To tackle data science challenges, think like an entrepreneur

October 2022

87%

of data science projects never make it into production (Source: VentureBeat)

What's more important than the customer? Nothing, but data and analytics are a close second.

IDC's spending and resiliency survey found that behind customer engagement, executive boards' "second highest strategic area of interest ... is leveraging data and improving decision making to remain competitive and/or seek to exploit changing market conditions."[1] Gartner® forecasts that by 2026, your competition will continue to increase investments in data and analytics by 45% to become more data driven, digital, and environmental, social and governance (ESG)-compliant.[2]

So why then have many organizations across sectors found the great promise of artificial intelligence (AI), machine learning (ML) and data analytics failing to meet expectations? We've reached an inflection point where Moore's Law and the associated exponentially accelerating power of technology is the baseline for keeping up with a rapidly changing technology ecosystem. So if 72% of data and analytics leaders are leading, or are heavily involved in, digital transformation initiatives, why aren't we seeing greater results in AI, ML and applied intelligence?

Consider:

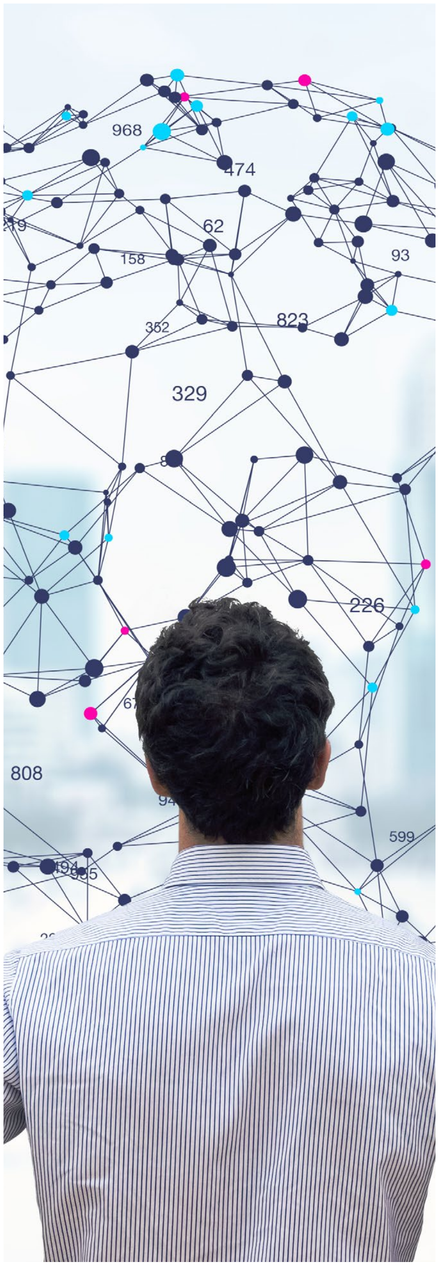
- Nearly 9 in 10 data science projects (87%) never make it into production (Source: VentureBeat).
- Only 1 in 5 analytics insights (20%) ultimately delivers business outcomes (Source: Gartner).
- Fewer than half of all ML models (45%) get deployed enterprise-wide, (Source: Algorithmia).

What's behind this dramatic shortfall? A primary factor is that AI projects are often initiated without an end goal in sight. Also, many organizations are unable to embed AI-driven insights into their legacy systems; others fail to drive agility in business process execution, and some even overlook the need for their AI projects to generate top-line growth or monetization.

Another reason is the lack of skilled data scientists. These professionals are costly to hire and difficult to retain. According to the U.S. Bureau of Labor Statistics, recent data science graduates earn over $100,000 a year, while experienced data science managers can earn much more — over $250,000 a year, according to a University of Wisconsin survey. And turnover is high: A mere 2% of data scientists hold the same job they did five years ago.

1. IDC Future Enterprise Resiliency and Spending Survey (Oct 2021)

2. 2022 Gartner Forecast Analysis: Data and Analytics Services, Worldwide. GARTNER is a registered trademark and service mark of Gartner, Inc. and/or its affiliates in the U.S. and internationally and is used herein with permission. All rights reserved.



