

Using External Data to Future-Proof Your Organization and Ensure Success Today and Tomorrow

Table of contents

Why BI and analytics are no longer up to the task	6
Data science is here to help	9
The pressures of acquiring better data	12
Hit a roadblock? Here's how to re-route	14
Are you speaking the same language?	16
Curiosity killed the cat, but will save your skin	18
Final thoughts: make yourself the hero of the hour	21
About Explorium	23



You'll have heard it said again and again in recent years: your organization needs to be more data-driven. And it's true; no business can compete in today's hyper-competitive commercial landscape unless they understand, in granular detail, what is happening under the hood, and why. Few people anywhere in the world would dispute that. The trouble isn't that this perspective is wrong, it's that it doesn't go far enough.

It's true that we're living through unprecedented times, and every

organization is scrambling to respond and get back on track. For data leaders, this means having answers today that can help organizations steer through uncertainty. Dealing with a black swan event requires us to think outside of the box, finding ways to expedite processes and find quick wins that can be replicated later. It's not just about putting on a band-aid; data leaders must build systems that solve today's problems, but that can also prepare their organizations for tomorrow's challenges.

Predicting a black swan event

2

You need to look beyond current and historical data, focusing on building predictive models that optimize your operations and continue to deliver real ROI going forward.

You know who was able to predict the current crisis? Scientists who built models that don't just look backwards, but which take into account a variety of factors that go beyond their immediate data. They take into account external factors beyond the immediate dataset. They don't restrict their theorizing to what happened in the past. They model what could happen in the future.

For your organization to have a hope of surviving whatever comes next, you need to start thinking like that, too. You need to look beyond current and historical data, focusing on building predictive models that optimize your operations and continue to deliver real ROI going forward.

In short, your organization doesn't need more data analysis. It needs more data science.

Over the next 23 pages, we'll show you why... and how.

Why BI and analytics are no longer up to the task

Right now, CDOs are in a tricky position. Everyone's looking to you for answers, and while the quality data you've been collecting and analyzing until this moment was fit for purpose just months ago, it's simply not enough to carry you through the chaos we're experiencing now. You can't solve tomorrow's - or even today's - problems with yesterday's data.

This is, in a nutshell, why you urgently need to upgrade from BI to sophisticated, AI-driven data science (if you haven't already). BI tends to draw on internal, historical data to give you a helpful picture of how things panned out in the past so that you can infer what that means for the business going forward. For example, your organization's marketing team might want to look at their ad conversion rate over the same quarter in previous years and use this to inform their forecasts this year.

Data science provides far more nuanced, well-evidenced ways of creating predictive models. You still look at that historical data, but you do so alongside contextual data, typically taken from external and/or alternative sources.

In the ad conversion example, that might include purchase history, consumer spending habits, detailed demographic or psychographic data, as well as up-to-date information on the wider economic picture and public perception of specific brands or products.

You would then use this to build a more reliable, dynamic, predictive model that gives you a good idea or how many conversions you can expect, or what kind of person is more likely to convert, guiding your targeting and ad spend decisions.

Right now, CDOs are in a tricky position. Everyone's looking to you for answers, and while the quality data you've been collecting and analyzing until this moment was fit for purpose just months ago, it's simply not enough to carry you through the chaos we're experiencing now. You can't solve tomorrow's - or even today's - problems with yesterday's data.



This isn't to say all BI is useless, but in today's increasingly uncertain landscape, yesterday's sure thing doesn't mean much. Your organization's multiple departments are navigating murky waters at best, and, at worst, flying blind. Your marketing team can't really predict what customers care about today with historic data — priorities have been so drastically reframed that it's no longer relevant. This is where data science can help.

Imagine a global retailer whose products aren't flying off the shelf as much as they'd like today. The company's BI tools say that their numbers are exactly where they should be, but recently sales have been slumping at multiple locations. The reality is that consumer demands have changed so much that simply looking at last year's sales is no longer enough.

