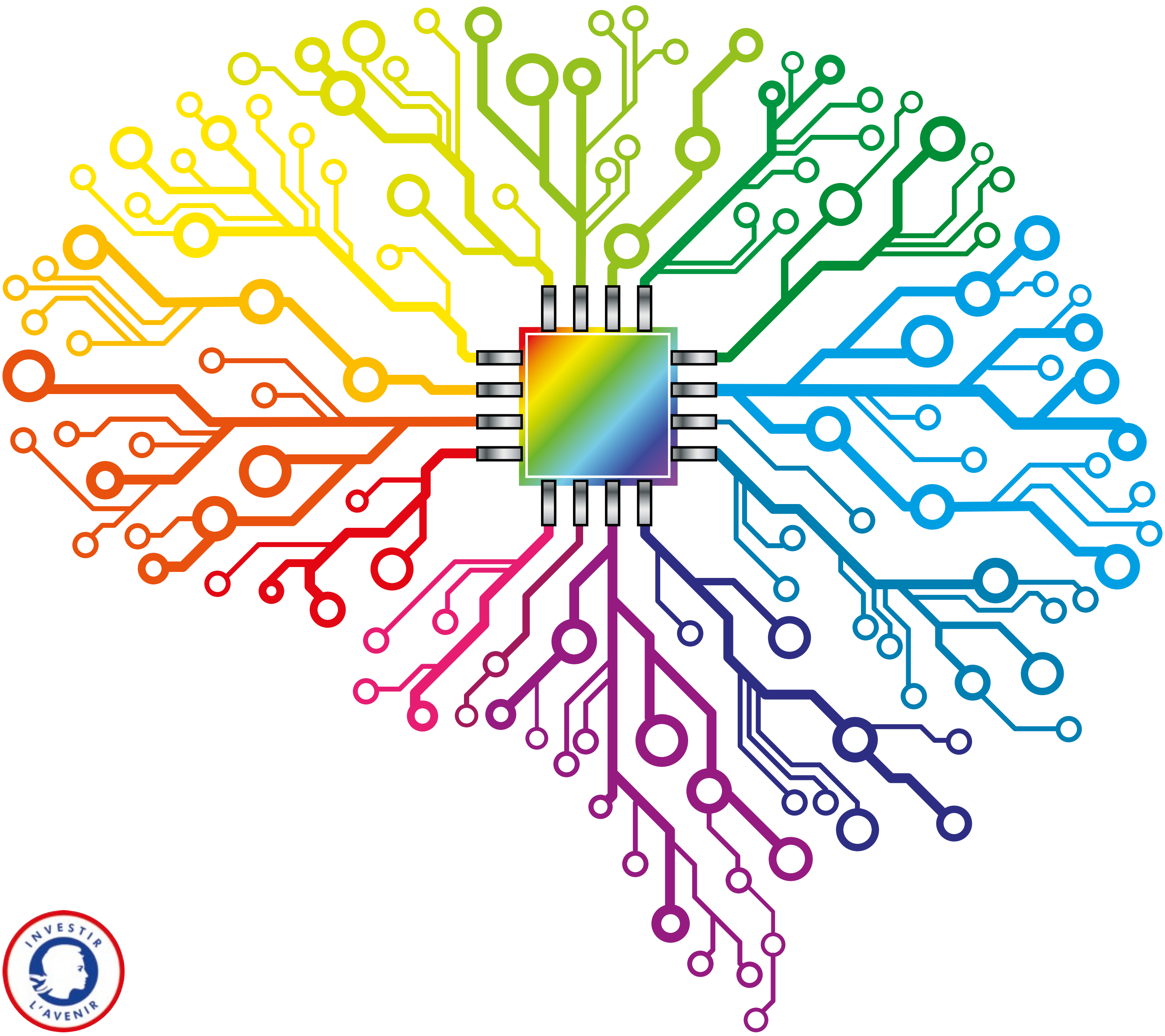
N°36 | APRIL 2020

Les CAHIERS LOUIS BACHELIER

Economic and financial news seen through research

The challenges of new technologies in finance

With Charles-Albert Lehalle, Driss Lamrani, Marie Brière, David Bounie and Winston Maxwell



INDEX

N° 36 | APRIL 2020

LOUIS BACHELIER .org

Economic and financial news seen through research

Publication of Institut Louis Bachelier – Palais Brongniart – 28 place de la bourse –75002 Paris
Tel. 01 73 01 93 40 – www.institutlouisbachelier.org – www.louisbachelier.org
Editorial Director: Jean-Michel Beacco
Editor-in-chief: Ryadh Benlahrech (ryadh.benlahrech@institutlouisbachelier.org)
Graphics Designer, Cover and Implementation:
Eux Production – Tel. 01 64 49 79 72 – www.euxproduction.com
Printer Kava : 42 rue Danton – 94270 Le Kremlin-Bicêtre – Tel. 06 14 32 96 87

PEFC
IMPRIM'VERT

EDITORIAL

The development of new technologies used in finance, particularly Artificial Intelligence (AI), has accelerated considerably over the last three years and is now entering a new phase. This welcome trend for the financial sector, which is concentrating on data-based businesses, is due in particular to the potential of deep learning and the emergence of meaningful data. The latter has taken over from Big Data due to the fact that it is more usable in AI systems applied to finance, as shown by the many projects currently being carried out by major groups and fintechs. The stated objective is to achieve increased human intelligence using digital technology.