The path to business growth through data and analytics is rife with friction, imbalance and bottlenecks, making it primed for disruption. After decades of investments in data and AI, companies are still dealing with a lack of scalability or flexibility, low trust in data and nonrepeatable experiments where integration is an afterthought. Most data scientists, for example, focus on the *efficacy* of AI/ML models, not on where in a business process those models should be embedded for the greatest efficiency and agility. It’s simply not part of their responsibility.

In short, the end-to-end value-chain mapping is rigid and deficient, but there is one way to overcome this and unlock nascent value: corporate boardrooms and data operations need to think more like entrepreneurs, with a relentless focus on customer experience. Here we present an organized perspective on common obstacles to success, and the immediate actions you can take to create value with every data science and AI endeavor you take on.

Three data science challenges

Many organizations seeking value from their AI/ML investments now face three daunting challenges:

- **Industrializing insights.** While most companies need to monetize their investments in AI/ML, many still struggle to scale to an enterprise level, manage the models (detecting bias and drift) and continually improve the models with feedback loops. To deliver a positive return on AI/ML investments, leading organizations will create a culture of continuous improvement.

- **The high cost of end-to-end AI/ML solutions.** Legacy hardware and software, including mainframe and enterprise applications, can make it difficult to run models seamlessly and efficiently. Still in use at many organizations, these systems are either too old or too complex (or both) and become an impediment to embedding AI and model scoring. These legacy integrations will require an astronomical financial expenditure in order to deliver any value. Furthermore, these systems don’t deal with API or service-oriented architecture, so they will become a point-to-point connection with matrixed integration — which itself is a very high-maintenance proposition.

- **Enterprise symbiosis.** Given today’s technology stack, tools are no longer the main impediment. Instead, it’s people and processes. Data scientists can create AI/ML models, but few understand the business processes and models that these models need in order to be fully integrated. Collaboration across functional silos is required for continuous innovation, yet few organizations embrace cross-functional collaboration as part of their culture. Having a common goal of operationalizing monetization, or optimizing a business process, can bring the whole enterprise together.

Figure 1. Hurdles to thinking like an entrepreneur

Leading organizations are adopting a mindset of continual transformation and productive disruption.

Entrepreneurial mindset

Embedding data-driven insights into business processes requires continuous disruption and transformation. Companies that do this well are those that think like an entrepreneur — and this leads them to continual transformation and productive disruption.

One important component of this entrepreneurial mindset is the creation of functional, multidisciplinary “pods”: work groups that cross functional silos to unify people and processes, empowering workers to take action on data-driven insights. Pods can also foster a culture of both collaboration and continual innovation, which are key to driving AI/ML projects across the “profitability border.” To ensure adoption in the applied intelligence space, it’s critical for the value stream to generate outcomes and also link to the next best action, rather than simply providing hypothetical insights or inferences.

Another component of an entrepreneurial attitude is a willingness to ask tough questions. For data science investments, these questions can include:

- How long will it take our organization to realize a return on our AI/ML investments?
- When can the organization expect our AI/ML investments to help us gain market share?
- How will our AI/ML projects increase consumer adoption of our end products?
- Is the AI/ML solution ethical, and are we creating responsible AI? Will the business trust the AI/ML solution?
- Are we creating a new product or service to create a competitive advantage?
- Are our AI/ML solutions optimal, manageable and sustainable? If not, what can we do to ensure that they are?
- How quickly will our AI/ML investments enhance collaboration, drive end user autonomy and accelerate our go-to-market plans?

Platform-driven thinking

There's another piece to the AI/ML puzzle and the need for entrepreneurial thinking. We've entered a new design concept era, involving a collection of tools, services, knowledge and support that empower teams to develop and enhance products, services or solutions at increased velocity. For this to be successful, you need a value stream-based focus on improving business agility, a product-based organization and an operating model aligned to the goal of end user adoption, creating a new flow of data-driven business transformation. Leading organizations are already engineering themselves around this new workflow, change and data, thereby accelerating their time to value and time to market.

