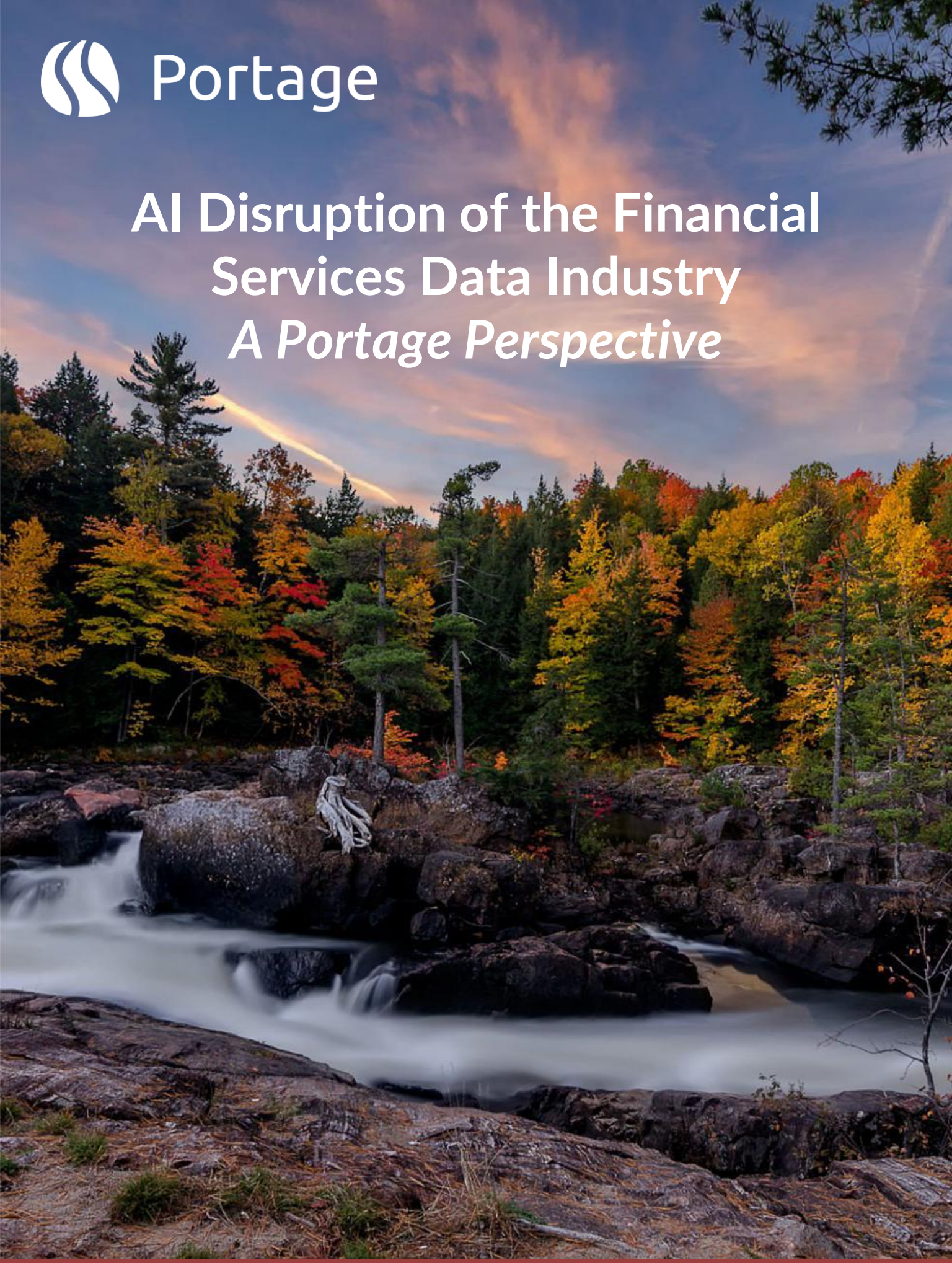
Portage

AI Disruption of the Financial Services Data Industry

A Portage Perspective



2026

AI Disruption of the Financial Services Data Industry: A Portage Perspective

Introduction: Data and the Challenge of Abundance

Data has become an abundant resource across every industry, with financial services standing out as one of its most intensive users. Each day, billions of dollars are put at risk based on insights derived from data feeds, relying on the assumption that this data is both accurate and available when needed. While data itself is no longer scarce, the real challenge lies in distinguishing meaningful signals from the growing volume of noise. As a result, the central issue is no longer how to collect data, it's how to transform it into information that is genuinely useful.

