



MARCH 2022

NEW EDGES: HOW TO MAKE THE MOST OF TRADING FLOW DATA

PREPARED FOR:



LEAPYEAR

TABLE OF CONTENTS

EXECUTIVE SUMMARY 2

INTRODUCTION..... 4

 METHODOLOGY 4

HOW TO LEVERAGE FLOW AND SENSITIVE DATA 6

USE CASES FOR SHARING FLOW AND SENSITIVE DATA..... 8

 AGENCY FLOW DATA FOR PRINCIPAL DESK..... 8

 PRIME BROKER AXE LIST 8

 INTERNALIZING THE FLOW 9

 MORTGAGE DATA FOR MUNICIPAL BONDS..... 9

 CLEARING LEADING TO TRADING STRATEGY10

 CORPORATE ACTIONS TO TRADING STRATEGY.....10

 INDUSTRY DATA AGGREGATION ACROSS UNITS11

 CROSS-ASSET CLASS TRADING DATA.....11

 CLIENT ONBOARDING12

 DEMAND FOR ANALYTICS13

AITE-NOVARICA GROUP'S ASSESSMENT 14

ABOUT AITE-NOVARICA GROUP16

 CONTACT16

 AUTHOR INFORMATION16

LIST OF FIGURES

FIGURE 1: SURVEY PARTICIPANT DEMOGRAPHICS..... 5

MARCH 2022

**NEW EDGES: HOW TO
MAKE THE MOST OF
TRADING FLOW DATA**

EXECUTIVE SUMMARY

New Edges: How to Make the Most of Trading Flow, commissioned by LeapYear and produced by Aite-Novarica Group, seeks to uncover use cases relevant to sharing sensitive data within financial institutions that generate new trading edges. The study focuses on areas where trading desks can create competitive differentiation and unlock fresh alpha. This study also explores how to share sensitive data while complying with regulations and protecting sensitive data.

The key takeaways from this study are as follows:

- **Share sensitive data safely in near-real-time:** Data's business value declines over time. Thus, sharing the data in near-real-time ensures data is complete for consumption and does not have to wait till end-of-day processing. Different data sharing techniques are available for firms to share sensitive data safely in near real time, and firms need to match the best tool and technique to the data and use case.
- **Increase internalization opportunities:** The area generating the most interest is how firms might leverage execution flow across business silos to increase trade internalization. Firms share flow data from external data providers easily, but they struggle to share it internally.
- **Generate new sources of alpha:** Trading desks are finding that sharing flow data provides additional data sources to help develop new trading edges and generate alpha as margins get tighter.
- **Explore established and emerging use cases:** The opportunities identified in the interviews performed for this paper extended well beyond the trading desk. Interviewed firms identified multiple use cases across business units for which sharing data can provide additional insight. Examples of areas firms have deployed or are exploring include enhancing municipal securities ratings with mortgage data, generating insights from clearing and corporate action to the trading desk, and sharing client data after onboarding.
- **Overcome the organizational, business, technology, and regulatory barriers:** Sophisticated technology solutions are available to solve compliance and regulatory constraints while sharing sensitive data. The biggest challenge is not the technology but achieving internal business alignment and compliance approval. Building use

cases and driving internal organizational approval are critical to successful data sharing initiatives.

- **Recognize that traditional methods no longer work:** Traditional data sharing techniques to ensure privacy, such as data masking, are not scalable. Interval-based flow data fails to provide all the data points that users can use to identify data irregularities and offer actionable data analysis. Effective flow data sharing needs a continuous stream of real-time data for analysis to be most valuable.
- **Leverage front-office data to manage risks:** The front office needs access to real-time flow data to react to changing market conditions and respond to an increase in requests for quotes (RFQs) in market-changing scenarios. According to a leading U.S. investment bank, financial institutions use statistics generated on an intraday basis to auto-adjust pricing and respond to market events. The objective is to attract more flow where required and reduce it when perceived as a risk.

INTRODUCTION

Data standardization in capital markets has created informational parity, making it difficult for firms to develop trading edges and generate alpha. Differentiation efforts have focused on tools, technology, and business processes around existing data sets and new data sources. Data within a firm is ideal, as it is unique to the firm and creates no external data license fees. Data classified as “most sensitive” has been challenging to share but often has the most value. Sharing “most sensitive” data across businesses can unlock alpha, create competitive differentiation, and reduce operational costs or risks.

