Technology at the service of market oversight

Iris Lucas,
Autorité des Marchés Financiers (AMF)

[special issue of *Réalités Industrielles*, February 2019]

Abstract:
Several changes during the past few years have left marks on the financial markets: fragmentation, digitization, algorithmic high-frequency trading. Concomitant with a considerable increase in the volume of market data, these phenomena have pushed regulatory authorities to overhaul and modernize their systems for detecting violations. Technology, undeniably now at the core of market oversight, is still an instrument requiring human expertise, at all levels ranging from the design of algorithms to the analysis of potential violations.

In 2007,[1] the liberalization of the monopoly that historical stock markets had over transactions in securities led to the emergence of alternative trading platforms. To offer ever more competitive bids on financial instruments, these platforms decreased "tick sizes" (*i.e.*, the minimal increment between two proposed prices in their order books) and thus increased the price spread.[2]

This splintering of the market among various platforms and the increased granularity of possible prices have multiplied the number of "messages" that strategies for trading, executing and routing orders have to take into account.[3] These trends push toward ever more automation and have underlain the development of algorithmic (in particular high-frequency) trading over the past ten years. Given this context, regulatory authorities have had to adapt to perform their assignment of overseeing the market.

Figure 1: Average monthly volume of messages, orders and transactions (data collected by the AMF during the third quarter 2018): 280 million messages, 130 million orders; 10 million transactions.
Source: AMF.

En moyenne par mois sur les valeurs du CAC40 sur Euronext Paris

[1] The first EU directive on markets in financial instruments (MIFID I), which came into effect on 1 November 2007, authorized investment services to fragment the flow of execution of their orders to buy and sell on different trading venues (creation of multilateral trading facilities and systematic internalizers).

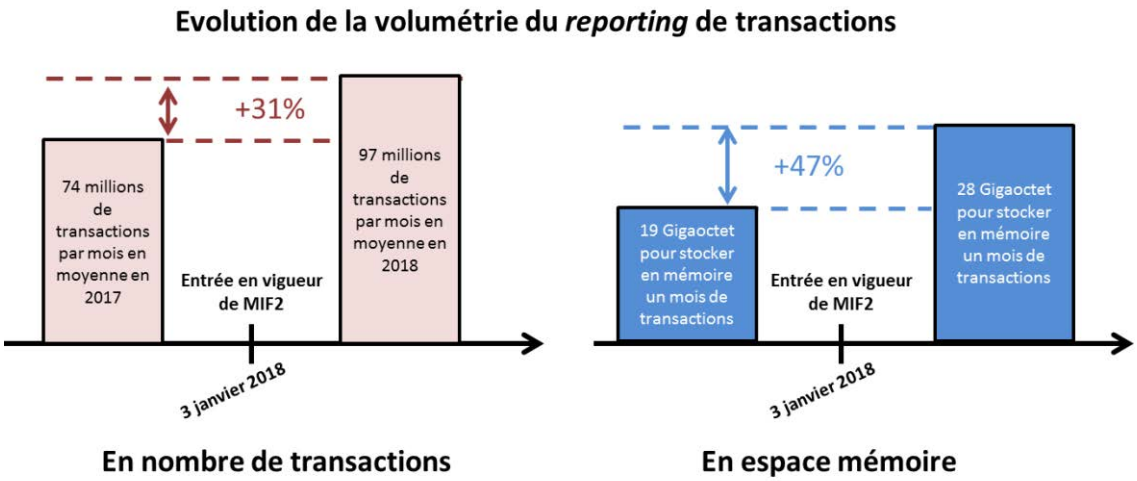
[2] This article has been translated from French by Noal Mellott (Omaha Beach, France). The translation into English has, with the editor's approval, completed a few references. All Web addresses have been consulted in May 2019.

[3] A "message" is any event recorded in an order book: the passing of a new order or the modification, execution or cancellation of an existing one. An order undergoes at least two events: an entry on the order book when it is passed and an entry for its execution, cancellation or expiration. A transaction corresponds to the execution of two orders (to buy and to sell).

