% of all organizations expected to maintain or increase the level of organizational focus on AI, with nearly a third boosting their AI investments as a direct result of the pandemic.[1]

In a recent survey, AI joined the Internet of Things (IoT) and cloud computing as the top 3 technologies CEOs expect will most help deliver results.[2] And 43% of IT professionals report that their company has accelerated its rollout of AI as a result of the COVID-19 pandemic.[3]

This translates into a compelling need for companies to build a strategic AI capability—from strategy to operating model to talent—and integrate that capability in the business.

Still, many AI projects languish, even after a promising proof-of-concept (PoC) phase. 90 percent of companies have difficulty scaling AI across their enterprises.[4] So, it's not surprising that about half of AI projects fail.[5]

Why? In a word, data. More than half of those responsible for AI strategy acknowledge that they are not clear on what their AI data needs are.[6] And 39% of IT professionals report that analyzing data to build and scale trusted AI is the most difficult part of their organization's AI journey, while 32% say that data complexity and silos are a top barrier to AI adoption.[7]

No wonder, then, that more than half of organizations cite data as the culprit for AI projects stalling, with data quality issues cited as the top factor (58%), followed by a lack of well-curated data (45%), and data governance issues (40%).[8] In fact, even general best practices for advanced data capabilities may prove insufficient for AI. While AI may only be one use case for data—albeit an important one—lessons learned need to be extended and expanded in the AI realm.

AI has several specific data considerations that make a difference in practice:

- *AI scale:* With AI, generally speaking, the more data available, the better the quality and accuracy of the results. So, the volume of data required for AI can be far higher than even for some advanced analytics.
- *AI speed:* Data must be current to achieve the speed of response needed for some AI insights and for the best predictions. At times it can even be real-time or very close to it.
- *Data variety:* AI results are often better, not just with more data, but with data that can add context. Yet an AI model's outcome can be altered by manipulating the data—maliciously or unintentionally—or even just from generalized "data drift," so making sure all data is handled properly becomes essential.
- *Data quality:* AI is highly sensitive to data, making it critical the data corresponds to the underlying reality. In some instances, AI pays close attention to spikes that other traditional analytic methods might consider anomalous, so accuracy is essential.
- *Human perspective:* How humans regard data—including the biases in our own experiences and the opacity in the "black boxes" of our own brains—often affects how we use it. Data has meaning in context and so must be viewed and understood that way. Without the proper context, it can be unwittingly misused or misinterpreted.

Too often, though, companies are overwhelmed by these complexities, becoming mired in well-meaning but misaligned approaches to solving the data challenges. To meet AI's unique data needs, an organization first needs a clear-eyed view of the business problem(s) it is trying to solve, then a pragmatic approach to solving them.

Keeping an eye on the business problem

Some classically trained data specialists still struggle with AI, focusing too much on the details of data science and engineering—*how* to do AI. Without sufficient understanding and regard for the larger business issues—*why* the company is doing AI—proofs of concepts and research-type projects can proliferate without benefiting the business. Moreover, data scientists and engineers are often tempted to embrace "big data solutions" to business problems that can sometimes be solved with high-quality, precision-targeted, even *qualitatively*-obtained "small data solutions".