At the Institut Louis Bachelier (ILB), the topic of AI in finance has been arousing considerable interest since 2014, a time when AI was still in its infancy. Initially, we focused on Big Data, with the scientific talks Meet the Data, in partnership with the University of Berkeley. We then gradually refined our thinking with the launch in 2015 of the ILB DataLab, an applied research department dedicated to the transformation of finance professions by AI. At the same time, many researchers in our network became interested in the impact of new technologies (AI and Blockchain), leading to the progressive creation of various research programmes to work on these topics, such as the interdisciplinary Finance and Insurance Reloaded (FaIR) programme and the Digital Finance Chair.

To illustrate the topics covered by the researchers and experts of our network, this new edition of the Cahiers Louis Bachelier deals with the digital transition in finance. It begins with an interview with Charles-Albert Lehalle, Scientific Director of the FaIR programme, whose aim is to bring together as many researchers as possible to collaborate on the impact of new technologies in finance. The second interview features Driss Lamrani, an AI specialist associated with ILB DataLab, who gives his point of view on the contribution of this technology to financial risk management. The third article, following an on-going study by Marie Brière, looks at the topic of robo advisors for individual investors, the use of which is fast developing. The last two articles, based on work by David Bounie and Winston Maxwell of the Digital Finance Chair, in affiliation with Telecom ParisTech, address algorithm bias and transparency, a highly topical and particularly sensitive issue for users and regulators at a time when these tools are increasingly common in everyday life.

Enjoy your reading!

Jean-Michel Beacco,
Delegate General
of the Institut Louis Bachelier

“THE CHOICE OF DATA IS AS CRUCIAL AS THE CHOICE OF MODELS”

While data sciences are already widely used in the financial sector, particularly in the automation of market transactions with high frequency trading, developments in machine learning and deep learning based on Artificial Intelligence (AI) have broadened the range of possible applications. Nevertheless, some on-going innovations raise new research questions.
The Institut Louis Bachelier (ILB) launched the FaIR (Finance and Insurance Reloaded) interdisciplinary research programme in early 2019. Its first objective is to help the financial industry, academics and regulators to better identify these crucial research questions. It also aims to help them find solutions to support the use of these new technologies (AI, blockchain) with regard to innovation. Three themes of application have been identified by this programme: the distribution of financial products, the improvement of risk intermediation, and a better connection to the real economy. In this interview, Charles-Albert Lehalle, Scientific Director of FaIR, discusses these innovations in finance, sometimes referred to as Finance 3.0.

ILB: What are the main upheavals in the financial sector caused by new technologies?

Charles-Albert Lehalle: As in many sectors, recent technologies such as data sciences, including AI and blockchain, are impacting the industry of financial markets in many ways. In terms of organizations and their governance, let me underline three major upheavals. The first involves a change in the mind-set about the role of data and data management. The new role of data requires a convergence between the front offices and their IT departments. While this is essential, it is also complicated to achieve for organizations that have often isolated their IT department, often putting them in a role close to the one of an external vendor, preventing them from easily provide innovations. The second upheaval concerns the modularization of financial services, which is significantly changing business models. Today, more than ever, there may be a blurring of roles between the providers and consumers of services. Clearly, there is less of an integrated end-to-end production chain than in the past, but rather separate functionality components, which may or may not be produced by the same actor. A typical example is Goldman Sachs’s “Marquee” platform, which is not yet a commercial success, but has some momentum and is definitely embracing modularity. Finally, third upheaval concerns modelling: in the absence of big data, formal modelling was the only solution, but when you use data sciences and in particular AI, data and not modelling is producing features somehow directly, potentially with little intermediation of an explicit model. It produces learning algorithms, which can be seen as “black boxes”. What is the best way to combine these new “empirical models” with more traditional “formal models”? It requires a better understanding

Innovations driven by new technologies in the financial sector call for rethinking the interactions between front offices and IT as well as the role of modelling.

Charles-Albert Lehalle is Head of Data Analytics at Capital Fund Management in Paris and Scientific Director of the FaIR (Finance and Insurance Reloaded) interdisciplinary research programme at the ILB. He is also Visiting Researcher at Imperial College, London.

of the implicit conditioning contained in the data, as well as more sophisticated analyses of these new black boxes, in order to make them more interpretable. It is a mistake to believe that data will replace models. Models allow operators to think about the relationship between applications and the real world, particularly in economics and finance, and we must continue to develop this way to think. It means we will have to combine reasoning with empirical, data-driven solutions.

How are the traditional actors in the financial sector setting about dealing with digitalisation, of which Artificial Intelligence (AI) is a part?

C-A L: There is general agreement on the need to innovate. To adopt the terms used by Philippe Aghion, professor at the College of France, AI is not an innovation, but a “general purpose technology” in the same way as the steam engine or electricity. On the basis of this observation, each sector has to invest, in order to identify and generate secondary innovations likely to be deployed. To produce secondary innovations market participants, including central banks, have set up small teams within AI or data science dedicated “labs”. These allow business experts to work for a period of several months with data science experts in the scope of well identified projects that will be later on fully developed in their parent departments. It is really important to understand that without a way to stimulate secondary innovations, nothing really new will come from data science and AI. Even though France has first class academics, it is not at the forefront of developments of this kind. Other countries, such as the UK or the US, have implemented this kind of approach long ago.

