

Why big data matters: using smart data and artificial intelligence to enhance FX trading.

The digital revolution has unleashed a wave of innovation in smart data and artificial intelligence (AI) that is greatly transforming the financial services industry. These technologies have the potential to deliver deep insights and improved operational efficiency to businesses. By 2035, global financial services are expected to see the third highest annual gross value-add driven by AI of any industry sector (4.3%).¹

Smart data and AI technologies have been in existence for close to six decades. Over the last ten years there has been a sharp fall in the cost of data storage, better processing capabilities and computing power, and more sophisticated analytics.

Recent regulation in financial services on data standards, specifically with respect to intraday liquidity management, risk reporting and data aggregation, have served as a catalyst for digital transformation, as financial services revise their strategies on how to respond to regulatory change. These developments have all contributed to making AI and smart data more in demand.

¹ Artificial Intelligence Will Enable 38% Profit Gains By 2035
Forbes, June 22, 2017

“It’s preferable to input higher-quality data into smart data and AI tools.”

Due to the fragmented nature of the FX market, latency and market moving events, market participants have lagged behind their peers in other asset classes in realizing the benefits. Both technologies have remained largely confined to niche applications because of cost and complexity, and a lack of sufficient or quality data.² Thanks to a rapid expansion in the availability of FX market data – this is now changing.

Data defined

In this document, we refer to ‘big data’ as significantly large sets of both unstructured and structured data that is high volume, high velocity and high variety. We use ‘smart data’ to refer to data sets from which signals and patterns have been extracted by intelligent algorithms.

We use AI to refer to the simulation of human intelligence processes by machines depending on objective parameters, and refer to machine learning (ML) as a subset of AI in which machines, given access to data, are able to learn for themselves.

The importance of big data

It’s preferable to input longer-term and higher-quality data into smart data and AI tools. For a complex, hybrid market such as FX, where liquidity is fragmented across multiple venues and there is no industry-wide tape, obtaining comprehensive FX trade data is difficult. Even the largest banks and trading venues see only a fraction of global FX flows. As a result, many financial institutions are on the quest for more representative trade data for better interpretation of activity in FX markets.

Source	Coverage**	Frequency
BIS triennial survey	USD5.1 trillion	3 years
ECNs*	USD300 billion	Daily
CME and other futures exchanges	USD100 billion	Intraday
FX brokerages and trading platforms	USD20 billion	Intraday
Investment banks and FX dealers	USD10 billion	Daily
CLS traded data	USD1.5 trillion	Intraday

Source: Quandl

* Thomson Reuters and ICAP

**Implies amount of activity by data publisher used to extrapolate estimate FX market metrics such as volume and order flow

With access to over 50% of global FX settlement volumes for 18 currencies and FX instruments (including swap, spot and outright forward) across 33 currency pairs, CLS has an unparalleled view of trading activity across the FX market. Since its inception in 2002, it has warehoused the details of every transaction to create the largest single source of executed FX trade data.

This data is now available to FX market participants – enabling them to enhance their trading strategies through smart data and AI technologies. CLS is also leveraging these technologies internally to support the systemically important payment-versus-payment infrastructure relied upon by the FX market globally to mitigate settlement risk – the largest risk in foreign exchange.

² Artificial Intelligence in FX: Harnessing the future now, eForex magazine, June 2017



Key driving forces are increased regulation and trade transparency requirements and the rapid expansion of available FX data and analytics. ”

Using smart data and AI to enhance FX trading strategies

Currency rates are determined by multiple factors, including macroeconomic events, geo-political developments, and sometimes simply the momentum of FX flows. All of these factors can lead to unexpected and short-term volatility. Smart data and AI tools can be used to analyze the enormous amounts of data produced in the currency market to help traders to spot patterns and correlations despite this volatility. FX trading systems can then be programmed to carry out user-defined algorithms, characterized by a set of rules based on parameters such as pricing, volume, and liquidity, to structure the trades that will be executed.