Over nine months, we asked product builders, data management experts and financial services leaders how they are dealing with the tidal wave of data.

Two common challenges emerged:

1. The need for better ways to access data
2. The need for better ways to draw insights from it

While data might seem unglamorous, the scale of the industries tasked with solving these issues is staggering. The data management industry generates $110.5B in annual revenue while data integration generates $15.2B.[1] Yet, in many ways, they are both stuck in the 1990s. The tools are outdated, slow and prohibitively expensive. (One major bank we spoke to was quoted $100,000 to add two columns to a spreadsheet file.)

The challenges are made more daunting by the idiosyncrasies of data management in the financial services world. In large financial institutions, nothing is standardized and, in many cases, it is siloed across different business lines. Financial institutions must mesh internal and external data in a complex, outdated and poorly documented web. As a result, data access is fundamentally difficult, requiring an understanding of business context, proprietary system behavior and highly specialized formats.

1. Source: Grandview Research.

AI Disruption of the Financial Services Data Industry: A Portage Perspective

Into this tangled landscape, enter artificial intelligence (AI) and large language models (LLMs). Their potential for disruption is clear. The current generation of AI is the first credible opportunity to simplify financial data access at scale. Modern AI tooling can make financial data radically easier to access and use for anyone, including by replacing decades-old legacy systems with auto-generated "glue code" for a financial institution's disparate formats.

Importantly, AI can also transform data quality from reactive to proactive—that is, problems like missing values, duplicates and inconsistencies can now be addressed automatically, without interrupting (human) workflow. Today's data quality services are largely diagnostic, but AI can build self-healing data systems, with tabular foundation models, synthetic data generation and a broader availability of structured and unstructured data allowing for the continuous improvement of data quality.

In short, AI has the potential to solve the two major data challenges—access and utility—now facing the financial services industry. That is why we believe that AI/LLM-based models will disrupt capital-intensive financial data incumbents. Significantly, these incumbents are rapidly unbundling their solutions, which—as we have seen in the past when fintech leaders unbundle—is creating fertile ground for startups to build.

1. Source: Grandview Research.

AI Disruption of the Financial Services Data Industry: A Portage Perspective

Our Investment Thesis

We believe that winning teams in financial data will be AI-first platforms. They will provide solutions that:

- Secure scarce, unique and often unstructured data such as documents, transcripts and overlooked sources
- Deliver and serve datasets via automated, developer-friendly application programming interfaces (APIs)
- Use LLM analytics to convert data into hyper-personalized, decision-grade insights in high-value niches
- Shift data quality improvement from diagnostic to proactive and from occasional to continuous

Four key themes support this thesis:

1. ***The financial data industry is monopolistic by niche:*** The firm that owns the reference layer of the data hierarchy effectively controls access, allowing it to tax the market. CUSIP, the alphanumerical code system for financial securities that was bought in 2024 by FactSet from S&P Global, is one example of a niche monopoly; emerging ones are Agio Ratings (in crypto risk) and Fiscal.ai (in their data mapping strategy), both Portage portfolio companies. The potential for monopolistic control of data access makes the space highly attractive.

2. ***Automation in financial data today is rising, but it is supervised:*** By our estimate, about three-quarters of data ingestion, normalization and drafting can be automated. Fully autonomous AI agents could weaken the price advantage of entrenched data providers. Data feeds are price sensitive—a clear opportunity for AI-first startups.

