

WHITE PAPER

# The Market Data Hypermarket

How the introduction of a standardized data catalogue  
could supercharge the market data industry

July 2023

By Benjamin Baxter, Jens Munthe, Eddie Molloy, Jordan Galhardo-Burnett

Foreword by Max Bowie

Max Bowie

*WatersTechnology*



## Foreword

**F**or years, market data professionals across the capital markets have struggled to manage the costs of licensing and delivering market data while meeting the demands of their in-house end-users.

The data industry is now experiencing a paradigm shift. The tools now exist to create data-agnostic platforms, or even to manage data assets in the cloud. Firms can now realize long-held ambitions to combine datasets and tools from different suppliers and create a truly best-of-breed data pipeline tailored to their firm's individual requirements. This allows them to save money on data management, while making more money by responding to new business opportunities.

But with more data available than ever before, and with ever-increasing use cases for that data, transparency, discoverability – i.e. what's available in the marketplace, what are all the assets you already have in-house, and what alternatives can be directly compared to existing datasets – and being able to monitor and manage it is even more important. We talk about reference data being the barcode for financial instruments; well, data has become more than just a price tag – it's an important asset in and of itself – and so data itself now needs a barcode.

That's where a catalogue comes in. A catalogue that normalizes data models and taxonomies across different datasets and vendors so that users can better manage existing assets but also obtain better insight into other datasets. Allowing data managers to not only see who is using what data, for what purpose, and how much it costs their firm, but to also see what comparable data exists that might better serve their needs – perhaps even at a better price. This will provide invaluable insight into a much broader and more comprehensive universe of data; not just the data they already know and use.

## Market Data: An Introduction

For over 40 years, financial services firms have been inextricably linked with data providers. Data informs decision-making across every industry segment. From pricing bonds, to modelling portfolio returns, to understanding the environmental credentials of a potential investment, the applications are broad.

This extensive landscape of data has led to a rapid increase in the number of vendors wanting to sell a fit-for-purpose dataset for the relevant business use case.

As digital processes in financial institutions have matured, the availability and diversity of data has matured alongside. The range and quantity of content has increased significantly, and service levels are being tuned for a variety of use cases from ultra-low latency market data for algorithmic trading to end of day pricing data for portfolio valuation.

Vendors handle billions of queries every day, churning through petabytes of data, and the largest financial services organisations work with hundreds of vendors to obtain the most suitable datasets. With an ever-growing list of vendors, it is no surprise that acquiring, organising, and maintaining a firm's data estate is becoming more difficult and costly.

Expand Research spoke to a range of financial institutions, technology providers and vendors to obtain an industry-spanning perspective on these challenges and investigate whether a **market data catalogue** could hold the key to bringing the industry to its next stage of maturity.

## A Complex and Intricate Web

Below is an outline of the key challenges facing market data teams in financial organizations.

### Discoverability of Offerings

At the time of writing, there is no comprehensive centralized search engine or index for discovering required datasets.

Firms generally utilize a combination of individual expertise and enquiries with vendors to identify what is available in the market. There are several obvious pitfalls to this approach:

- **Inefficient** - Laborious desk research and calling up vendors is time consuming and inefficient; having to send queries and convey data requirements takes time.
- **Lack of a comprehensive view** - Relying on individuals almost guarantees that suitable vendors and datasets are overlooked. No one person is likely to have a complete view across all possible options and any vendor data coverage update may be missed.

- **Key person dependency** - If firms have been relying on one person's knowledge of data availability, there is always the risk that person could leave the organization.

## Inconsistency in Naming Conventions

One of the reasons it is difficult to discover data and understand it without detailed review and investigation, is a lack of consistent naming conventions across the industry. Up to now, there has been no defined taxonomy for products, which becomes problematic when looking for more specific or granular types of data or associated delivery mechanisms (e.g. streaming vs. end of day).

This creates inefficiency in discovering, organizing, and understanding an enterprise database.

Common inventory systems are used throughout the industry to provide a line-item view of licensing agreements and consumption, but firms utilise these granular inventories in different ways with inconsistent categorizations.

## Consumption Reporting and Monitoring

Reporting and monitoring data consumption and commercial agreements is also a difficult endeavour. Vendors often have multiple touchpoints throughout organisations, speaking to different individuals about different products.

Alistair Brooker from Calero MDSL sees this challenge regularly at their clients, where their product helps collate data usage: "Calero believes there's a huge amount more that Market Data Vendors/Providers could be delivering to their customers to improve the management of data inventory."

Carrying out enterprise-wide reviews of commercial agreements with a vendor often requires pulling together multiple agreements or reference points. It is also difficult to analyse the data around consumption to understand where common trends exist. For example, how data consumption scales with the increase in the size of front office teams.

