

# Lesson 47: Financial Data Vendors

## Module 4: Traditional Digital Finance

Digital Finance Course

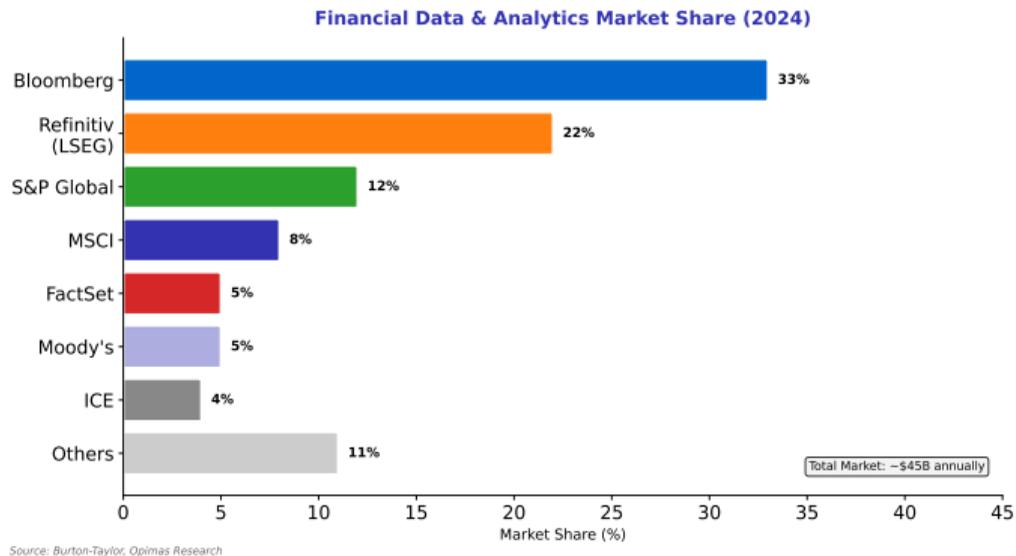
2025

# Learning Objectives

- Understand financial data vendor landscape and business models
- Analyze Bloomberg Terminal, LSEG (Refinitiv), and FactSet platforms
- Examine market data types: real-time, historical, reference, alternative
- Evaluate data licensing models and exchange fees
- Assess alternative data sources and emerging providers

---

Source: **Financial industry data and regulatory publications**



**Bloomberg and Refinitiv dominate the financial data vendor market.**

## Market Size and Growth:

- Global financial data market: **\$35B+** annual revenue (2024)
- Growth rate: 5-7% CAGR
- Drivers: Regulatory reporting, algorithmic trading, alternative data
- Consolidation trend (Refinitiv-LSEG, IHS Markit-S&P)

## Top Vendors by Revenue (2024):

- 1 **Bloomberg LP**: **\$12B+** (**\$24-27k**/terminal/year)
- 2 **LSEG (Refinitiv)**: **\$7B+** (acquired by LSEG 2021 for **\$27B**)
- 3 **S&P Global (incl. IHS Markit)**: **\$4B+** (merged 2022)
- 4 **FactSet**: **\$2B+**
- 5 **Moody's Analytics**: **\$1.5B+**
- 6 **Morningstar**: **\$1.5B+**

## Data Categories:

### 1. Market Data:

- Real-time prices (equities, bonds, derivatives, FX, commodities)
- Historical time series
- Order book depth (Level 2, Level 3)
- Indices and benchmarks

### 2. Reference Data:

- Security master (ISINs, CUSIPs, identifiers)
- Corporate actions (dividends, splits, mergers)
- Company fundamentals (financials, ratios)
- Legal entity data (LEI, org structure)

### 3. Analytics and Research:

- Analyst estimates and forecasts
- Credit ratings and models
- Risk analytics (VaR, stress tests)
- News and sentiment

---

Quality data is the foundation for effective machine learning models. [Source: Stanford AI Index 2024]

## Terminal-Based Model (Bloomberg):

- **Pricing:** \$24-27k per terminal per year
- **Users:** 350,000+ terminals globally (2024)
- **Revenue:** \$12B+ annually (95%+ from terminals)
- **Value Prop:** Integrated platform (data + analytics + communication)
- **Lock-In:** Workflow integration, chat network