3. ***Specialization matters:*** Many buyers of financial data solutions view generalist tools as incomplete. That leaves them having to build complete solutions piece-by-piece. Our research suggests they are waiting for game-changing end-to-end AI to arrive.

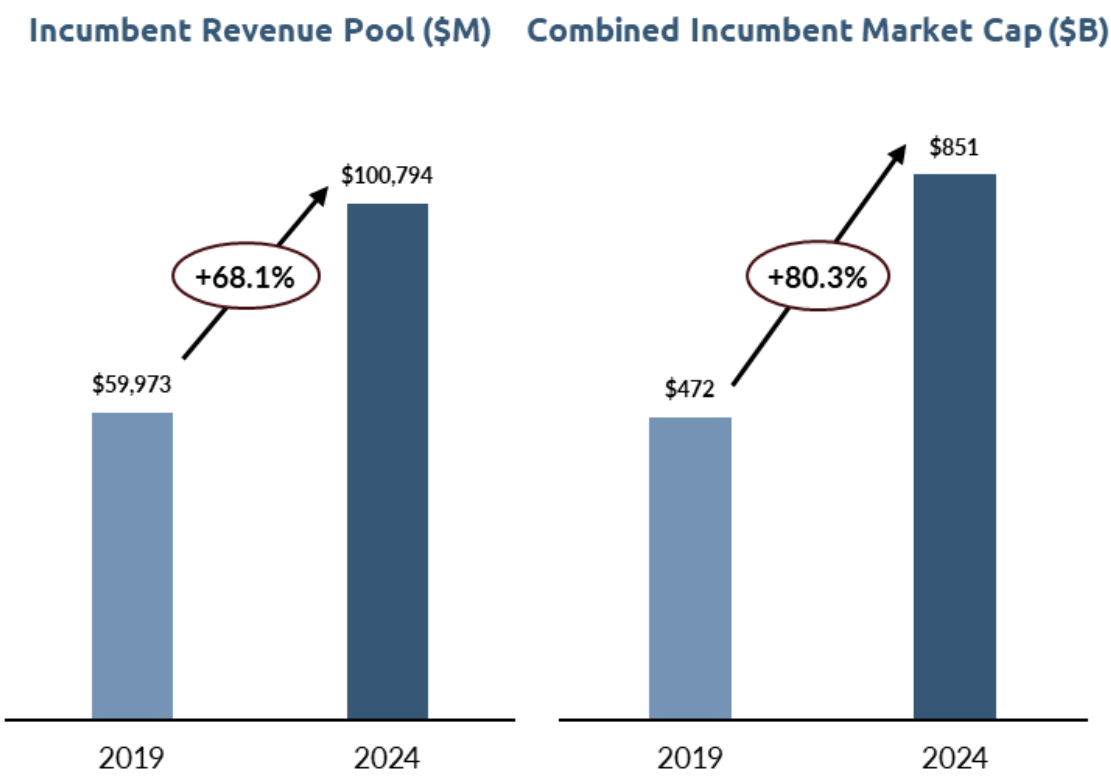
4. ***Data quality improvement needs to shift from passive to active, from occasional to continuous:*** While the new "market standard" has yet to be built, the data quality hurdle on more frontier datasets must be solved.

AI Disruption of the Financial Services Data Industry: A Portage Perspective

Market Overview

The universe of financial services data is large and diverse. It includes real-time pricing and trading data, reference/identification data, financial fundamentals and analyst estimates, news and transcripts, index/benchmark data and alternative datasets (such as geotargeting, satellite and web data), as well as the analytic and delivery layers such as terminals, feeds/APIs, cloud shares, and proprietary analytics such as those provided by MSCI.

Buyers are varied, as well, and they range from financial institutions to tech companies. In 2024, global spending on financial market data and news reached a record US$44.3B, driven by expansion in analytics, reference/pricing data and cloud delivery.[1] A small number of scaled platforms (for instance, Bloomberg, LSEG/Refinitiv, S&P Global, ICE, FactSet and Morningstar) control the space; meanwhile, securities exchanges have been increasingly pursuing strategies to monetize proprietary data, analytics and technology.[2] The result: incumbents have pricing power, and they can quickly roll up new entrants through mergers and acquisitions, which is a tailwind for earnings growth.

