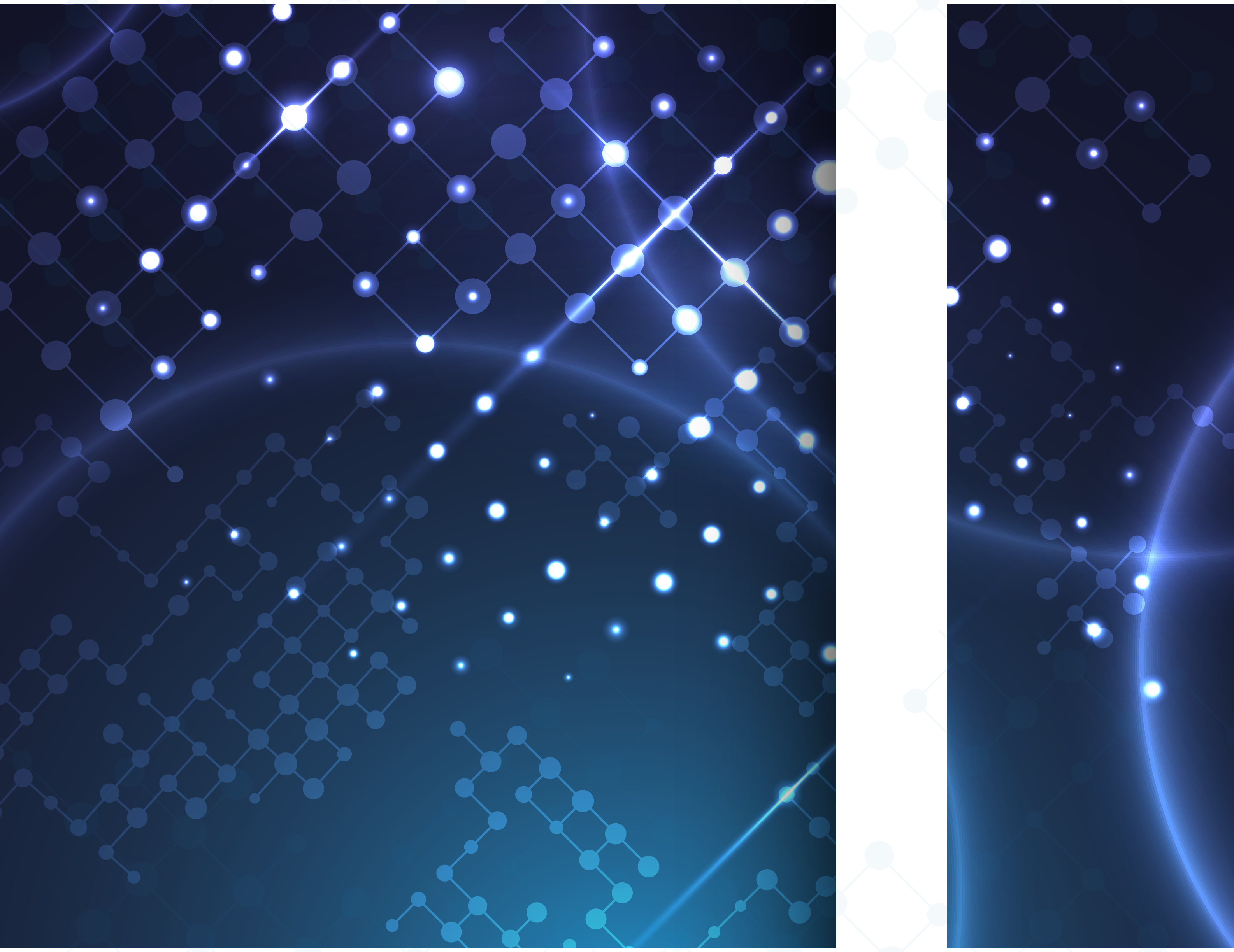
FaIR ADVANCES REPORT



FaIR
Finance and Insurance Reloaded

an Institut Louis Bachelier interdisciplinary program

PUBLISHED in 2021

FAIR IS AN ILB FLAGSHIP

The core of the Louis Bachelier Group (ILB, Fondation du Risque, Institut Europlace de Finance) and its surrounding network is research, that aims to promote sustainable development in Economics and Finance.

We currently host more than 60 sponsored research programs divided in four societal transitions: environmental, digital, demographic, and financial.

More than ever, there is a need for cooperation between academia, business, and the public authorities. The Louis Bachelier network exists to promote the public-private-academic collaborations, enhance exchange of knowledge, and develop scientific contributions to societal challenges. As of today, more than 25 academic institutions, 100 corporate as well as public authorities has joined the network. Together, we seek answers for a world in transition!

The interdisciplinary program FaIR is one of our fine examples of this and more than ever an absolute priority at ILB.

We take a transitional challenge, combine it with our strengths and mix French and foreign experts. The result of that recipe can be found in this report on FaIR Advances. In it you will find contributions from ILB chair researchers as well as from the International Scientific Council and external partners such as the Turing Institute in London, with whom FaIR has had its ties since its creation in 2019.

With the emergence of new technologies in the banking and insurance sectors in particular, finance is undergoing a major change – hence the name of this interdisciplinary program. We need investments in interdisciplinary programs that focus on machine learning and artificial intelligence (AI) to learn, understand and advance in a sustainable way and at ILB we are more than happy to make our contribution.

The ILB Finance and Insurance Reloaded (FaIR) program supports and encourages innovations resulting from new technologies (ranging from Artificial Intelligence to blockchain) in the finance and insurance sectors. FaIR also functions as meeting place for all ILB initiatives that unite industrial as well as research partners around the impact of these new technologies. The FaIR program has brought a lot of people and ideas together during its first years of existence and created the FaIR advances of this publication. We look forward to the next phase and more experts to join.

I especially wish to thank Charles-Albert Lehalle who is scientifically responsible for FaIR and who has put in a lot of hours, ideas and network to build what FaIR is today. I also wish to thank Louis Bertucci who has joined more recently and is the editor in chief of this publication. And finally, I warmly thank all other contributors to the FaIR Advances whether they are round table participants, text contributors or part of the ILB team that helped this impressing collection of insights come to life.

JEAN-MICHEL BEACCO
CEO ILB

AN ON-GOING PROCESS OF ASSESSING THE IMPACT OF NEW TECHNOLOGIES ON THE FINANCIAL AND INSURANCE SECTORS

Three years ago, Institut Louis Bachelier saw an increase of requests coming from the industry to finance collaborations with academia on different forms of machine learning and blockchain topics. This appetite to explore the potential added value of these new technologies on financial markets and for the insurance sector was not well organised; it seems closer to curiosity than the will of conducting strong research collaborations. It nevertheless fostered the interest of the scientific board of the Institute on these domains, raised the question of helping the industry to focus collaborative efforts on impacting projects. It has been the starting point of the “Finance and Insurance Reloaded” transverse research program.

We are currently in the first phase of this project, structured around more transparency and information sharing. We first investigated what could be the main fields of applications in a workshop at the Collège de France. The emerging themes have been: usages oriented towards clients of these sectors, internal usages of new technologies to improve risk intermediation, and usages improving the connection between the “real economy” and the financial and insurance sectors, essentially via “alternative data”. We also had to define with more accuracy the scope of these “new technologies”. Hopefully Institut Louis Bachelier is in contact with international experts, and discussions started with overseas academics, some of them joining our International Scientific Board later on. Moreover, we could leverage on reports issued by other organisations; I have in mind the report by Finance Innovation on “Intelligence artificielle, blockchain et technologies quantiques au service de la finance de demain”, i.e. “Artificial intelligence, blockchain and quantum computing for finance”. We decided to focus on A.I. and blockchain, because we had the sentiment that they would have more short term transformative impacts than quantum computing.

A big progress in our understanding of A.I. for finance stems from the observation of what the Turing Institute or the Center for Data Science of NYU put in place. This was in line with theoretical contributions of economists like Philippe Aghion, supporting the idea that A.I. is a “generic technology”. These technologies, like electricity or steam engines, need investments to generate “secondary innovations”. Without these supplementary efforts, one can only observe progress provided by A.I. for other usages or in other sectors, and fail to transfer them at minimal costs to her or his own domain. This reasoning supported the creation of “Instituts interdisciplinaires d’intelligence artificielle (3IA)” in France: one of the best way to foster the emergence of secondary innovations from A.I. is to put together experts of data sciences with invited experts of a specific domain with an identified topic to work together on, during a specified time to build an operative proof of concept. This proof of concept is then taken by the invited experts back in their department; not only this instance will spread, but these experts will now be able to have more ideas about potential innovations from A.I. around them. On the blockchain side, our views have been shaped in several steps: first it is obvious for any professional that the blockchain itself has direct applications in the “post trade” space, it can prevent operational errors since it allows to operate simultaneously all the steps involved in a transaction. It will have an impact on the operation of middle and back offices of financial and insurance companies. Thinking about such a future raises questions about how many “blockchains” should sustain this new ecosystem; the answer clearly lays in the realm of the Network Economy. A third component of the impact of blockchain is to create new networks of trust, that can be for instance very useful for insurance companies or to assess some ESG scores. And the last consequence of the emergence of the technology of blockchains is the need of a stable coin: to leverage at one hundred percent the efficiency of a blockchain, it is needed to be able to operate money in it. It is under optimal to host this money outside of the blockchain, since it will introduce a potential source of failure of a transaction, hence one needs to find a way to host "coins" with a value very close to a basket of external currencies inside the blockchain.

Last but not least, Institut Louis Bachelier wanted to keep a focus on risk management and regulation.

During the last two years; our main tools to improve the "advanced knowledge" on A.I. and blockchains for finance and insurance have been seminars and roundtables. A standalone workshop hosting academic talks and roundtables is co-hosted by FaIR and the French regulator ACPR; it gave birth to a series of transcriptions of "FaIR roundtables" that are widely used in this report to shed light on different topics of interest. Other roundtables have been generally hosted by conferences co-organised by Institut Louis Bachelier or by institutions close to it (like Euronext or the AFGAP). Frequent exchanges with the finance department of the Turing Institute, under the auspices of a partnership linking FaIR to this great Institute, has been very fruitful in the process of delivering insight about A.I..

This report gives an overview of the different topics that have been addressed during the last 18 months. The covid pandemic added some inertia to the work of putting together extracts of content produced by FaIR. Nevertheless, thanks to the effort of the ILB staff, readers will be able to make their own views on the implications of these technologies for the finance and insurance sectors. These contributions come from academics and professionals involved in designing innovations using these technologies. The choice of topics and speakers has been positively influenced during discussions with the International Scientific Committee of FaIR, whose members gently provided contributions to this report.

CHARLES-ALBERT LEHALLE

SENIOR RESEARCH ADVISOR, CAPITAL FUND MANAGEMENT,
SCIENTIFIC DIRECTOR OF FaIR AND VISITING RESEARCHER,
IMPERIAL COLLEGE LONDON

A FOREWORD FROM THE EDITOR-IN-CHIEF

This edition of FaIR's Advances is the first annual report of the Finance and Insurance Reloaded (FaIR) interdisciplinary program hosted at the Institut Louis Bachelier.

This report is a collection of contributions of world renowned experts that I have had the chance to organise into this 3-part report. They present their point of view on how the finance and insurance landscape is, and will be, impacted by new technologies. From robo-advisors to machine learning and blockchains, this report explores some deep revolutions currently reshaping the distribution of financial products, the intermediation of risk and how the financial sector gets insights from the real economy.

First and foremost, I would like to thank all the contributors of this report. The members of the International Scientific Committee - ISC - (Darrell Duffie, Alexander Lipton, Peter Kolm and Cris Doloc) have taken the time to share their valuable insights on FaIR's topics of interest. Other contributors include researchers with close relationship to FaIR (Lukasz Spruch, Blanka Horvath, Michel Crouhy and Aimé Lachapelle) as well as researchers with direct relationship to Institut Louis Bachelier (Christine Balagué and Christian Robert).

Second, I would like to thank Charles-Albert Lehalle, scientific director of the FaIR's program, for giving me the opportunity to take part in the editing process of this report. Lastly, I also take the opportunity to thank the entire editing team at Institut Louis Bachelier for their dedicated work.

