OPINION

Noh, Sungho

Outlook and Development Priorities for the Market Data Industry

The global market data industry has continued to grow, driven by the proliferation of AI-based analytic tools and a broadening investor base, with total spending reaching a record USD 49.2 billion in 2025. Alongside this quantitative growth, the industry's tendency to generate increasing returns as data accumulates is expected to concentrate the benefits among a handful of large providers, while value diverges between substitutable, generic raw data and proprietary, exclusively supplied IPs. In response, global providers are integrating key stages of the value chain to secure competitive advantages, primarily in two directions: vertical integration of the data source and distribution, or horizontal expansion at the data processing layer. Korea's market data industry has strengths in the upstream including, consistent, high-quality source data. However, its processing and distribution layers, which lack benchmark IP with global-standard status, could be relative weaknesses. This suggests a need to redefine market data as a core asset in the AI era and the data business as an independent growth segment by (1) commercializing source data and improving distribution channels, (2) achieving economies of scale through the combination of data assets and partnerships, and (3) designing a regulatory framework that balances data accessibility with production incentives.

Market data—such as information on quotes, trade executions, and trading volumes that result from securities trading—reflects the market's price discovery function and is highly valuable for investment decisions and execution, risk management, and other uses.[1)] The global

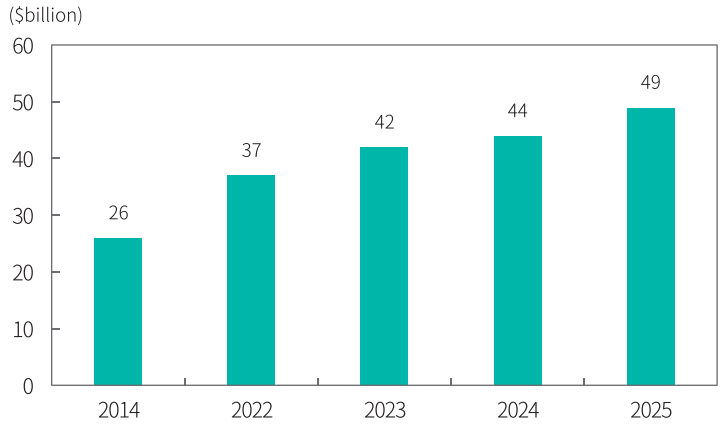
* All opinions expressed in this paper represent the author's personal views and thus should not be interpreted as Korea Capital Market Institute's official position.

* Ph.D., Research Fellow, Financial Services Industry Department, Tel: 82-2-3771-0629, E-mail: nohs@kcmi.re.kr

1) In a broad sense, market data may include information—such as disclosures, news, and credit ratings—that is valuable when used in conjunction with those generated in the trading process. This article includes such information only to the limited extent necessary for the discussion.

market data analytics and sales industry grew substantially from USD 26.5 billion in 2014 to USD 44.3 billion in 2024, reaching a record USD 49.2 billion in 2025. Growth appears likely to continue, as the compound annual growth rate rose from 4.4% in 2014–2022 to 9.7% in 2022–2025 (Figure 1).

<Figure 1> Size of the Global Market Data Industry



Note: Total annual spending in the financial data industry
Source: Burton-Taylor

Drivers of Growth in the Market Data Industry

The growth of the market data industry can be attributed to two main factors. First, sophisticated AI-based data analytics tools have boosted demand for high-frequency, large-volume market data. Of the 131 global asset managers surveyed, 55% have already integrated AI into their investment processes, and 91% plan to expand its use over the next 12 months.[2] Data providers are also competing to establish an early foothold at the intersection of data and AI, as exemplified by the LSEG–Microsoft partnership[3] and S&P Global's launch of the Kensho LLM-ready API.[4] Such a competition implies that, due to the changes in how data is used, growing emphasis is being placed on machine-readable data that is consumed directly by algorithms and AI systems—i.e, "AI-ready" data—rather than by humans.[5]

2) Mercer, 2026. 2, "Moving Beyond the AI Pitch: Asset Managers' Use of AI"