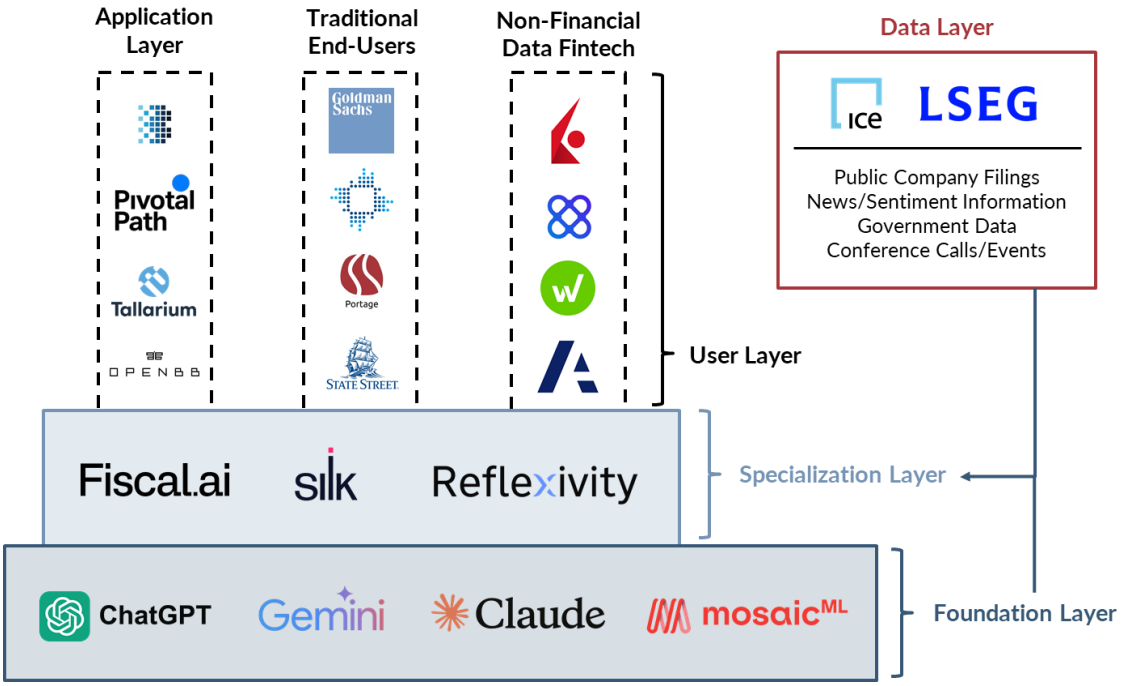


1. Burton-Taylor Consulting via Traders Magazine
2. LSEG Annual Report 2024
3. Note(s): Incumbent peer group included FactSet, S&P Global, MSCI, Morningstar, Intercontinental Exchange, London Stock Exchange, CBOE, Nasdaq, Envestnet, Broadridge, CME Group, SS&C, Moody's, Experian, Equifax, TransUnion, Fair Isaac Corporation, and Verisk. Source(s): S&P Capital IQ, Boosted AI

AI Disruption of the Financial Services Data Industry: A Portage Perspective

When we look more closely at the industry, incumbents can be seen as falling along a gradient. At one end are data generators (exchanges, regulatory filing warehouses, and so on); in the middle are data distributors, such as Capital IQ, FactSet, Bloomberg and others; at the far end are pure analytical layers, such as MSCI, Moody's and so on. We think of this as a gradient because these are not rigid categories: incumbents often move across the gradient as they grow, release new products and sign new partnerships. They do that frequently, embracing inorganic growth and strategic acquisitions to achieve new tech culture, new product capabilities, new data sets and even the removal of nascent challengers. It is, in short, a highly acquisitive environment.