This report, like the FaIR program, is organized around 3 chapters. The first chapter focuses on how new technologies will impact the distribution of financial products to end-users. It is oriented toward the outside of the financial system. The second chapter is concerned with the inside of the financial, namely how financial actors will intermediate the risk within an increasingly digital world. The third chapter explores how financial institutions can leverage recent technologies to better understand the underlying economy, to better connect to the real economy.

Overall, each chapter is composed of all the relevant contributions from FaIR's ISC, other researcher's contributions, and excerpts from FaIR's round tables. Indeed, during its first year of existence, six round tables have been organized by FaIR. This report presents some excerpts from the transcript of those round tables, put in the context of each of the three chapters.

The editing team hopes this report will shed light on important transformations that the finance and insurance industries are currently facing. FaIR's research community is dedicated to delivering a comprehensive understanding of those transformations, and is committed to supporting all the parties involved, financial actors as well as national and supra-national agencies.

LOUIS BERTUCCI

RESEARCHER AT INSTITUT LOUIS BACHELIER
EDITOR-IN-CHIEF OF THIS REPORT

THE FAIR INTERNATIONAL SCIENTIFIC COMMITTEE (ISC)

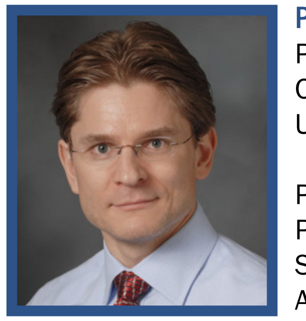
The Finance and Insurance Reloaded Transverse Program leverages on the academic network of Institut Louis Bachelier. Since it is of paramount importance to have advices from experts outside academic working under the umbrella of the Institute, the FaIR program is fortunate to be in touch with experts from the best universities in the US (Stanford, New York University, University of Chicago and the Massachusetts Institute of Technology) in the domains of post-trade, Blockchain, robo-advising and the use of new technologies on financial markets.

Cris DOLOC

Cris Doloc is an accomplished Quantitative & Computational technologist with more than 25 years of experience in Enterprise Software Architecture, Machine Learning and High Performance Computing. He holds a PhD in Computational Physics and he has spent the last two decades in the field of Quantitative and Computational Finance working for several top-tier financial firms. As such He has been the Chief Technology Officer of Terra-Nova Financial, the Head of Valuation Infrastructure at Chicago Trading Company and the Founder & Principal of Quantras Research Ltd. He is currently developing FintelligeX which is an Ed-Tech platform and Quantitative and Computational Learning Academy. He is also teaching at the University of Chicago in the Program of Financial Mathematics.

Darrell DUFFIE

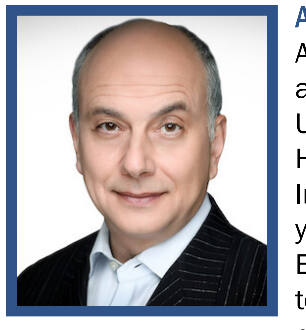
Darrell Duffie is the Adams Distinguished Professor of Management and Professor of Finance at Stanford University's Graduate School of Business and Professor by courtesy in the Department of Economics.
Duffie is a Research Fellow of the National Bureau of Economic Research and a Fellow of the American Academy of Arts and Sciences. He was the 2009 president of the American Finance Association. Duffie is an independent director on the board of U.S. Dimensional Funds. He chaired the Financial Stability Board's Market Participants Group on Reference Rate Reform. Duffie is a Project Advisor of The G30 Working Group on Digital Currencies and a member of the Systemic Risk Council.
Duffie's research focus is the design and regulation of financial markets. His most recent books include How Big Banks Fail (Princeton University Press, 2010), Measuring Corporate Default Risk (Oxford University Press, 2011), Dark Markets (Princeton University Press, 2012), and Fragmenting Markets: Post-Crisis Bank Regulations and Financial Market Liquidity (forthcoming, 2020).



Petter KOLM

Petter Kolm, Director of the Mathematics in Finance Master's Program and Clinical Professor, Courant Institute of Mathematical Sciences, New York University.

Petter Kolm is the Director of the Mathematics in Finance Master's Program and Clinical Professor at the Courant Institute of Mathematical Sciences, New York University. Previously, Petter worked at Goldman Sachs Asset Management where his responsibilities included researching and developing new quantitative investment strategies. Petter has coauthored many articles and books on quantitative finance and financial data science and financial data science, serves on several editorial boards for academic journals, professional associations, and company advisory boards. He holds a Ph.D. in Mathematics from Yale, an M.Phil. in Applied Mathematics from the Royal Institute of Technology, and an M.S. in Mathematics from ETH Zurich.



Alexander LIPTON

Alexander Lipton is Co-Founder and Chief Information Officer at Sila, Partner at Numeraire Financial, Visiting Professor and Dean's Fellow at the Hebrew University of Jerusalem, and Connection Science Fellow at MIT.
He is an Advisory Board Member at numerous Fintech companies worldwide. In 2016 he left Bank of America Merrill Lynch, where he served for ten years in various senior managerial roles including Quantitative Solutions Executive and Co-Head of the Global Quantitative Group. Before switching to finance, Alex was a Full Professor of Mathematics at the University of Illinois and a Consultant at Los Alamos National Laboratory.
His current professional interests include FinTech, particularly, applications of distributed ledger technology to payments and banking, digital currencies, including stablecoins and asset-backed cryptocurrencies, robust asset allocation, automated investing, balance sheet optimization and industrial-strength risk management systems for large systemically important financial institutions.
In 2000 Alex was awarded the first ever Quant of the Year Award by Risk Magazine. Alex published eight books and more than a hundred scientific papers. He is currently finishing his next book (with Adrien Treccani) "Blockchain and Distributed Ledgers: Mathematics, Technology, and Economics" which will be published in the beginning of 2021.

Moreover, the FaIR Transverse Program has access to experts from the Alan Turing Institute in the UK thanks to a partnership with this institute.

GENERAL SUMMARY

Scientific Director of FaIR: Charles-Albert Lehalle
Produced by: ILB Communication & Engagement, Paris 2021
Editor in chief: Louis Bertucci
Editorial committee: Charles-Albert Lehalle, Feriel Boushaki and Stine Hansen
Design: Myriam Kasmi
To cite this report : «FaIR's Advances Report, edited by Louis Bertucci, Institut Louis Bachelier, 2021»

DISTRIBUTION OF PRODUCTS
TOWARD THE OUTSIDE OF THE FINANCIAL SYSTEM

NEW TECHNOLOGIES PROVIDE A BETTER INSIGHT ON CLIENTS' NEEDS

One of the most important challenges of the financial system is to leverage new technologies in order to provide a tailored experience to end-users. Like many other industries, the financial system is subject to heavy competition from small new businesses (startups - fintech) as well as from new forms of infrastructures.

The increase in competition is partly due to the advent of new technologies and it has led to a tremendous amount of innovation. To be more specific, new technologies refer to the advent of artificial intelligence and process automation but also, maybe more importantly, to the wide availability of communication networks through internet-connected devices (smartphones and high speed internet, but also secure communications, online banking, etc.). This has opened the way for more direct innovation toward the outside of the financial system, toward end-users.

Financial companies, including those from the insurance sector, have expertise in transforming and transferring risks, however small and new companies are coming after the user experience. Large companies have to adapt and find a way to innovate not to be hit too much by newcomers.

On the one hand, the client experience is put at the center of the decision and the development process. Along with other sectors, clients of the financial sector are being more and more accustomed to a personalised experience when they interact with the financial system. One of the most important instances of such a personalized experience are robo-advisors. Here "advisor" has to be understood in the general sense. The goal is not only to advise clients on the investments they should make, but the experience can be much more personalized. From tax loss harvesting to investment based on life events and life goals, all aspects of the investment process can be tailored to fit the needs of each client.

On the other hand, the whole process may be automated with limited human intervention, which induces important cost savings. AI, coupled with automation, allows new companies to provide a better user experience with less costs. Nonetheless, financial actors are not currently replacing every human with algorithms but rather use algorithms to leverage the capabilities of humans. This allows portfolios to grow in size beyond what a single individual can manage by himself, and allows internal trading decisions to be assisted by technology.

The impact on the insurance sector has also been very important. Despite the insurance business being somewhat different, clients usually interact a lot with their insurance provider. Therefore, a straightforward way to improve user experience is with claim processes. Because of their very nature, insurance companies are also using new technologies like AI to refine their segmentation models. Segmentation is a real challenge because the regulatory framework mostly does not allow it to be performed at a very granular scale, even though insurance providers (or even other companies, like GAFA) may have enough data to do this.

On the infrastructure side, blockchains represent a big competitor to the financial sector. Tokenization allows the creation of new asset classes that have the ability to reach an even larger set of investors and issuers. Financial institutions need to adapt to this new infrastructure that may become a new paradigm.

This poses important challenges to regulatory agencies. Indeed, fintech are currently attacking several functions of the traditional financial system at the same time. They are able to use the technology in such a way that it induces an unbundling of the core functions of banking. The heavy regulation imposed on banks could be one reason for which fintech are able to gain some market share in the competitive sector.

This chapter features a set of two short papers including one from a member of the International Scientific Committee (ISC) and one from a member of one of FaIR's research programs, as well as a condensate of relevant excerpts from FaIR's round tables. First, Petter Kolm (ISC) from NYU shares his view of the current and future state of robo-advisors. Second, Christine Balagué from Institut Mines-Telecom business school presents her view on personalisation and improvement of the user experience.

ISC CONTRIBUTION

INVESTMENT ADVICE IS ONLY AN APP AWAY : ROBS-ADVISORS TODAY AND TOMORROW

Petter KOLM,
Director of the Mathematics in Finance Master's Program and Clinical Professor, Courant Institute of Mathematical Sciences, New York University.

1. WHAT ARE ROBO-ADVISORS?

Advances in fintech have led to the development of easy-to-use online platforms referred to as robo-advisors or digital-advisors, offering automated investment and portfolio management services to retail investors.[1] By leveraging algorithms embodying well-established investment principles and the availability of liquid ETFs for many asset classes, robo-advisors automatically manage client portfolios that deliver similar or better investment performance at a lower cost as compared to traditional financial retail services.

As of 2020, assets under management (AUM) by robo-advisors world-wide is about $990B and is expected to grow at an annual rate of about 26% to a total of $2.5T by 2024 (Statista, 2020). U.S. is by far the leading robo-market with about $680B in AUM and 8.8M users, where the largest firms include Vanguard Per sonal Advisor Services with $140B in AUM, Schwab Intelligent Portfolios ($40B), Wealthfront ($20B), and Betterment ($18B) (Statista, 2020; Backend Benchmark ing, 2020c).[2]

First entering the investing scene after the financial crisis in 2008, robo-advisors represent perhaps the greatest innovation in retail investing since the introduction of index funds and discount brokerages in the seventies and ETFs in the early nineties (Kador, 2002; Lynch and Rothchild, 1996; Wiandt and McClatchy, 2001). While their initial success is often attributed to the interest they received amongst millennials, who are open to new online technologies and mobile apps, today's average robo-advisor client in the U.S. is in their 40s and has an account balance around $20K, suggesting a broadening of the client base (Kaya, 2017; Hayes, 2019).

2. HOW DO ROBO-ADVISORS WORK?

By and large, investors are responsible for the choices they make when selecting investments. However, because of time constraints and complexity of investment decisions, many retail investors rely on professional advice and services, whether from more traditional registered financial advisors and broker-dealers, or from the more recent robo-advisors. The majority of robo-advisors in the U.S. limit their advice to portfolio management, including that of allocating investments amongst asset classes, and may otherwise provide only limited advice about other aspects of financial planning, such as estate and retirement planning, and cash-flow management.

2.1. Core Principles of Retail Investing. Broadly speaking, whether a novice or seasoned retail investor, there are several core investing principles that apply to them, including:
• Establish an investment plan and objectives,
• Seek broad diversification,
• Weight investment cost and value,
• Account for taxes.
Establishing an investment plan involves much more than picking a few stocks, mutual funds or ETFs to invest in. For instance, it is important to consider one's current financial situation and

goals, investment timeline and risk appetite.
First coined by Miguel de Cervantes, the author of Don Quixote, in the early 1600s, “don’t put all your eggs in one basket” is perhaps the most well-know metaphor expressing the investment advice of the benefits of diversification. Many decades of academic research and centuries of practical experience have shown the importance of diversifying across asset classes and investment themes. Seeking broad diversification should be one of the main goals of any retail investment strategy. In the absence of views about asset returns, the most diversified portfolio is the market portfolio.
Any investment process needs to weigh the cost of exposure to the assets with their added value. While this is common sense, it is “easier said than done” as costs come in many different forms, such as transaction and trading costs, management and advisor fees, and taxes to name a few.
For retail investors tax considerations and tax optimization are critical as a significant part of their investments are held in taxable accounts. We will come back to this topic below.