But it looks like a huge challenge…

C-A L: Indeed, to be able to produce innovations, the financial sector has to change its way to think about innovation strategies. Again, bringing together front office quantitative teams and IT teams around data to produce innovative features is a prerequisite. Nevertheless, traditional financial players are somehow protected from disruption in the short term because of their size, and also because they benefit from a kind of “protection through regulation”, which introduces an high fixed cost that is a barrier for start-ups. Therefore they have a little more time than their equivalent in other sectors. Nevertheless, think about Intercontinental Exchange buying the New York Stock Exchange and Euronext in 2012. This is a typical case of a data and techno-driven company buying an historical one; it can happen in the financial industry.

The GAFAM companies seem to be ahead of traditional actors in terms of AI. What are the potential risks?

C-A L: It’s a remarkable fact that financial actors have been absent from the podiums of competitions around AI. Insurers are not involved in Natural Language Processing (NLP), even though this should be part of their expertise, given the billions of texts at their disposal. Similarly banks have only very recently embarked on reinforcement learning, which consists in optimizing strategies with AI. Yet they are not to be seen on the winners’ podiums, even though discovery investment and hedging strategies is part of their core business! Returning to the initial question about potential risks, some AI-driven libraries are provided only by the Web giants. As they are open source everyone uses these tools. However, a bug that would be introduced in the release of a new version may produce a systemic risk, due to “bad synchronization” of the users of a corrupted library. Regulators are aware of this possibility and are thinking about it.

What about the competitive risks by Web players for traditional finance?

C-A L: For the moment, the Web giants are sticking to their original positioning, with their own regulators providing services (the cloud, AI, and soon blockchain). I don’t think they want to provide financial services in the short term and thus be regulated by the same entities as banks. The real competition is rather at the level of fintechs, but they are subject to the financial regulation, which is stricter than in other sectors.

Let’s now turn to one of the topics addressed in the FaIR programme: improving the client experience, which is one of the areas of development for AI in finance. What are the benefits for banks and customers?

C-A L: The personalisation of the customer experience is an interesting challenge, as it allows financial products to be adapted to the customer, while reducing costs. In addition, the functionalities provided by AI can be used to better manage the balance sheets of financial institutions. In this way financial institutions can better identify customers with a risk appetite compatible with the products banks are ready to sell at a lower cost in order to balance their books.

Another topic that FaIR is working on is better risk intermediation, for example with blockchain. What are the implications for the sector?

C-A L: Blockchain, which differs from AI, facilitates the automatic resolution of contracts. So the use of blockchain is very convenient and would avoid many end-of-day account reconciliations. If you really want to use 100% of the power of blockchain for automatic resolution, it will prevent market participants to buy or sell without having the funds available beforehand, whereas it is not the case today. Currently financing can be found during the course of the day, once you know that you will need cash. In short: requiring the pre-financing of a transaction is a revolution for financial processes, which are currently settled at the end of the day or on D+1. In order to fully exploit blockchain, ➜

it will be necessary to have euros or dollars on a blockchain. This is the subject of "stable coins" that have been much in the news since Facebook, with a consortium of other actors, announced its Libra model for a stable-coin. Other major issues regarding the introduction of blockchain for financial markets are energy consumption and the speed of transactions. On this last point, some recent developments are very encouraging.

The third theme that the FaIR programme is studying concerns how finance can be best connected to the real economy through AI, whereas in the popular imagination, finance has no connection with the real economy. Can you elaborate on this?

C-A L: The feeling of disconnection between finance and the real economy may stem from the fact that prices of financial assets are formed through the matching of supply and demand in financial marketplaces, where buyers and sellers are financial actors. This may give the impression that prices are formed "by finance, for finance". But this is not true. To generate returns, market participants need to anticipate the health of companies or of the economy, and therefore they try to gather as much relevant information as possible to make informed decisions. AI and the availability of new data from the real world, known as "alternative data" (satellite images, texts from websites, job postings, etc.) make possible to construct new indicators on the health of companies and states. By exploiting these alternative data, asset prices integrate more real-world information. With better financial dashboards generated by this data and by learning algorithms, the connection between finance and the real economy will become stronger, closer to the reality. This trend has taken off in the last five years or so and is continuing.

Do you have any recommendations or at least watch points to suggest to financial actors?

C-A L: As I mentioned earlier, the role of modelling should not be neglected. It is very important to be aware that using data to tune an algorithm actually implicitly requires a lot of modelling choices. To take a simple example, the one of a credit granting algorithm feeds on data. If the data comes from a specific geographical area or if there are 75% men and 25% women in the database, the results will be influenced by these imbalances. In other words, they will be biased. In reality, the choice of data, and therefore knowledge implicitly stored in the data, that is used has an influential role. It should be explicit and approached like a choice of model. Modelling remains crucial for understanding the economic and financial mechanisms at work.

To conclude, what topics will the FaIR programme be working on next?