## Enterprise Data Licensing (Refinitiv, FactSet):

- **Pricing:** Per-user or enterprise-wide licenses
- **Customization:** APIs, feeds, data integration
- **Flexibility:** Choose specific datasets
- **Cost:** \$10-50k+ per user depending on entitlements

## Exchange Data Licensing:

- **Direct from Exchanges:** NYSE, Nasdaq, CME, Eurex
- **Professional vs Non-Professional:** Different fee tiers
- **Display vs Non-Display:** Machine-only usage often higher fees
- **Redistribution:** Additional fees for reselling data

## Alternative Data (Emerging):

- **Pricing:** Subscription (monthly/annual) or per-use
- **Range:** \$10k-500k+ annually depending on dataset
- **Examples:** Satellite imagery, credit card data, web scraping
- **Providers:** Specialty vendors (not traditional giants)

*Typical hedge fund data spending: \$1-5M/year; large investment bank: \$50-100M+/year*

---

Source: Financial industry data and regulatory publications

# Bloomberg Terminal: Core Features

## Market Data and Analytics:

- **Real-Time Pricing:** 35M+ securities across all asset classes
- **Charts:** Advanced technical analysis, custom indicators
- **News:** Bloomberg News + 1000+ third-party sources
- **Economic Data:** Global macro indicators, central bank calendars
- **Derivatives Pricing:** DLIB library, implied vol surfaces

## Equity Functions:

- **DES:** Security description and fundamentals
- **FA:** Financial analysis (income statement, balance sheet)
- **COMP:** Comparables and peer analysis
- **EQS:** Equity screening
- **GP:** Price charting
- **HDS:** Historical data download

## Fixed Income and Rates:

- **YAS:** Yield and spread analysis
- **WCDS:** Credit default swap pricing
- **IRSB:** Interest rate swap builder
- **FIHC:** Corporate bond screening
- **SWPM:** Swap portfolio manager

## Communication:

- **IB (Instant Bloomberg):** Secure messaging (350k users)
- **MSG:** Market color and trade ideas
- **SDEV:** Syndicate desk for new issues

## Execution:

- **EMSX:** Multi-asset execution management
- **FIT:** Fixed income trading
- **FXGO:** FX spot and derivatives trading

---

Source: Financial industry data and regulatory publications

# Bloomberg APIs and Data Services

## API Offerings:

- **B-PIPE:** Real-time streaming, sub-ms latency, \$2-5k/mo
- **SAPI:** Server-side (C++, Python, .NET); risk/analytics
- **DAPI:** Excel (=BDP, =BDH); ad-hoc analysis

## Enterprise Services:

- **Data License:** Redistribution, EOD, reference data
- **BEAP:** Cloud-hosted, RESTful/WebSocket APIs
- **Advantages:** Coverage, IB network, 24/7 support

---

Bloomberg APIs power trading, risk, and analytics workflows globally. [Source: Bloomberg 2024]

## Refinitiv (Acquired by LSEG 2021):

### Eikon/Workspace Platforms:

- Eikon: \$12-22k/user/year, 400k+ users, MS Office integration
- Workspace: Cloud-native successor (2021), browser-based, API-first

### Key Datasets:

- Real-Time: 650+ exchanges and OTC markets
- Tick History: 18+ years intraday data
- Fundamentals: 85,000+ companies, I/B/E/S estimates

## Refinitiv Data Platform (RDP):

- RESTful/WebSocket APIs for real-time, historical, ESG
- Consumption-based pricing model

## Elektron Data Feed:

- Low-latency feed competing with Bloomberg B-PIPE

## Content Strengths:

- FX: Market leader in foreign exchange
- Fixed Income: Extensive corporate bond coverage
- News: Reuters (120+ years), ESG metrics, Ownership data

---

LSEG/Refinitiv provides comprehensive multi-asset data with strong FX and fixed income coverage. [Source: LSEG Annual Report 2024]

## Post-Trade Services:

### LCH (London Clearing House):

- SwapClear: \$500T+ IRS notional
- RepoClear: Repo and securities lending
- CDSClear: Credit default swaps
- EquityClear, FX Clear