Leveraging trade flow data to source new trading edges and alpha created the most interest and potential use cases from the firms interviewed for this study. Three main use cases to integrate flow data into the trading desk follow:

- Supplement commercially available flow data sources to identify trading strategies.
- Use order flow data to identify opportunities for internalization, riskless principal, or crossing of trades.
- Leverage post-trade execution volume and data across businesses to identify trends not apparent in other flow data sources.

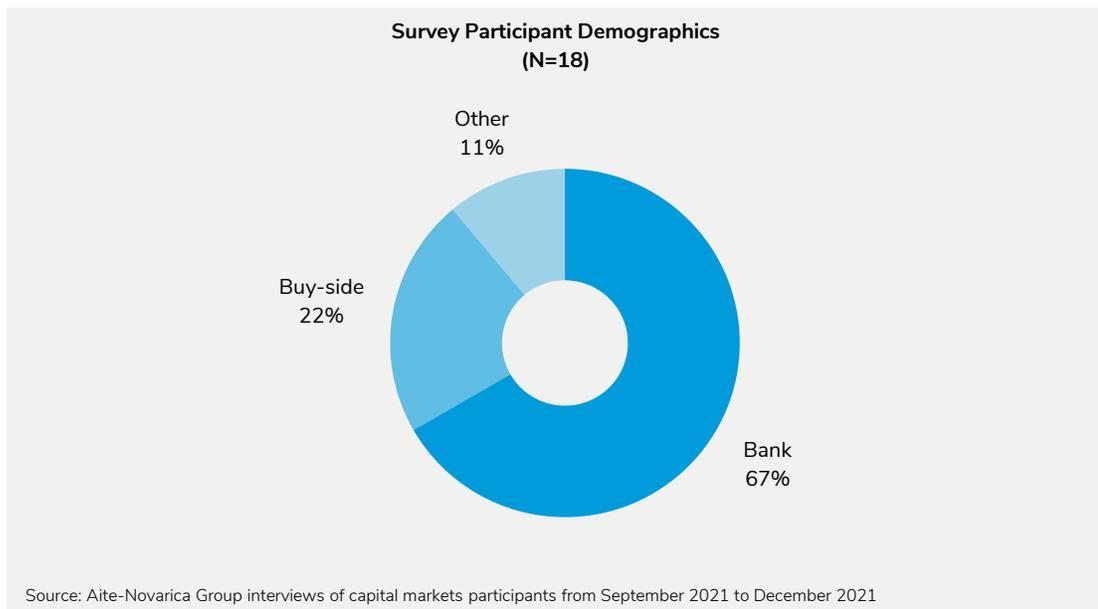
Business and regulatory requirements have built data silos, leading to data duplication issues. Data duplication generates additional operational costs and risks and makes it challenging to manage the client relationships. Firms report that these issues emerge with client onboarding and sharing documents across lines of business. Other opportunities to leverage data sharing include poking compliant holes in the informational walls that separate business units.

METHODOLOGY

LeapYear commissioned Aite-Novarica Group to survey market participants about use cases for sharing data inside a firm. Aite-Novarica Group conducted a qualitative survey of market participants from September 2021 to December 2021 to gather their views on sharing flow data within a firm. Participants in this study are primarily from front-office sell-side desks across asset classes (67%) and from the front office at asset management firms (22%; Figure 1).

Survey results demonstrate the demand for and challenges of data sharing and identify use case opportunities to share additional data across the enterprise. Lastly, Aite-Novarica Group considered the techniques and approaches firms use to manage sharing sensitive data across the enterprise.

FIGURE 1: SURVEY PARTICIPANT DEMOGRAPHICS



HOW TO LEVERAGE FLOW AND SENSITIVE DATA

This section considers how to leverage flow data across the enterprise. These considerations come from Aite-Novarica Group's discussions with market participants and internal research.

- **Understand your personally identifiable information (PII), sensitive PII, and contractually and regulatorily restricted data:** Sensitive data is the first set of data surrounded by regulations and contracts with a firm, as sharing it offers the greatest business integration value. The challenge is that hundreds of laws, regulations, and client contracts guard it. Identifying those data elements, the value they might present to other businesses, and the constraints under which the organization can share them is the right starting place. Getting compliance and business approval and support is step two. Only then should firms consider how to apply technology to share transactions related to that data.
- **Determine whether to use real-time, near real-time, or batch data:** Understand the time value of sharing sensitive data. Not all use cases require real-time data sharing. Real-time data is typically most valuable for flow data. Sharing flow data over discrete time intervals, such as the end of the day or weekly, would not provide all the data points useful for data analysis nor capture any leading irregularity indicators.
- **Choose data sharing techniques based on goals:** Not all data sharing techniques are equal. It is not easy to identify or tag data and ensure the integrity of confidential metadata tags across business lines due to the large volume of data, disparate systems, and legacy processes. Multiple techniques for data sharing are available that maintain data confidentiality, and the right one depends on the data, regulatory constraints, and business goals. Some of the key techniques for data sharing appear below:
 - **Differential privacy:** Differential privacy is a mathematical standard for privacy. An analysis is differentially private if one cannot tell by looking at the output whether the original data set included an individual's data. Differential privacy offers other benefits, such as defense against future privacy attacks and a privacy budget to measure how much privacy is lost over multiple queries.