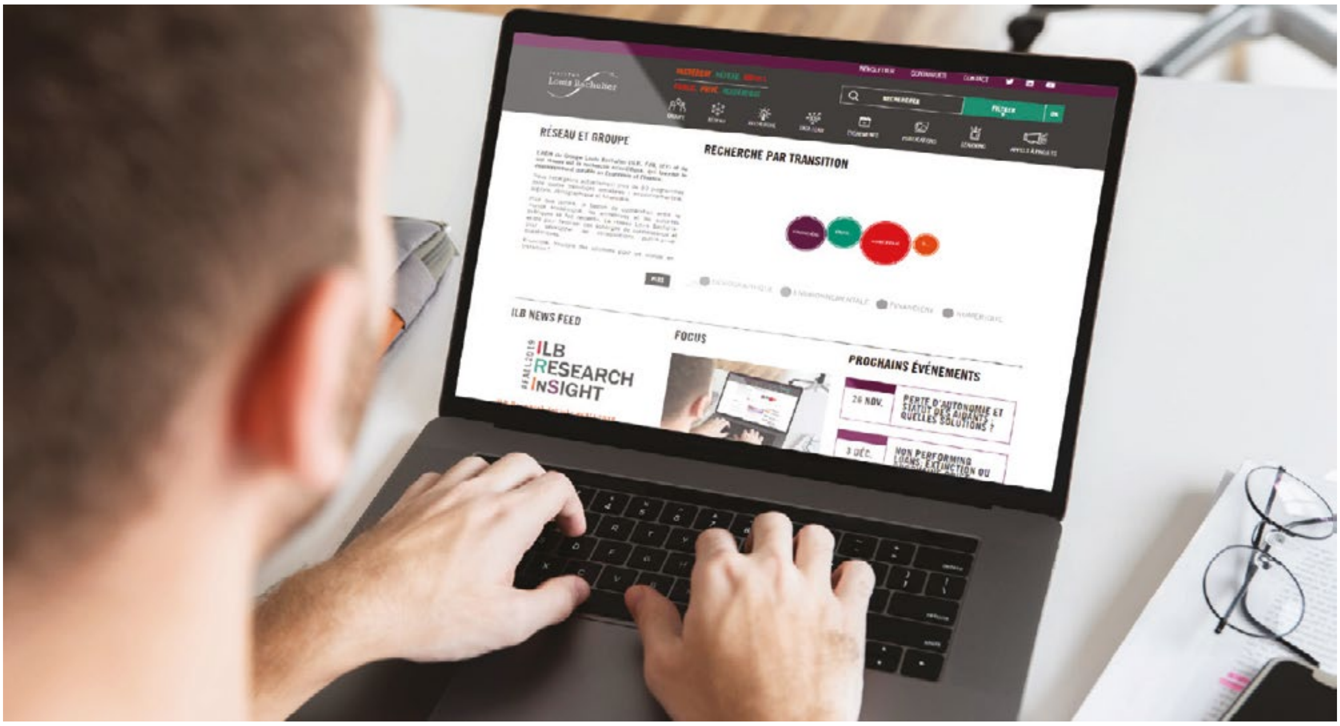
C-A L: The subject of regulation is crucial and raises many questions. How should algorithms be regulated? What are the chains of responsibility? What political choices will be made regarding the ethics of algorithms? And above all, how will algorithms be deployed operationally? How will self-learning algorithms be certified and supervised? We would like to organize seminars and workshops on these last three issues. ●

INSTITUT
Louis Bachelier

VISIT

THE ILB PORTAL

INSTITUTLOUISBACHELIER.ORG



A one-point entry to partnership-driven research

RESEARCH. NETWORK. IMPACT.

PUBLIC. PRIVATE. ACADEMIC.

“THE TRADITIONAL APPROACH TO RISK MANAGEMENT MUST EVOLVE TO ANTICIPATE NEW CRISES”

Since the financial crisis of 2008, the financial sector has seen the emergence of new risks that had not necessarily been identified in the past, such as risks related to global warming, risks of technological disruption or the risk of a global pandemic. In this changing context, the traditional approach to risk management may prove inadequate for detecting these new risks, thus necessitating the development of new tools. In this regard, certain artificial intelligence (AI) technologies offer new prospects for managing financial risks. To clarify the picture, Driss Lamrani, an expert in economics, finance and AI, answered questions from the Institut Louis Bachelier (ILB).

ILB: What are the limitations of traditional approaches to risk management?

Driss Lamrani: For more than three decades, traditional risk management has considered that the risks occurring in the future will be of the same nature as those of the past, in terms of magnitude and frequency. However, the 2008 crisis and today’s unprecedented and sudden health crisis show the limits of this approach. In order to incorporate extreme events into traditional models, modellers have been introducing some disruptive factors, but this has clearly proved insufficient. Moreover, traditional risk management is based on two assumptions: the absence of arbitrage opportunities, and the completeness of financial and insurance markets. However, these assumptions have very much been called into question and are being undermined by the digital transition, which in particular generates a mass of important information that needs to be processed to inform risk anticipation. Moreover, since traditional models are increasingly less forward-looking, regulators often react after events. Finally, the period of change we are in now, with the emergence of new risks and the transformation of many industries, does not lend itself to the traditional approach to risk management. The example of the coal industry illustrates this point: the use of coal is set to shrink dramatically due to global warming or even end altogether, but many companies in the sector have been financed by banks, underwritten by insurance companies and invested in by asset managers looking for value creation. Consequently, new scenarios need to be devised and constructed by risk professionals. New thinking on risk control is clearly needed, and the traditional approach to risk management must evolve if it is to anticipate and manage new crises.