Data science is here to help

Instead, a data-science driven CDO could direct their team to look at a broader cross-section of data: social media interactions, search engine queries, foot traffic and geospatial data, and even demographic data. This new predictive model can more accurately pinpoint where demand will be highest, and which products to stock in order to maximize profits.

The same CDO could also have trouble with their supply chain, especially when it comes to procurement. Economic shocks tend to alter business as usual, causing price distortions and market shortages that can severely impact production costs. Even so, historic data tells them that they should keep doing what they were doing.

A more enterprising head of data could look at the prevailing external conditions to get a better idea. For instance, they could compare their own data to real time weather conditions, commodities market signals, company and business data, and even comparative pricing information for similar products, they could understand their new pricing reality and alter their procurement processes accordingly. More importantly, they can set up new data tools that can help future-proof organizations' analytics by precisely forecasting potential dangers.

Elsewhere, a CEO could be tasking their CDO with understanding their new revenue potential during the current crisis. The company's models are all built on data collected during an extended boom cycle, and thus can't really help explain how revenues will react with a marketplace and economy that is nothing like the one they were in just three months prior.

A smart predictive revenue model could forecast company profits by comparing them to weather patterns, stock market prices, broader economic indicators such as confidence and manufacturing indices, news events, and company data collected from across the industry. This could help better prepare for a potential financial winter, reduce costs, and make it through tough times mostly intact.



The pressures of acquiring better data

All this talk about data science is nice but the truth is, there's no data science without the quality data. And yet Gartner estimates that "65% of organizations do not measure the financial cost of poor-quality data." This is a problem, Especially in unpredictable times like today, internal data silos no longer maintain the quality or relevancy they did even months ago.

So, to ensure your organization is able to use quality data to make business decisions, data leaders, who already spend a large amount of time on acquiring quality data, need to spend even more time on this task — and do it rapidly.

Acquiring data, internally or externally, is a time-consuming and resource-heavy process. It might look a little something like this:

 Understand the issue the business is trying to solve. This is critical in order to proceed with researching the external data sources that may improve your model.

2. Research in order to find the right type of data from the right partner or provider.

 Perform your due-diligence in order to check a whole host of possible issues including compliance, latency, and coverage.
Test this new data. For external vendors, data POCs are timeconsuming so make sure you choose the providers you test carefully. Be sure to go through NDAs, commercial terms, and budget allocation that will enable you to enrich your data with the provider's data.

 Calculate ROI with the understanding that Some initiatives will end with negative ROI. That's the nature of this process — you never know beforehand what data will actually have an impact.

So how can data leaders and their teams bypass this process? By looking for tools that instantly connect them with thousands of external and impactful datasets. More on that a little later.



Hit a roadblock? Here's how to re-route

You can't work out where you're going until you know where you're starting from. You need to get to grips with how your organization makes decisions now so that you can create a strategy for ushering your colleagues towards a more data science-driven approach.

For example, does your company already invest significant resources in data analysis or BI? If so, they are probably already used to basing their decisions on insights extracted from high-quality data and will understand why this is so crucial to their success. You can now build on this, explaining the pitfalls of limiting their analysis to historical or internal data, rather than taking into account external factors and information that will enrich their data-driven decisions.

You may discover, on the other hand, that senior management takes a more instinctive approach to making and enacting final decisions.

Perhaps they view their insights from data analysis more as a resource to support or further interrogate their assumptions, rather than as a guiding principle. If so, you have a very different task at hand.

Rather than making a case for pivoting towards predictive modeling, you will need to persuade your organization that AI and data science should play a more serious role in their decision-making process. That could, for example, involve framing predictive modeling as a way of accounting for a much broader range of previously unimaginable outcomes, complementing rather than negating the value of their lived experience and expertise.

The point is, your existing corporate culture will dictate the approach you need to take (and the arguments you need to make) to bring the organization from where it is now to where it needs to be. Make sure you take this seriously and work your findings into your transformation strategy.