## Competitive Sourcing

Ultimately, all the issues above combine to make it more difficult to find the best value outcome for a consumer, the problem statement can be summarized as follows:

**How do firms find the right content, with the right level of service & quality, at the most competitive price?**



# A New Approach: A Market Data Catalogue

How might the industry look to solve these challenges? Many have proposed the idea of an industry-spanning catalogue or index.

The concept is not new; it mirrors the idea of a carefully curated index of any well-organised library, where alongside the name and author, books are categorized by topic and location. A market data catalogue would provide a vendor-neutral index of data sources from a broad range of vendors and other important data providers unimpaired by any commercial bias.

In setting up this capability, a data catalogue would need to take an existing array of individual catalogues from a critical mass of relevant vendors and map them into a common language, or taxonomy, categorizing and indexing them based on their data characteristics such as asset class, content type, delivery mechanism and frequency of data.

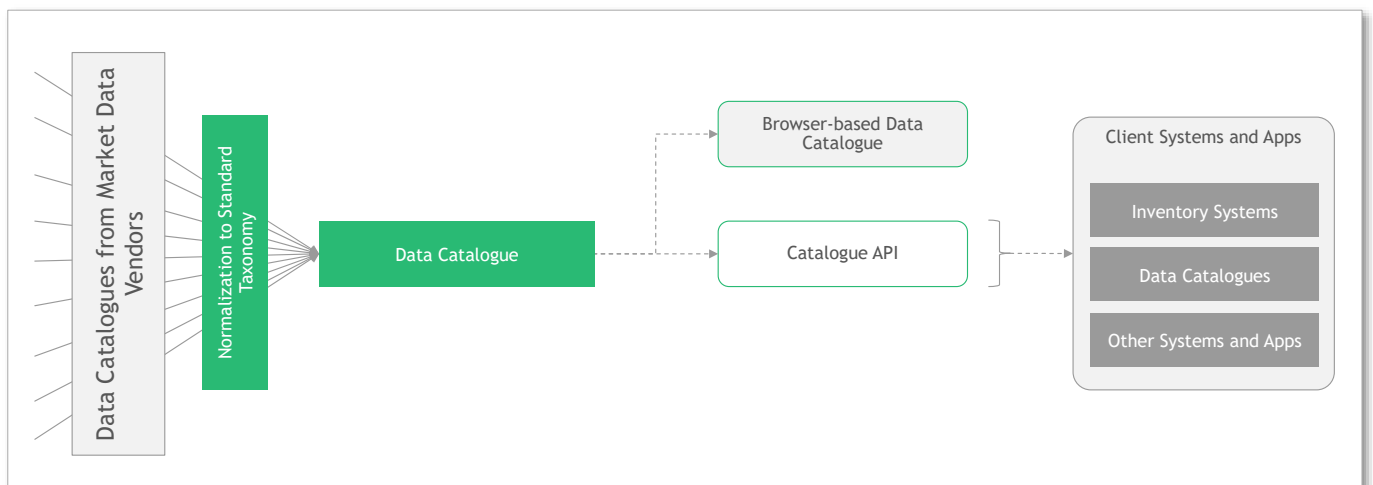
This catalogue could help firms perform two key functions:

- Map their existing inventory of market data products to gain a comprehensive industry standard categorization and highlight alternative providers.
- Leverage a searchable and filterable database to easily discover new datasets and their vendors.

A leading stakeholder from a major tier one bank sees that an integrated market data catalogue could be particularly helpful: “Demand for new niche datasets is increasing and it’s very difficult for [Market Data] teams to currently facilitate that demand. So, the timing is right for an integrated catalogue.”

In essence, this type of cataloguing could take the industry to the next phase of its maturity, in a similar way to open banking platforms or price comparison sites for retail businesses: increasing competition, but also reducing friction in commercial engagements.

## Exhibit 1: The flow of data through a data cataloguing architecture would provide a ‘one-stop shop’ for market data purchasers



Source: Expand Research

## Accessing the Catalogue

We can envisage that a catalogue should be delivered through two channels.

- 1** The first should be an Application Programming Interface (API). By using an API, a firm could ping queries to the catalogue database through a set of simple REST based procedures. This would allow integration of information from a catalogue into existing inventory and reporting systems as well as any firm-based catalogues which keep a record of both internal and external data sources.
- 2** A web-based interface would allow market data teams to directly search and query the catalogue - providing a low-tech, easily adoptable access to this potentially valuable resource.