Successful organizations will embrace the "law of accelerating returns," in which technology acceleration results in the most potent innovation, such as quantum computers, 3D printing, AR/VR and nanotechnology. One complication of the new platform thinking is that it involves technologies that may be advancing on differing schedules. Where one technology is fully realized, another may be just emerging. Yet to gain the benefits of platform thinking, all technologies need to be integrated to enable highly interoperable systems, yet loosely coupled components, that can then be scaled independently.

Platform thinking also emphasizes the need for optimizing end-to-end costs. This requires strong leadership, high levels of discipline and agile execution. Tactics and tools for optimizing costs include a collaboration workbench, training for greater AI and data literacy, and a solid change management program.

For some leading organizations, this also involves using what are known as "AI citizen data scientists" (see next page). This group, which can include everyday application developers, does not typically possess ML or analytics expertise. But these AI citizens can create AI models, if they have the right tools, including software development platforms for AI and ML, as well as collaboration workbenches that enable software reuse and eliminate duplications of effort and asset development.

Multidisciplinary teams are another key element of platform thinking. As mentioned above, cross-functional pods can empower workers to cut across business domains with a common definition of success and outcomes, ranging from project incubation to industrialization. Pods can also foster a product mindset. This focuses pod members on delivering tangible outcomes to both internal and external customers. Pod teams can also use Agile development practices to foster test- and behavior-driven development, focused intently on the user experience and related outcomes.

Why leading organizations enlist AI citizen data scientists

Some of the heavy lifting of the AI/ML life cycle done by costly data scientists can be offloaded to "AI citizen data scientists" (known alternatively as "AI citizen developers"). The key to success is equipping these data science enthusiasts with standard sets of packaged model libraries, cataloged models and algorithms, and automated/parameterized data pipelines, for repeatable execution.

While these AI citizen developers may not be able to improvise and drive higher accuracy to models, they can combine training models with various datasets and parameters. They can also incorporate models into business processes and enable AI operations, thereby lowering your total cost of ownership (TCO). Furthermore, with platform-driven thinking and the right setup and tools they can contribute in-depth domain knowledge.

For example, if you had a perception neural network trained on certain data, an AI citizen developer could create a data pipeline and bring it to a certain level of accuracy (say, 80% to 87%). An experienced data scientist could then apply their expertise to tune and "hyperparameterize" the model, taking it to a higher level of efficacy.

AI citizen developers can help with all three main phases of AI modeling:

- **Experimentation.** During this phase, as much as 60% of the work can be dedicated to initial data management tasks, such as data cleansing. With "transfer learning" mechanisms adopted by AI citizen developers, pretrained models can speed time to market.
- **Pilot.** In this phase, models are incorporated into a pilot for a specific region or business process. By utilizing model libraries and understanding the usability of algorithms in other use cases, AI citizen developers can rapidly operationalize pilots, especially when data science experts might be tied to other mature use cases. The faster time to market can also accelerate the adoption of applied intelligence in the organization.
- **Feedback loop.** In this phase, once the AI models are deployed at scale, continuous improvement with retraining might not be available until full realization of the benefits is established. AI citizen developers can be leveraged to orchestrate the retraining with feedback to keep the model current, thereby avoiding the impact of drift.

To empower citizen data scientists to work productively, leading organizations first do some preliminary work. This can include (1) creating a marketplace for reusable algorithms and models; (2) establishing a library/catalog of models using already-hardened algorithms; (3) acquiring collaboration tools that aid model "explainability" with documentation and reproducibility; (4) acquiring production tools for application integration, change management and scaling.

The organizations then work to evolve a high-performing AI group that includes both expert and citizen data scientists.