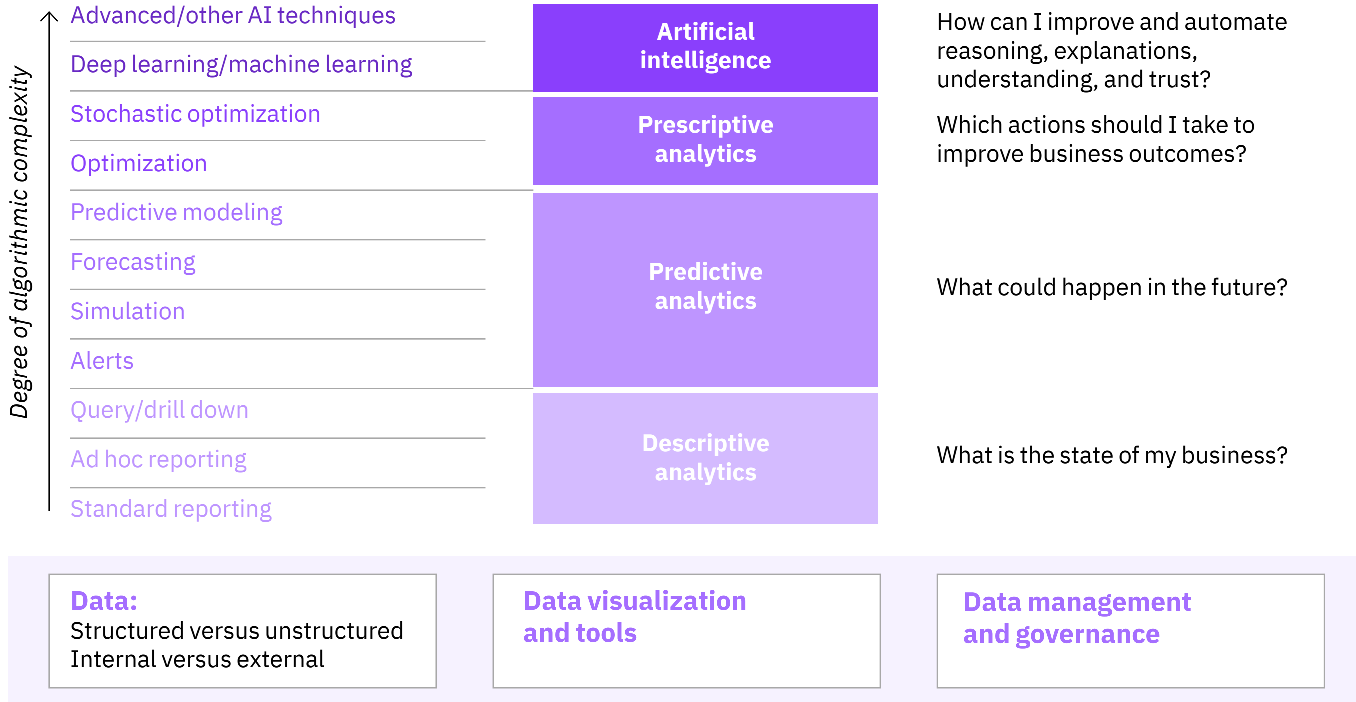
Teams first need to ask two basic questions: What business problem are we trying to solve? And how should we best solve it? Sometimes, the most sophisticated AI may not be the best answer (see Figure 1). This can also help clarify when and how business workflow interventions should take place based on what insights the data reveals and actions AI recommends or takes.

A quick win that shows business results will keep the business interested and focused.

Figure 1

Needs matching

Selecting the right tool from the AI-analytics continuum.



A pragmatic approach to value

Given the importance of data for AI, and the many challenges in obtaining, integrating, preparing, and properly managing it, a common impulse for many organizations is to launch a single project to get all its data organized. Often this involves putting it into a large data lake and trying to iron out all of the organization's data issues once and for all. Of course, that rarely works.

This ambition is impractical because of its ungainly scope, preventing a clear path to a reasonable return on investment (ROI). But even more important, data and business needs change too quickly to make such a massive undertaking advisable.

Instead, organizations building a solid set of AI capabilities will learn to think pragmatically (see "The Weather Company: Lessons learned from the forecasting business"). What data is readily available, accessible, and clean? A quick win that shows business results will keep the business interested and focused. Using data with quality or availability issues risks turning a potential quick win into a lengthy and unproductive data cleansing exercise.

Successful organizations will target business-oriented use cases with a short- or medium-term focus, but will maintain a long-term focus on continuous delivery of AI value.

The Weather Company: Lessons learned from the forecasting business

The Weather Company (TWC) is built on putting data—"tons" of it—to work. More specifically, applying AI and models of all sorts to generate forecasts and insights. The company has been working on and with AI for two decades and is a global leader in applying it at a massive scale. Along the way, it's learned many lessons about meeting the peculiar data requirements of AI.