2.2. From Investing Principles to Algorithms. It is not surprising that robo advisors have incorporated the core principles from Section 2.1 in their services and product offerings. From a broad perspective, most robo-advisors have integrated them as a three step process:
(1) Client assessment and onboarding,
(2) Implementation of the investment strategy, and
(3) Ongoing management of the investment strategy, where each step is fully, or at least to a large extent, automated.

2.2.1. *Client Assessment and Onboarding.* Most commonly, a robo-advisor provides a new client with an automated online survey to collect general and investment specific information, including their age, net worth, investment goals, risk capacity and risk tolerance (Andrus, 2017). Goals may include generating income, saving for retirement, planning for large future expenditures (such as the purchase of a house or car) and/or establishing a financial safety net for emergencies. This automated evaluation represents a big change as compared to traditional wealth management services that rely on an initial in-person consultation. While less personal, the online assessments are less time-consuming and offer great cost-savings over the traditional approach. It is no wonder that many clients prefer the simplicity and ease of those robo-advisors who have onboarding processes that take no more than fifteen minutes to complete (Lo, Campfield, and Brodeur, 2018).

2.2.2. *Implementation of the Investment Strategy.* Robo-advisors use algorithms embodying well-established investment principles to automatically construct and periodically rebalance a diversified portfolio of the largest and most liquid ETFs. The most commonly used approach by robo-advisors to build and manage portfolios is based on modern portfolio theory (MPT), also referred to as as mean-variance optimization (MVO), introduced by Markowitz (1952). In its simplest form, MPT provides a framework for constructing a portfolio by choosing the amount to invest in each asset such that the expected return of the resulting portfolio is maximized at a prespecified level of risk as measured by portfolio volatility. Some robo-advisors use various extensions of MPT to incorporate additional features in their mod els such as transaction costs, tax lots, other risk specifications and the ability to incorporate subjective views in the portfolio construction process.
Robo-advisors in the U.S. predominately use passive ETFs in their offerings. Passive ETFs can be traded at low cost throughout the day and be used efficiently for tax-loss harvesting (TLH). The purpose of TLH is to realize current losses that can later be offset against future gains, as described below.
Moreover, the robo-advisors’ usage of automation and passive investment strategies reduce risks of internal agency conflicts and conflicts of interest that traditionally could arise between

financial advisors and their clients (Inderst and Ottaviani, 2009). Needless to say, conflicts of interest cannot be completely eliminated as they naturally arise in situations of asymmetric information, including that of robo advisory.

2.2.3. *Ongoing Management of the Investment Strategy.* An important aspect of robo-advisory services is the ongoing monitoring and rebalancing of client portfolios. The algorithms perform rebalancing in three main situations: (a) when due to market movements, portfolio holdings drift too far away from their desired target allocation; (b) when tax-loss harvesting opportunities are identified; and (c) when clients update their preferences. However, as robo-advisors design their portfolio strategies for the long-term, annual turnover is low.
Robo-advisors permit clients to override their algorithms in several ways, including changing portfolio allocations and modifying their risk profile and preferences. Unnecessary client overrides can trigger significant trading and tax consequences that are far from optimal. Many robo-advisors use realtime automatic messages on their platforms and mobile apps to warn their clients against this behavior.

3. AUTOMATED TAX MANAGEMENT FOR RETAIL INVESTORS

One of the core principles of retail investing is that any portfolio strategy should be tax optimal. There are many ways in which to address taxes, some simpler and others more complex. Besides diversification and automatic rebalancing, one of the most valuable services robo-advisors provide is automated tax management.

Constantinides (1984) showed that it is optimal to realize capital losses in stocks immediately and to defer capital gains for as long as possible. A common approach is to take the strategic asset allocation as given and overlay “tax algorithms” either at the account level or the trading level to generate so-called tax alpha through tax-loss harvesting and asset location.

3.1. Tax-loss harvesting. Tax-loss harvesting (TLH) is a strategy where assets held in the portfolio at a loss are sold opportunistically (Jeffrey and Arnott, 1993; Stein and Narasimhan, 1999). When losses are realized, they can be used to off set future capital gains and income from dividends. While TLH traditionally has been performed on a quarterly or annual frequency, computer algorithms can con tinuously monitor and act on opportunities as they arise. Hence, robo-advisors can significantly increase the number of TLH opportunities relative to traditional investment managers.
The so-called wash-sale rule can make TLH a complex endeavor. A wash sale occurs when you sell an asset at a loss and then buy that same asset or “substantially identical” assets within a window of thirty days before or after the sale date. The wash-sale rule was designed to discourage investors from selling an asset at a loss to claim a tax benefit, again and again.
One can navigate the wash-sale rule in a few different ways. A basic approach is to sell a ticker at a loss and buy a “dual” ticker to replicate the exposure of the original ticker. For instance, one might sell the holdings in an ETF tracking the Russell 1000 index and later buy back an ETF tracking the S&P 500.
Using ETFs for TLH works well for several reasons. First, ETFs require autho rized participants to create and redeem shares in kind, thereby enabling the ETF to avoid selling securities and realizing capital gains taxes to meet redemptions. Second, the ETF manager can reduce the ETF’s tax liability by providing the au thorized participant with the tax lots that have the lowest cost basis. With the dramatic increase of highly liquid ETFs in different sub-asset classes, identifying dual tickers is easy.
On the flip-side of TLH is the need to realize long-term capital gains to reset the cost basis, thereby preparing for higher short-term capital losses in the future. Naturally, there is a tax cost to recognizing the capital gain, but doing so restarts the holding period and provides

the investor with the option to recognize future capital losses (Stein, Vadlamudi, and Bouchey, 2008; Yang and Meziani, 2012).

3.2. Asset Location. While maybe not as broadly known as it should be and often misunderstood, asset location is another important tax overlay strategy. Asset location should not be confused with asset allocation.

The idea of asset location is that assets have different after-tax profiles independent of whether they are held in taxable or tax-advantaged accounts. In other words, the after-tax return of an asset can be very different if held in a tax-deferred, tax-exempt or taxable account. For example, the coupon payment on a bond held in a taxable account is taxed as ordinary income. In this instance, if possible, an investor should instead hold the bond in a tax-exempt account. Consequently, as retail investors may have a number of taxable, tax-deferred and tax-exempt accounts, some robo-advisors can assist clients by optimally allocate assets preferentially to the different accounts so as to maximize the after-tax return while the overall strategic allocation is maintained (Khentov, 2016; Huang and Kolm, 2019).

4. TRENDS AND DEVELOPMENTS

The fintech landscape continues to evolve rapidly. Main trends in the robo advisory space include greater personalization and customization, improved inte gration of services and platforms, and increased automation. In the modern world, autonomous finance has become the new norm, where algorithms assist us in mak ing financial decisions, invest in a more disciplined fashion in line with our prefer ences, ensure we save long-term for retirement and smartly manage our outstanding debts. A recent study suggests that close to 60% of the U.S. population will be using robo-advisors by 2025 (Schwab, 2018).

4.1. Expanding Service Offerings. In this new world where robo-advisors, and technology companies more broadly, may become the gatekeepers of the access to banking services, an expansion of services is crucial in order to compete with traditional banks. Many robo-advisors have moved in this direction and are of fering services including cash and checking accounts, debit cards, lending and retirement services (McCann, 2020). Recognizing that automation cannot replace human touch everywhere, some robo-advisors are offering customers financial advice from certified financial planners on staff who can assist in making decisions such as how to start investing; address significant life events (changing job, having a child, purchasing a home, etc.); plan for college, marriage and retirement, to name a few.

4.2. Is Smart Beta Investing Ripe for Robo-Advisors? A well-trodden path in investment management is that innovative products are first introduced in institutional contexts, and only with a significant delay are they later, gradually, made available in the retail space. Such has been the case with smart beta offerings. However, today, liquid smart beta ETFs are more broadly available. Surprisingly though, there are few smart beta products available amongst robo-advisors.
In the institutional smart beta space, smart beta ETFs have seen increased interest due to improved technology, reduced costs and an evergrowing body of empirical evidence of what drives underlying risk premia. Because of this evolution, the retail advice market today has a strong foundation to build upon when implementing smart beta solutions. Huang and Kolm (2019) argue that smart beta is ripe for the retail audience and discuss some of the challenges in implementing smart beta in robo-advisory offerings.
Based on a recent survey, Agather and Gunthorp (2018) suggest that smart beta products are increasingly popular amongst financial advisors across Canada, U.K. and the U.S. for the purpose of diversifying client portfolios and to express strategic views. We believe that

the continued increase in liquid smart beta ETFs will provide robo-advisors an opportunity to increase portfolio customization to better fit investment objectives and financial goals of retail investors.

4.3. Goals-Based Investing. As discussed earlier, maintaining a broadly diversified investment portfolio is a core investment principle for retail investors. As each individual has different financial goals and risk profiles, there is no “one size fits all.”
Financial theory prescribes that an individual’s total assets and liabilities should be considered when constructing their investment portfolio. Also the more illiquid assets, such as that of an individual’s home, should be included in the robo-advisor’s asset allocation decision. This will result in a portfolio of assets with lower corre lation to the housing market.
It is well-known that individuals do not treat all of their investments the same, but rather practice what is referred to as mental accounting (Thaler, 1985; Thaler, 1999). In particular, it has been demonstrated that individuals will attach different “chunks” of money to different risk-return preferences, depending on how they see that money being used in the future. An individual may view their homeownership differently from that of their stock portfolio. For example, they may tolerate a larger loss in their stock portfolio, but not willing to risk losing their home.
Goals-based wealth management is an investment and portfolio management approach that focuses directly on investors’ financial goals. Shefrin and Statman (2000) suggest, in behavioral portfolio theory (BPT), that investors behave as if they have multiple mental accounts. Each mental account has varying levels of as piration, depending on its goals. BPT proposes a portfolio management framework where investors are goal-seeking (aspirational) while remaining concerned about downside risk. Specifically, rather than to trade off return versus risk as in MVO, investors should trade off goals versus safety. Not surprisingly, BPT leads to nor matively different statements about the optimal portfolio than those based MPT (see, for example, Das, Ostrov, Radhakrishnan, and Srivastav (2018)).
Based on its intuitive appeal and ability to model individual investor’s financial goals in a flexible and highly customizable way, we believe that goals-based wealth management will continue to evolve and may emerge as the predominant approach in retail investment management.

4.4. Responsible Investing. When considering investment choices, individuals have in recent years gradually become more concerned about socially responsible investing (SRI) and pay greater attention to the environmental, social and gover nance (ESG) practices underlying an investment.
There are several challenges for robo-advisors here. First is to find liquid ETFs that can be used for constructing SRI/ESG-aware portfolios. Second, as investor’s SRI/ESG preferences vary, it is difficult for robo-advisors to provide the customiza tion needed with the currently limited number of SRI/ESG ETFs. Some people argue that SRI/ESG preferences may be best customized at the stock level rather than through ETFs.

4.5. Financial Literacy and Investor Education. Rich with technical terminology and jargon, for many of us financial services and products can be intimidatingly complex. In a modern society investor education is needed in order to improve individuals’ financial literacy. At the core of financial advice and planning is communication (Lusardi, 2015).
Needless to say, there are costs to society for a financially uniformed population as we may end up paying for our mistakes. Unmistakably, there are clear benefits in having citizens that have healthy financial habits and plan for their future. Who should provide such education? First, perhaps middle and high-schools could adopt course content on basic personal finance. Second, financial advisors and other fiduciaries could educate their clients about investment management and planning for retirement. Third, employers could play a role in providing employees with information and continued education about personal finance.