The uptake of smart data and AI trading has been most notable among asset managers and hedge funds, who have been excited by the prospect of high returns from algorithm-centric strategies. These include quant funds, which are characterized by their reliance on systematic quantitative investment strategies and, increasingly, machine-learning systems. Assets under management in these funds are soon expected to top USD1 trillion – nearly double their volume in 2010.³

There are two stages of the trade lifecycle where AI and smart data technologies have the potential to enhance trading strategies. One is pre-trade analysis, in which participants are looking to understand when the market is most liquid, who's buying or selling, and the flow of volumes. The other is historical data which can feed into execution algorithms and trading models for scenario testing and cost analysis to ensure best execution.

The value of volume data in FX trading

A 2017 academic study by the University of Melbourne and Cass University based on CLS data has identified correlations between volume and price. The findings indicate that “FX volume embeds predictive information for both the time-series and cross-section of currency returns”. In short: if prices go up on low volumes, then they are likely to reverse. If prices go up on high volumes, then they are likely to sustain.⁴ This predictive information, as evidenced in the study, could therefore improve trading decisions, timing to market as well as profitability.

By applying smart data driven tools to the FX spot, swap and outright forward volumes submitted to its settlement service, CLS aggregates and segregates total spot volume data for the market on an intraday hourly basis. Traders can input this data into algorithmic or non-algorithmic trading tools to detect potential price movements and depth of liquidity to determine the best time to trade. They can also use ML trading algorithms which ‘learn’ from previous success, to exploit volume patterns and generate investment gains.

Furthermore, ML models have been shown to reduce price slippage significantly. A smart algorithm might seek to reduce slippage by executing trades in different ways or at different times, such as selling the order in clips over a given period. Having a comprehensive view of volumes in the market, therefore, can improve overall trade decisions and timing to access the market.

Leveraging volume-derived data sets and AI

Accurate volume data can empower traders with the insight to generate substantial value. Smart data and AI-driven derivations of CLS's volume data can be used to support the market.

³ Quant hedge funds set to surpass \$1tn management mark, Financial Times, January 2018

⁴ The Value of Volume in Foreign Exchange, The University of Melbourne, and Cass Business School and the Centre for Economic Policy and Research, 2017.

Managing the flow – using order flow data for more informed trade decisions

Order flow can be an important mechanism for both dealers and individual FX traders to track the flow and volume of trades made by banks and institutions, and to detect or generate trading signals. A lack of transparency in the FX market means that it has historically been difficult for market participants to gauge the relative predominance of informed and uninformed traders in any given currency price movement. Leveraging order flow data can reveal market participation, giving traders valuable insights on underlying market dynamics and support for alpha generation analysis.

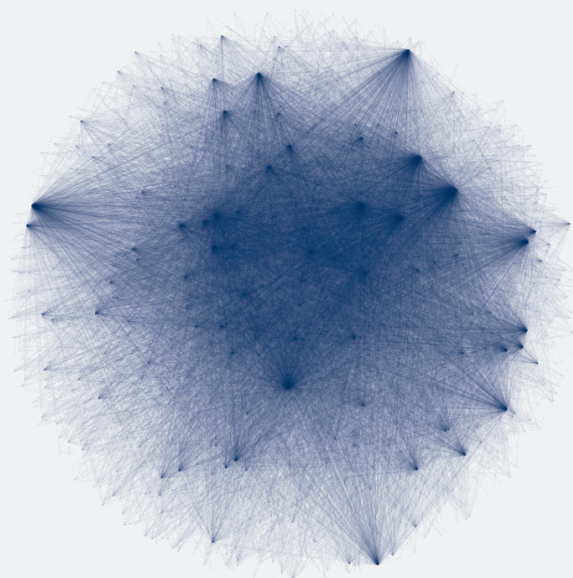
CLS took a derivation of its spot volume data and applied smart data to identify FX price takers and market makers based on a statistical analysis of network data. The nodes were then separated into two groups using the concept of ‘coreness’. Coreness is a measure that can help identify tightly interlinked groups within a network. The FX price makers are assumed to be those nodes that maintain a consistently high coreness over time.