- **K-anonymity:** This term refers to combining data sets with similar attributes. This technique obscures any identifying information about individuals contributing to that data.
- **Data masking:** This term refers to substituting the data element at a transactional level with other random characters.
- **Pseudonymization:** Like data masking, this technique substitutes the data element at the transaction level with other characters of the same data type.
- **Synthetic data:** This technique uses machine learning to analyze real data sets and generate fake ones that mimic their characteristics and diversity.
- **Multiparty computation:** This technique uses cryptography to help parties with limited trust share and analyze data without revealing the underlying data set.
- **Protect sensitive data from attackers:** Strict controls and security protocols must support data sharing needs to minimize the risk of privacy leaks and regulatory action. Attackers can combine two data sets to identify information about the raw data. According to a leading European investment bank operating in the U.S., new initiatives to share flow data are scrutinized by top management and compliance to ensure no confidential data is shared. Many firms have limited compliance bandwidth and may be unable to evaluate whether sensitive data is already shared or undertake new initiatives to share sensitive data. Unfortunately, most firms default to a more conservative approach to data sharing, with additional compliance scrutiny instead of reviewing processes and deploying technology to ensure data protection.

USE CASES FOR SHARING FLOW AND SENSITIVE DATA

This section of the paper highlights the use cases that Aite-Novarica Group's discussions with market participants revealed. Each use case outlines how sharing flow data can create alpha or enhance operational efficiency.

AGENCY FLOW DATA FOR PRINCIPAL DESK

The principal desk can obtain business value from the client execution flow data to the agency desk. A real-time view into the transaction details of the client executions (e.g., security traded, place of venue, order types) provides insights to help adjust principal trading strategies.

Insights:

- The agency desk sees customers' trading indications while trying to find a counterparty willing to buy or sell the security for the same price. At the same time, the principal desks are always looking for information to increase their trading profitability.
- Aggregate security-level buys and sells of securities or sectors from the agency desk could indicate to the principal desk whether it should accumulate or sell off individual positions.
- The principle of sharing flow data does not compromise the agency desk, which offers best-execution services to its client. The principal desk can offer or match the best price by using the agency flow data.

PRIME BROKER AXE LIST

The ability to integrate, aggregate, and disseminate all the inventory data across all clients and principals is significant to prime brokers. Prime brokers publish this list of axe and inventory for their clients.

Insights:

- Aggregating all the proposed levels at which the prime broker can trade attracts more client orders. The challenge is that the prime broker constructs the axe list from all the prospective levels at which a client can execute the order through the prime broker. Thus, prime brokers must balance their desire for the longest axe list, which

leads to an opportunity for revenue maximization, with protecting the confidentiality of their clients' intent to trade.

- Organizations can build the axe list with technology solutions offering controls that can aggregate all the clients' data, apply data access controls, and optimize the output for accuracy. Market participants share that such tools are widely used in the prime brokerage industry. Participants certainly demand the ability to aggregate the list across asset classes.
- Additionally, the compliance aspect involves the client data since the data falls within the same business division. According to one leading investment bank, more industry tools aggregate the information across sell-side trading desks and provide a consolidated view to buy-side firms.

INTERNALIZING THE FLOW

Firms are looking for opportunities to internalize the order, avoid paying bid-ask spreads, and reduce the risk and the operational cost of processing trades. Internalization allows firms to avoid disclosing liquidity to outsiders for competitors to benefit.

Insights:

- According to a U.S.-based investment bank, firms easily consume flow data from external data providers. Still, such arrangements do not allow firms to differentiate their data analyses; other participants can also consume the data from the external provider.
- Aggregation of data on quotes and executions across trading desks and business units would highlight opportunities to internalize trades.

MORTGAGE DATA FOR MUNICIPAL BONDS

Firms offering their clients mortgage services and trading capabilities in municipal bonds can redesign their municipal bonds trading strategies based on the demand for a mortgage within a specific municipal area.

Insights:

- The municipal trading desk could adjust its trading strategy based on the insights that it can gain on the mortgages issued in a specific ZIP code. Market participants

believe this ability is only beneficial for firms offering mortgage and municipal trading desk services.