5. HOW DO ROBO-ADVISORS AND THEIR CLIENTS REACT IN BEAR MARKETS?

The introduction of the first robo-advisors in 2008 and their subsequent expansion has coincided with the longest bull market in history. Consequently, over the years there has been concerns about how robo-advisors and their clients will react when markets goes bearish. We have had to wait for more than ten years for a market shock such as that caused by the COVID-19 pandemic during the spring of 2020 where the S&P 500 index fell more than 25%. Specifically, Backend Benchmarking (2020a) reports that from January through mid-March in 2020, robo-advisors declined on average 17.6%, as their equity and fixed income holdings fell about 29.2% and 2.3% on average, respectively. Not surprisingly, the robo-advisors with broader diversification towards treasuries and investment-grade fixed income fared better than their peers.

The significant market drop provided tax-loss harvesting opportunities for robo advisors. For example, Betterment and Schwab traded over 60% of their accounts to realize losses on behalf of clients (Backend Benchmarking, 2020).

During this period, some robo-advisors notably increased their direct client en gagements in several ways, including the usage of blog posts and online videos, to educate them about market volatility and the importance of investing with a long-term perspective, thereby discouraging clients from moving assets or overrid ing the algorithms. Those robo-advisors that followed this path appear to have been successful in helping their client stick to their investment objectives.

Some robo-advisors reported that an increased number of clients reached out for advice from financial advisors on their staff. For instance, client engagement with Betterment advisors was 50% higher in March as compared to January in the same year (Carlson, 2020).

6. CONCLUSIONS

More than ten years in the making, robo-advisors have continued to grow both in size and product offerings. The main reasons contributing to the continued success of the robo-advisors include:

• Low cost. Fully automatic algorithm-driven management of client port folios that significantly lowers the cost of financial advice and wealth management.

• Personalization and customization. By providing a general investment management framework and a suite of financial services, robo-advisors can in a highly scalable fashion customize investment strategies and provide a digital banking experience that suits the specific needs of each individual investor.

• Anywhere, anytime convenience. People have gotten used to accessing their digital lives and beyond through mobile apps on their phones and laptops. Robo-advisors provide their clients with this convenience for their investment portfolios and other aspects of their financial life; anywhere, anytime.

• Wealth management services for the masses. Robo-advisors are making many sophisticated investment and financial advisory services, that in the past were only accessible to high net worth individuals, available to all retail investors at low cost.

In closing, during the spring of 2020 robo-advisors demonstrated they can stay on course and support the needs of their clients during a period of significant market turmoil. This recent experience demonstrates that the robo-advisory industry is maturing and able to offer a broad suite of investment management and financial advisory services to their clients, whether in up or down markets.

REFERENCES

Agather, Rolf and Peter Gunthorp (2018). Smart Beta: 2018 Global Survey Findings from Asset Owners. Tech. rep. FTSE Russell. url: https://investmentnews.co.nz/wp-content/uploads/Smartbeta18.pdf

Andrus, Danielle (2017). 4 Ways Robo-Advisors Improve Client Onboarding. url: https://www.thinkadvisor.com/2017/06/16/4-ways-robo- advisors-improve-client-onboarding/.

Backend Benchmarking (2020a). How Do Robos React During a Bear Market? url: https://www.backendbenchmarking.com/blog/2020/03/how-do-robos-react during-a-bear-market/.

(2020b). Robo Advisor News Q1 2020. url: https://www.backendbenchmarking.com/blog/2020/03/robo-advisor-news-q1-2020/.

(2020c). The Robo Report - First Quarter 2020. url: https://www.backendbenchmarking.com/the-robo-report/.

Robo-Advisors Kept Investors Calm During Tumult - With Help From Humans at Times (2020). url : https://www.barrons.com/articles/tesla-stock-has-surged-kimbal-musk-made-his-first-sale-of-shares-in-2020-51591441237

Constantinides, George M. (1984). "Optimal Stock Trading with Personal Taxes: Implications for Prices and the Abnormal January Returns". In: Journal of Fi nancial Economics 13.1, pp. 65–89. Das, Sanjiv R., Daniel Ostrov, Anand Radhakrishnan, and Deep Srivastav (2018). "A New Approach to Goals-Based Wealth Management". In: Journal Of Invest ment Management 16.3, pp. 1–27.

Hayes, Adam S. (2019). "The Active Construction of Passive Investors: Roboadvi sors and Algorithmic 'Low-Finance'". In: Socio-Economic Review, pp. 1–28.

Huang, Lisa and Petter N. Kolm (2019). "Smart Beta Investing for the Masses: The Case for a Retail Offering". In: Equity Smart Beta and Factor Investing for Practitioners. Ed. by Khalid Ghayur, Ronan G. Heaney, and Stephen C. Platt. Wiley. Inderst,

Roman and Marco Ottaviani (2009). "Misselling Through Agents". In: American Economic Review 99.3, pp. 883–908.

Jeffrey, Robert H. and Robert D. Arnott (1993). "Is Your Alpha Big Enough to Cover Its Taxes?" In: Journal of Portfolio Management 19.3, pp. 15–25.

Kador, John (2002). Charles Schwab: How One Company Beat Wall Street and Reinvented the Brokerage Industry.

John Wiley & Sons. Robo-Advice – A True Innovation in Asset Management (2017). Deutsche Bank Research.

Khentov, Boris (2016). Asset Location Methodology. url: https://www.betterment.com/resources/asset-location-methodology/

Onboarding: The Imperative to Improve the First Experience (2018). Broadridge. url: https://www.broadridge.com/white-paper/onboarding-the-imperative-to-improve-the-first-experience

Lusardi, Annamaria (2015). "Financial Literacy: Do People Know the ABCs of Finance?" In: Public Understanding of Science 24.3, pp. 260–271. Lynch, Peter and John Rothchild (1996). Beating the Street. Simon & Schuster. Markowitz, Harry M. (1952). "Portfolio Selection". In: Journal of Finance 7.1, pp. 77–91.

McCann, Bailey (2020). "Robo Advisers Keep Adding On Services". In: Wall Street Journal. url : https://www.wsj.com/articles/robo-advisers-keep-adding-on-arms-11583331556?shareToken=st1c2c56b0fc8e4db3b2fe78860caf2490

The Rise of Robo: Americans' Perspectives and Predictions on the Use of Digital Advice (2018). Charles Schwab.

Shefrin, Hersh and Meir Statman (2000). "Behavioral Portfolio Theory". In: Journal Of Financial And Quantitative Analysis 35.2, pp. 127–151.

Statista (2020). Robo-Advisors - Worldwide. url: https://www.statista.com/outlook/337/100/robo-advisors/

Stein, David M. and Premkumar Narasimhan (1999). "Of Passive and Active Equity Portfolios in the Presence of Taxes". In: The Journal of Wealth Management 2.2, pp. 55–63.

Stein, David M., Hemambara Vadlamudi, and Paul W. Bouchey (2008). "Enhancing Active Tax Management Through the Realization of Capital Gains". In: The Journal of Wealth Management 10.4, pp. 9–16.

Thaler, Richard H. (1985). "Mental Accounting and Consumer Choice". In: Marketing Science 4.3, pp. 199–214. – (1999). "Mental Accounting Matters". In: Journal Of Behavioral Decision Making 12.3, pp. 183–206.

FaIR NETWORK CONTRIBUTOR

Christine BALAGUÉ

Christine Balagué is professor and chairholder of the Good in Tech Chair (www.goodintech.org) at Institut Mines-Télécom Business School on responsible digital innovation. Her research focuses on modeling the behavior of connected individuals (social networks, artificial intelligence algorithms, connected objects), as well as on technology ethics and more specifically artificial intelligence.

Christine Balagué is also a member of the CSA expert committee on online disinformation, a member of the defense ethics committee and a member of the impact commission on the recommendations of the Haute Autorité de Santé. She is also a member of the board of the Dataia convergence institute (Data Sciences, Intelligence and Society) and of the Board of Directors of Cap Digital. As Vice-President of the French National Digital Council from 2013 to 2016, she has participated in various projects submitted to the French government on the major digital challenges (Net Neutrality, Platform

Neutrality, e-inclusion, e-education, e-health, Digital Ambition). She is also the author of more than 90 publications in international scientific journals and conferences, as well as several books on digital and society. Christine Balagué has a PhD in Management Sciences from HEC, a degree from ESSEC and a Master's degree in Econometrics from ENSAE. She is also recipient of the French National Order of Merit.

FaIR NETWORK CONTRIBUTION

GOOD IN TECH'S VISION ON «CUSTOMIZATION OF FINANCIAL PRODUCTS AND CUSTOMER EXPERIENCE IMPROVEMENT».

Christine BALAGUÉ,
Professor and chairholder of the Good in Tech Chair (www.goodintech.org) at Institut Mines-Télécom Business School on responsible digital innovation.

Interactions between customers and companies in the financial sector are increasingly covered by new technologies. Artificial intelligence and data analysis tools are playing an growingly important role, aiming to improve customer experience and develop financial product customization. However, these technologies raise ethical concerns. For instance, there are currently many automatic tools for granting bank credit using machine learning. These tools create a score for each credit applicant using increasingly customized and targeted data. A growing academic literature in social sciences shows that these tools can be biased and discriminate against certain minorities. In addition, credit allocation decisions made by a Machine Learning algorithm can be opaque, rendering it difficult to provide an explanation to the credit applicant following an attribution decision. Lastly, the increasingly personalized targeting of these automatic tools requires the acquisition of more and more data on bank customers, which raises privacy and confidentiality concerns. Regarding this issue, one of the winning projects of the Good in tech February open call develops an «ethics by design» Machine Learning algorithm interface for the attribution of bank credits. It will consist in creating an interface enabling arbitration between these different ethical issues (fairness, interpretability and privacy) in order to create efficient credit attribution algorithms. In addition, it will create a workshop-type approach to open a debate on these algorithms by associating technical and non-technical bank staff in the configuration of these tools.

Another subject concerns the customization of financial services through intelligent voice assistants such as Google Home or Alexa. These systems, which dominate their market, are increasingly being adopted (75% of American households will be equipped with them by 2024) because of their user-friendliness. This variable has been known as a major factor in the field of technology acceptance models (TAM models) for more than thirty years now. However, voice assistants are nowadays black boxes. This raises ethical issues of explicability, transparency, and non-discrimination measures through their response selection algorithms.
Similarly, many ethical concerns are also raised by recommendation systems based on machine learning or deep learning, which create barriers and present endogenous biases.

Finally, technologies such as blockchain deserve to be studied in terms of ethical and responsible implications. Indeed, this technology based on distributed systems could be considered as more ethical and responsible. However, the opacity of these systems, and the algorithms making up the blockchain, also raise ethical concerns that need to be addressed.

Several works in progress of the Good in Tech Chair deal with these different topics. The contribution of the Good in Tech Chair to the FaIR program would help deepen these research themes, participate in interdisciplinary work, and develop collaborations with other existing Chairs within the Institut Louis Bachelier.

Contacts :
- Web site : goodintech.org
- Christine Balagué, Professor at Institut Mines-Télécom Business School, Good in Tech Co-Chair, christine.balague@imt-bs.eu
- Marie-Laure Salles-Djelic, Professor, Dean of Sciences Po School of Management and Innovation Good in Tech Co-Chair, marielaure.djelic@sciencespo.fr

ROUND TABLE SPEAKERS

Pierre-Alain DE MALLERAY
CEO, Santiane

Aimé LACHAPELLE
Aimé is co-founder and Managing Partner of Emerton Data, a company specialized in AI and data transformation. He has over 13 years of experience in initiating and leading large data & AI projects, in multiple environments from start-ups to large groups.
Before founding Emerton Data, Aimé was successively Head of Pricing Innovation with Axa and Director Data Science with Capgemini Invent. He conducted end-to-end projects involving data and technology, with a focus on value delivery, mainly in the financial and industrial sectors. Aimé has developed a deep expertise in insurance, and functional expertise in innovation, data-driven pricing, advanced analytics and machine learning. Aimé holds a PhD in Applied Mathematics from Université Paris-Dauphine.

Philippe TRAINAR
Philippe Trainar is Professor of the insurance chair at the Conservatoire National des Arts et Métiers (National Conservatory of Arts and Crafts) and Director of the SCOR Foundation for Science.
He is a member of the Boards of Directors of the SCOR Foundation for Science, the Toulouse School of Economics and Humensis and of the Scientific Committee of the French Prudential Control and Regulatory Authority and of the Scientific Council of the Maurice Allais Foundation.
He is President of the Risk Commission of APREF and member of the editorial boards of the Revue Française d'Economie, the magazines «Commentaire, Risques et Sociétal», and the «Revue d'Economie Financière».