How has AI changed this dynamic? The diagram below shows how we view the emerging layers of AI in financial services data. The raw data sources remain unchanged but specialization layers sitting on top of foundational models become the new primary data access points to brand new AI-first analytical applications, traditional end-users, and fintechs. These companies still operate on the same market gradient described above, but sit within a brand new tech stack, allowing them to compete and deliver value to customers in brand new ways. In particular, we expect the specialization and application layers described below to blur boundaries as consolidation and product expansion occurs.



AI Disruption of the Financial Services Data Industry: A Portage Perspective

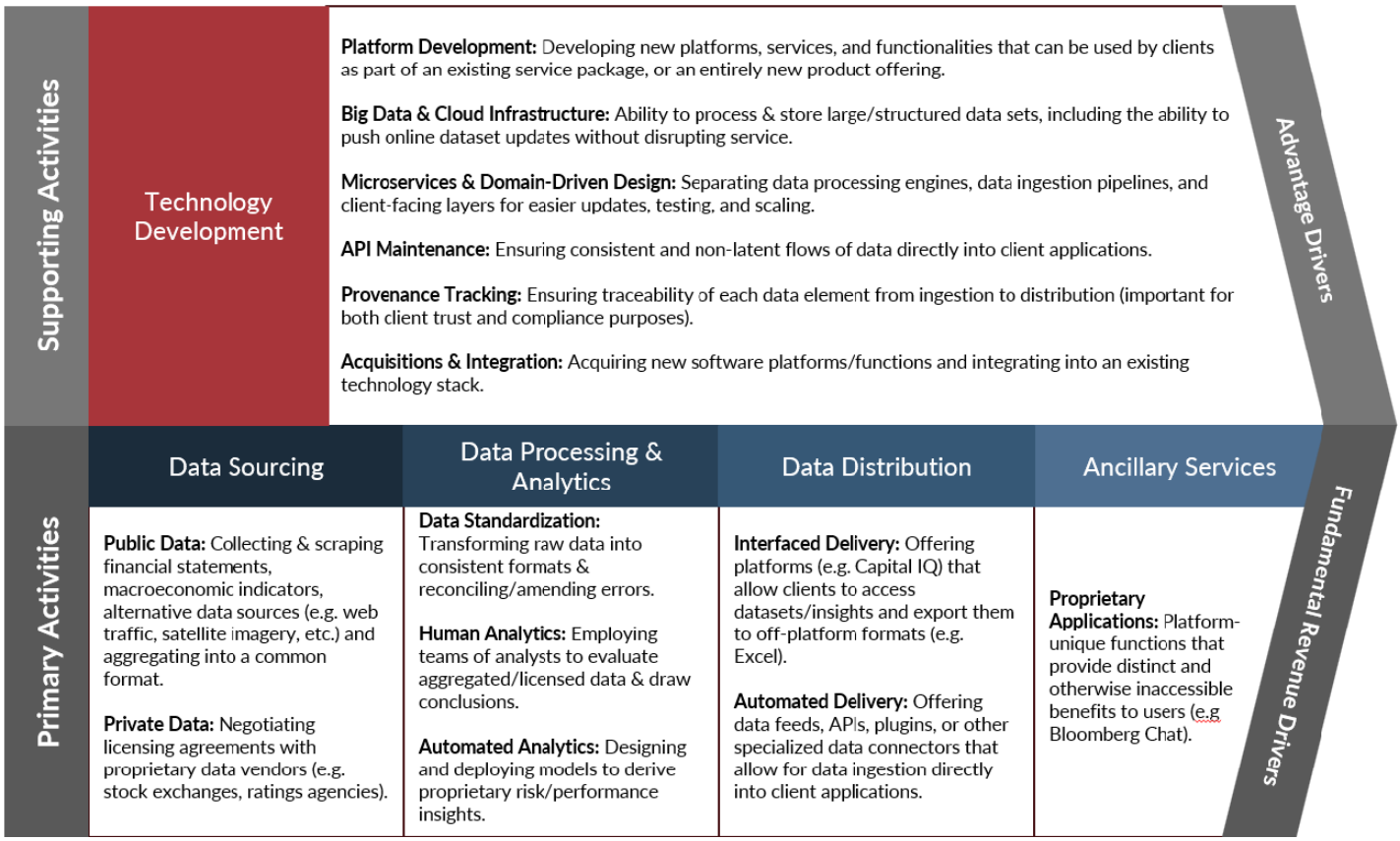
Revenue models are diverse. Players generate revenue through a blend of per-seat terminal subscriptions, enterprise feed/API licenses, usage- or asset-linked fees (notably, in indices,) and data-share/usage constructs in cloud channels. Price sensitivity is low for better solutions, but AI is also allowing for new entrants to provide similar solutions at steep discounts, resulting in new avenues of competition. As well, switching costs are significant; data rights and entitlements, identifier dependencies, audit requirements and established links to downstream risk, reporting and executing systems work to tie up customers by raising the costs of replacement, which skews the market in favor of incumbents.[1]