## What are the Benefits?

The benefits of the creation of a catalogue can be categorised as follows:

- **Greater transparency:** Simply being able to identify a data type by both content and delivery type and find the associated vendor that can deliver will bring significant time savings. The idea that you would walk into any library and find the relevant book without an index would seem crazy today, so why shouldn't that apply to the world of market data?
- **Easier cost management:** With greater transparency will come the opportunity to compare external vendors and their associated costs for a particular data type. While the catalogue will not go as far as being a price comparison site, it will allow firms to identify competing vendors and products. Internally within a firm, the catalogue will also help with cost allocation as the content (and therefore the associated cost) will be more easily allocated to business lines.
- **More impactful reporting:** Ultimately, understanding the use of data and the associated business use case allows firms to understand how they can optimize their efficiency and maturity with respect to the use of financial data. The catalogue will deliver this by bringing a common taxonomy for each data type and delivery mechanism.

## How do we get there?

Firms have been grappling with the implementation of a catalogue for some time now, but several obstacles have hampered progress in the past.

**Vendor incentives:** A key problem to overcome is fragmentation of information in the marketplace. Vendors hold the data on their own various product offerings; their clients have their inventory data. To fully map that inventory (and catalogue products outside of the inventory), the client would have to obtain comprehensive product information from every vendor in the industry.

This is difficult to achieve both due to the number of vendors but also the motivation of vendors to provide such information. Until now, some vendors have viewed increased transparency as a threat to their market position that would reduce the barriers for clients to switch data providers. However, increased transparency is hugely beneficial to the industry. On the vendor side, it means greater data discoverability, but also efficiencies as they and their client base will be speaking the same language.

Some vendors are already looking at better ways of servicing their clients with catalogues and more transparent pricing, which will help influence some of their peers to take the same route. The alternative is the threat of more regulation from the EU/US. If this eventually comes, it is likely to be more rigid and less vendor friendly, so an industry led solution is the better option.

**Change within a firm:** As with any project that is not revenue-generating, socializing the need for a catalogue inside financial institutions can be another challenge. As the requirements are technical and the benefits hard to quantify, getting buy-in from the business to allocate resources and budgets to the creation of a catalogue can be hard to justify.

There is inevitably some internal education required for a catalogue to succeed. While there may not be immediate revenue benefits, a catalogue will enable cost savings and save time - justifying the case for change.

**Technical challenges:** The technical challenges of implementing a catalogue are relatively few and will vary on a firm-by-firm basis. The most cited challenge is the work that needs to be done to map firms' own inventory of market data products to a standardized catalogue.

Particularly, integration of a catalogue API layer into an existing inventory system will require the knitting together of two databases into one uniform reporting tool.

While this may require some one-off investment, the maintenance and upkeep of such a system would be nominal, providing a platform which can be developed and enhanced to consumers' needs.

## The Future for Market Data

The current data challenge is well understood and while the solution seems obvious, the route to getting there is less clear. The misalignment of incentives has created barriers that make it difficult for vendors and clients to bring an industry-wide catalogue to life.

The key to unlocking this capability must be through broader industry collaboration. Co-ordinating this effort requires a neutral third party that can bring all relevant parties to the table. This includes not just the vendors and consumers, but also technology and platform providers, consultants, and firms that supply intelligence in market data management.

For further information on how Expand are supporting their clients in this space, please contact Eddie Molloy at [Eddie.Molloy@expandresearch.com](mailto:Eddie.Molloy@expandresearch.com).

# Contacts

**Benjamin Baxter** is a Senior Principal in Expand's London office.  
[Benjamin.Baxter@expandresearch.com](mailto:Benjamin.Baxter@expandresearch.com)

**Eddie Molloy** is a Director in Expand's London office  
[Eddie.Molloy@expandresearch.com](mailto:Eddie.Molloy@expandresearch.com)

**Jens Munthe** is a Project Leader in Expand's London office  
[Jens.Munthe@expandresearch.com](mailto:Jens.Munthe@expandresearch.com)

**Jordan Galhardo-Burnett** is a Principal in Expand's London office  
[Jordan.Burnett@expandresearch.com](mailto:Jordan.Burnett@expandresearch.com)

## About Expand Research

Expand Research is a wholly owned subsidiary of the Boston Consulting Group. Expand supports financial services organisations by performing detailed diagnostics in areas including market data, technology, operations, and corporate functions as well as business performance measures such as volumes and revenue.

Expand's extensive database contains proprietary data from many of the world's largest financial institutions. This provides an opportunity for tailored comparisons versus bespoke peer groups and flexibility in the way data is leveraged to provide insights into specific areas of focus.

Expand is also a facilitator of community events such as roundtables and forums. These events provide a platform for C-level executives and divisional heads to deliberate on common industry challenges with like-minded peers.





[expandresearch.com](https://expandresearch.com)