David DUROUCHOUX
Deputy CEO Société Générale FORGE - Client relationship, business development, CBDC & central banks coverage.

Mark SINSHEIMER
Mark is an independent consultant and trainer to asset management firms, institutional investors and financial regulators on international best practices. He is also Associate Professor of Finance at Skema and EDHEC Business Schools. Previously, Mark had board level responsibilities with global asset management companies from 1996 to 2002 as Head of Marketing and Sales for Credit Lyonnais Asset Management and Head of Business Development for CDC Asset Management. Before, he was Product specialist for Paribas Asset Management and Trader for Crédit Lyonnais.
Mark has a BA in History, MA in International Relations and a Graduate Degree in Law from Paris V University. He is a CFA charterholder and passed CAIA level 1 & 2.
Mark founded CFA France and was its Advocacy Chair until 2011. He is a an active volunteer for the CFA Institute where he served on the Annual Conference Committee, Global Advisory Committee, European Advocacy Committee, Candidate Curriculum Committee, Standards and Practice Council and Advisory Council of INSEAD's "Global Investors Workshop". He was a member of the Governing Council of the European Asset Management Association where he chaired EAMA' working group on index construction best practices. He is regularly interviewed by the financial press or invited to speak in forums and conferences.

Philippe MAUPAS
Co-founder of Alpha & K and Blogger

Matteo RAVA
Matteo contributes to ESMA's work on the business conduct aspects relating to the provision of investment services and activities by investment firms and credit institutions in the retail financial services markets, and on investor protection more generally.

Jean-Philippe BARJON
Head of Coordination for Sales Management Supervision,ACPR, Autorité de Contrôle Prudentiel et de Résolution

Sandrine LEMRY
Sandrine Lemery is a CNAM professor holding the Chair of Actuarial Science and Vice-President of the Institute of Actuaries. She is also the non-executive President of the Supervisory Board of the Pension Reserve Fund.

NEW INFRASTRUCTURES TO BETTER SERVE CLIENTS

Throughout it's one year of existence, the FaIR program has had the opportunity to gather a unique set of round tables, featuring some of the world's best researchers. Here is a list of the round tables that are relevant to this chapter :
- Insurance, **05/09/2019, Conservatoire National des Arts et Métiers, Paris, organized by Toulouse School of Economics**
- Robo-advisor, **16/01/2020, ACPR, PARIS,** Organized by Fair
- Resolution of post-trades via a distributed ledger, **20/05/2019, ACPR, PARIS,** organized by Fair

New technologies and infrastructure are putting a lot of pressure on traditional actors from the financial system. However, different sectors react differently to such competition. While the insurance sector follows mostly a slow-paced process of innovation, banks have had to find ways to stimulate innovation. Here follows a discussion on both of these topics. The first one is from the Insurance round table where Pierre-Alain de Malleray and Aimé Lachapelle discuss about how new technology has impacted the insurance business. The second one is from the tokenisation round table in which David Durouchoux explains how Societé General has been trying to stimulate internal innovation.

ROUND TABLE: EXCERPT 1
NEW CHALLENGES IN INSURANCE

05/09/2019,
CONSERVATOIRE NATIONAL DES ARTS ET MÉTIERS, PARIS,
ORGANIZED BY TOULOUSE SCHOOL OF ECONOMICS

Pierre-Alain DE MALLERAY

"Hello everyone, my name is Pierre Alain de Malleray. I am the current CEO of a brokerage company. I speak here on behalf of the distribution point of view. Prior to that, I worked in the public sector, similarly to Sandrine, and then I actually worked a few years in reinsurance with Philippe. So, perhaps let's start with a short answer to the very vast and important question: has insurance been disrupted? I think that this is a no-brainer. I believe the answer is definitely no. Yet, when we look a little bit more closely at what disruption could be like in our sector, we can zoom in on different aspects of the value chain, distribution and tariffs segmentation. There is a craze on a very hype subject as to other big data, as well as statistical learning, deep learning, and data in general that people leave on the internet with their smartphones etc. The question is : will insurances be able to approach the risk more closely than today?

Distribution, tariff, segmentation and user experience is another means to disrupt an economic sector. There are a lot of sectors around that are disrupted by user experience and more vastly by the product itself. Can the product be really different than what it is today, apart from the pricing? My answer to all these sub-questions is the following: when we look at the distribution issue, we have to answer differently according to the country, e.g. France or the United Kingdom. The US is a bit different. But let's take these two examples which have taken really different paths. When you look at what is the weight of the distribution over the internet, and here

I am talking about insurance for individuals, I am not talking about insurance for companies. But the first policy was sold in the UK in the year 2000 over the internet, and when I say over the internet, it is a mix generally between either 100% internet or internet+telephone. I globalised those two means of distribution channels under the same umbrella. The first policy was sold in the year 2000 and by the end of the 2010 decade, 70% of motor insurance policies were sold over the internet. So there was really a huge massive disruption of the market regarding distribution.
The UK market is very much brokerage driven. Insurance policies are generally fixed term ones, so individuals must comparatively yearly review on the internet to make quotes, etc. Furthermore, aggregators have played a crucial role in this evolution because they were literally nonexistent during the 2000s decade. They have progressively invested a lot in branding, press commercials, movie commercials, on television, etc. In the middle of the decade, around 2005, the investment of aggregators in commercials were higher than the investments of incumbent insurance companies. They waged a war and they won that war regarding direct access to consumers. So this market has literally shifted. When we look back at our country, i.e. France, there has been very slow disruption, so I don't think the word disruption is the right one. It is more like a slow evolution, regarding the market share of each distribution channel. It's around 70% in the UK. What is the figure in France? Well it is more or less 10-12%. So 70% corresponds to 10-12% in France. When you look at other industrialized countries it is more or less around 15%. This is not really a surprise, because when you look at the other economic sectors, such as retail, other goods' distribution, or the market share of direct internet channels, it's more or less around 15%. For your information, when I say internet, it is either on a smartphone or the desktop, or a mix between these channels. So, market shares of these means of distribution, like new direct internet, is around the same as other economic sectors. Yet, because the insurance product is highly dematerialized, e.g. you don't need to go to a store to buy it, we could assume that the internet distribution growth is going to grow at quite a high pace. The actual growth rate of these channels in France is more or less 10-15% a year, so the number of people who go on the internet to buy insurance products has a growth rate of roughly 10-15%. That is what I wanted to say regarding distribution. Let me briefly explain what we do here at Santiane. Santiane is a broker on the French health insurance market. We exclusively sell Santiane over the internet with a mix of internet and telephone. We are now the number one e-broker. I think one interesting thing to know is that you don't buy insurance on the internet like other products. Insurance products are complexe. They usually can scare people away, especially when you talk about issues such as health, you want to know exactly what you are buying.

You could go to a hospital where you need to have medical care from a specialist doctor, etc. You need to understand the terms. You need to compare one contract to another. Consequently, you don't just buy it just with a simple click on your smartphone. The consumer, and French consumers, need to be coached by an advisor. If this advisor is independent from the insurance company, then it works best because they inspire trust. And here at Santiane, we have a team of advisors who coach new clients. They stay on the phone for a long time, for an average of 1 hour and 15 minutes per call. So it is not the call center you can imagine with short conversations, etc. It is really an intensive pedagogical kind of call. As far as the other topics are concerned, perhaps we can address them later on, with products such as big data, etc."

Aimé LACHAPELLE

I have been working for around 5 to 6 years in the insurance sector, mainly in pricing with Axa global direct, direct insurance in France. Since then, I have been working in the consulting business, where I basically spend half of my time on AI and data transformation topics in the insurance and banking industry and the other half on the same topic but for other industries.

So, I'm going to tackle this question very quickly here. I have seen very profoundly what has happened within Axa, and also for slightly more than 10 clients in the last 3 years, in 10 insurance companies. Therefore, what I have observed is that insurance is an information and technology business by essence. It is a fair tight ground but then the segment is the same, i.e. it is shared, there is no disruption that was announced or expected. Perhaps it's because it was wrong to announce or expect it. Finally it is continuously evolving, like in many other industries. For instance, there is a similar situation in other sectors, such as the banking industry, the automotive industry, etc. People are doing things, step by step, with all of the usual barriers, i.e. how to industrialize a new technology that is used to doing AI and data. It is not the same technology. We need to go for the clud. Clud for insurance is not an easy question because there are a lot of personal data. Yet something is happening. As Microsoft announced, there was a big deal in Switzerland involving UBS and Swissfree. These are huge moves of these companies for the clud that will facilitate the transformation. Then there is a clear opportunity at all levels of the industry. At the core business level, even if it is tricky, for instance for claims : how to simplify and make more cost effective claims processes? There is obviously pricing augmentation, and then we go for a new paradigm with more incentive schemes with telematics, for instance. There are also all things customer experience-related. So, I heard Pierre Alain that you mentioned that the byclick is not the right thing, but maybe there is a claimclick that could be done. You are in a car, you just push the claim button and everything is done and settled by your risk partner who is your insurer. So the opportunity is clear, i.e. more disruption, no revolution. Regarding strong barriers, we already mentioned some of them. So for new entrants they need reserves to come, it is a very complexe product, so many things need to be managed. You have all the risk management aspects. You have technological barriers but still some changes, which we will come back to later. I would also like to jump in on what Pierre Alain said on the UK market, that there was a clear disruption on the brokering side. I think that the figures are even higher now. The latest figures are more than 90%, with motor figures that are bought through aggregators, i.e. internet aggregators. You mentioned several factors that allowed for this disruption, so I think there is probably also a more open regulation in the UK that encouraged this type of evolution. Something happened 3 or 4 years ago. There was a very sophisticated insurer with Admiral Group who announced that they would use Facebook information data in their pricing, and the regulator didn't block this, Facebook did. So I think this tells us a lot about what is going on in this country.

Philippe TRAINAR

Yes, thank you very much. In fact let me come back to the background. What is digitalisation and what is digitalisation in insurance? When looking at digitalisation, it is economically interesting because it reduces costs, and in turn reduces costs through former channels, the first one being big data. We know that data is all of the current economy and it reduces the cost of tracking and searching data. The second dimension is the internet of things. What is the novelty of the internet of things? It can provide a data cluster, i.e. reducing largely the cost of surge in tracking again. You have a bijection between the different data in the cluster. It is accepted not to do any statistical analysis to infer bijection. Then you have the third pillar that is robotics and artificial intelligence, which replicates young behavior and which in other words causes verification, replication, etc. It is very dangerous for the most skilled labels in our economies. Finally, you have the last pillar, which is very interesting because it is there when disruption happens in different economies. It is a kind of comeback of the economy which brings them closer and closer to supply on an industrial level. It has been seen in the disruption of different arising sectors, yet we did not see it in insurance in France. The reason is because distribution in France is largely integrated to the agency system that is dependent from one company to another and integrated in the insurance banking system. In some health insurance companies, employees directly sell products. In France, brokers have no place or no important place

in the retail market, except the dynamic actor which is probably the future of the market with Pierre-Alain. Consequently, we have not seen any revolution within the retail distribution in France for it is very difficult for it to happen. We have to indeed change everything before a revolution can happen.
May I just complete this very quickly. I think effectively that the digital revolution is some kind of silent revolution. What we see is that insurance companies are using more and more big data and the internet of things, as well as robotics. They are beginning to use artificial intelligence, and I don't know what Aimé would say about that. But it is more progressive. It is really progressive because at the same time, insurers are cautious, as there are a lot of liability problems when using, for instance, artificial intelligence or internet of things and therefore they just progressively advance, step by step, in order to be sure they are on secure ground.

ROUND TABLE: EXCERPT 2
RESOLUTION OF POST-TRADES VIA A DISTRIBUTED LEDGER

ROUND TABLE « ADDING VALUE TO FINANCIAL MARKETS WITH THE BLOCKCHAIN »:
PROMOTE BREAKTHROUGH INNOVATION BY DAVID DUROUCHOUX
20/05/2019,
ACPR, PARIS
ORGANIZED BY FaIR

David DUROUCHOUX

There's a press release dated from mid April, regarding a Security Token, for which we worked in partnership with PwC . We are very happy with this participation and we had a lot of feedback from them. Thank you very much Sébastien. The initial idea originated when the bank's General Management decided to launch a call for projects to all employees, looking for disruptive business. Everything was eligible.