1. Burton-Taylor Index Industry Benchmark Report 2025

AI Disruption of the Financial Services Data Industry: A Portage Perspective

Disruption in the Financial Data Value Chain

We view the value chain in financial data as breaking down into three core layers. Primary value accrues to owners of scarce, rights-controlled inputs in data sourcing, then to vendors that convert and govern those inputs in data processing and analytics, and finally to distributors that embed outputs in workflows. Beneath those three levels—Data Sourcing, Data Processing & Analytics, and Data Distribution—the supporting layers (including platform development, cloud delivery, provenance and entitlements, and M&A integration) raise switching costs and determine which point solutions can scale or become standards.



AI Disruption of the Financial Services Data Industry: A Portage Perspective

If we take each major layer in turn, we see AI disrupting this value chain in the following ways:

Data Sourcing

- Automated web scraping and extraction: Models that parse and extract relevant information while adapting to changing website structures.
- Red flag detection: Automated identification of unusual entries or inconsistencies in source data.
- Advanced data normalization: Identifying and correcting errors in source data prior to database admission.

Data Processing & Analytics

- Low-latency model updates: Using AI to immediately incorporate new data to update financial models and forecasts.
- Advanced data signalling: Models that analyze search trends, website traffic, social media trends and other non-traditional data sources to deliver advanced, client-personalized insights and signals.

Data Distribution/APIs

- Developer tooling: AI-powered code generation and automated software development kits for streamlined integration.
- Analyst efficiency: Automated query suggestions and pipeline orchestration for rapid insights.
- No-code interfaces: Visual builders and conversational APIs to support non-technical users.
- Volume-based pricing: More effective pricing mechanisms and improved monetization.

AI Disruption of the Financial Services Data Industry: A Portage Perspective

Where We See Opportunities

Our excitement about the disruptive potential of AI-first financial data entrants is driven by the structural vulnerabilities of incumbent operating and revenue models and by the clear need for better solutions. The current generation of AI offers an opportunity to simplify financial data access at scale, making it easier to use, and to shift data quality efforts from reactive/occasional to proactive/continuous. At the same time, a window of opportunity is opening as incumbents unbundle their solutions: new entrants are building better, more specialized products than traditional data distributors and better applications that can act as go-to-market wedges. Finally, our discussions with buyers confirmed that price-sensitivity is low if an advantage over existing solutions and a positive return on investment are clear. We believe that AI-first solutions can tick both of those boxes.