What are the impacts of the digital transition on the management of financial risks?

DL: The digital transition is mainly manifested by the mass and variety of available data to be processed and analysed in order to assess future scenarios likelihood and the consequences regarding the exposure of financial institutions, over and beyond the usual financial data taken into consideration. This new typology of information, based on alternative data, is contributing to and will go on contributing to the renewal of risk management. By way of example are the announcements by the World Health Organization (WHO), which are not considered in traditional models. However, this implies working on massive amounts of unstructured data, unlike ratings established by agencies on the basis of credit scoring.

Driss Lamrani is CEO of the Fintech A2IM-Roboadvisors, which specializes in Artificial Intelligence software for the financial sector edition. He regularly collaborates with the ILB DataLab. A graduate of Ecole Centrale de Paris, majoring in economics and finance, and of the French Society of Financial Analysts, Driss Lamrani has more than 20 years' experience in the banking and financial sector.

Alternative data is and will continue to be increasingly important in the analysis and management of financial risks.

How can the digital transition, particularly AI, help financial institutions manage their risks?

DL: As I have mentioned, alternative data is and will continue to be increasingly important in the analysis and management of financial risks. The use of AI is therefore recommended for monitoring and analysing this data, as it calls for a large amount of reprocessing, especially with regard to natural language processing and understanding. Consequently, it is very difficult for this work to be carried out by human beings. Consider once again the example of WHO announcements, in 99 cases out of 100, these will not be relevant for financial risk management. AI is useful for cost optimization to identify the one case where such alerts would be important for financial institutions risk management.

What are the most promising AI tools available for improving financial risk management?

DL: At the moment AI can automatically perform several human cognitive actions with high relevance rates. To analyse the numerous alternative data, we can emphasize three important tools:

– Natural Language Processing enables texts to be read and understood;

– Image recognition enables information to be recognised and data to be organized on photographic supports;

– Inference analysis (deep reasoning and natural language understanding) allows the consequences of a set of data to be deduced.

The development of algorithms that replicate human brains saves a considerable amount of time. For example, when JP Morgan publishes its quarterly financial results, more than 100,000 news reports appear the following day. With AI, it would be possible to reduce this number to the 10 most relevant articles in less than an hour, whereas it would take an experienced human analyst 30 days to go through the 100,000 articles and achieve the same result. As well as these various tools taken separately, the comparative advantage of AI is that it can build a complete cognitive process to address the specific issues of the financial sector.

What constraints need to be removed for AI to develop more widely in risk management?

DL: The current situation should prompt banks, asset managers and insurance companies to analyse risks that have only recently become apparent, such as liquidity risk resulting from a global pandemic. Short-term thinking is a problem for many financial institutions, because they do not invest sufficiently in AI. Yet this technology can help to better anticipate risks, especially systemic risks, by broadening the sources of data analysed and optimizing human experts' analysis time. ●

DO ROBO-ADVISORS IMPROVE INDIVIDUAL INVESTORS' DECISION-MAKING?

To address the relative reluctance of households to access the financial markets and help them further diversify their investments, robo-advisors provide a promising alternative to human advisors by offering more personalized advice. A recent scientific research has focused on this tool, which is rapidly developing among asset managers and financial institutions.

Based on *Robo advising, attention, and long-term investment,* by Milo Bianchi and Marie Brière, and on an interview with Marie Brière.

The robo-advisor market has been booming since the first such tool was introduced in the United States in 2008, a period marked by the devastating subprime financial crisis. In fact, assets under the management of robo-advisors rose by 47.1% during 2019 to $1,277 billion ($1.3 trillion) and involved more than 70 million users worldwide, according to Statista. By 2023, the market is expected to be growing by 21 per cent annually, according to this German-based data provider's forecast. "Even though the penetration rate of robo-advisors is still low, it is expected to go on growing in the future, especially among the younger generation in developed countries. In developing countries, such as China, digital financial tools are a means of rapidly financializing certain segments of the population," Marie Brière says.

IMPROVING THE DECISIONS OF INDIVIDUAL INVESTORS

The growing development of robo-advisors is based on various needs on the part of both management professionals and individual investors. On the one hand, reducing costs, improving client relations with more personalized offers and limiting conflicts of interest for human advisors are key areas for improvement. On the other hand, investment decisions made by households are generally tainted by numerous biases, as the scientific literature has amply demonstrated. In fact, individual investors exhibit a number of shortcomings, often linked to a lack of financial education, such as less participation in financial markets (particularly the stock market), lack of attention to their investments, poor diversification of their investments, and familiarity bias (national preference, etc.) in their decision-making.

However, in conjunction with these findings, in more than 7 out of 10 cases, retail investors in Europe and the United States consult a human advisor when purchasing financial products. A number of research questions consequently emerge. Are there synergies between robots and traditional human advisors? What are the impacts of these tools on the decisions of retail investors? What are the characteristics of users? It should be noted that robo-advisors guide investors in accordance with their respective profiles after a detailed questionnaire (personal financial situation, savings objectives, investment horizon, risk appetite, etc.).