It was a pretty nice mix of formalwear and casual wear. The goal was to bring out and directly sponsor by executive committee members ideas that would drive the bank's businesses in the upcoming years. We have a very high regulatory influence and offer solutions with a very high technical debt, at a time when global competition is heightened. It is the consequence of the business strategic projections in decline. The idea behind this initiative at Société Générale was to see how, through new technologies, particularly the Blockchain and artificial intelligence, it would be feasible to develop innovative solutions.

It was also an opportunity to give people the chance to step out of their usual teams' box, to give them a budget, and to see what they could come up with in terms of «start-up mode». These volunteers received support and coaching from bank's managers, who consequently were also confronted with their own business choices. The idea was to see if they too were capable of going beyond their habits and to push them towards breakthrough innovations.
Approximately 1,000 projects have been submitted with 60 to 70 funded. And we're one of about 20 or so that have come through, with two start-ups that have merged. Our project was based on the observation that the technical Proof of Concept on the Blockchain part era was outdated. So the question was how to move into commercial production, in order to bring out real business issues. It was therefore a question of understanding how to provide the bank with a commercial offer based on digital assets. Our aim is above all to understand how to continue to sell products in such a context. Whether the technology is R3, Ethereum, H-Graph, etc, it doesn't matter, since, as Jean Safar pointed out earlier, technologies are in fact evolving very rapidly. Our goal was to understand how concretely speaking our traditional origination processes could benefit from these new technologies.

Since it's complicated, when you're not a bank, to be a Fintech that truly operates in a regulated market, we chose to position ourselves within Société Générale. This way we could understand, from the inside, the regulatory complexity that digital assets' marketing could face. Today, our areas of work are based on a regulation context and on the understanding of the adapted commercial positioning.

Our objective is to provide our customers with new offers, based on Blockchain of course, and to see how smart contracts can help us improve. For instance, there is obviously transparency provided to the customer. An Automated Transfer Agent would risk disintermediating us ourselves. But we need to tackle this issue head on. We could theoretically talk about the Blockchain for hours on end. Some people do it very well. The problem is the organization. You have to get something out in production, that's what's very challenging. That's why we wanted to take advantage of start-ups to release a viable product with just a small but meaningful number of features.

With respect to the mid-April bond issue, as described by Sébastien, it is Société Générale, a refinancing subsidiary that issues a $100 million bond amount, and Société Générale parent company that buys back these bonds.

It is nevertheless a real test that we practice widely. Innovation in relation to what we have seen elsewhere is firstly on a Public Blockchain. We're not on R3, Ripple, etc., or the Settlement Blockchain. Secondly, the token carries the legal privilege, i.e., the holder of the token obviously owns the rights to the bond.

With respect to the mid-April bond issue, as described by Sébastien, it is Société Générale, a refinancing subsidiary that issues a $100 million bond amount, and Société Générale parent company that buys back these bonds.

It is nevertheless a real test that we practice widely. Innovation in relation to what we have seen elsewhere is firstly on a Public Blockchain. We're not on R3, Ripple, etc., or the Settlement Blockchain. Secondly, the token carries the legal privilege, i.e., the holder of the token obviously owns the rights to the bond.

NEW CHALLENGES FOR RISK MANAGEMENT AND REGULATORS

Although the focus is usually made on financial actors, new technologies are also a great challenge for regulators. The main concern is that the regulatory framework should put strong barriers in order to protect the consumer while giving the ability to financial institutions to keep innovating. This question is particularly important for the robo-advisor business. Indeed, since the goal is to advise consumers on their investment, the regulator really wants to make sure the robot acts in the best interest of the client. On the insurance side, it is more a question of privacy. In this case, the regulator wants to prevent misuse of personal information in the move to segmentation.

Here follows two excerpts of FaIR round tables. Each addressing the regulatory challenges regarding robo-advisors, STO and insurance respectively. The first one is from the Robo-advisor round table. The second one is from the Insurance round table in which Philippe, Aimé Lachapelle and Sandrine Lemry talk about segmentation and the corresponding regulatory issues.

ROUND TABLE: EXCERPT 3 ROBO-ADVISORS

16/01/2020,
ACPR, PARIS,
ORGANIZED BY FaIR

Mark SINSHEIMER

I'll speak as a regulator, but I'm really curious as to what other panelists have to say. We looked at this quite carefully. And teaming for best protection, we really put ourselves on the client side, rather than on the firm side. So, how we approached it was to regulate at least in Europe, and I'm thinking specifically of MIFID for the security sector. It sets a series of controls and elements that have to be executed by an intermediary when it performs certain investment services such as investment advice or portfolio management. What we didn't want to do was lower some of these controls or simplify them based on the technology you are using because that is not relevant for the client, contrary to the nature of the service they are receiving. Therefore, for us, the issue is not really to set different rules or requirements according to whether you are providing human advice, or robo advice, because as previously mentioned this can change tomorrow. It's not even robo-advice, it's an app, or something else. It can be done over the phone, you can have a million models. What we find challenging as regulators and what we are trying to do is instead, is to guide robo-advisors and explain how to apply current regulation. We acknowledge that this regulation has been, willingly or not, drafted with the traditional banking model in mind. So, sometimes it can be challenging for the robo-advisor because they don't think that way, they're not structured as a typical bank. To sum up, once again, from a regulatory perspective, the issue is not really about setting different rules but about working together, with the industry, to provide guidance and help them, and make the regulatory aspect not an surmountable barrier but help them comply with existing requirements.

In the United States there is this issue regarding fiduciary responsibility, especially when it relates to DC pension funds, etc. For instance, regulators have roughly taken your approach but the department of Labor has been a little bit embarrassed because it's hard to trust a robot. So, here is where we define whether it's pure play, or not. If you trust an individual, who knows how to use a robot, then it's business as usual. Yet, if it's something that is purely algo-driven or it pretends to be so, then it could be business as usual, but some people can feel uncomfortable about it all.

Matteo RAVA

In Europe, at least we are one step ahead on that topic, because there is a clear obligation for intermediaries to act in the best interest of the client. This is not the case in the US. And not for all services, and that is why they tried to fight. Industry is resisting. I think the advantage in the US is probably that the level of financial knowledge, i.e. the average financial understanding of clients is much higher than the average retail client. I come from Italy and let me give you the following example. My mother has a degree in Nuclear Physics, so quantum physics comes easily to her but if you asked her the difference between a bond and a share, she wouldn't know what to answer. So, again, compared to the US, it's quite striking I would say.

Philippe MAUPAS

I don't really agree with your reference to the US. They have a fiduciary standard which doesn't exist here, and registered financial advisors who are rarely subject to this standard. Broker dealers aren't held to suitability standards which, unless I am very wrong, is the standard that prevails in Europe. Even we have the expression «intermediaries have to work in the best interest of the client», provided they provide suitable advice.

Matteo RAVA

Yes, but the best interest is legally a step before and you have to, of course, provide in Europe suitable advice which means matching the investment objective. Yet, there is always the higher presumption that, however you apply those details and requirements, there is always the best interest for the client, which, to me, is always broader in Europe in terms of legal requirements than in the US, in terms of actors to whom it applies to.

Mark SINSHEIMER

Do you see a problem regarding authorization and monitoring robo-advisors? How is inspection done for instance?

Matteo RAVA

My short answer is that the phenomenon in Europe is still very small in terms of numbers. The reality is that, for them, it still has very little impact in the overall supervisor activity, and they tend to supervise more by services, again, rather than by specific business models.

Jean-Philippe BARJON

Yes, I can confirm what Mateo said. Clearly, the general stance of supervisors is to be technologically neutral. I mean our mandate is to monitor and to supervise implementation of the rules regarding consumer protection. So, we are focused on this subject. The question of whether a kind of player will flourish in the next decade is not exactly our subject.

Indeed, this is not an issue for us, as long as these players, will comply with the rules regarding consumer protection.

Mark SINSHEIMER

The main issue is supervision fraud, negligence, etc. if it's human, it's like «business as usual», but if it's purely algo robo, I mean, honestly, you've got to evaluate the model risks. So what is the supervision model, how it is tested and monitored? How is it researched? You need to look at data risk, contingency risk, continuity risk, coding risk, etc. The best risk manager in the world, possibly Barr Rosenberg, who received numerous medals for his achievements, was banned from markets, because of a coding error in his firm. So, small errors, big consequences.

Jean-Philippe Barjon

Yes, I agree with you. But this is the discussion, automation creates a kind of systematic risk. I agree with you, it's obvious. On the other end, it is clear as well that it is more easily auditable.

ROUND TABLE: EXCERPT 4
NEW CHALLENGES IN INSURANCE

05/09/2019,
CONSERVATOIRE NATIONAL DES ARTS ET MÉTIERS, PARIS,
ORGANIZED BY TOULOUSE SCHOOL OF ECONOMICS

Philippe TRAINAR

Yes of course the segmentation question mentioned by Pierre Alain is largely political. The real question is : if you don't know your risk, I don't know my risk . What is the most profitable behavior of an insurer? It's not up to my segmentation to just calculate the mean contribution you should pay. Using big data for market segmentation in this case is just a loss of money. It doesn't change the premium level. Therefore I think that isn't the question. The right one is : do you know the risk, and do I know my risk ? Obviously, if I am drunk and driving, I suppose you will never be favorable to an insurer who would ask you to pay for me. You would say ok, I am in shape and he is a bad driver, by definition. Then we know that if we don't do this segmentation, and you and I know my risk, normally the insurance market should obviously be destroyed. Therefore, this is the reason why insurers are fearing the fact that we know our risk, but the insurer does not know your risk. In that case, it is very important. You have a very well known economics theorem that has been applied, not in insurance, but in banking and it is actually the same. You cannot differentiate with interest rates a good payer from a bad payer and who will pay or not their debt. Therefore, the only consequence is either that the market is happy or that you are very cautious offering very small credit, and in insurance very small covers. Therefore, maybe you should segment, and the challenge is not for insurers. The challenge is potentially for the government. It is therefore the redistribution policy of the government that is impacted, where the boundaries between insurance and redistribution are not so clear. Hence, here you have a real challenge that you have to address, and of course big data will later help you do what you can't. You know your risk, I don't know your risk. Will big data help me solve this problem? I say, in most cases NO. To me, the problem lies in my answer.

Therefore why isn't it a question of big data? For instance if you are using google for health issues, I don't know if you are looking at data for yourself, for your wife, for your kids, for your grandmother, your grandfather etc. Therefore, if I offer you, and not, say, your grandfather, the wrong price, you will refuse, the deal will be off and I will have lost a client. Consequently, because there is no bijection between most of the data, I will be able to gather some data and your risk profile, and most of the time I will not be able to use it correctly. However, it is very costly to gather big data. Even if we have seen the reduction in gathering data in France, it remains very costly. Therefore, I have no incentive in terms of profitability if I am not sure of the bijection between data and your risk profile. Hence, I think that the segmentation question is not well understood, especially as politicians are confusing the question of redistribution and insurance. Therefore, at this point in time we are not correctly dealing with the question of segmentation. We should believe in separating the efficiency of the insurance market and the question of the efficiency of redistribution that is the consistency of redistribution with political targets and deals in different ways.

Aimé LACHAPELLE

Ok, thank you. I would like to react to what you just said Philippe. I think it depends on, first the paradigm, what are we insuring and what are the risks? I think that this is completely true when you talk about short term risks, i.e. you are the best at knowing your risks. For example, you know if you drink or not. My perspective on the yearly contract insurance is that I know less about my risk than what the machine learning model will say about my risk. Because it is just impossible for my mind to aggregate all the behavior I will have during a year when I am driving my car. Therefore, I am unsure if we stay with the scheme where we have a yearly contract. Insurers can disappear because we don't know our own risk as well as complexe algorithm processing, even with imperfect data. The second point is segmentation - and you say this is a question of efficiency, and I fully agree - because now this is a market scheme, so players are following these rules. Hence, segmentation also remains because all players are going along with it. There is probably a limit point. Today the capitalistic short term view is that the limit point is slightly better than others. The long term view could be different. As I mentioned earlier, with IoT and telematics, and scores that are impacting prices, you can downsize prices. The grading scores are better, you have incentives, and you will reduce the overall claim level. So, you forget about segmentation when you are doing this. Therefore, my point is that this limit point depends on the point of view.