To do this assignment, the AMF (Autorité des Marchés Financiers, the French regulatory authority of the stock market in particular) collects a large volume of market data from various sources. Besides public databases (data vendors, such as Thomson Reuters or Bloomberg), the AMF receives data on transactions from the stock markets and on orders (which, though not always executed, do influence prices) and from market intermediaries (the brokers subject to European regulations). Under the obligations imposed by MIFIR, the EU regulation accompanying MIFID II, the AMF receives from market intermediaries in France their reports of transactions, and from European regulatory authorities, the reports from foreign intermediaries when their activities involve French financial instruments. Likewise, the AMF reroutes the reports from French brokers on transactions in foreign securities toward the appropriate parties outside France.

Reporting requirements thus provide the AMF with a huge set of data on transactions in French securities and their derivatives (whether listed or not) — transactions on the principal market (Euronext Paris) or on alternative markets and platforms (including those located in the United Kingdom) as well as over the counter. Since 3 January 2018, when MIFID II came into effect, the volume of the data thus received has considerably grown. MIFIR, the associated regulation, has both extended the range of financial instruments subject to reporting requirements and enhanced the quality of information about end-clients and the persons (or algorithms) implicated in the chain for making decisions and executing transactions.

Figure 2: The impact of MIFID II on the number of transactions and on the memory needed to manage the reports collected from professional market intermediaries.
Source: AMF.



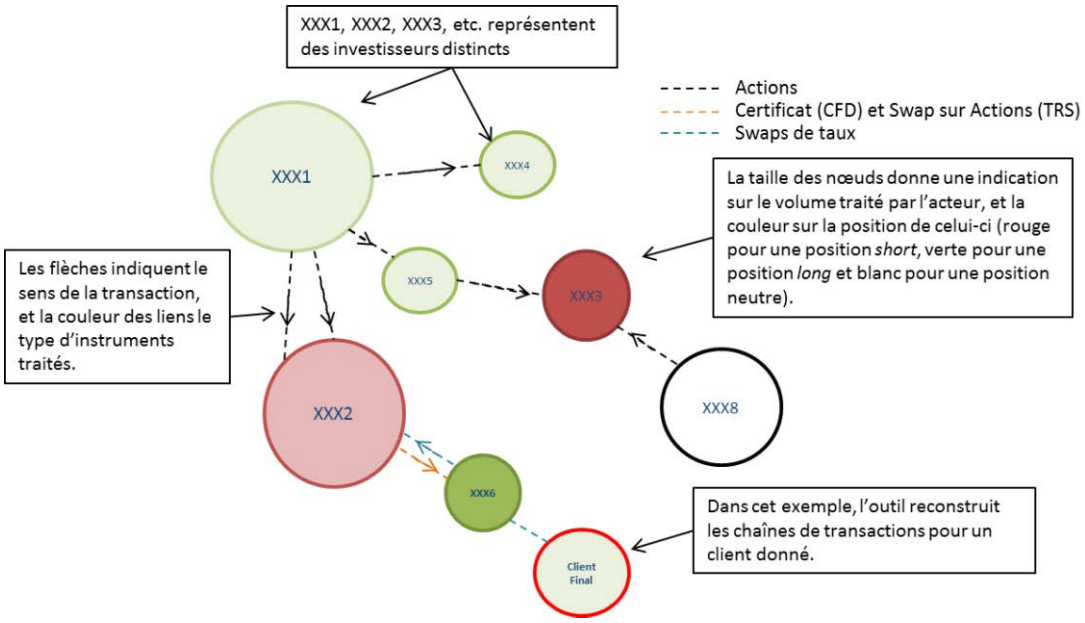
Market oversight entails processing the data reported

This huge set of data is a mine with information that the AMF uses to oversee the financial markets and, in particular, look for violations. The AMF has devised its own system for detecting violations that uses algorithms to analyze the collected market data and set off warnings to signal situations with a risk of misconduct or market abuse (under the EU's market abuse regulation, MAR). Just as structural changes in the markets have pushed many firms to automate their trading strategies for executing orders, the growing volume of data received by the AMF and the variety of potential market manipulations have led this authority to increasingly automate its tools of detection so that its regulatory platform remains state-of-the-art.