One of the first is not underestimating the task at hand. It might seem simple to merge weather and sales data, but it's anything but. Insight gained from experience is needed when extracting data, knowing how to perform necessary data joins, and understanding the implications of how results will be run and delivered. For example, will the models need to run centrally, in the cloud, or "on the edge" in an end-user's device? Will they be trained online—and thus, updated continuously—or offline and updated periodically? The answers have implications for how to handle the data.

In fact, a huge lesson from TWC is that companies need to plan realistically—anticipate the cost, resources, and time needed—for data prep and integration. Experience shows 80% of a team's time may be spent data wrangling, relegating the cool work of actually running models, tuning them, and interrogating results to a much smaller fraction of their time. Automation can help.

Consider just one area of TWC's work: forecasting seasonal flu and allergy impacts. The unique challenges of merging all the necessary data means constructing custom data pipelines, services—even creating a special data lake for modeling. This requires deep data skills, as does being able to answer a deceptively simple question: can we slice this data any way we want—or need to? Fortunately, when first confronting this, TWC already had capable engineers and data scientists who were up to the task.

Equally important? The team was backed by executives who understood, bought into the project, and supported the investment of time and resources it would take to accomplish the objectives and achieve the desired outcomes.

That said, many of the generalized focus areas for data are not new; some are even decades old but acquire newfound urgency, importance, and opportunity for improvement in the AI era:

1. *Integration.* Building a capability that taps data for AI from across all parts of an organization and beyond its walls.
2. *Governance.* Applying the latest approaches to governance when managing data for AI to build trust in the insights that emerge.
3. *Tools.* Providing the necessary tools to the teams that need them.

With the right degree of business acumen and pragmatism, companies can make progress on continuing to build the institutional data "muscles" to address some of the unique characteristics of AI.

The integration imperative

The COVID-19 pandemic has vividly demonstrated how the standard disclaimer that "past performance is not indicative of future results" applies to data and the insights drawn from it. Companies that relied on analysis of previous years' data and historical models found their forecasts useless.[9] To survive, they had to begin ingesting short-term data—for example, recent sales and weather data, even oil prices—and update AI models for more relevant forecasts. In fact, AI quickly became even more popular because of its ability to ingest volatile short-term data to augment existing data, adapt quickly and yield relevant predictions.

Much of this shorter-term, quickly refreshable data originates outside a company, so a robust approach to third-party and even public data integration is an increasingly necessary capability for AI.

The ability to trace a path from a major event that's unfolding to its subsequent effects, such as unemployment or reduced disposable income, gives management insight.

If data sources that provide the right levels of granularity are effectively tapped, the insights derived from them can then be highly localized, and therefore, truly useful. The ability to trace a path from a major event that's unfolding—like the pandemic—to its subsequent effects, such as unemployment, reduced disposable income, and the consequent effects on shopping patterns, gives management insight.

An example of an approach that supports this kind of localization is the IBM COVID-19 Lockdown Index. It allows an understanding—updated daily based on the latest data—of the current state of COVID-19's impact on economic activity, county-by-county, across the US. Companies can use this index in harmony with their own data to create useful, actionable forecasts, then plan and adjust accordingly (see "Perspective: Integrate disparate data sources to yield local COVID-19 insights.")

Trust from governance and provenance

The use of external or third-party data especially illustrates the need for organizations to know their data well. Who owns it? Who can approve its use—and for how long? Commercialization of AI assets based on third-party data can raise even more complex issues. If an AI asset has been trained on data that a company no longer has access to or approval to use, what then? And while these questions obviously apply to external data, similar problems can be faced with internal data.