A STUDY ON THE BEHAVIOUR OF INVESTORS USING ROBO-ADVISORS

To address these issues, the researchers conducted a specific study on a sample of approximately 20,000 users of robo-advisors in France over a period of two years. "The analysis of the impact of robo-advisors is especially interesting because it is one of the only areas of finance in which we can study the interactions between humans and machines," Marie Brière points out. "Our study focused on a robo-advisor specialized in employee savings, a very large market in France." In concrete terms, each company with employee savings plans offers its employees the opportunity to invest their savings in a range of dedicated funds. Under a Company Savings Plan, the money is locked-up for at least five years, and until retirement in a Retirement Savings Plan, except in the case of

Marie Brière is Head of the Investor Research Centre at Amundi, Affiliate Professor at PSL Paris-Dauphine University and Chair of the Scientific Committee of the International Financial Risk Forum. She co-authored the paper Robo advising, attention, and long-term investment with Milo Bianchi, Professor of Finance at the Toulouse School of Economics (TSE) and member of the Sustainable Finance and Digital Finance Centers at TSE and of LTI@UniTO.

Methodology

The researchers studied the impact of robo-advisors on the investment performance and characteristics of their users. To do so, they accessed a database of around 20,000 clients of an employee savings robo-advisor provided by a major European asset manager over a two-year period, as well as data from a test population. After dividing their sample into a test group and a control group, they analysed the different variables using the difference-in-difference statistical estimation method and examined the differential behaviour of investors before and after subscribing to the robo-advisor, compared with the test population.

early release. The savings are therefore more oriented towards long-term objectives. Asset management is then delegated to a management company. From 2017 the robo-advisor service's input was gradually introduced to employees, who had the opportunity to decide whether or not to subscribe to the service. "We looked at the differential behaviour of investors on several variables before and after using the robo-advisor, by comparing them to a test population which had not been exposed to the robot," Marie Brière says.

ROBO-ADVISORS INCREASE INVESTORS' ATTENTION AND RESPONSIVENESS…

While numerous studies have shown that individual investors are not very attentive to their investments due to lack of time or to avoid unpleasant surprises regarding their financial performance, the use of a robo-advisor seems to have a positive impact on investor attention. The latter has been measured in particular by the number of times per month investors access their savings accounts or the number of transactions carried out. "Investor attention is particularly high during the first months following subscription to the robo-advisor service. This finding is very interesting. It is thus possible to imagine a complementarity between human decisions and those guided by the robot, without necessarily placing them in opposition to each other," Marie Brière explains.

… AND THEIR FINANCIAL PERFORMANCE

In addition to improving attention, the robo-advisor increased diversification of people's investments, which is one of investors' historically observed shortcomings. Such diversification is also reflected in the holding of more risky investments, including diversified funds, compared to investors who do not use the robo-advisor. Furthermore, financial performance is on average higher among users of a robo-advisor. "Our work is not finished. We intend pursuing it further with a view to refining and confirming these initial results. Perhaps in the current market turmoil our findings would be less or more pronounced, given that from 2017 to 2019 stock markets rose a lot," Marie Brière says.
Finally, the profile of the users was not necessarily what might be expected in advance. On average, they were young, male and had above-average assets (in the context of employee savings schemes). They were individuals who were already relatively attentive to their savings. Consequently, it is not clear that robo-advisors can attract people who are more remote from financial markets, such as those with modest incomes or with more limited financial education. Pending the potential democratization of robo-advisors, the public should acquire familiarity with the concept, given that aversion to algorithms is particularly acute.

Key points

- Subscribing to a robo-advisor increases investor activity and attention. This finding suggests a certain complementarity between human judgements and those emanating from machines.
- On average, the robo-advisor studied improved the decisions of individual investors in terms of performance and diversification.
- Existing users of robo-advisors are mostly young, male, have more savings and pay greater attention to their assets than the average population. It is not yet clear whether this type of service can reach segments of the population that are more remote from the financial markets.

HOW CAN ALGORITHM BIASES BE LIMITED?

Although decision-support algorithms are now part of everyday life and are used in all sectors, especially finance, they are still criticized for their lack of transparency and their biased results. Researchers have studied this issue and make recommendations to reverse the situation.

Based on *Algorithmes: biais, discrimination et équité,* by Patrice Bertail, David Bounie, Stephan Clémençon and Patrick Waelbroeck, and on an interview with David Bounie.

Educational orientation, medical diagnosis, and the granting of credits are just a few examples of the fields that use algorithms in France. However, the problems experienced by Parcoursup in assigning graduates to higher education streams have revealed the shortcomings of the algorithms to the general public and the authorities. On the other side of the Atlantic, algorithms go even further: they filter applications during recruitment processes and help judges reach their decisions. In the financial sector, the use of algorithms is commonplace. A study by the European Securities and Markets Authority (ESMA) published in 2017 estimates that almost 45% of European stock market volumes are generated by high-frequency trading algorithms. In the United States, the proportion is higher than 70%. “These tools have developed rapidly in recent years. Some associations in the U.S. have realized that algorithms are not fair, and this gives rise to societal problems,” David Bounie says. Among the grievances voiced are discrimination against women and ethnic minorities. “The goal of our scientific work has been to better understand how biases in algorithms are formed and whether they are avoidable. We have drawn on various disciplines, including mathematics, computer science, law and philosophy,” he adds.