The algorithms developed for market oversight use various criteria to detect atypical sequences (abnormal volumes, sudden price variations, volatility, etc.) characteristic of market manipulations (such as layering[4] or insider trading). These algorithms rely on data-processing techniques (such as crossing data sources or performing iterative calculations or computations on very short time scales), which are costly in machine resources. To eventually turn up patterns of events that suggest layering for example, the state of an order book is recalculated after each event during the session. For a security on the CAC 40, there are now, on the average, 17 events per second, but this frequency can soar up to 2426.[3] To look for cases of insider trading, all of a trader's activities have to be brought together under consideration on spot markets (securities, bonds, etc.) and on markets for derivatives (options, futures, CFDs, equity swaps, etc.), any venue where orders are executed (standard stock exchanges, multilateral systems, internalizers, over-the-counter markets, etc.).

The reporting of order book events is often complicated, since an investor might have recourse to several intermediaries, who might, in turn, transfer the execution of a client's order to other intermediaries. The difficulty is to reconstitute the chain of intermediaries and thus identify the ultimate beneficiaries (instead of the intermediaries who simply transmitted the orders).

Figure 3: Example from a tool for visualizing chains of intermediaries.
Source: AMF.

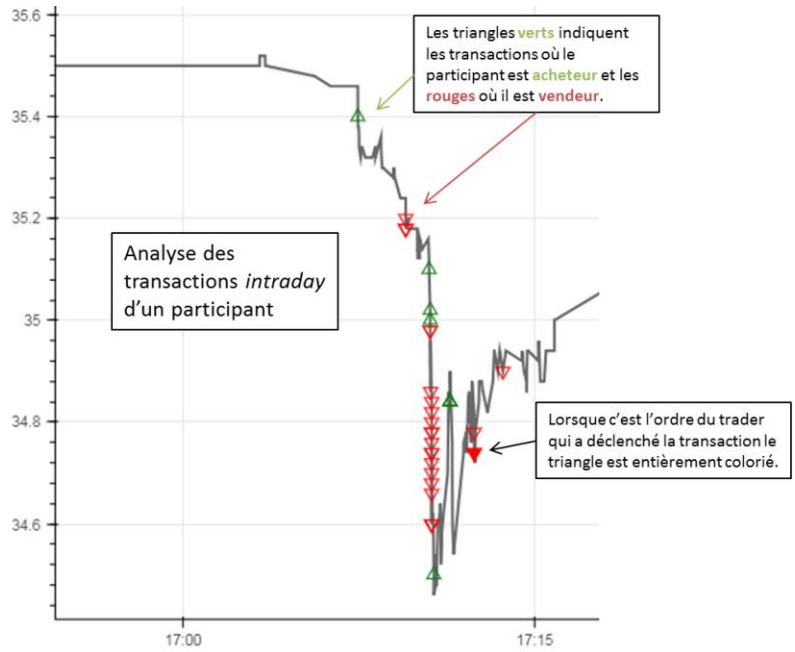


The algorithms at the core of market oversight do not supplant human analysis

Although data-processing is now essential for detecting misconduct, it cannot supplant human analysis for confirming cases of fraudulent behavior. For this purpose, our teams rely on tools developed in house for calculating a wide panel of indicators, such as a trader's impact on a security's price, the changes in his positioning, and his gains and losses over time. Analysts also have a library of graphs for visualizing event by event changes in an order book or displaying a trader's operations on different time scales (intraday or over several days).

[4] When "layering", a trader baits the market by making and canceling orders in order to simulate fictive interest (*e.g.*, of potential buyers) in a security and thus influence its price (upwards). He then profits from these manipulations by staking out a position in the opposite direction, as a seller of the securities in question.

Figure 4: Example from a tool for analyzing a trader's intraday activities, transaction by transaction.
Source: AMF.