Philippe TRAINAR

May I just react very quickly. My point is that, if you don't know your risk, and if I can't know your risk, I have no incentive to know it. Why? Because I will spend money trying to know when I could simply effortlessly apply the mean tarif. You will accept it because you don't know your risk. And Pierre Alain will also accept it because he does not know your risk. The only problem is if a significant fraction of my clients know their risk very well, then I have to segment in order to be sure that I am efficient. Sandrine?

Sandrine LEMRY

And if I may add, there are already segmentations which are very known, for instance in men and women. But they have been forbidden by European regulation, so you are not allowed to discriminate on death or lifespan for men and women. So actually segmentation has not always been used and sometimes the regulation wants indeed to prevent this because the social contract is important. What you were saying, Philippe, on redistribution vs. mutualisation is not so obvious. So there has already been a lot of very well known segmentation. A case against segmentation, that has already been partly addressed here, is whether its cost is worth it. At one point it wasn't. For instance, the insurance stereotype is that actuaries set the tarif. They say, for instance, well it is 100, and then the sales rep wants to sell the product, and they say «ok, but I can sell it 70», and then the real price of the product is not worth as much compared to what you can sell it for. So lots of effort has been put into selling more, and then deciding the right cost of the product. The other thing I've seen on segmentation is that all insurers, when you look at their business plans, have twice the market. It is because everyone makes huge projections using the same strategy, i.e. I have clients on this line of product, and my new strategy is to equipe my client with all kinds of insurance. That is what we have seen for instance with bank insurers that are now coming from health into retail insurance, etc. So the strategy of an insurer is somehow segmentation, but it is also having the entire risk in order to deliver a good service and also to compensate the cost of risks between one branch and another. Therefore, what I would say is that currently segmentation can be worth only as a niche which is not what we have observed. On the contrary, I would say what we previously stated, i.e. that insurance is a national market. The European commission has tried time and again to make it a European market, but insurance is really dependent on taxes for savings. It is dependent on lots of issues, for health, etc. For instance, I would take the example of construction. In construction, it is exactly the opposite that has happened. In France, we have a construction market where insurers sell construction contracts to companies. Construction companies that carry a bad risk, ask them for a tariff that is very high or they refuse. Then we have foreign insurers, who came from Gibraltar and lots of other places, who started selling at the same price, or a bit less, than the other companies with huge commissions that were not obvious to the clients and we sort of had a segmentation at the opposite ends: bad risk going to companies in free service provision. These companies that were not supervisors, if they were in France, would go completely bust, and now we have a huge problem with this issue.

It is an interesting topic, because we seem to say in France that nothing is going on, but is it dependent on the French market. Actually, we could already have some European actors coming from abroad and selling with all these capacities. What Pierre Alain was saying about how GAFAs are not ready to enter the insurance market, is that they really understand the product and insurance products are not really standardizable. We speak a lot about determining the right price, but the thing is, how to design the right product for the client? This is at the time being not only a question of data. It is more complicated than that, and perhaps it will change. There are huge possibilities for change, but nothing is shifting yet.

INTERMEDIATION OF RISK
INSIDE THE FINANCIAL SYSTEM

IMPROVING RISK INTERMEDIATION THANKS TO NEW TECHNOLOGIES

The main goal of the financial system is to transfer risks in the economy. With recent technological innovations, financial actors and the financial system as a whole may find ways to better assess and move risks around in the economy. On the one hand, an increased digitalisation of financial products mainly due to recent developments in computer science and distributed ledger technologies creates new kinds of interactions between financial agents. With this, comes new regulatory possibilities and concerns. On the other hand, the recent boom in artificial intelligence offers new potential tools for financial actors to better extract structure in data that has the potential to allow a greater level of automation and understanding of risks.

Along those transformations may very well come an even more important change. Improved risk intermediation may also be accompanied with disintermediation. As competition increases, (from new financial actors, from GAFAM, as well as from alternative global borderless infrastructures), current financial actors may be at the risk of being disintermediate. With the correct technological design, the system as a whole could be more efficient at transferring risk with less intermediaries. This has the potential to fundamentally alter the value chain of risks.

Distributed ledger technologies (DLT) are only at their infancy, but they represent an important regulatory perspective. Most regulators around the globe are currently actively looking at the potential use cases of DLT. The main promise is a more efficient financial system, especially regarding the reduction of transaction costs that would be achieved with increased interoperability of existing systems. A regulator willing to implement such a system would have two options: 1) adding an interoperability layer that would connect to existing systems and make seamless transfers from one system to the others, and 2) creating a new global system that would be interoperable by design. The latter would most likely come with some degree of disintermediation. Current intermediaries may not have sufficient incentives to innovate in this direction (because they are intermediaries), so the regulators would have to tackle this question in the near future.

However, there are also new players (GAFAM) that are moving more and more toward financial applications (see examples in China : AliPay and Weechat - payment service providers). As they have been very effective at building global interoperable systems in the field of information transfer and communication, there is the possibility they will also be very effective at building a global financial infrastructure. With those tech companies but also smaller one, referred to as fintech, the traditional financial sector is under heavy competition on each of its functions independently. This unbundling of functions is very characteristic of the current context. Regulators will likely have to take decisions in this challenging context.

Alongside developments in infrastructure, the growing use of artificial intelligence in the financial sector can also have an important impact on risk intermediation. As more and more data are being collected (by both financial and non-financial actors) more complex models can be used to have supposedly better representation of the economy. The recent developments in generative deep learning models may even be used to leverage the amount of data collected. Indeed, generative models will soon be able to have a very good representation of the underlying distributions of real data so that they could be used to generate several other alternative histories of data that could then be fed to other deep learning models.

With a large amount of data, two of the main applications are Data-driven trading decisions and Deep hedging. While data analysis, and in particular machine learning, have been used to extract useful information from large amounts of data, being able to create actionable insights that can be used to automate trading decisions or hedging strategies is crucial for a lot of financial actors. Automation in decision making is not new of course, but big data and machine learning enables algorithms to exploit much more complex patterns.

Like almost any technological breakthrough, artificial intelligence induces a conceptual revolution. In many sectors, including financial applications, research in this area has been quite dense in the last decade. We are to a point at which new types of jobs are created in the financial sector that require new skills and therefore new education programs. The next decade will likely settle on the direction this field will take and the true useful applications enabled by this technological revolution.

In the remaining of this chapter, there are a total of five short papers, including three from the International Scientific Committee (ISC) and two from members of one of FaIR's research programs, as well as an extended abstract of a research paper and a condensate of relevant excerpts from FaIR's round tables. First for the ISC, Darrell Duffie (Stanford) shares his view on the interoperability of International payment systems notably through the use of a Central Bank Digital Currency. Second, Alexander Lipton (MIT) explains how DLT as technological innovations can potentially help modernise the current financial ecosystem as a whole. Third, Cris Doloc (University of Chicago) explores the impact of machine learning techniques improving to the point where it is possible to generate actionable insights for automated models in trading. This chapter then features two papers from the FaIR research community. First Lukasz Spruch (Alan Turing Institute) gives a technical overview of the recent developments in model based machine learning in finance. Lastly, Blanka Horvath presents her view on the transformative role of new technologies on market models, especially on market generators. In an extended abstract, Michel Crouhy, Dan Galaia and Zvi Wiener explore through a functional approach what is the fintech's impact on financial intermediation. In the last part of this chapter, some excerpts from FaIR's round tables are presented in the context of artificial intelligence and risk intermediation.

ISC CONTRIBUTIONS

INTEROPERABLE PAYMENT SYSTEMS AND THE ROLE OF CENTRAL BANK DIGITAL CURRENCIES

Darrell DUFFIE,
Adams Distinguished Professor of Management and Professor of Finance at Stanford University's Graduate School of Business and Professor by courtesy in the Department of Economics

Abstract: I explain the meaning of an interoperable payment system and why interoperability is crucial for efficiency. I review some alternative approaches to interoperability, including central bank digital currencies (CBDCs).

Keywords: Digital currency, payment system, interoperability

1. INTRODUCTION

Innovative payment technologies are transforming monetary systems, commerce, and banking. When new payment systems lack interoperability with each other or with important legacy payment systems, however, the result can be highly inefficient. Whether customer-to-business, business-to-business, or peer-to-peer, user costs and delays rise with the multiplicity of non interoperable payment methodologies. Infrastructure costs grow. Complexity increases. Financial intermediaries and financial market infrastructure may be unable to achieve significant netting of inflows against outflows when using different payment systems, and may therefore require inefficiently high cash buffers.

In this short note, I explore the essential meaning of interoperability and its implications for innovative payment systems, including hybrid or synthetic CBDCs. For brevity and focus, I abstract from many important policy factors such as financial inclusion, privacy, anti-money laundering and other legal issues, competition for payment services, monetary policy transmission, financial stability, and disruption of banking franchises.

While market forces associated with scale and network economies create incentives for convergence onto common or interoperable payment platforms, there are also "walled-garden" incentives for firms to limit interoperability, sacrificing payment-system efficiency in order to raise customer switching costs. Interoperability is a public good whose benefits are not fully internalized by each market participant. This presents an opportunity for central banks and other official-sector players to regulate standards for interoperability, or to provide their own general-purpose payment systems.

When every agent in the economy makes and receives payments in a common safe digital currency, interoperability is more easily achieved. For example, with a general-purpose CBDC in the form of central bank deposits, interoperability is dramatically simplified. Alternatively, with a CBDC that is transferred on payment systems operated by private-sector payment service providers, the central bank can enforce standards for maintaining interoperability. Yet the introduction of a CBDC raises a host of tradeoffs that have caused central banks to hesitate. Among these concerns are the disruption of the legacy commercial banking sector and the responsibility of the central bank for monitoring the legality of transactions and for securing private information. I will outline alternative feasible approaches for a highly interoperable and efficient payment system.

2. THE MEANING OF INTEROPERABILITY

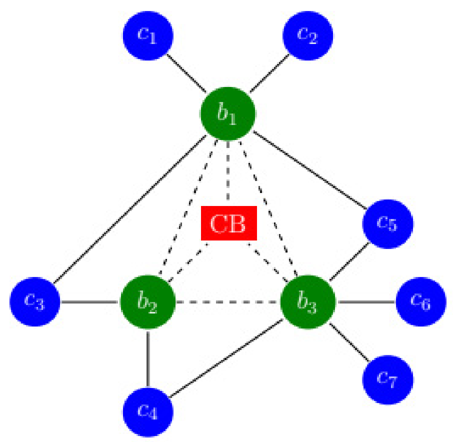
To get at the meaning of interoperability, we can think of a payment system as a collection of multi-account ledgers. Each ledger is capable of instant transfers between any two accounts on that ledger. Two different ledgers are interoperable if there is always at least one intermediary holding accounts on each of the two ledgers that automatically meets legal requests to transfer funds from any account on one of the ledgers to any account on the other. For example, a request to transfer from an account on ledger A to an account on ledger B is met by a transfer from the source account on A to an intermediary's account on A, and an equal transfer from the intermediary's account on B to the destination account on B. For interoperability to be effective, transfers should be done at negligible latency and user cost. The definition of interoperability is extended by calling two different ledgers interoperable if they are elements of a network of interoperable ledgers. (A network is a connected graph. In our setting, the links of the graph are defined by pair-wise interoper- ability)

A conventional two-tiered bank-based payment-system, as depicted in Figure 1, offers a degree of interoperability. The inner tier consists of a central bank (CB) and banks holding accounts at the central bank. In the inner tier, banks make most of their payments to each other by transferring deposits held in their accounts at the central bank. The outer tier consists of the banks and their customers, who can direct payments from their own bank accounts to accounts at other banks. The account ledgers of two different banks are thus interoperable via an inner tier settlement system, such as Fedwire in the U.S. or the Eurozone's Target2. In practice, however, the degree of interoperability is reduced, often significantly, by user costs and fees of various types, significant latencies (often more than a day), and limited time-of-day access. Some payment system authorities, including those of the United Kingdom and the Eurozone, have reacted by introducing "fast payment systems" that offer almost instant transactions, around the clock, with low user costs. Still, however, fast payments are not dominant in the U.K. and the Eurozone, and are in a much earlier stage of development in the United States.

A CBDC that is based on central bank deposit accounts for every user has only one ledger. With a broadly used CBDC, interoperability is therefore not an issue, given the relative ease of arranging for low-cost instant transfers across the entire economy.