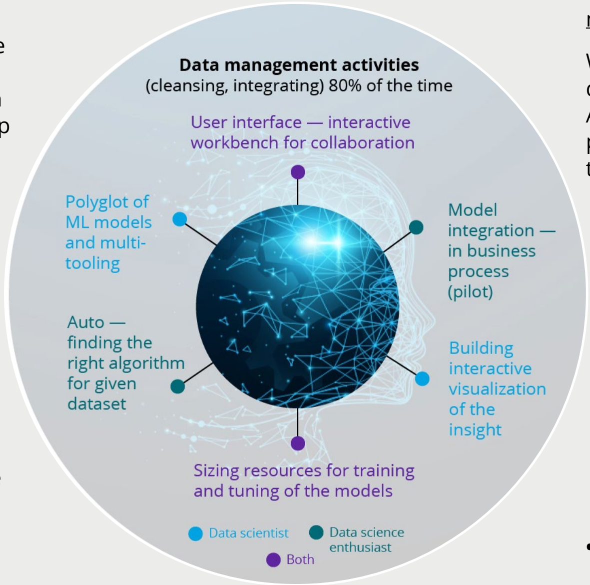


Figure 2. Talent and data management in the AI development pipeline

This evolution typically proceeds in three distinct steps:

Step 1: Coordinated AI experiments. Here, groups collaborate on development AI models. While all projects are owned by the data product owners, business units (BUs) and functional leaders, resources are provided by a central AI workbench.

Step 2: Coordinated detection and observability. The models developed in Step 1 are now deployed with automation into production and are then proactively monitored for anomalous data that can create drift. Portions of this phase are owned by the BUs and functional leaders. End-to-end orchestration is provided by the central AI utility.

Step 3: Coordinated innovation. In this step, "moonshot" programs are launched that aim to extend AI capabilities beyond what's currently possible. These corporate-wide initiatives are defined by the central AI utility, but executed by local BUs and corporate functions. (For an example of a DXC moonshot, see Dr. Peter Scott-Morgan becomes Peter 2.0.)

These organizations also employ tools and processes to correct and filter anomalous data. For example, a drop in revenue might be due to some factor outside the organization's control, such as extreme weather. If that's the case, the organization needs to filter out that data, making sure it does not skew the data with training drift. (For more, read two DXC perspectives: "Defining a data strategy" and "Boosting data metabolism to improve decision making.")

Why should you consider developing citizen data scientists? Because not only do these AI enthusiasts help mind meld and create potential innovation, but the organizations that already use them at scale report three valuable benefits — and you could too:

- **Higher productivity.** Citizen data scientists use collaboration workbenches, which act as software development platforms that can host different products and allow for experimentation. This enables AI models to be shared across traditional silos, empowering their groups to collaborate on engineering features and model-tuning. This also encourages the next generation of aspiring data scientists.
- **Optimized costs.** Organizations can start with an efficient "blended" model that's controlled by enabling people other than high-paid data scientists to develop AI models. Also, citizen data scientists can create a continuous integration/continuous delivery (CI/CD) pipeline to maintain agility and deploy on a standard AI Ops platform to detect model drift. This can enhance the business' trust of the model, clearing the way for automation.
- **Faster go-to-market.** Organizations using citizen data scientists and a collaboration workbench can get a jump start on the initial phase of embedding AI models into their business processes. They can then continuously monitor these processes from the process optimization and avoid model drift by correcting for anomalous data. The goal is not only to collect data, but also learn from it — and then react quickly, correctly and effectively.

Companies that fully exploit today's new technologies, embrace an entrepreneurial mindset and — most importantly — leverage AI citizen developers effectively will experience exponential growth.

How DXC Technology can help

Wherever you are in your AI/ML journey, DXC Technology can meet you there and help you develop your own entrepreneurial approach to data science.

First, you can benefit from our full range of data science accelerators and services. These include advisory, design and engineering services augmented by our intellectual property and a deep pool of experienced professionals who have developed AI models at scale.

You can also tap into our deep experience helping organizations embed AI/ML into their business processes (see **Figure 3**). At automaker BMW, for example, the R&D teams working on autonomous vehicles now use DXC's digital solution to collect, store and manage vehicle sensor data in seconds, rather than in days or weeks, resulting in faster development cycles.