To fulfill their assignment of market oversight, regulatory authorities invest in the quality of data. Much of the data collected comes from brokers who often have their own reporting systems, a potential source of errors ranging from mistakes about the amount, date or time of a transaction to the absence of declarations. For detecting market abuses, the quality of data is crucial, since reporting errors might alter the results yielded by algorithms, such that sequences of events are detected that should not have been. A "false positive" alert is less serious than the reverse, when the system fails to detect suspicious conduct because it lacks data or because the data in the reports filed are inaccurate. To handle such cases, the AMF systematically processes data so as to trace potentially erroneous data. Alerts of this sort are then analyzed manually. When an error is confirmed, a demand for correction is sent to the source of the report.

Market oversight has to stay on the cutting edge of technology

Ranking among the major international authorities that very early decided to invest massively in market data, the AMF tries to keep its digital tools technologically up to date so as to fully benefit from the data procured under the new directives and regulations. In 2016, it launched an ambitious program for overhauling its information system, in particular its new platform, ICY ("I see why"), which relies on big data for market oversight. Overhauling this infrastructure is not just the occasion to increase its storage and computational capacities but also, and even more, to make it evolve and adapt to changing needs.

Immersed in a system based on Hadoop, our data scientists benefit from a distributed architecture that enables them to manage the storage of data and computations in parallel. Using Python for developing experiments (a new algorithm of *ad hoc* detection and analysis) and Spark for an industrialization of data-processing, these work teams are organized to benefit as much as possible from this new environment. They must take up many challenges, *e.g.*, profit from this computing power to make advances in the algorithms for detecting market abuses or review data storage techniques and choose the technology best adapted for optimized processing. The possibilities

stemming from integrated products in Hadoop, such as Apache Hive and Apache Phoenix, are being examined.

The AMF is conducting experiments on the utility of artificial intelligence (AI), in particular, of applications based on algorithms of a machine learning type. Initial feedback from experiments in RegTech — digital technology at the service of regulation — let us glimpse interesting prospects, as the regulatory authority of financial markets in Quebec has pointed out. This authority has announced a supervised self-learning algorithm for detecting transactions that do not meet up to mandatory clearing and settlement requirements.[5] In another experiment, the US Security Exchange Commission (SEC) is trying to know whether a textual analysis of the articles mentioning credit default swaps (CDSs) would have predicted the 2008 meltdown.[6] The results of its natural language processing programs have convincingly shown that, just before the crisis, the number of articles on swaps rose tenfold.

Artificial intelligence at the service of regulatory authorities is no longer excessively utopian. Promising programs are in the pipeline, but they are few in number because they usually come with a high price tag in terms of time, resources and skills. In 2017, the Financial Stability Board (FSB) released its review of existing and coming AI applications in market finances.[7] This report covered all fields of market finance, ranging from hedge funds' trading strategies through back-office activities to regulatory compliance and fraud detection. Among the many possible applications for market oversight are: the cluster analysis of market intermediaries to establish a typology and facilitate analysis;[8] supervised learning for improving the quality of data; the analysis of messages on the social media in order to detect manipulations involving the dissemination of fake information; the identification of cases of collusion (misconduct by organized groups or networks); and so forth.

Though full of promises, AI is still in an experimental phase for regulatory authorities. It will be necessary to wait a few more years before a RegTech rolls it out. AI will not threaten jobs in activities related to market oversight, but its potential for analytics holds genuine prospects. Convinced that these new techniques will be the key to performing its assignments tomorrow, the AMF is running in the race toward "oversight 2.0" with the goal of being at the service of the operation and integrity of the financial markets.

[5] Press release of 27 April 2017: "AMF creates FinTech lab and signs partnership with R3" available at http://www.fil-information.gouv.qc.ca/Pages/Article.aspx?aiguillage=diffuseurs&lang=en&listeDiff=75&idArticle=2504273498.

[6] BAUGUESS S.W. (2017) "The role of big data, machine learning and AI in assessing risks: A regulatory perspective", speech on 21 June available at https://www.sec.gov/news/speech/bauguess-big-data-ai.

[7] BOARD F.S. (2017) "Artificial intelligence and machine learning in financial services: Market developments and financial stability implications", November, 45p., available via https://www.fsb.org/wp-content/uploads/P011117.pdf.

[8] The clustering method divides a set of data into homogenous groups so that each element in a subset has certain characteristics in common.