More than half of organizations struggle with data integration in general, while a third lack confidence in their capabilities to connect multiple data sources.[11]

Then there's the issue of regulated data. Country borders can affect data availability and how it can be used. Data quality can also vary from country to country, meaning data across a geography may not be consistent or usable in the same way. For example, in some countries, credit card transaction data is captured in great detail because of advanced microchips and card readers; in others that lack that hardware, credit card machines collect less transaction information.

The right approach to data governance can help with these issues, including knowing and being able to trace the data's provenance. In its simplest form, data governance refers to the norms, principles and rules applied to managing different types of data. Data governance can apply at a very localized level—within an organization—to handle data properly and maintain its integrity and usefulness at each stage of its lifecycle. It can also apply to cooperation between organizations, across ecosystems and even among countries—how they will share data.

Perhaps more important than the right technology are the right rules and culture to make getting access to data easy and well-managed. A data culture well-acquainted with standard terms and conditions and the organization's way of handling data will be much more likely to share data and make good use of it.

Perspective: Integrate disparate data sources to yield local COVID-19 insights

Knowing a huge storm is bearing down on the eastern half of a country is interesting in a general way, but knowing it will hit your state, your county, or your town—or pass you by—is better. Localized insights can be powerful.

The IBM COVID-19 Lockdown Index quantifies the degree of disruption currently in place in each locale. It melds longitudinal information with real-time daily updates on the disease's progression, hospital bed and infection statistics, local community restrictions, and overall US market volatility measures. It then projects when the peak density of the disease might be reached county-by-county and the distinctive shape of the down-curve in each location, then assigns a risk score and identifies the pace at which a local lockdown may lift.

Its user-friendly interface visually illuminates the degree of lockdown in each U.S county and how that lockdown is most likely to evolve over the coming four weeks (see Figure 2).

The index uses publicly available data, including: epidemiological measures of state hospital demand requirements; county-by-county COVID-19 caseload; hospital bed use; intensive-care-unit bed use; and ventilators in use. The system also scrapes local news reports for updates on school open/closure dates and other potential community restrictions.

Another similar initiative is The Emergent Alliance, a not-for-profit collaboration of IBM, Rolls-Royce, Microsoft, and dozens of global companies. The alliance is working to bring accurate and up-to-date regional pictures of Covid-19 cases to help local authorities mount a more effective response to coronavirus outbreaks.

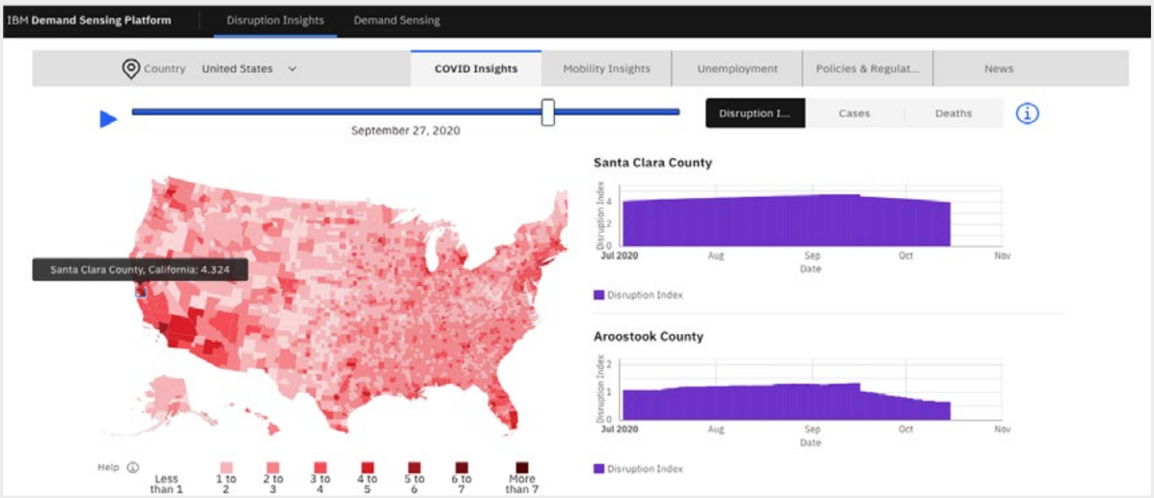
A localized Risk Index combines data from infection rates, social media, news, Airbnb data, and more. The analysis includes the disease's impact on health, how governments are reacting, how public behavior is changing, and the effect of all this on the economy.