We are particularly interested in these opportunities in the financial data value chain:

- ***Proprietary data for an AI-driven world (Data Sourcing):*** The buyers and experts we consulted all agreed that AI is creating new categories of useful data ("data sourcing" on our value chain above). For instance, AI can deliver direct model weighting data, provide private and alternative (collected from non-traditional sources) data, and develop data niches that can be derived from these categories.
- ***Monopolistic data and analytic layers (Data Processing & Analytics):*** The financial data environment features more monopolies than a casual observer might assume. On the "data processing and analytics" level of our value chain, analytical layers like Moody's and MSCI today function as industry standards, while data mapping tools like CUSIP are true monopolies—and therefore highly cash accretive. Several upstarts have arisen in this space with no clear winner, but the opportunity is attractive and represents a true asymmetric return opportunity
- ***Next-generation data distribution and insights (Data Distribution):*** Distribution is the layer in the financial data value chain most ripe for disruption as incumbents unbundle their products. Current solutions are clunky; API-based platforms that feature low- or no-code interfaces represent the future for buyers and how they want to manage data.

AI Disruption of the Financial Services Data Industry: A Portage Perspective

Investable Themes

Against the backdrop of those macro considerations in the financial data space, we see three broad investable themes:

Standard-setting niches

As we noted above, monopolies abound in the financial data industry, and AI-first entrants have the potential to seize similar status. Our research has also brought to light areas of financial data where the landscape is fractured and there is no standard set for analytics or mapping; these "dead zones" include fixed income, AI model weighting, unstructured data, crypto and private markets. Because of the monopoly-friendly nature of the industry, the winners in these spaces could be able to effectively set the toll.

Examples in the Portage portfolio are **Fiscal.ai**'s Fiscal ID, which is a practical identifier that reduces reconciliation requirements across sources, and **Agio Ratings**, which has the potential to become the reference for financial risk in digital assets.

AI-ready data infrastructure

Buyers today must contend with unstructured data; what they want is clean, rights-compliant, model-ready data with clear provenance. There is demand for APIs that can deliver filings, calls, contracts and sector insights that meet those criteria. Our research suggests that data infrastructure is an area where competition with incumbents will be most intense, but also where incumbents may have the most blind spots in terms of the threats AI poses to this business segment.

Entrants that align with this opportunity include **AlphaSense** (research retrieval), **Hebbia** (document intelligence for due diligence and onboarding) and **Fiscal.ai** (near-real-time financials from filings). In more niche areas, **Theia Insights** delivers AI-driven industry classification and thematic models that help ETF and index providers build precise, forward-looking investment products; **Thematic** offers an AI-powered platform that enables asset managers to research, back-test and launch custom indices and thematic ETFs with speed and precision; and **Naro IQ** offers API-first infrastructure that lets financial firms easily create, launch and manage white-label ETFs and funds.

