



The External Data Maturity Model

The External Data Maturity Model

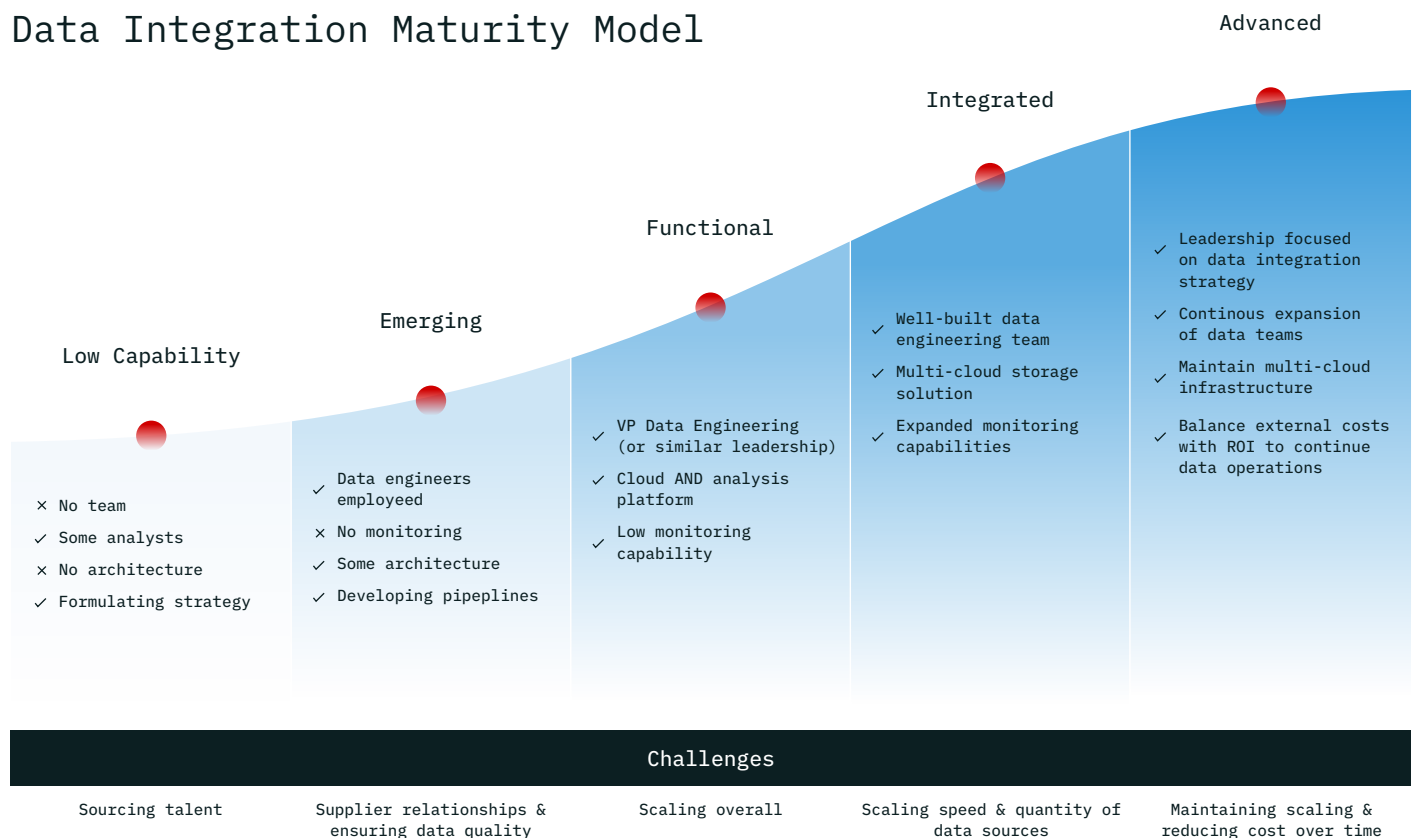
External forces
are all around us...