In following this process, CLS was able to generate a flow data set showing the directional volume by counterparty type, market maker versus price taker and/or net directional volume for all parties. The data can be delivered at varying frequencies – intraday, daily and monthly.

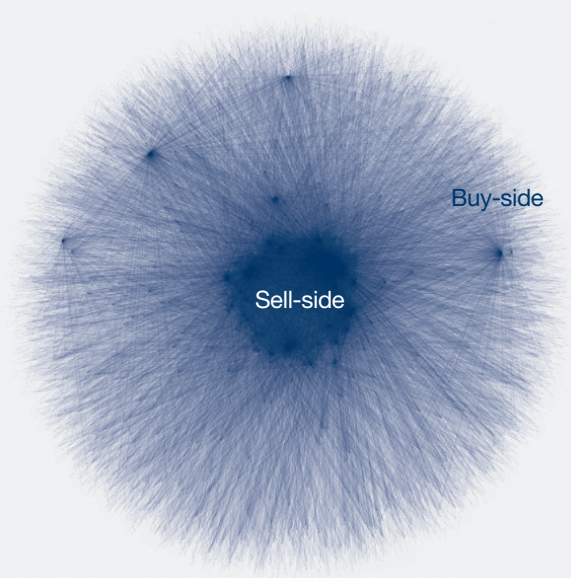
Network data

The two diagrams below represent network data. In each diagram, the nodes represent parties to a trade. One or more trades between the same pair of parties are represented by a single link. The diagrams illustrate how nodes in a randomly initialized network (left) are re-positioned by coreness into two layers (right).

Those with a high coreness are deemed market makers. All other parties are included in the second group, the price takers. Results show FX price makers concentrated in the center and price takers positioned at periphery.



A randomly initialized network of trade parties into two layers



Parties re-positioned by coreness



Applying time-series analysis in modelling and forecasting

CLS internal case study

Processing an average of 500,000 FX trades every day since 2002, worth over USD1.55 trillion, CLS needed a way to analyze its own data and the volume data available for the currency markets to determine if its settlement system – CLS Settlement – would have sufficient capacity to handle any sudden high-volume surges in the market. Turning to smart data and AI, it developed an internal volume forecast model leveraging its deep and rich data history to enable scenario analysis and stress testing against known or scheduled events, intraday, intraweek and monthly patterns.

While AI technologies cannot predict market-moving news, they can have a degree of success in quantifying how big an impact a significant market-moving event can have. The sudden decision by the Swiss National Bank (SNB) to unpeg the Swiss franc from the euro on 15 January 2015 (which resulted in an almost immediate 15% appreciation in value) is a case in point. The action by the SNB took financial markets largely by surprise, but CLS's capacity volume forecast model enabled the organization to assess the impact the event might have on subsequent volume flows ensuring the firm responded to and managed the increase in transactions quickly and efficiently.

This foresight enables CLS to predict changing demand of its processing capacity and protect against the impact of market moving events. Using the model, CLS's operations team was able to anticipate the resultant settlement volumes on 20 January 2015 and activate its additional reserve capacity to support peak trading settlement volumes. CLS settled a record 2.26 million transactions on 20 January 2015 totalling USD9.2 trillion – 99.5% of which were settled within 55 minutes.

CLS realized that having an accurate volume forecast in the market has significant relevance for traders. An accurate volume forecast provides traders with an enhanced view of potential market trading activity scenarios which allows for improved risk management, reduced market impact and minimized price slippage.

The CLS forecast model is trained against anticipated events that are expected ahead of time, such as national holidays and scheduled economic events, such as inflation, GDP and interest rate announcements by central banks. Transaction data for all historical peaks in volume CLS has witnessed is also fed into the model. With this large quantity of data, the model is able to capture complex, non-linear relationships between factors such as currency calendars and macro-economic events impacting the market.