EXPLORIUM



Are you speaking the same language?

One major challenge when shifting to a corporate culture that's based on data science is this: your business colleagues probably don't feel confident speaking the language of data. They likely know what broad questions they need to be answered from a business point of view, and what gaps in their knowledge they need to fill, but they may not know how to frame this as a data problem.

That means it's your job to help them become more data literate. Part of this involves developing some domain knowledge of your own, so you can explain the purpose of your projects in terms they can appreciate and relate to easily. However, you also need to engage all stakeholders in the design of your predictive models and structure of your data investigations, right from the outset.

You will also need to carefully audit, assess, and catalog the data you hold internally, making it easier for others in the organization to

understand what they already have and how it might be used. This should also illuminate opportunities for data enrichment, highlighting what you currently lack that might be useful, and ways these datasets could be augmented with carefully selected external data.

As you explain and demonstrate what you can do and how it connects to their goals and priorities, your colleagues will become more comfortable asking questions about data science and will be able to request or suggest new projects and ideas.

In other words, they'll become increasingly data literate, facilitating more productive partnerships throughout your organization - and encouraging a more data science-driven approach across the board.



Curiosity killed the cat... but will save your skin

To stay ahead of potential black swan events, you need to be the kind of person - the kind of organization, in fact - that never stops asking "what if..?"

That includes questioning the limitations of the data you collect inside the company. Don't stop at asking what this data can tell you: ask what it can't tell you, too. What other factors could impact on the replicability of historical results and patterns? How would that feed into your model? Crucially, where might you find the datasets capable of filling in these great unknowns?

External data can be incredibly valuable, giving you the insights and competitive edge your rivals could only dream of. Sources range from public records, open-source projects, and government databases through to data harvested from social networks, blogs, and websites.

A word of warning, though. If you're seeking out individual vendors in the data marketplace, you may be dealing with hundreds of different formatting standards, not to mention ethical standards governing the way the data was collected. To avoid harmonization headaches, quality quandaries and crises of conscience, it's well worth exploring whether to opt for a single platform that is already set up to connect to multiple, properly vetted, external data sources that are pre-cleaned and consistent.





Today's CDOs are tasked with a huge and continually evolving set of challenges and responsibilities. As the Harvard Business Review wrote earlier this month, this one job title covers seven distinct roles. Sure, you can be categorized as the data and analytics officer-in-chief, but you might also be:

Final thoughts: make yourself the hero of the hour

- \rightarrow An entrepreneur in charge of identifying new ways to monetize data
- \rightarrow A data-centered application developer
- \rightarrow A defender of data security
- → An architect of data environments that clean and aggregate datasets \rightarrow A data governor working with senior business managers to establish responsibility for data domains
- \rightarrow The resident ethics officer, ensuring all data is gathered, treated and deployed with integrity and care

Taking on all these facets would be an unbelievably complex and demanding role, and it's certainly too broad a scope for one person alone. Yet all these issues do need to be covered if you are to transform your organization into something truly data science-driven. That's why it's so important to have the right technology, strategy, and processes in place to support you.

You need a secure, top-notch platform that links to up-to-the-minute external data. That automates data preparation tasks, allowing you to jump right in and use those datasets without wasting time cleaning it up first. That offers the transparency you need to be sure this data has been collected the right way. That helps you enrich data swiftly, filling in the gaps and identifying new and emerging revenue streams.

By streamlining your workload in this way, you can focus on the bigger, strategic challenge: winning over your business colleagues and getting them excited and invested in what data science can do for them.



EXPLORIUM

About Explorium

Our automated data discovery and feature generation platform automatically connects a company's data to thousands of relevant premium, partner, and open data sources to extract an optimal feature set based on model impact. We're creating a new category as the first company to empower and service business leaders and data scientists with end-to-end automation of data discovery and feature generation —

fueling superior decision-making and driving real business impact.