Anyone who has attended business school, or even joined a singular session on the basics of business has likely come across Porter's five forces.

This framework defines how to analyze the competition for a business, and originated in the Harvard Business Review back in 1979. In this framework, he states the five forces are:

1. The threat of new entrants
2. The threat of established rivals
3. The threat of substitutes
4. The bargaining power of suppliers
5. The bargaining power of consumers

Data Integration Maturity Model



While Porter has said that this framework alone should not be used to evaluate the potential profitability of a business, it's a great starting point to establish industry-specific knowledge. In today's market, it's safe to say that 99% of companies out there use internal data to drive decisions. Technology has advanced to create HR tools that gather and track employee sentiment, demographics, and more that are capable of delivering that information to analysts, who in turn modify it to glean insights before handing it off to leadership teams to guide decision-making.

And that same workflow is how external data should be used—data on competitors, suppliers, and consumers—just like Porter theorized to provide the best view of the macroenvironment surrounding a business. The concept of using external data is the same as using internal data, but the process is far more complex. The internal data industry is far more advanced than its external counterpart, and many companies aren't even on the map when it comes to having successfully implemented and integrated external data into their business processes.

Companies struggle to determine whether to build or buy third-party data solutions, but may not stop to consider their own capabilities when it comes to external data. The maturity model for external data shown below can help determine how far your organization is from benefiting from external data integration and gaining true insight into Porter's five forces.

The external data maturity model is designed as a tool to assess where an organization falls on the curve and provides a roadmap to help companies plan for their future state.

At each stage, organizations must evaluate their current state against a specific team, architecture, and objective requirements while also considering the challenges of maintaining each status.

The External Data Maturity Model

Low Maturity: A Vision on Paper

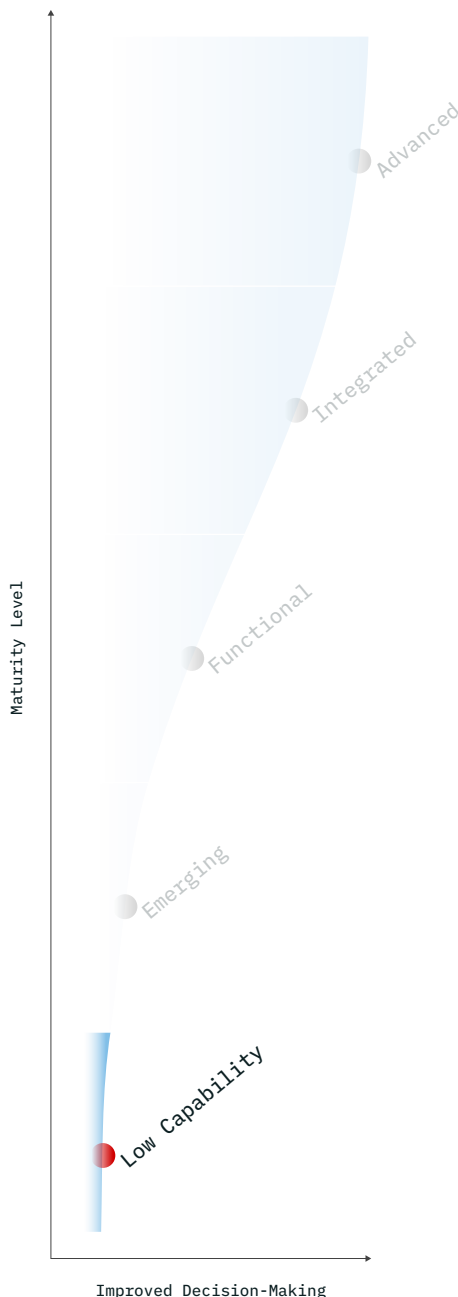
While the word low may carry negative connotations, it's best to think of this maturity stage as a starting point. Your organization is likely in this stage if:

- There is no data engineering team
- There are very few analysts
- There is no external data architecture
- An external data strategy is being formulated but isn't robust or finalized

It's critical to begin evaluating your organization from the specific viewpoint of external data integration to incorporate it with internal data to gain new insights and better outcomes. A company can be placed low on the model for external data maturity, but be fully advanced in its internal data capabilities. The benefits of each are related, but the processes for integrating each are unique.