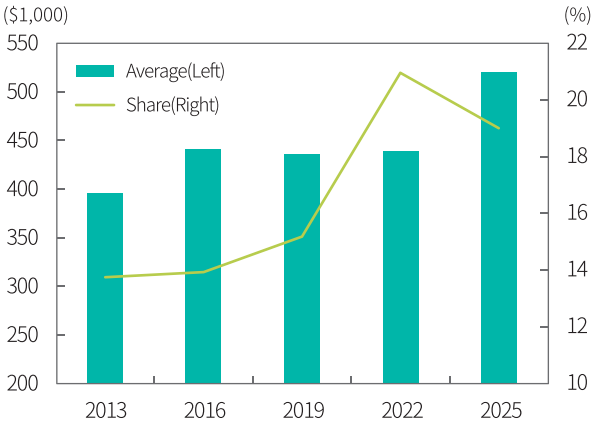
3) LSEG, 2022. 12. 12, "LSEG and Microsoft Launch 10-Year Strategic Partnership for Next-Generation Data and Analytics and Cloud Infrastructure Solutions"

4) S&P Global, 2024. 11. 13, "S&P Global Launches Kensho LLM-ready API (beta)"

5) For instance, Noh, S., 2024, "The Significance of Structured Corporate Financial Disclosures in the AI Era," Capital Market Focus 2024-06, notes that structured financial disclosure information is becoming increasingly important as demand for AI-based corporate analysis grows.

A growing number of market participants is another driver of growth in the market data industry. In the United States, both the share of households that own stocks directly and the average value of their holdings show a long-term upward trend (Figure 2). In Korea, the number of stockholders also rose substantially, from 6.18 million in 2019 to 14.23 million in 2024 (Figure 3). A broader participant base supports growth in the data industry by increasing the number of fee-paying units—accounts, terminals, apps, and the like—that consume data such as market prices and indices. In addition, higher trading volumes increase the production of source data and demand for real-time processing, potentially expanding demand for investment in infrastructure for market data production and distribution.

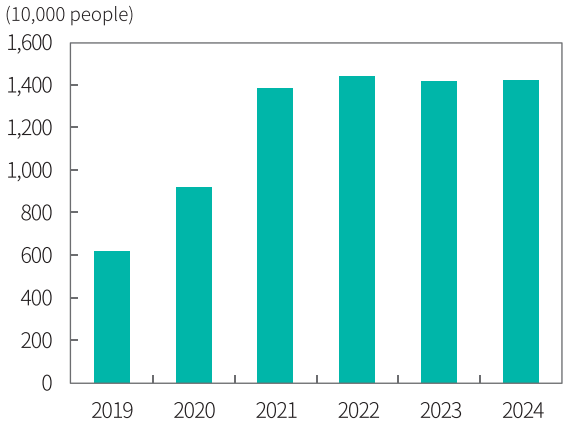
<Figure 2> US Households Owning Stocks



Note: Share of surveyed households directly owning stock accounts and the average value of their holdings

Source: Survey of Consumer Finances

<Figure 3> Trends in Stock Ownership in Korea



Note: Based on shares of listed companies with December fiscal year-ends; the 2024 figure is estimated from the announced number of individual stockholders

Source: Korea Securities Depository

Economic Characteristics of the Data Industry

Combining the economic characteristics of the data industry with the growth drivers described above shows prospects for the industry's competitive landscape. First, since data is non-rival and has near-zero replication cost, data production exhibits increasing returns and accumulation of data further strengthens both the incentive and the capacity to collect more of it.[6] At the same time, because value of information depends on its scarcity, returns generated

6) Farboodi, M. and L. Veldkamp, 2020, "Long-Run Growth of Financial Data Technology," American Economic Review 110(8), 2485–2523

by known signals decline as analytical tools become widespread.[7] Such characteristics help explain recent changes in the market data industry: increasing returns, combined with expanding AI-driven demand, encourage growth centered on a small number of large providers, while scarcity can depress the price of generic raw data and raise the value of proprietary source data and processing IP. As related data resources proliferate—including those associated with multi-venue market structures and alternative data—the bargaining power of generic raw data is likely to weaken, shifting competition toward establishing workflow standards ahead of rivals.

Other characteristics of data include externalities and public-good attributes. When transactions in customer data reveal information about others, the price of data falls below the socially optimal level, resulting in oversupply and failing to ensure efficient allocation.[8] Moreover, market data, as a product of price discovery, has quasi-public-good characteristics that create a conflict between incentives for exclusive provision and equitable access.[9] Consequently, data neutrality regulation may promote competition but at the cost of weakening incentives to produce data. These characteristics underpin data access and pricing regulations, including Europe's consolidated tape (CT) and the US Market Data Infrastructure Rule.