The benefits of the project, though, extend beyond dealing with the pandemic. Sentiment data on topics in the news can help anticipate new behavioral patterns. Seeing an uptick in news around outdoor sports for example, or for Airbnb bookings in the mountains, might prompt a campaign around hiking or other related goods and services.[10]

Figure 2

Visual insight

The COVID-19 Lockdown Index



If recipients of insights don’t trust the data they’re based on, the most powerful and sophisticated AI algorithms are unlikely to have much business impact.

Automating as much as possible saves time and resources—while encouraging broader adoption of AI methods—since getting the data is no longer a herculean chore. It can help teams avoid being bogged down for weeks or months pursuing permission to use data.

Nothing in an organization’s culture matters more than trust. While essential within an organization, it is even more important when an organization shares or accepts data from outside its borders (see Figure 3). It doesn’t matter how good an AI team is with gathering data and building solutions that demonstrate value. If recipients of the insights generated don’t trust the data—where it’s coming from, how it’s used, that its biases are detected and mitigated or are otherwise transparent, and that it conforms to regulations and laws—then the results will likely be suboptimal. In fact, in such circumstances, the most powerful and sophisticated AI algorithms are unlikely to have much business impact.

Lessons gleaned from recent conceptual and technological advances in blockchain provide a potential source of help in establishing trust. One of blockchain’s most important principles is that it brings the right parties together from the beginning—the critical few that would make or break a network and can answer questions such as, “What data will be shared? With whom?” This minimally viable ecosystem, which can also include regulators, decides on the incentive structure, monetization frameworks, and governance rulebook for the network.

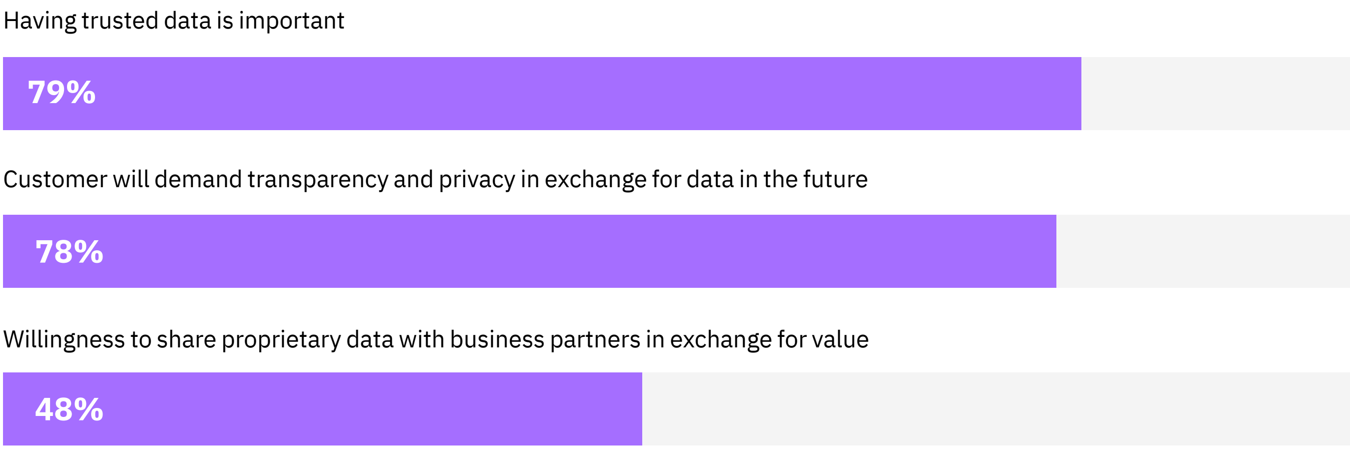
An important data governance concept is to preserve aggregate knowledge and insights, but not sensitive data that may have regulatory implications. A blockchain principle can help here, too: a permission-based approach and network validation preserves transparency, data integrity, data lineage, and provenance. This can help address a top concern voiced by 66% of IT professionals about mitigating the lack of clarity on provenance of AI training data.[12]

Of course, good governance doesn’t end with tracing the data an AI model is trained on. It also includes examining the decisions made by a human based on that data, which is essential to explainability, especially when those decisions may be contested.