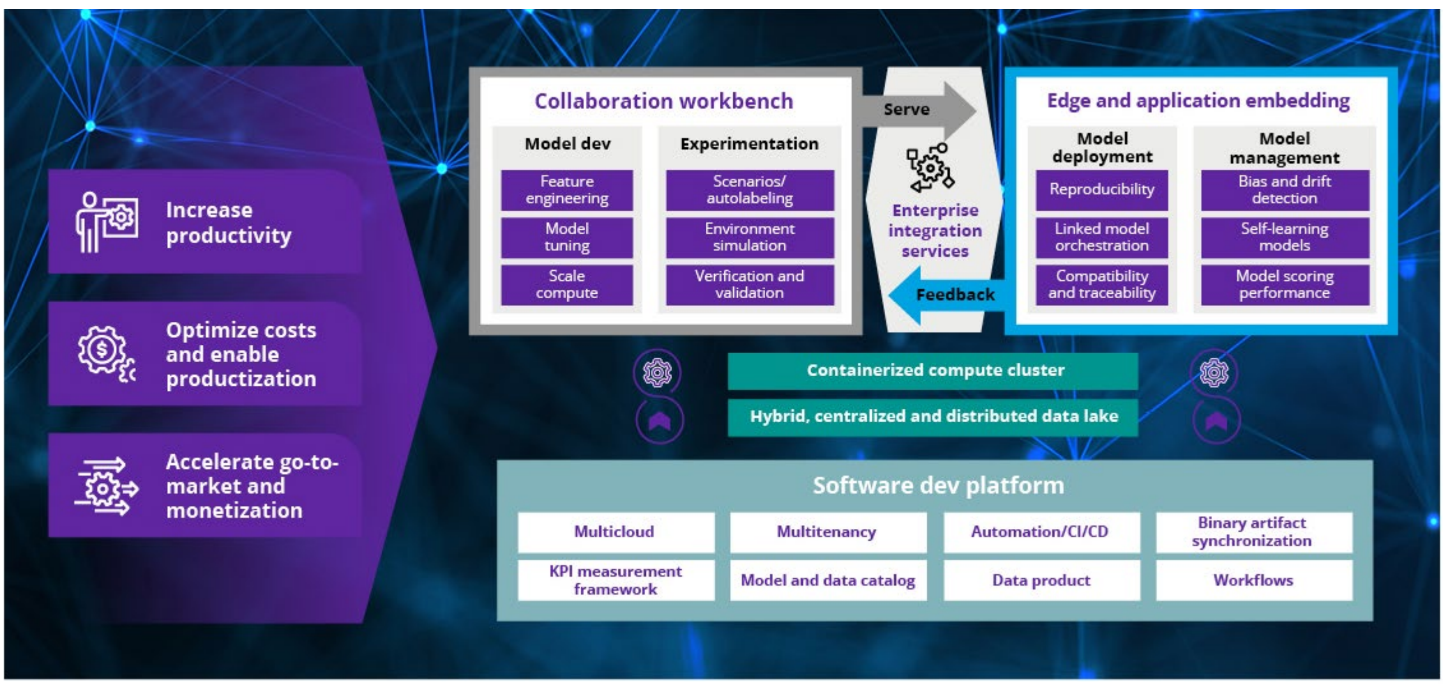


Figure 3. Key building blocks of hyperscaling the realization of your AI

Are you ready to apply an entrepreneurial mindset to your data science projects? **Get started with these three actions:**

- **Action 1.** Itemize the project's expected outcomes. These can be quantitative (such as monetization) or qualitative (such as improved business agility or business value) by introducing a new product or service to the market (see "Addressing the autonomous vehicle data problem").
- **Action 2.** Evaluate whether you are set up with the right people, processes and technologies to provide the required throughput.
- **Action 3.** Determine that you are doing more than driving model efficacy for a business problem, and that you are well positioned to realize the end-to-end user experience through a solution or platform.

Once these first actions are complete, get in touch with DXC, and we'll help you turn your data science challenges into business success.

About the author

Srijani Dey is chief data and analytics officer at DXC Technology and a DXC Distinguished Technologist.

Learn more at
dxc.com/data-analytics

Get the insights that matter.
dxc.com/optin

About DXC Technology

DXC Technology (NYSE: DXC) helps global companies run their mission-critical systems and operations while modernizing IT, optimizing data architectures, and ensuring security and scalability across public, private and hybrid clouds. The world's largest companies and public sector organizations trust DXC to deploy services to drive new levels of performance, competitiveness, and customer experience across their IT estates. Learn more about how we deliver excellence for our customers and colleagues at **DXC.com**.