ALGORITHMS HAVE VARIOUS BIASES

In addition to the technical difficulties of algorithms in predicting outcomes and reducing the risk of error, they also contain biases that can lead to unfair and/or discriminatory decisions. Researchers have identified three main groups of biases.

– Cognitive biases come from the programmers who wrote the codes for the algorithms or from users who self-select into a category that

Clearly, if the data are biased, the results of the algorithm will be biased as well.

then feeds the algorithm. However, a number of studies in psychology and cognitive science have established that there are many cognitive biases emanating from humans. Stereotype bias, for example, has been known about for many years. It can occur on employment websites when women and men apply or do not apply for particular jobs, and this in turn leads to bias in the algorithm.

– Statistical biases mainly concern the quality of data. Clearly, if the data are biased, the results of the algorithm will be biased as well. “With biased or poor quality data, the algorithm cannot train itself properly and automatically causes errors. Collecting vast quantities of data to put into an algorithm does not imply that the results are satisfactory. Quantity should not be confused with quality,” David Bounie says.

– Voluntary or involuntary economic biases are due to financial factors, according to several studies. The economic context must therefore be taken into account when constructing an algorithm.

PROMISING SOLUTIONS TO REDUCE BIAS

To address the above-mentioned issues, researchers have identified two broad categories of solutions. Firstly, statistical techniques that make it possible to verify, correct and supplement the information related to the data. “Before producing results, understanding the

David Bounie is Professor and Head of the Economic and Social Sciences Department at Telecom Paris. He is also co-founder of the Digital Finance Chair. His research, which has led to publications in international journals, focuses on the impact of new technologies on the financial sector.

Methodology

The researchers have produced what they call a "position paper" on the problems of bias and fairness of algorithms. Accordingly they have had discussions with the various stakeholders (political, academic, industrial) within a multidisciplinary logic (mathematics, computer science, social sciences, etc.), in order to summarize the state of the art in this field and to make scientific recommendations for correcting algorithmic biases.

data is essential, as it allows us, for example, to calculate the probability of inclusion of a specific category of individuals in the database. With statistics, it is also possible to model the characteristics of individuals and then adjust the algorithms," David Bounie explains. The second category aggregates data trails, so that equity or non-discriminatory legal provisions can be introduced into the algorithms. "Computationally, it is almost impossible to produce a fair and universal algorithm. Therefore, equity concepts and constraints must be implemented from the very beginning of programming an algorithm," says Bounie.

THE PUBLIC AUTHORITIES NEED TO REGULATE HIGH-RISK ALGORITHMS

As well as the statistical and computer trails allowing the biases of algorithms to be corrected, legal provisions are recommended to promote their fairness, transparency and auditability. "Fairness is a subjective concept that depends on the culture and political system of each country. Therefore, algorithmic fairness must be a political choice, such as giving (or failing to give) the same chance to each individual according to criteria such as gender, ethnicity or social origins," David Bounie says. Moreover, supervision by the public authorities of certain "high-risk" (critical) algorithms is indispensable. In this regard, the European Commission recently provided guidance on its position by defining certain high-risk algorithms for society. It is important to say that fundamental rights, in particular the right to non-discrimination, must be respected by algorithms. For its part, the General Data Protection Regulation (GDPR), which came into force in May 2018, provides a framework for the decisions taken by algorithms. Nevertheless, discussions are on-going and legislation is likely to continue evolving in parallel with the growing use of these decision-support tools. "Vigilance must be maintained. In principle, impact studies should be carried out on algorithms posing a high risk for society," David Bounie concludes.

Fundamental rights, in particular the right to non-discrimination, must be respected by algorithms.

Key points

- Algorithms contain several bias (cognitive, statistical, economic). They therefore do not produce neutral results.
- In order to correct algorithm biases, there are statistical and computational techniques that must be integrated at the design stage of the algorithm conception (ethics by design).
- Algorithmic fairness is a political choice because it is impossible to design universal algorithms for all cultures and political systems. Impact studies must be conducted beforehand on algorithms with a high risk for society (health, educational orientation, judicial decisions, etc.).

IMPROVING THE EXPLAINABILITY AND ACCOUNTABILITY OF ALGORITHMS

Thanks to advances in deep learning, decision support algorithms are emerging in almost all sectors, but many of these artificial intelligence (AI) tools remain "black boxes".
A multidisciplinary group of researchers has been working on this problem of opacity in order to remedy it and thus promote public confidence in algorithms.

Based on the new working paper *Flexible and Context-Specific AI Explainability: A Multidisciplinary Approach,* by Valérie Beaudouin, Isabelle Bloch, David Bounie, Stephan Clemençon, Florence d'Alché-Buc, James Eagan, Winston Maxwell, Pavlo Mozharovskyi and Jayneel Parekh, and on an interview with Winston Maxwell.