### UnaVista:

- Trade reporting (EMIR, MiFID II, SFTR)
- Regulatory data management
- Transaction reporting hub

### FXall and Tradeweb (Affiliates):

- Multi-dealer FX trading platform
- Fixed income electronic trading
- Integration with LSEG data feeds

## Risk and Analytics:

### Refinitiv Risk Platform:

- Enterprise risk management
- Counterparty credit risk (CVA, PFE)
- Market risk (VaR, stress testing)
- Regulatory capital (Basel III, FRTB)

### StarMine Quantitative Analytics:

- Equity valuation models
- Earnings quality and momentum scores
- Analyst revisions and sentiment
- Used by quant funds for alpha research

### LSEG Synergies (Post-Acquisition):

- Integration of exchange data (LSE, Borsa Italiana)
- Combined index business (FTSE Russell + Refinitiv)
- Cross-selling post-trade and data services

---

**Risk management is essential for financial stability and profitability.**

## FactSet Overview:

- 200,000+ users, \$2B+ revenue
- Pricing: \$12-18k/user/year
- Target: Asset managers, wealth managers

## Core Strengths:

- Portfolio analytics (attribution, risk)
- Equity fundamentals (20+ years)
- Research aggregation (MiFID II)

## Platform Features:

- Fixed income (bonds, ABS/MBS)
- Ownership data (institutional)
- RESTful APIs + Python SDK

## vs Bloomberg:

- Better portfolio workflows
- More affordable for data-only
- Weaker on sell-side trading

---

FactSet leads in buy-side portfolio analytics. [Source: FactSet 2024]

# Specialized Data Vendors

## Credit and Fixed Income:

### S&P Global Market Intelligence:

- Credit ratings (S&P Ratings)
- Capital IQ platform (private company data)
- Compustat (standardized financials)
- Merged with IHS Markit (2022): Added derivatives pricing (MarkitServ)

### Moody's Analytics:

- Credit ratings and research
- RiskCalc (PD models for private firms)
- CreditLens (IFRS 9 ECL)
- Economic scenarios and forecasting

### TRACE (FINRA):

- US corporate bond trade reporting
- Post-trade transparency (15-minute delay)
- Free access via FINRA website

## Equity and Indices:

### Morningstar:

- Mutual fund and ETF data (leader)
- Morningstar Direct (institutional platform)
- Equity research and analyst ratings
- ESG and sustainability ratings

### MSCI:

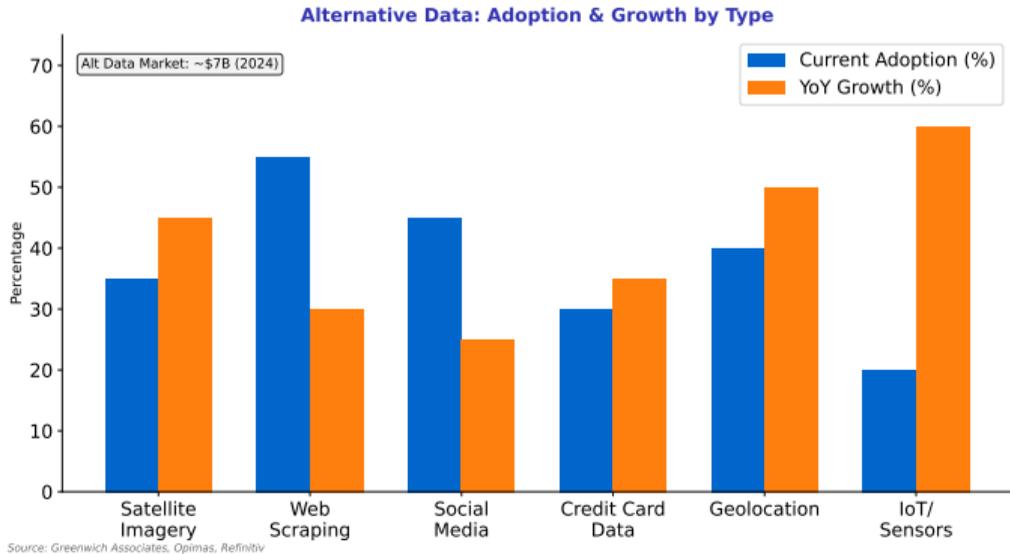
- Global equity indices (MSCI World, EM)
- Factor indices (value, momentum, quality)
- ESG ratings and climate metrics
- Barra risk models (factor-based)