- Consumers buying properties via mortgages in a specific ZIP code will indicate more tax revenue for the town municipal. The increased tax revenue would suggest a higher credit rating to securities that municipal corporation issues. At the same time, a default by multiple mortgage borrowers in a specific ZIP code could provide an early alert to the mortgage trading desk to adjust its trading strategies.

CLEARING LEADING TO TRADING STRATEGY

Firms offering execution and clearing services often clear trades they did not execute. A real-time view of the trades for which the firm is providing only clearing services to the execution desk could provide additional flow data, enabling them to adjust their trading strategies.

Insights:

- Market participants agree that they offer clearing services to clients under agreements that prevent them from sharing client-specific, clearing-services-related data with other entities within the firm.
- The trading desk can get additional flow data from the trades taken up for clearing. Similarly, options trading flow via the clearing agency indicates direction and volume regarding the underlying security to the desk trading.

CORPORATE ACTIONS TO TRADING STRATEGY

Insight into the clients' corporate actions elections for securities or cash is valuable to the trading desk. It provides trends on the future availability of securities or liquidity in the market.

Insights:

- Corporate action processing is a complex process involving the securities holder's decision by a specific date. Sometimes the decision concerns how security holders receive the benefits from a corporate action—as cash or securities.
- One market participant from a Canadian investment bank shared that his team pitched an idea whereby client elections on corporate actions could be used to

identify the impact on the stock price. The firm appreciated the idea of deriving value from corporate action data, but compliance could not agree to the data sharing arrangements.

- A research participant mentioned that the trading activity that vendor solutions offering optimization services trigger in corporate actions could also provide insights into the election securities holders make to receive corporate action.

INDUSTRY DATA AGGREGATION ACROSS UNITS

Internally aggregating data from market aggregators can be valuable to firms without compromising client confidentiality. Some of the firm's business divisions contribute trade details to market aggregators and receive aggregated views to analyze market depth. This can often be desk specific without subsequent aggregation across trading desks.

Insights:

- According to one market participant, the contributing division analyzes and consumes most of the insights and does not share them across the firm. Each division finds its own suitable market aggregator.
- Another market participant shared that it aggregates trading data and shares it anonymously with liquidity-takers. Liquidity-takers get market insight on client flow executed with other liquidity-takers but do not share the details with other internal groups.
- Some asset classes (e.g., stock exchanges for equity stocks) have no principal place that can provide real-time market depth for assets such as over-the-counter derivatives or cryptocurrency.

CROSS-ASSET CLASS TRADING DATA

Firms often organize trading desks by asset class; few desks have insight into others' trading activities. Opportunities exist for data aggregation services to bring insights by combining flow data across asset classes.

Insights:

- According to a U.S.-based investment bank, a corporate bond trading desk dealing in fixed income securities from an issuer could benefit from data on the trading activity from the equity desk. Similarly, the securities lending/borrowing desk could get more insight into a security if redemptions of the exchange-traded fund include that security.
- Industry initiatives for real-time reporting of over-the-counter derivatives trading activity under the Dodd-Frank Act in the U.S., the revised Markets in Financial Instruments Directive (MiFID II) in Europe, and similar regulations across the globe have limited firms' ability to consume real-time published data. This is predominantly because of the lack of high-quality data firms are reporting and because, under Dodd-Frank, the swap data repository does not always publish the traded quantity.
- The flow data under cryptocurrency trading is even more segregated across multiple venues and off venues. The order flow for the spot market, where each broker and bank have different liquidity, is similar.

CLIENT ONBOARDING

Organizational and technology silos make it difficult for firms establishing client relationships to share data across organizational silos. Consider a fund administrator onboarding a hedge fund client with distribution, administration, and money management occurring in three countries. Sharing something as simple as the results of a Know Your Customer/anti-money laundering check on the firm and its principals becomes a complex endeavor.

Insights:

- According to a U.S.-based investment bank, a centralized client team onboards customers across business lines. It then manages client data sharing across multiple business users. It processes each request for client data via multiple layers of approval across management and uses data masking and filtering techniques to share the data. It does not share much in real time except a general message when a change is updated in the client record.

DEMAND FOR ANALYTICS

The increase in trading volume across various markets has increased the need for data analytics to gain insight. Analytics beget insights, and insights, in turn, create distinguished parameters and lead to alpha generation. Firms have adopted advanced technology for non sensitive data but struggle to skim the sensitive data analytically.