"Everybody agrees algorithms should be explainable, especially in safety-critical areas such as health and air transport. There's a real consensus," Winston Maxwell says. For instance, the European Commission recently published a white paper on its AI strategy, which emphasizes the explainability, transparency and accountability of algorithmic decisions. In France, the Villani report on AI, published in 2018, also insisted on the need to make algorithms more transparent and understandable. However, this policy objective runs up against numerous implementation problems related to the concept of explainability, such as ethics, the definition of the right level of explanation, the technical characteristics of the AI methods used, the preservation of trade secrets, the additional costs generated by explanations, and the lack of clear legal definitions regarding what explanations mean. Existing legislation already requires explainability, particularly for algorithms used by administrative authorities in France, but the law generally leaves a great deal of leeway, thus adding further difficulties for developers, users and regulators of these tools.
"Our research work has been carried out using a multidisciplinary approach, bringing together data science, applied mathematics, computer science, economics, statistics, law and sociology, so as to provide in-depth thinking with regard to definitions, techniques and the need for explainability, which are integrated into the broader notions of transparency and accountability," says David Bounie, a co-author of the report. In short, explainability may be used, for example, to help users understand how a search engine works, to help investigators find out why an autonomous vehicle crashed, or to detect possible discrimination in loan approval processes.

EXPLAINABILITY DEPENDS ON FOUR CONTEXTUAL FACTORS

To attain their goal of demystifying explainability, the researchers developed an original methodology whose starting point is contextual. They identified four important contextual factors.
– The audience, i.e. the persons targeted by the explanation. The level of the explanation will differ depending on whether it is addressed to a user or a regulator, for example.
– The level of impact of the algorithm. Explaining why an autonomous car crashed is of greater importance to society than explaining the mechanisms behind an advertising or video recommendation algorithm.
– The legal and regulatory framework, which varies according to different geographical areas, for example in Europe with General Data Protection Regulation (GDPR).
– The operational environment surrounding explainability, such as its mandatory status for certain critical applications, the need for certification prior to deployment, or ease of use by users.
"Taking into account the four contextual factors of explainability is essential. For business users and developers, explainability is primarily driven by operational requirements, and those are quite different from the legal requirements," says Winston Maxwell.

Winston Maxwell is director of law and digital technology studies in the economics and social science department of Telecom Paris. Previously, he was a partner at the law firm Hogan Lovells, specializing in data law. He is a graduate of Cornell Law School and obtained a PhD in economics from Telecom Paris (*Smarter Internet Regulation Through Cost-Benefit Analysis,* published by Presses des Mines in 2017). His research work is mainly focused on the regulation of artificial intelligence.

Methodology

In producing their "position paper" on the problems of algorithm explainability and accountability, the researchers discussed the subject with the various stakeholders (political, academic, business) using a multidisciplinary approach (mathematics, computer science, social sciences, etc.). The aim is to determine the state of the art and to identify scientific and policy recommendations for improving the explainability of algorithms.

EXPLAINABILITY'S COSTS AND BENEFITS TO SOCIETY

After the first stage related to the different contexts of explainability, the researchers studied technical explainability solutions for different AI algorithms. In broad terms, they drew up an inventory of the different methods used to render different machine-learning models more transparent. These methods include hybrid AI approaches, which combine the best of several AI techniques including symbolic, knowledge-based, AI. Hybrid approaches are particularly promising, according to Isabelle Bloch, one of the study's co-authors: "These approaches could reduce the gap between algorithm performance and explainability. Ultimately, explainability will be an integral part of performance characteristics."

The researchers have made another major innovation by incorporating a cost-benefit analysis into explainability. In other words, the level of explainability will be driven in part by comparing the costs and benefits of explainability for society. As mentioned above, the explanation of an autonomous car accident and the explanation of a search engine result do not have the same impact on society; the benefits of explanation will be different in each case. The researchers identified several categories of costs related to explainability, in particular costs relating to the storage of data logs in dedicated registers, which will prove essential to permit ex post explainability of decisions. The researchers point out, however, that the GDPR limits the storage of personal data. "The issue of data storage and explainability will necessarily involve political choices, because the GDPR discourages data retention, in particular with regard to biometric and facial recognition data. Thinking on this subject is still in its infancy and regulators will certainly have make their decisions based on the particular application and its impact on society", Winston Maxwell points out.

LOCAL AND GLOBAL EXPLAINABILITY

In addition to contextual factors and cost-benefit analyses, the researchers also observed that the right level of explainability must take into account both global and local considerations. Global explainability involves describing the algorithm as a whole and how it should be used (or not, as the case may be). "It is like a user's manual and a warning notice, which includes the type of data used to train the algorithm, and the situations in which the algorithm should or should not be used. The European Commission has adopted this approach in its white paper on AI strategy," notes Winston Maxwell. As for local explicability, this involves explaining particular algorithmic decisions, such as a loan refusal. "Both these dimensions of explainability are necessary, even if they are aimed at completely different things and depend on the four contextual factors we identified," Winston Maxwell emphasizes. There is no doubt that, in the coming months, the issue of the explainability of AI, and in particular of deep learning algorithms, will become more prominent as the policy debate on AI evolves in Brussels. ●

Key points

- Explainability of an algorithm depends on four contextual factors: who the explanation is addressed to, the impact of the algorithmic application, the legal and regulatory environment, and the operational framework. In addition, the global (general operation of the algorithm) and local (specific decision-making) level of explainability must be taken into account.
- The explainability of an algorithm must be considered in light of the costs and benefits for society. In particular, the storage of data logs for algorithmic decisions will require a political choice, as it is costly and is not always compatible with the GDPR.
- Explainability is generally at odds with the performance of algorithms, because machine learning models are often built with only one performance objective in mind. However, with the development of hybrid AI techniques, explainability will become an integral part of performance parameters.

INSTITUT
Louis Bachelier