### FTSE Russell (LSEG):

- FTSE 100, Russell 2000 indices
- Fixed income indices
- Smart beta and factor indices

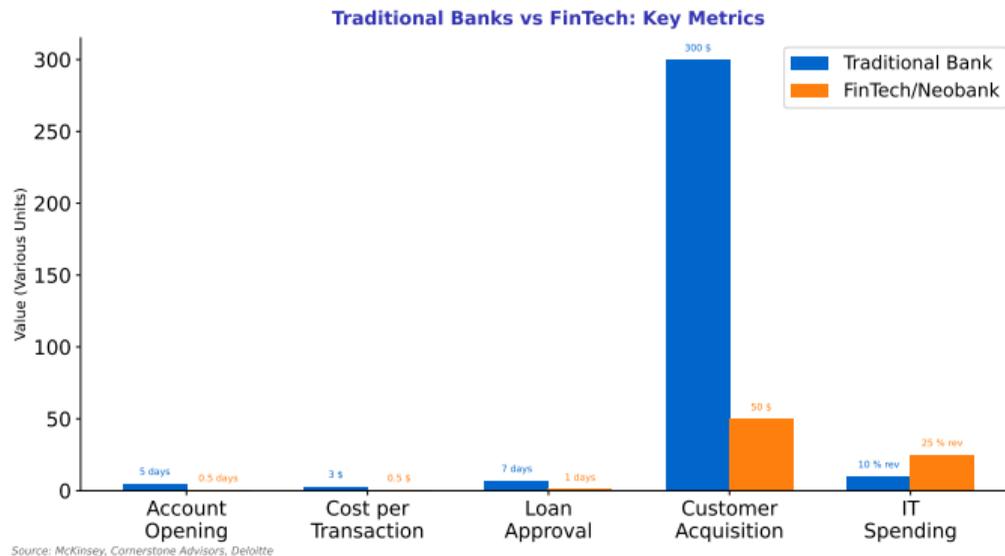
---

Quality data is the foundation for effective machine learning models. [Source: McKinsey, Gartner 2024]



Alternative data sources provide unique insights beyond traditional datasets.

# FinTech vs Traditional Data Providers



**FinTech data providers challenge incumbents with modern APIs and pricing.**

# Alternative Data Providers

## Alternative Data Categories:

### 1. Satellite and Geolocation:

- **Orbital Insight:** Satellite imagery (parking lots, oil tanks)
- **RS Metrics:** Construction activity, crop yields
- **Advan (SafeGraph):** Foot traffic to retail stores
- Use case: Predict retail sales, oil inventory

### 2. Credit Card and Transaction Data:

- **Earnest Analytics:** Consumer spending by category
- **Second Measure (Bloomberg):** Transaction panels
- **Facteus:** Anonymized credit/debit card data
- Use case: Estimate company revenues in real-time

### 3. Web Scraping and App Data:

- **Thinknum Alternative Data:** Company job postings, pricing changes
- **App Annie (data.ai):** Mobile app downloads and usage
- **SimilarWeb:** Website traffic and engagement

### 4. Sentiment and NLP:

- **RavenPack:** News sentiment analytics
- **Dataminr:** Real-time event detection
- **Brain Company:** AI-generated market insights
- Use case: Predict earnings surprises, market reactions

### 5. ESG and Climate:

- **Truvalue Labs (FactSet):** ESG sentiment from news
- **Sustainalytics (Morningstar):** ESG risk ratings
- **CDP (Carbon Disclosure Project):** Climate data

## Alternative Data Market:

- Market size: **\$2B+** (2024), growing 30%+ annually
- Hedge fund adoption: 80%+ use alt data (2024)
- Challenges: Data quality, survivorship bias, regulatory (PII)
- Pricing: **\$25k-500k+** per dataset annually