Companies at this level have to overcome personnel, strategy, and integration challenges to move up the maturity model. The External Data Maturity Model Talented engineers and analysts may be hard to find, and competition for hiring is fierce in today's market. If the company is small enough, developing architecture and strategy plans may fall on current employees without much foresight into what the setup should look like months, quarters, or even years down the road. Minimal viable products are often a staple here when it comes to setting up an initial cloud integration to begin importing external data.

In the same vein, the current data and analytical personnel within a low maturity organization are likely spread thin, and may not have the time to dedicate to researching and applying the current best practices for external data integration, which can result in issues when it's time to scale.



The External Data Maturity Model

Emerging Maturity: Adequate for Occasional Use

To be plotted as emerging on the maturity model, your company will likely fit within the following criteria:

- Data engineers employed
- There are no external data monitoring capabilities
- There is some architecture in place
- External data pipelines are being developed

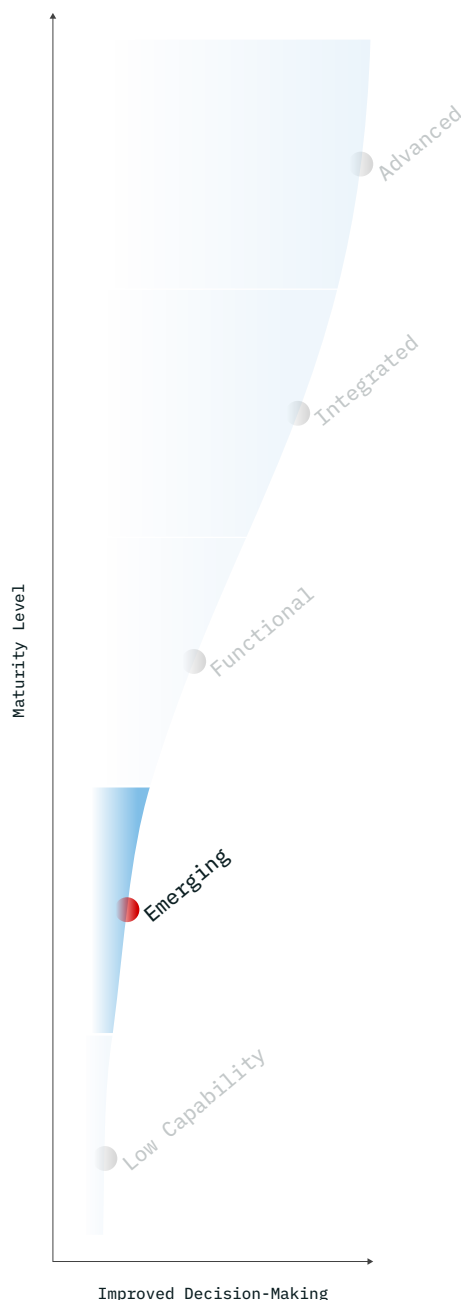
It's important to note that at this stage, it's not the number of data engineers present that matters, but rather their skill and capabilities. While each stage outlines criteria for various levels of data integration, it's less about numbers and truly about the capability of an organization's data vertical at each stage.

The key differences between emerging and low maturity stages are the objectives and challenges at each. The personnel and architecture requirements at each should be trending upward, but in the emerging stage, companies are scaling their data discovery process, building their external data pipeline, and migrating to the cloud.

Additionally at this stage, two new challenges are introduced: ensuring the quality of data pipelines, and maintaining supplier relationships. If governance over external data is not maintained from ingestion through each migration, transformation, and application, errors can slip through the cracks and negatively impact models, analysis, or other use cases where external data is critical.

And as data pipelines begin to become part of the architecture, organizations need data to validate their technology, so data suppliers become integrated into this complex process. Successful organizations will develop a standardized, objective evaluation for data suppliers to help regulate what flows into their tech stack and ensure its quality is high—and that it remains that way.