Figure 3

The sharing dilemma

Trusted data is highly valued, but too closely guarded.



Source: Pureswaran, Veena; Parm Sanha; Smitha Soman. “Advancing Global Trade with Blockchain: How to unleash value from trusted, interconnected marketplaces,” Q: To what extent do you agree with the following statements about trusted data?” IBM Institute for Business Value. May 2020.

Standard data methods and tools can’t be applied unchanged to AI.

The right tools for the job

Many AI applications transform raw data into signals and find patterns and insight in time-series and other very large data sets. Often, to reveal these signals, the data sets have to be processed hundreds of times.

Imagine a large retailer that has over 600 million SKUs. Then imagine what it takes to process them and the capacity to do so daily *hundreds of times*. Certainly not a trivial problem, and one that requires advanced data skills for sure, but one that also incurs financial costs—including environmental costs—for that degree of compute power.

And then a metadata challenge looms that can be larger than the original data challenge. Think of a single image capture for a smart vehicle and all the contextual data it generates—date, time, location, objects in the image and their speed relative to each other, as well as the world outside the image (environmental context)—and the list gets quite long very quickly.

Standard data methods and tools cannot be applied unchanged to AI. An organization needs proper tooling to prepare, refine, clean, combine, and reuse data for AI (see “Perspective: IBM’s Chief Data Office—Tools to automate data governance”).

For example:

- *Reusable data pipelines and technology to support data cataloging,* which can automate some aspects of data governance and lineage validation. This makes it easier to reuse data and avoid re-establishing all the approvals and initial groundwork to obtain and use it, since many licenses stipulate what part of the data may be used and how.
- *Data virtualization tools* that lower cost and simplify some aspects of using data, allowing data to be represented and manipulated in a new environment without actually moving it.
- *“No-code” modeling tools* that enable visual, almost drag-and-drop application of AI models to data sets for business users and others without AI engineering expertise.

Closely related is the kind of environment companies should cultivate. People throughout the business—data scientists, data analysts, business owners—should easily be able to obtain the data they need. A feature store—think of it as a user-friendly interface that sits on top of a traditional data catalog—can make it easy to locate trusted data and AI models that have already been catalogued and vetted. This supports reuse throughout the enterprise, not just for technical users, helping a company realize the most value from its data and AI investments.

Since data bias can undermine AI results—and even disenfranchise key stakeholder groups—tools that can automatically detect and help mitigate bias should be part of the company’s enabling platform, along with data visualization tools that can highlight potential issues with data.

Well-designed tools can also engender trust. For example, imagine an oil and gas company with a treasure trove of geographic, topographic, and seismic data. Its team of field workers, including scientists and engineers from various disciplines, is ready to make use of it—but lacks the programming skills needed to work with AI models that the data scientists have built. Integrating the right tools with the models and data could give the field crews responsible for everyday operations the ability to fully engage with the AI without being coders skilled in data science.

Choosing the right analytic approaches and tools can help facilitate a focus on value: for example, tying algorithmic performance to business metrics. In a real-world illustration, a European grocer was using AI to predict where best to invest in future locations. In analyzing the current context of where existing stores were, it needed to be able to distinguish between small stores and very large ones.