Insights:

- According to a leading investment bank, financial institutions demand relevant data analysis due to the large volume of data they manage. However, identifying relevant data for analysis is challenging. For example, data from trade negotiation with manual intervention and auto-generated data with configured parameters sometimes look the same. Therefore, the investment bank added new indicators to compare the data set from manual trade negotiation with trades executed via an automated process.
- Buy-side portfolio managers hire data scientists to generate opportunities to discover alpha, but data scientists are spending considerable time on data collection. Data scientists need constantly available data—in real time, intraday, and at the end of the day.
- The value of data sourced on a real-time basis, such as the depth of the trading order book, can sometimes erode quickly. Heavy reliance on the end-of-the-day data can delay business decisions.

AITE-NOVARICA GROUP'S ASSESSMENT

This section details Aite-Novarica Group's assessment of flow data.

- **Create a new edge with flow data:** Margins are getting tighter due to electronic trading and data availability, but trading desks are under pressure to look for additional data sources to develop more sophisticated trading strategies and generate alpha. Data-first approaches are driving firms' business strategies, given the growing amount of data in the financial industry. The basis of a data-first approach is in-depth insight into underlying transactions. Sharing flow data provides firms with opportunities to derive more business value, as it does not result in external data license fees.
- **Explore opportunities across business use cases and units:** Based on Aite-Novarica Group's discussions with market participants, firms can explore flow data sharing across multiple use cases: from front- to back-office functions and across business units and asset classes. Real-time flow data sharing provides the required edge to respond quickly to market variation and generate the desired alpha return.
- **Choose high-value data sets:** The value of the data set should be greater than the cost to share the data. Data aggregation and sharing should have more value than the cost involved. There is real value to flow data shared on a real-time basis. Sharing flow data should generate trading alpha. Identifying the high-value data sets around trading volatility and order flow requires significant data analysis and resources.
- **Consider the value of internal flow data over external flow data:** Firms have easily adopted flow data sharing from external data providers, as it transfers the risk to the data provider. However, external data sources do not provide firms opportunities to differentiate their data analyses. Thus, firms are exploring opportunities to close the gaps and address trades that they could internalize but that get executed away due to a lack of information.
- **Track the market:** The front office needs access to real-time flow data to analyze rapidly changing market conditions and respond to a significant increase in RFQs in market-changing scenarios. According to a leading U.S. investment bank, the front office uses statistics generated on an intraday basis to auto-adjust pricing and respond to market events and remain competitive. The objective is to attract more

flow where required and reduce the flow where it is perceived as a risk. According to another investment bank, gaining access to data is a challenge and consumes substantial time and resources.

- **Solve for compliance and profit:** Sophisticated technology solutions are available to address the compliance requirement that firms refrain from sharing client-specific data. Firms that do not share flow data due to organizational structure issues can benefit tremendously from data sharing. A careful analysis of the data technique ensures that data privacy is uncompromised, allowing firms to take advantage of the potential use cases for sharing flow data.
- **Use technology tools to address compliance concerns:** Firms deploy multiple layers of control checks around confidential data to preserve it from unauthorized access. On the other side, business insights are hidden in the transaction data. Market participants have adopted technology solutions that enable real-time flow data sharing to generate revenue and meet the need for business insights and compliance. The traditional techniques of ensuring privacy with data sharing, such as data masking, are not scalable, especially during market volatility. Firms should maintain a scalable solution adopted for flow data sharing to derive continuous value.
- **Blend internal and external flow data:** Combine the internal flow data analysis with external flow trading data analysis to capture market sentiments and identify the reaction entry and exit points. Internal data analysis will identify how the flow data can be most valuable to the firm and prepare the firm's strategy to react to the market.

ABOUT AITE-NOVARICA GROUP

Aite-Novarica Group is an advisory firm providing mission-critical insights on technology, regulations, strategy, and operations to hundreds of banks, insurers, payments providers, and investment firms—as well as the technology and service providers that support them. Comprising former senior technology, strategy, and operations executives as well as experienced researchers and consultants, our experts provide actionable advice to our client base, leveraging deep insights developed via our extensive network of clients and other industry contacts.

CONTACT

Research and consulting services:

Aite-Novarica Group Sales
+1.617.338.6050
sales@aite-novarica.com

Press and conference inquiries:

Aite-Novarica Group PR
+1.617.398.5048
pr@aite-novarica.com

For all other inquiries, contact:

info@aite-novarica.com

Global Headquarters:

280 Summer Street, 6th Floor
Boston, MA 02210
www.aite-novarica.com

AUTHOR INFORMATION

Vinod Jain
+1.908.224.7770
vjain@aite-novarica.com