While the low and emerging maturity stages are similar, they lead up to the critical decision point. Deciding to build versus buy external data solutions can prevent your organization from maturing in the next stages, and hold it back from transitioning from a peripheral to a 360-degree view of the environment surrounding it. We'll dive into the decision point in part two of this three-part series.



The External Data Maturity Model

Functional Maturity: Use Cases with Low Demand & Low Criticality

Internally, a functional stage company has:

- A VP (or similar leadership role) of data engineering
- Both a cloud AND data analysis platform
- Some monitoring capabilities

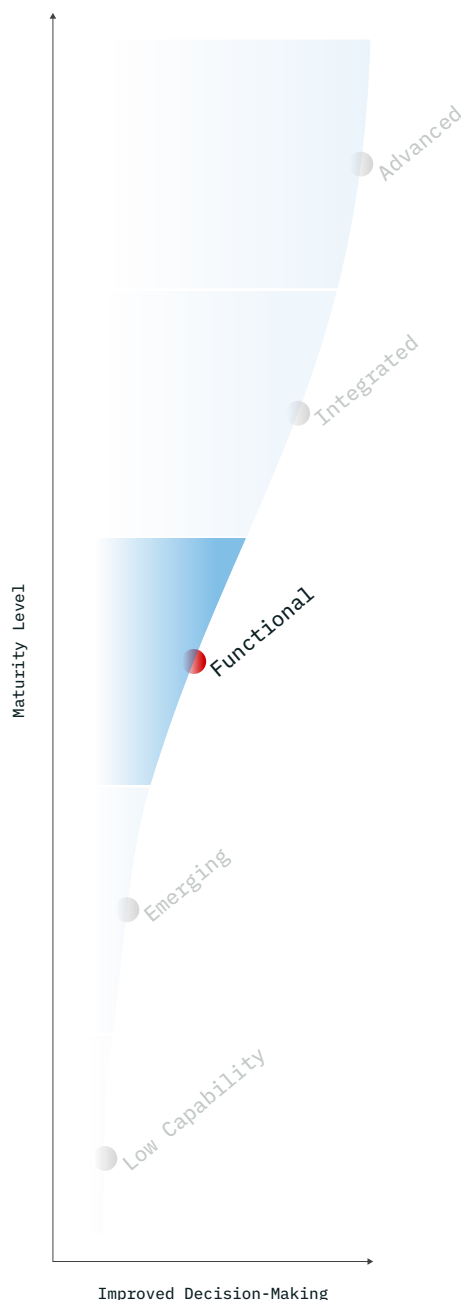
The biggest challenge at this stage is scaling. From here on out, everything your company does must support the scaling of external data, which directly impacts resources and technology. As external data becomes more commonplace in business decision-making, more datasets will be onboarded for data scientists and analysts to use. To scale, the data accessibility should not be siloed, so that costs and efforts are not duplicated across the organization.

Your data team should grow as well. Not only will more analysts be necessary, but data engineers and scientists should be onboarding and growing their roles and responsibilities as well. When scaling your human labor, it's important to consider how they spend their time. Consider this common statistic for today's data engineers—they spend 80% of their time working on administrative, tedious data-related tasks, and only 20% of their time on work that provides a direct value-add for your organization. Scaling should support the reversal of this, so consider the capabilities of the technology that's being built or bought and how it supports the day-to-day tasks of your data team.

As the demand for the number of datasets you are tackling increases, your external data solution must scale to meet that. Any solution should have capabilities to integrate, transform, and monitor data, and those capabilities must constantly evolve to meet changing needs and requirements while also meeting your security requirements. The number of attack surfaces, compliance adherence, and infrastructure setup are all critical considerations that data leaders must consider for their homegrown or purchased solutions.

Your external data strategy should also support your company's growth plans. Will new markets be explored? New industries? Are new products coming that expand your audience? A capable external data solution will support all this growth by adapting to new standards, regulations, and customer needs as the company continues on an upward trajectory.