Accurate volume data can empower traders with the insight to generate substantial value. ”

Using forecast data to determine available liquidity and optimize execution

Corporate treasurers, asset managers and other market participants are sensitive to friction in trade entry or exit, and are always looking for ways to improve best execution. One way to do this is to trade when liquidity is plentiful, however, structural changes in the FX market largely driven by regulation and best practice, have created an ever-growing challenge of operating in a changing and increasingly fragmented landscape. As a consequence, market participants are on a constant search for liquidity. While a greater diversity of prices is now available thanks to an increase in venues and participants, the execution environment has become more complex to navigate.⁵

Traders can face allocation and execution challenges when trading in less liquid currencies or trying to maximize alpha capture without creating a negative price impact. The advancement of big data, smart data and AI technologies is helping to automate more elements of liquidity management, leading to a better view of trading capacity and better risk management.

It is critical for FX traders to be able to source accurate trade data to feed into cash flows and collateral positions, in order to build a real-time data structure to identify trading and liquidity risks. While liquidity conditions generally show some consistency at certain times of day, they can differ considerably in response to events, such as data releases, central bank decisions and elections. AI systems can perform analyses of volume surges in the market, taking into account market impacts of past and future scheduled economic events, to transform historical insight into foresight.

FX forecast data provides a forward-looking view of FX markets over short, medium and long-term horizons to identify times to trade with greater liquidity. Traders can use this information to optimize strategies in liquid and illiquid pairs. Incorporating volume forecasts in either algorithmic or non-algorithmic trading strategies can reduce execution costs and diminish price slippage and market impact.

Conclusion

Smart data and AI technologies are advancing at an unprecedented pace. Combined with exponential increases in computer processing power and increased storage enabled by cloud computing, the key driving forces are the rapid expansion in the availability of data and regulation. CLS is utilizing this data to support a reliably robust infrastructure at the center of the FX market. As evidenced in this document, it is also harnessing its big data to support market need around reliable trade data at a higher frequency for better visibility into FX trade activity and trade decision support.

Applying AI and smart data technologies to big data makes it possible for FX traders to gain insights into and subsequently automate trading strategies depending on market volumes and pricing, order flow and liquidity. CLS is currently exploring ways to apply unsupervised ML to its order flow data to classify market participants and their trading patterns which will further enhance its data offering. FX participants can therefore improve trading outcomes, realize operational efficiencies and enhance FX liquidity and risk management.

⁵ FX market fragmentation – addressing the liquidity challenge, FXMM May 2017



CLSMarketData

	FX Volume	FX Flow	FX Price	FX Forecast
Description	Transaction volumes for 33 currency pairs representing 50% of global FX activity	Transaction volumes aggregated by type of market participant and side of trade taken	Transaction prices weighted by volume (VWAP) and by time (TWAP)	Transaction volumes projected hourly for the week ahead
History	From 2011	From 2012	From 2015	From 2018
Granularity	Hourly aggregates	Hourly aggregates	5-minute aggregates	120 forecast hours
Delivery frequency	Availability in hourly or end-of-day variants			Daily
Reporting lag	Within 30 minutes of the end of each hour			N/A
Coverage	33 FX pairs covering 50% of global FX volume			8 major currency pairs
Data source	Direct from executed trade confirmations received by CLS for settlement			Direct from executed trade confirmations received by CLS for settlement
Currencies included	AUD, CAD, CHF, DKK, EUR, GBP, HKD, HUF, ILS, JPY, KRW, MXN, NOK, NZD, SEK, SGD, USD, ZAR			AUD/USD; EUR/GBP; EUR/JPY; EUR/USD; USD/CAD; USD/CHF; USD/JPY; GBP/USD
Trade types	Spot, swap, forward	Spot	Spot	Spot
Key fields	Time, currency pair, trade count, trade volume (USD)	Time, buyer & seller type, currency, volume, count	Time, currency pair, VWAP, TWAP	Currency pair, product, forecast data, hour, volume