AI Disruption of the Financial Services Data Industry: A Portage Perspective

Agentic workflow modules inside high-friction processes

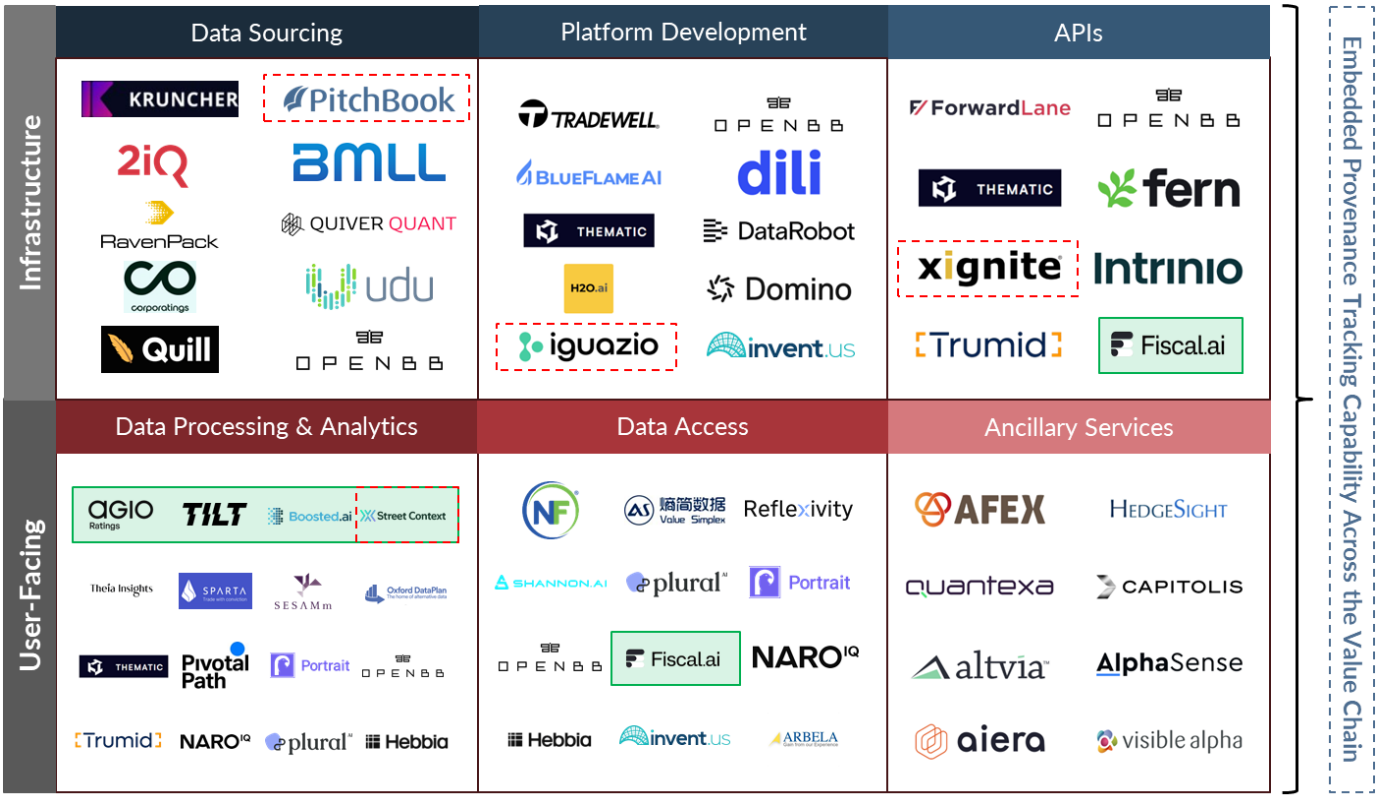
While AI agents are poised to transform and/or replace human agency in several sectors, the extent of their deployment in financial services and its related sectors is likely to be more limited. Within the industry, agents are viewed as enhancements rather than substitutes which enables demand as productivity accelerators. This also serves as an advantage within the sales cycle as keeping humans in the loop quickens procurement. ROI must be real and measurable, and financial analysts need tools that solve for their real-world challenges. Our research indicates that the greatest demand is for agentic tools that embed within and enhance existing analysts' workflows. The key here is effective orchestration, and we see future alignments with next-generation distribution layers creating a more holistic financial analysis experience.

Among the entrants of interest in this space, **Invent.us** ships wealth management workflows that connect to advisor data warehouses; **Trumid** provides data solutions for pre- and post-trade intelligence that plug into existing systems; and **Plural AI** offers an AI-powered semantic knowledge engine that lets finance professionals search and analyze structured and unstructured data.

AI Disruption of the Financial Services Data Industry: A Portage Perspective

Market map

Companies that we believe align with our investable themes



This document does not constitute or form part of an offer to issue or sell, or of a solicitation of an offer to subscribe or buy, any securities or other financial instruments, nor does it constitute a financial promotion, investment advice or an inducement or incitement to participate in any product, offering or investment. Any offer to purchase or buy securities or other financial instruments will only be made pursuant to an offering document and the subscription documents, which will be furnished to qualified investors on a confidential basis or their request for their consideration in connection with any such offering.

Certain statements herein are based on the subjective views and analysis of Portage and have not been independently verified. Any forecasts contained herein are necessarily speculative in nature and are based upon certain assumptions and the best judgement of Portage. Some, or all, of such assumptions may not materialize or may vary significantly from actual results.