---

Quality data is the foundation for effective machine learning models. [Source: McKinsey, Gartner 2024]

# Exchange Data Licensing Models

## Professional vs Non-Professional:

### Non-Professional:

- Individual investors (not paid for trading)
- Lower fees: **\$1-5/month** per exchange
- Examples: Retail brokerage users, students

### Professional:

- Anyone compensated for securities-related work
- Higher fees: **\$50-200/month** per exchange
- Institutional traders, analysts, advisors

## Display vs Non-Display:

### Display:

- Human consumption (screens, reports)
- Priced per user per month

### Non-Display:

- Algorithmic trading, risk systems
- Enterprise-level pricing (often higher than display)
- Audit rights for exchanges (verify usage)

## Depth of Book Fees:

### Level 1 (Top of Book):

- Best bid/ask only
- Lowest fees: **\$5-20/month** (professional)

### Level 2 (Full Book):

- All price levels and depth
- Higher fees: **\$50-150/month**

### Level 3 (Order-by-Order):

- Individual order IDs and attribution
- Premium fees: **\$500-2000/month**
- Used for market making and order flow analysis

## Enterprise Fees (Non-Display):

- Flat fees: **\$50k-500k+/year** per exchange
- Query-based: **\$0.01-1.00** per query
- Device-based: Fee per server/instance
- Redistribution: Additional fees for reselling data

---

Quality data is the foundation for effective machine learning models. [Source: McKinsey, Gartner 2024]

# Vendor Data Redistribution Agreements

## Vendor Licensing Arrangements:

### Bloomberg, Refinitiv, FactSet:

- Bundle exchange fees into subscription
- Negotiate enterprise agreements with exchanges
- Pass-through fees or inclusive pricing
- Simplifies client billing (single invoice)

## Controlled vs Derived Data:

### Controlled (Raw):

- Real-time prices, order book data
- Subject to exchange fees and terms
- Cannot redistribute without exchange permission

### Derived (Analytics):

- Calculations based on raw data (indices, signals)
- More flexible licensing (vendor IP)
- Still may require exchange approval for some uses

## Data Licensing Challenges:

### Compliance and Audits:

- Exchanges audit vendor and end-user compliance
- User counts, usage types (display/non-display)
- Penalties for under-reporting: 3-5x back fees
- Automated monitoring tools (e.g., Vela, MDDM)

### Cost Optimization:

- Audit internal user entitlements
- Deactivate unused subscriptions (15-30% savings typical)
- Negotiate enterprise caps with vendors
- Shift to derived data where permissible

### Regulatory Data (Free/Cheap):

- **SEC EDGAR**: US public company filings
- **FINRA TRACE**: Corporate bond trades
- **Treasury.gov**: US government bonds
- **ESMA FIRDS**: EU financial instruments reference

---

Quality data is the foundation for effective machine learning models.

## Common Issues:

- Stale/missing data (illiquid securities)
- Incorrect reference data (ISINs, hierarchies)
- Survivorship bias (delisted removed)
- Point-in-time errors (look-ahead bias)

## Governance Framework:

- Data validation (outlier checks, cross-vendor)
- Data lineage and version control
- Master data management (golden sources)
- Vendor SLAs (99.9%+ accuracy)

Issue	Impact
Survivorship bias	Overstates backtest returns
Point-in-time errors	Look-ahead bias in models
Stale pricing	Wrong valuations

---

Data quality is critical for reliable financial analysis and backtesting.

## Enterprise Data Platforms:

### Snowflake Financial Services Data Cloud:

- Centralized data lake for financial data
- Pre-integrated market data (Bloomberg, Refinitiv)
- Data sharing and collaboration
- Used by hedge funds and asset managers

### Databricks:

- Unified analytics on lakehouse architecture
- Delta Lake for time-series data
- Machine learning and AI integration

### Cloudera / Hadoop:

- Big data storage and processing
- Legacy choice (declining vs cloud)

## Data Management Tools:

### Atoti (ActiveViam):

- In-memory OLAP for financial data
- Real-time risk aggregation
- Used by investment banks for trading risk

### Kdb+/Q (KX Systems):

- Time-series database (columnar)
- Extremely fast for tick data (microsecond queries)
- Used by HFT firms and quant funds
- Pricing: \$100k+ per server annually