To advance beyond the functional stage, companies must keep scaling at the forefront of any changes made to their external data solution. If scaling cannot be maintained, then your company cannot progress to the next level of the external data maturity model.



Integrated Maturity: for Critical, Moderate Demand

The integrated maturity stage is the second-most sophisticated stage for external data integration your company can maintain. At this stage, a company will have:

- A well-built data engineering team with depth and breadth in experience and capabilities
- Multi-cloud storage solutions
- Expanding monitoring capabilities

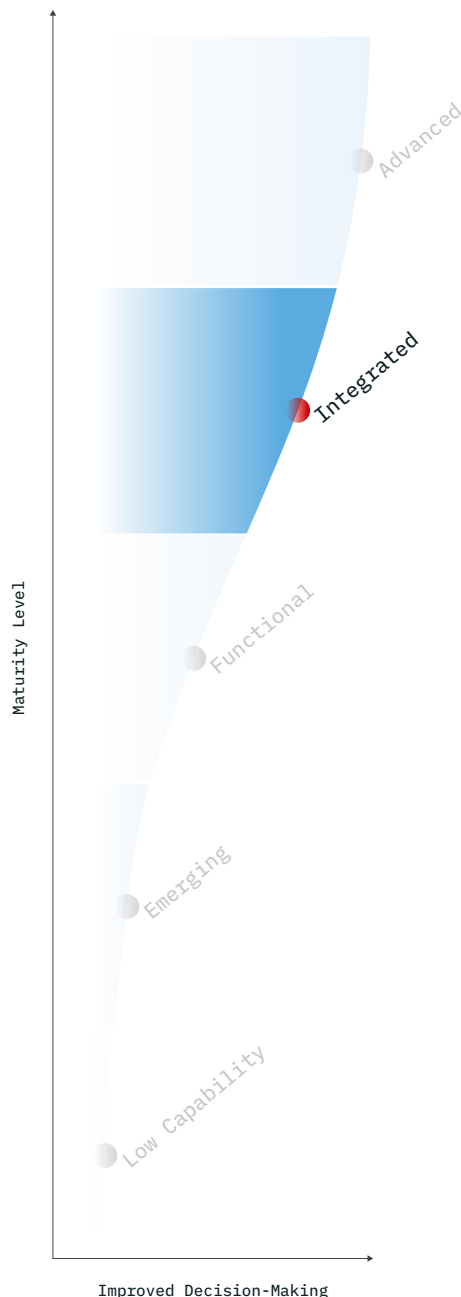
Functional companies face scaling as a broad problem that applies to a variety of resources. But at the integrated maturity stage, the challenge of scaling becomes laser-focused on data sources. At this stage, your organization is typically onboarding loads of new datasets, and the external pipelines for each cannot be built fast enough to keep up with the demand.

These datasets must also be manipulated post-onboarding so that end-users can access and utilize the data as quickly as possible, which becomes a secondary challenge. The work that goes into making a dataset usable for an end-user is often underestimated and can cause delays and inefficiencies if it's not mastered efficiently.

The external data platform has to be robust and able to handle many different processes to make the data usable. The list is quite extensive to meet the needs of internal consumers:

- **Data curation** abilities that reduce the time and expertise required to source and trial new datasets, quickly understand the formatting and structure of the dataset, and include time-series data.
- **Data quality** checks that repeatedly and efficiently monitor data, notify stakeholders of issues, identify anomalies, and provide machine-readable data quality calculations.
- **Data transformation** automation that can parse, format, filter, and modify datasets as needed while maintaining a full catalog of changes and historical information.
- **Entity matching** capabilities that transcend naming conventions and provide accurate context within similar datasets.

While this is not a simple list, providing your data engineers and analysts with these capabilities will allow them to ingest as many datasets as needed to fuel their work, while doing so more quickly and more efficiently than before. Each of these needs must be met and your company has to onboard new datasets at a pace that keeps up with demand before it can become advanced on the data maturity model, which is the final stage for external data integration.