To select the right metrics, the team needed to marry a deep understanding of the data science with a broad view of the business objectives. For those so inclined, Logarithmic-Root-Mean-Square-Error proved a far better approach than more traditional Root-Mean-Square-Error to capture the massive differences in scale—a clear case of understanding the business intent and the best tool to match it.[13]

The right tools matter, but successful use of AI requires integrating these tools as part of a broader, strategic data platform. With such a platform in place, teams also benefit from having cross-functional skills at the outset of a project, including strategists, business analysts, functional subject matter experts (SMEs), data engineers, data scientists, data stewards, and product and project managers.

Successful AI deployment also demands that these teams have organizational structures aligned with their approach to data. The way an organization handles data tends to mimic the organization's structure. A company firmly rooted in fiefdoms and silos will struggle with sharing data. Companies building a successful AI data capability would do well to follow an operating model where AI expertise and practice, wherever it is focused, aligns with many of the characteristics of the business.

Ultimately, though, it's not enough to make sure data practices are up to the task of supporting today's AI technologies. Data management approaches must evolve ahead of what's on the AI horizon—such as neuro-symbolic techniques—and anticipate their unique needs.

Perspective: IBM's Chief Data Office—Tools to automate data governance

Onboarding, cataloging, and governing data assets for a company over 100 years old, with approximately 600 known entities and businesses, is a huge challenge. A key tool to help is the IBM glossary, a set of business terms that describes the data that is the foundation for governance. Initially, using it was almost completely manual, requiring extensive communication with the governance teams and the data source owners.

Long term, this was not an extensible approach and would require hundreds of governance leads matched with an equal number of SMEs to standardize and onboard data. In an attempt to help automate the curation stage and provide metadata to help experts, the teams turned to AI to label the data and automate some of the standard data functions. The initiative, called Automated Metadata Generation, helped reduce the huge backlog of data assets in need of onboarding and curation from 6-8 weeks to under 10 days, with significant decreases in demand on team members' time.

The approach also improved data quality, further simplifying data governance and regulatory requirements. Expanding the use of the technology greatly accelerated the company-wide ability to govern data. Its success led to a request to harden and turn it into a commercial offering.

Governance was an essential part of that hardening. Faulty governance spawns serious issues, so the team enabled the tool to capture the immutable data that could validate the data's provenance and lineage to provide assurances that the models based on it could be easily maintained in a sustainable, proactive manner.

Action guide

Dealing with the data dilemma

While getting AI out of the lab and into full production is far from a trivial undertaking, we have identified key actions businesses can take to speed the path to scaling AI.

Leading practices for less-established AI adopters:
Companies in the considering, evaluating, and piloting phases

Business value

Focus on the problem—and whether AI can solve it. Develop and adapt AI use cases to demonstrate the benefit to others who may co-own or have useful data—including business partners and other third parties. Design Thinking workshops, in particular, can include diverse stakeholders from the start. Including those who control important data sets can give stakeholders a sense of voice, shared ownership, and interest in the outcome of the AI project —and help open up access to valuable data.

Integration

Begin with what's at hand. Focus on what data is actually available. Don't be distracted by "toy data" or even synthetic data to manipulate for experiments. At the other extreme, don't be tempted to integrate data that you can't easily obtain or aren't likely to receive permission to use.

Governance

Prune judiciously. Place into data curation mode or, if warranted, shut down AI projects that can't show a steady increase in training data or real buy-in from stakeholders to share data.

Tools

Don't go it alone. No one organization has all the data, tools, or platforms it needs. The complexity of integrating and governing data from disparate sources with different owners and use restrictions can be overwhelming. Carefully pick partners you can collaborate with who can provide useable data, top-tier tools, and an integrated platform. That way, you don't spend the majority of your resources sourcing, preparing, and managing it all on your own.

Leading practices for more established AI adopters:
Companies in the implementing, operating, and optimizing phases

Business value

Invest in "data journalism" to communicate value. Build an effective culture for trusted AI and data. Regularly report on both project results and potential via events and active data journalism. Interview stakeholders to help other groups understand their perspective and needs. Build stories that illustrate good data practices and the benefits of data harnessed for AI.