### InfluxDB, TimescaleDB:

- Open-source time-series databases
- Lower cost alternatives to Kdb+
- Suitable for smaller-scale deployments

*Trend: Shift from on-premise to cloud data platforms (Snowflake, Databricks) for flexibility and cost*

---

**Quality data is the foundation for effective machine learning models.**

## Cloud Data Delivery:

### Vendor Cloud Platforms:

- Bloomberg Enterprise Access Point (BEAP)
- Refinitiv Data Platform (RDP)
- FactSet Cloud (AWS-hosted)

### Benefits:

- No on-premise infrastructure
- Elastic scaling (pay per use)
- Faster time-to-market
- Global availability and redundancy

### Consumption-Based Pricing:

- Pay per API call or data volume
- Shift from per-user to usage-based
- Better aligns cost with value
- Challenges: Unpredictable costs, audit complexity

## API Standardization:

### OpenFIGI (Bloomberg):

- Financial Instrument Global Identifier
- Free API for security identification
- Maps ISINs, CUSIPs, tickers to FIGI

### FINOS (Fintech Open Source):

- Open-source financial data standards
- Common Data Model (CDM) collaboration
- Interoperability across vendors

### Real-Time Streaming:

- WebSocket APIs for live data
- Kafka-based message buses
- Sub-second latency for cloud delivery
- Use case: Real-time dashboards, algos

*Major shift: Terminal-based → API-first data consumption (especially for quant firms)*

---

Source: Financial industry data and regulatory publications

# Summary: Financial Data Vendors

## Top Vendors:

- Bloomberg: \$12B, 350k terminals
- LSEG/Refinitiv: \$7B, FX/FI leader
- FactSet: \$2B, buy-side focus

## Alternative Data:

- \$2B market, 30%+ CAGR
- 80% hedge fund adoption

## Exchange Fees:

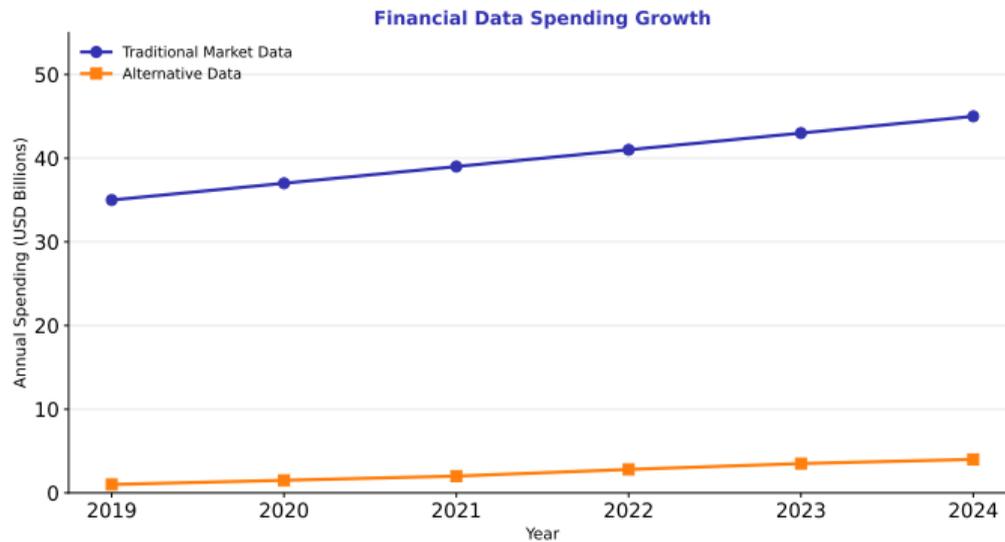
- Professional: \$50-200/mo/exchange
- Non-display: \$50k-500k+/year

## Trends:

- Cloud/API-first delivery
- Consumption-based pricing
- Time-series DB (Kdb+, Snowflake)

---

Data spending continues to grow as firms seek competitive advantage. [Source: Burton-Taylor 2024]



Source: Burton-Taylor, Opimas, FISD

**Data spending growing faster than technology budgets.**