The External Data Maturity Model

Advanced Maturity: for Critical, Moderate Demand

Advanced maturity is the highest level of maturity on the external data integration maturity model. The biggest challenges at this level are maintaining scaling capabilities and reducing cost over time in a way that doesn't limit the benefits external data integration provides.

Typically, advanced organizations:

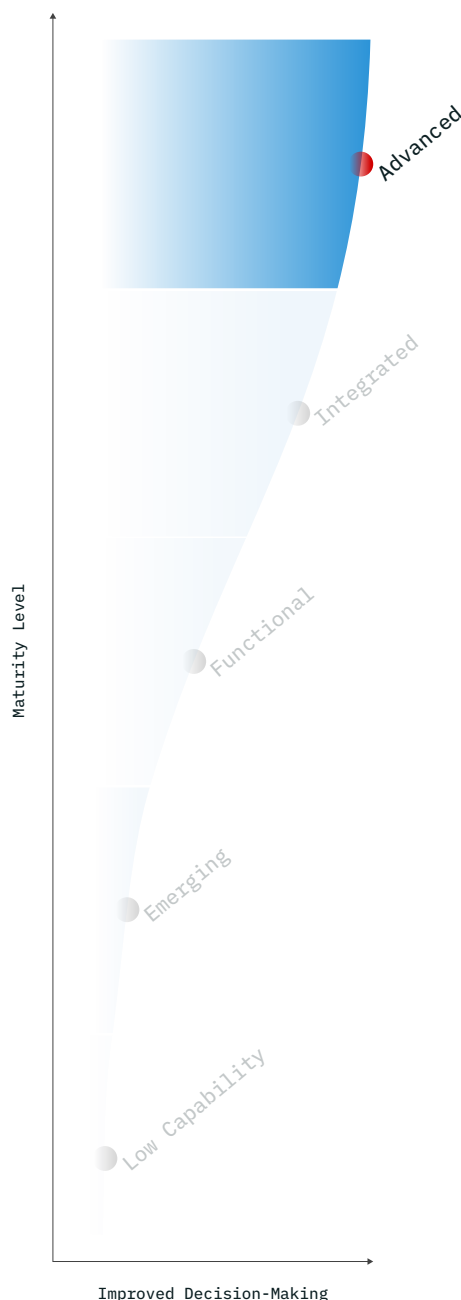
- Have various data engineering leadership roles such as CTO, CSO, and related VPs with a focus on external data integration as part of their overall strategy
- Continue to expand their data engineering, science, and analytical teams with additional personnel and technology
- Maintain multi-cloud infrastructure
- Balance external data costs with ROI to profitability continue data operations

At the advanced maturity stage, maintaining the status quo and adjusting for new changes or requirements is how your company continues to grow its external data solutions. Most companies want to infinitely continue their growth and upward trajectory, and any good external data integration solution needs to support this goal.

At this stage, your company has likely been supported by its external data solution for a few years. Your organization's leadership team should evaluate the cost of the solution versus the company's growing needs, and ask themselves:

- Is there more work that can be automated?
- What new tools complement existing functionality and can provide more robust solutions?
- What can be done to make engineers' and analysts' jobs easier and more meaningful to the business?
- What ROI are these current solutions providing?

Continuing to successfully navigate the needs of a growing external data solution without compromise is critical in this stage.



The Benefits of Continued Maturation

Making changes to your organization's current external data integration solutions to improve time to insights is not an easy task. But improving even a small amount can provide significant value within your organization from resources to productivity to efficiency, and those insights can drive competitive advantage and better business outcomes.

More directly, accurate external data that is efficiently onboarded and utilized can promote better, quicker decision-making, support and justify use cases and processes that were previously unidentified. This creates stronger technology solutions for modeling, machine learning, and artificial intelligence to improve products, customer experience, and employee productivity.

Ready to determine your company's external data maturity status?

Contact Crux today to discuss how we can help overcome some of your biggest external data challenges by reaching out [here](#).