Examples of Value Chain Strategies in the Market Data Industry

The market data industry's value chain can be broadly divided into six stages (Figure 4). Institutions such as exchanges and issuers produce quotes, trade execution information, and disclosure materials (①), which are then collected in real time through feeds, tickers, and other channels (②). The collected data subsequently undergoes cleansing and standardization, including error correction and reference data management (③), before being processed into indices, evaluated prices, and credit ratings, or serving as inputs for signal and risk analysis (④). Data processed into usable forms is then distributed to users through platforms (⑤). Further value can be created as users integrate the data into their operational workflows, such as trading, investment management, and regulatory reporting (⑥). The recent proliferation

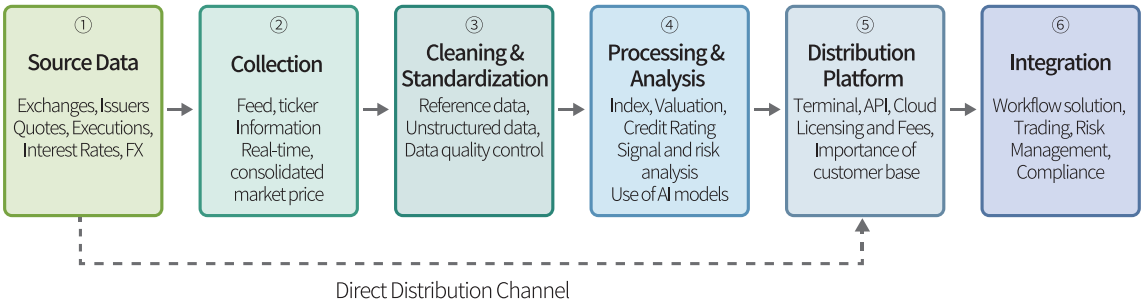
7) Grossman, S. and J. Stiglitz, 1980, "On the Impossibility of Informationally Efficient Markets," American Economic Review 70(3), 393–408.

8) Acemoglu, D., A. Makhdoumi, A. Malekian, and A. Ozdaglar, 2022, "Too Much Data: Prices and Inefficiencies in Data Markets," American Economic Journal: Microeconomics 14(4), 218–256.

9) Fang, H. and S.J. Kim, 2026, "Data Neutrality, Data Supply, and Market Competition," NBER Working Paper No. 35159.

of cloud services and APIs has also created channels through which source data producers can distribute data directly, bypassing intermediate stages (①→⑤). The value added at each stage of the value chain is determined by the scarcity and quality of the data secured by that function, as well as the service provider's status as an industry standard and customers' switching costs.

<Figure 4> Structure of the Market Data Value Chain



Source: Author

Although publicly available statistics on revenue by stage are limited, global data services revenue—approximately USD 45 billion in 2022—is estimated to comprise about 17% from source producers, such as exchanges (USD 7.4 billion), and 83% from vendors (USD 37.3 billion). The downstream distribution and platform market is oligopolistic, with Bloomberg and LSEG together accounting for approximately 50%.[10)]

Major market data vendors are pursuing growth strategies that integrate key elements of the value chain, primarily in two directions. Vertical integration combines adjacent upstream and downstream functions to gain control over the trade lifecycle, while horizontal integration broadens the range of products and domains within the same layer to achieve economies of scale and scope. Representative examples and issues associated with each direction are presented in Table 1.

10) DataBP & Quinlan, 2023, Reclaiming the Market Data Value Chain: Leveraging the Cloud to Optimise Distribution and Product Innovation

<Table 1> Comparison of Value Chain Integration Cases Among Data Vendors

Category	Vertical Integration	Horizontal Integration
Objective	Vertical combination of an exchange with data and distribution platforms	Horizontal expansion within the processing layer (benchmarks and data)
Representative Example	LSEG's acquisition of Refinitiv (2021, approximately USD 27 billion)	Merger of S&P Global and IHS Markit (2022, approximately USD 44 billion)
Expected Benefits	Control over the trade lifecycle, transition to subscription revenue, internalization of distribution networks	Economies of scope, cross-selling, stronger benchmark IP portfolio
Constraints	Integration costs, platform dependence, and switching costs	Costs of ensuring data consistency, forecasting demand for combined data
Policy Considerations	Transmission of upstream monopoly power downstream, data neutrality	Concentration in benchmark and valuation data markets