Bring to life the human element and positive impacts that AI projects are having on people. And don't forget to envision the future: what will a major business or technical publication be saying about your company (and team) in five years' time?

Integration

Iterate, integrate, and automate. Apply agile methods—iterative sprints, in particular—to identify early on issues with data that may hamper success. Identify other data sets—external and internal—that can enrich insights and achieve business objectives. Build in mechanisms to reuse data and tools to automate and avoid rework.

Governance

Scale data practices to scale AI. During pilots, find answers to operational questions. For example, how you will invoke your models, how often the models will need to be tuned, the rate of data ingestion—streaming versus a scheduled job?—the data volume in production, and the size of hardware required. To deploy into production in the real world, you'll need to consider other data requirements related to governance, volume, and the role of data stewards.

Tools

Empower "citizen data scientists" (carefully). A feature store environment and "no-/ low-code" modeling tools enable business users and others without AI engineering expertise to take advantage of AI on their own. But robust governance is essential to prevent well-meaning citizens from "running amok."

Notes and sources

1 Unpublished data from IBM Institute for Business Value survey on AI Value (n=2765 C-level business executives in major industries, functional areas, and geographies).

2 "2021 CEO Study: Find your essential." IBM Institute for Business Value. February 2021. https://www.ibm.com/thought-leadership/institute-business-value/c-suite-study/ceo

3 "Global AI Adoption Index 2021." Morning Consult on behalf of IBM. April 2021. https://newsroom.ibm.com/IBMs-Global-AI-Adoption-Index-2021

4 Linthwaite, Rachel. "Overcome Obstacles To Get To AI At Scale." Forrester. January 2020. https://www.ibm.com/downloads/cas/VBMPEQLN; "IDC Survey Finds Artificial Intelligence to be a Priority for Organizations But Few Have Implemented an Enterprise-Wide Strategy." IDC. July 8, 2019. https://www.idc.com/getdoc.jsp?containerId=prUS46534820

5 Ibid.

6 Ibid.

7 "Global AI Adoption Index 2021." Morning Consult on behalf of IBM. April 2021. https://newsroom.ibm.com/IBMs-Global-AI-Adoption-Index-2021

8 "AI Global Survey." IDC. May 2019; Linthwaite, Rachel. "Overcome Obstacles To Get To AI At Scale." Forrester. January 2020. https://www.ibm.com/downloads/cas/VBMPEQLN

9 Poole, Erin. "The new frontier: Industries navigate supply chain challenges quicker with demand sensing." IBM Smarter Business Review. May 21, 2020. https://www.ibm.com/blogs/services/2020/05/21/the-new-frontier-industries-navigate-supply-chain-challenges-quicker-with-demand-sensing/

10 "How AI and data models help governments fight Covid-19." Financial Times. 2020. https://www.ft.com/partnercontent/ibm/how-ai-and-data-models-help-governments-fight-covid-19.html ; https://emergentalliance.org/about

11 "AI Global Survey." IDC. May 2019; Linthwaite, Rachel. "Overcome Obstacles To Get To AI At Scale." Forrester. January 2020. https://www.ibm.com/downloads/cas/VBMPEQLN

12 "Global AI Adoption Index 2021." Morning Consult on behalf of IBM. April 2021. https://newsroom.ibm.com/IBMs-Global-AI-Adoption-Index-2021

13 Experts' experience. For a discussion of the two methods, see Saxena, Sharoon. "What's the Difference Between RMSE and RMSLE?" Analytics Vidhya. June 26, 2019. https://medium.com/analytics-vidhya/root-mean-square-log-error-rmse-vs-rmlse-935c6cc1802a.

About Expert Insights

Expert Insights represent the opinions of thought leaders on newsworthy business and related technology topics. They are based upon conversations with leading subject matter experts from around the globe. For more information, contact the IBM Institute for Business Value at iibv@us.ibm.com.

IBM Corporation
New Orchard Road
Armonk, NY 10504
Produced in the United States of America
June 2021

11038811USEN-00