Source: Author, based on disclosures by LSEG and S&P Global

LSEG exemplifies an exchange combining data production and distribution functions. Through its approximately USD 27 billion all-share acquisition of Refinitiv in January 2021,[11] the exchange, as a source data producer, secured a global distribution network together with index and analytics capabilities, integrating the value chain as a whole. The stated objective of the acquisition was to shift from trading volume-linked revenue to subscription revenue, and data-related businesses have since come to account for approximately 61% of group revenue.[12]

S&P Global exemplifies horizontal expansion of the product range covered at the data processing stage. In addition to its existing benchmark assets, such as credit ratings, indices, and commodity price assessments, the company added energy, supply chain, and private market data through its 2022 merger with IHS Markit. It subsequently optimized its portfolio through selective acquisitions, divestitures, and spin-offs. Its revenue structure, in which recurring revenue accounts for 79.5%,[13] and its AUM-linked monetization in the index business can be viewed as outcomes of its strategy of combining benchmark IP.

11) LSEG, January 29, 2021, "LSEG all-share acquisition of Refinitiv transaction - completion and admission to trading," Press Release

12) The number is based on LSEG's 2025 Preliminary Results, released on February 26, 2026.

13) The number is based on S&P Global's 2025 Annual Report.

Priorities for Korea's Market Data Business

The preceding cases suggest that a key success factor in the growing market data industry is establishing a virtuous cycle of irreplaceable data assets → integration with platforms and workflows → switching costs and subscription revenue → reinvestment. Applying these lessons to Korea's market environment yields three broad development priorities.

First, source data should be productized and distribution channels improved. Rather than simply converting all data into "AI-ready" form, priorities should be established, starting with products for which rights and demand are clear. A licensing framework that clearly defines usage rights by purpose, such as simple viewing or AI training, could be introduced first for cleansed data, including reliable quote and trade execution data and consolidated market prices across multiple venues. At the distribution stage, direct sales through cloud services and APIs are expanding, but may not always be cost-efficient. A practical approach would therefore be to diversify distribution channels by customer segment, weighing the costs of sales, support, and redistribution management against the returns from partnerships with existing vendors.

Second, economies of scale should be secured by combining domestic data assets and forming partnerships. Combining data dispersed across major market infrastructure institutions enables analysis and operational support that individual institutions would find difficult to provide. However, a fragmented structure does not necessarily imply inefficiency, and it is necessary to assess whether the (user) benefits from combining data exceed the costs of standardization, operations, and reconciling rights. Overseas cases suggest that the key to generating stable revenue lies not in enlarging organizations, but in connecting data production, processing, and distribution functions in ways that enable a virtuous cycle.

Third, a regulatory framework should be designed to balance data accessibility with incentives for data production. When determining the intensity of regulation for the market data industry, the trade-off between equitable access and incentives to invest in data production and cleansing infrastructure must be considered. Furthermore, for data containing personal information, a separate framework for data rights and privacy management should be established to address externalities arising from information exposure. The outcomes of such regulatory policies will require ongoing evaluation and management, independently of the data industry's quantitative growth.

Korea’s market data industry has a layered structure organized by security types: data production by the Korea Exchange, Nextrade (for stocks), and Korea Financial Investment Association (for funds and bonds); collection and processing by infrastructure institutions such as the Korea Securities Depository and Koscom; and analysis and distribution by private vendors and valuation firms. Positive factors include scarce, high-quality source assets—such as highly consistent market data, high-frequency data, and standardized electronic disclosures (e.g., DART)—alongside rising data revenue and a high proportion of overseas customers. However, these coexist with threats: the processing layer lacks globally renowned index and benchmark IP, while global vendors are expanding their customer base in the distribution and platform layer. Market data needs to be redefined as a core asset in the AI era rather than a by-product of trading, and the data business as an independent growth industry rather than an auxiliary function, so that the strengths upstream of Korea’s market data value chain can translate into added value in its mid- and downstream segments.