Alternatively, a high degree of interoperability can be achieved by upgrading a conventional two-tiered bank-account-based payment system, of the sort depicted in Figure 1, by introducing a single common account ledger for the customer accounts of all banks in the system. This single outer-tier ledger is interoperable with a single inner-tier ledger, containing the accounts of all banks at the central bank. With this two-ledger approach, each bank is responsible for

Figure 1: A schematic of a two-tiered payment system. The inner tier consists of a central bank and banks. Banks make most of their payments to each other by transferring reserves held in their deposit accounts at the central bank. The outer tier of the payment system consists of banks and their customers, who can direct payments from their own bank accounts to the bank accounts of others



meeting its own deposit liabilities and is able to observe only the account information of its own customers. Although this two-ledger approach has yet to be adopted, it is already technically feasible and seems to satisfy the criteria of central banks that wish to improve the efficiency of their payment systems without introducing a general-purpose CBDC.

3. INTEROPERABILITY WITH CBDC

Until now, most central banks have shied away from providing CBDC directly to everyone in the economy. But there have been serious prototypes and proposals for "hybrid" or "synthetic" CBDC, by which private-sector actors would be responsible for payment services associated with one or more digital currencies backed by the central bank. Of the proposed variants of this model, two are prominent:

1. The central bank issues CBDC "tokens" to one or more payment service providers who redistribute the tokens to a broad set of customers, along with payment apps and other infrastructure. The CBDC tokens are analogous to paper currency, as direct claims on the central bank, but are transferred electronically. Token holdings are recorded in ledger accounts maintained by the central bank or by payment service providers. The payment service providers could be commercial banks or fintech firms. Interoperability requires that all actors in the economy are able to seemlessly transfer the CBDC to each other, implying a common payment technology or strong standardization. The Peoples Bank of China has begun to test a related form of CBDC called DC/EP (digital currency/electronic payment). The Bank of England has published its principles for a similar approach to CBDC.
2. As an alternative, which Adrian and Mancini-Griffoli (2019) call "synthetic CBDC," payment service providers can be permitted to back their own private-sector digital currencies 100% with deposit accounts at the central bank. In effect, each of the resulting private sector digital currency is in the form of narrow payment-bank deposits (I am a member of the board of directors of a proposed narrow bank, TNB Inc., which is not a payments narrow bank.). In this case, interoperability requires not only interoperable payment technologies, but also perfect fungibility of the various resulting private digital currencies. This raises various addi tional technical challenges that deserve a lengthier discussion.

4. FINAL REMARKS

I believe that fast highly interoperable payment technologies of some form will dominate some major economies within the next decade or so. These new technologies can be based on next generation bank-account-based payment systems, central bank digital currencies, or some hybrid or synthetic form of CBDC.

Most developed-market central banks continue to show a preference for increasing the efficiency of bank-account-based payment systems over the deployment of CBDCs, but have also become more open to discussing the potential benefits of the introduction of a CBDC.

STABLE COINS AND THEIR ROLE IN TRANSFORMING THE FINANCIAL ECOSYSTEM

Alexander LIPTON,
Co-Founder and Chief Information Officer at Sila, Partner at Numeraire Financial, Visiting Professor and Dean's Fellow at the Hebrew University of Jerusalem, and Connection Science Fellow at MIT

1. INTRODUCTION

It became painfully obvious during and after the Global Financial Crisis (GFC) that the existing banking and payment systems, are unstable, obsolete and no longer aligned with the constantly changing requirements of the real economy. Unfortunately, the GFC turned into a wasted opportunity to reorganize the world financial ecosystem. Too-big-to-fail banks dramatically increased (rather than decreased!) in size by amplifying their share of the banking business at the expense of their failed competitors, while the number of banks has shrunk. Superficially, banks are better capitalized, but their complexity increased to such a degree that their true financial positions cannot be ascertained with neither by regulators nor by depositors, investors, and, somewhat alarmingly, by their own management. Balance sheets of large banks which were opaque before the GFC became even more so and have added risks of unquantifiable complexity. As a result, many banks and other financial institutions have become too-big-to-manage. In addition, clearing and settlement of many over-the-counter derivatives was mandatory moved to central clearing counterparties (CCPs), making CCPs potential points of failure for the system as a whole. A high level of interconnectedness of CCPs due to the fact that they have many common general clearing members adds an extra degree of instability.

Not surprisingly, the general public is frustrated with the status quo. This frustration manifests itself in various aspects of economic and social life and almost universal aversion to banks and banking. Thankfully, companies harnessing new technologies, including blockchains and distributed ledgers, can put much needed competitive pressures on the incumbents, help newly formed fintech companies to enter the market, and provide considerable benefits to the general public. Assuming that regulators encourage competition and properly adapt regulation, we expect to see contests between fractional reserve and narrow banks; physical and digital cash; fiat currencies and asset-backed cryptocurrencies; and, most importantly, centralized and distributed payment systems. The outcome of these contests will reshape the entire financial ecosystem going forward, and, in a few years, might change it beyond recognition. If properly executed and regulated, a financial system of the future will be much better at payments, products issuance and personalization of financial services, and much closer connected to the real economy. In addition, it will become significantly less risky.

2. THE BANKING SYSTEM

2.1. Overview

While Internet Protocols (TCP/IP) have been responsible for the phenomenal proliferation of the Internet, which inspired and promoted completely new business models in numerous fields, such as digital commerce, social media, communications, etc., banking lags way behind. One of the biggest impediments in unleashing creativity and developing new business models in banking and finance, rather than mending existing legacy businesses, is the absence of successful TCP/IP for money and identity. In particular, in the 21st century, we still do not have a perfect electronic analogue of paper currency, which allows for an instantaneous and anonymous transfer of value without involvement of a middleman.

2.2. Credit Money Creation and Annihilation

In order to improve the system, we need to understand how it operates at present, and, first and foremost, the true nature of the process of money creation and annihilation. In our opinion the so-called modern monetary circuit theory, which claims that money is created by commercial banks when it is lent to their clients and destroyed by banks when it is repaid, provides the most convincing explanation of the process, see [1], [6]. The corresponding interest comes from the next round of borrowing and stays in the system for good. In case when a borrower defaults, money is not repaid and destroyed as it should, and remains in the system forever, which is tantamount to forgery.

2.3. Bookkeeping and Transactions

As a result of the above process, the preponderance of money in circulation is just a sequence of transactions, organized in ledgers, which are maintained by various private banks, and by central banks providing means (central bank cash) and tools (various money transfer systems) for reconciling these ledgers. As part of ledger-maintaining and transactional functions for their clients, private banks play two very important roles, which central banks are not equipped to perform. They serve as the system gatekeepers, via know your customer (KYC) services, and system policemen, via anti-money laundering (AML) efforts. It is clear that, in addition to the better known areas of application of distributed ledger technology (DLT), for instance, in digital currencies (DCs), including central bank issued digital currencies (CBDCs), DLT can be used to solve such complex issues as trust and identity, thus helping to address the KYC and AML requirements. Moreover, since all banking activities boil down to maintaining a ledger, judicious applications of DLT can enable development of the distributed Financial Market Infrastructure (dFMI), with numerous promising applications in trading, clearing and settlement triad, payments, trade finance, and so on. However, if applied for its own sake, DLT can make the situation even less satisfactory than it already is.

2.4. Payments

The hallmark of the current financial system (if not its raison d'etre) is a long chain of middlemen engaged in moving money and securities between a buyer and a seller, not least in the form of correspondent banks. As usual, middlemen thrive in most situations, particularly when foreign exchange transactions are involved. The amount of various fees can easily reach 3% or more of the transactional amount, which is a very meaningful amount for most participants.

2.5. What is Wrong with the Current Setup?

The biggest problem with the existing banking system is that it is too complex for its own good, predominantly due to the fact that it commingles three distinct activities:
- creation and annihilation of credit money through lending;
- record keeping;
- execution of transactions.

Obviously, separating these activities is highly desirable in order to make banking less risky, nimbler and more efficient. Setting lending aside, we emphasize that judicious usage of DLT can be very helpful in bringing record keeping and transaction execution into the 21st century.

3. DISTRIBUTED LEDGERS

3.1. General Considerations

IBM defines a distributed ledger as follows, [3]: "Blockchain is a shared, distributed ledger that facilitates the process of recording transactions and tracking assets in a business network. An asset can be tangible - a house, a car, cash, land - or intangible like intellectual property, such as patents, copyrights, or branding. Virtually anything of value can be tracked and traded on a blockchain network, reducing risk and cutting costs for all involved."
While the idea of a BC is old, since BCs naturally occur whenever power, land, or property change hands, modern technology gives the old idea of BC a new lease of life [6]. DLT opens exciting opportunities for making conventional banking and trading activities less expensive and more efficient by removing frictions. Moreover, it has the potential for restructuring the whole financial system on new principles of dFMI. Obviously, achieving this goal requires overcoming not only technical but also epistemological and political obstacles.

In theory, payment systems using DLT can significantly reduce transaction costs to below 1% as they involve significantly fewer middlemen and have shorter paths for information flows. However, to build such systems in practice one needs to overcome several challenges, including regulation, scalability, privacy, and ease of use for all the relevant parties.
Potentially DLT can have numerous applications outside of payments, for example, in clearance and settlement, trade finance, rehypothecation, syndicated loans, and other areas where frictions are particularly high. However, before applying DLT to all these areas, one needs to appreciate that some current systems are shaped by business and other considerations much more than by pure technological reasons.

3.2. Cryptocurrency Creation and Transactions

Cryptocurrencies took the world by storm. Bitcoin, proposed in a seminal paper published in 2008 by Satoshi Nakamoto, [10], was the first in a very long line of cryptocurrencies. Some of Bitcoin's initial appeal was due to the fact that its creation was purely algorithmic in nature in sharp contrast to fiat. Shortly afterward, Ethereum, a second-generation, secure, decentralized computing system, was introduced by Buterin, [2]. Since 2008, several thousand of native cryptocurrencies operating on their own purpose-built ledgers, as well as tokens built on top of existing blockchain platforms, such as Ethereum, have been launched with varying degrees of success.
Bitcoin, Ethereum and, to a lesser degree, Ripple all tried to play the role of money but ended up as a new speculative asset class. While technically very impressive, existing cryptos cannot be money in a traditional sense for at least three reasons:
- they are not stable enough, even for short transaction times, which prevents their usage for purposes of conventional commerce,
- they are operationally inconvenient to use for an average economic agent, which is a prerequisite for their wide adoption,
- their built-in algorithmic monetary policies are too mechanical and inflexible to be practically useful.

Not surprisingly, the status quo in cryptocurrency land is not satisfactory and encourages further developments.

4. STABLE COINS AND THEIR TAXONOMY

4.1. Overview

One of the most exciting developments in this regard is the introduction of stable coins. The main advantage of a properly design stable coin is that it can serve as the cornerstone of a fully digital ecosystem facilitating fast and regulatory compliant payments on a commercial scale with minimal reliance on the existing banking, [11].
We feel that rather than designing monetary policy of its own, which is a challenging task in itself, innovators should initially aim at tokenizing existing stable financial instruments denominated in fiat currencies, such as US dollars. Later on, if these efforts prove to be successful, they should be able to tokenize more diverse asset pools, which can serve as a counterbalance to fiat currencies and will be very useful for cross-border transfers.
Contrary to numerous claims, at present it is not possible to build a truly decentralized, stable coin. In fact, the mere estimation of the price on a crypto coin in terms of a fiat currency requires an oracle and automatically breaks decentralization. Ideally, it should be possible to directly observe this price from the distributed ledger. This is not quite possible for both conceptual and technical reasons. Hence, any potentially successful stable coin has to combine centralized and decentralized features.

4.2. Coins fully collateralized with fiat

Custodial coins that are fully collateralized with fiat are relatively centralized as their creation and annihilation is performed by a single party, which is also responsible for maintaining collateral.

However, once created, coins can freely move on the corresponding blockchain. Due to their semi-centralized design, custodial coins are particularly prone to regulatory influences and must be regulatory compliant in order to be able to survive. Several coins of this nature, such as Theter, TrueUSD, Sila, etc., either already exist or are currently being developed. It is easy to see that crypto exchanges, which are not fully integrated with the existing banking system, have to rely on stable coins in order to do business. However, building stable coins suitable for more broad commercial applications is a much more exciting goal.

All fully collateralized coins have inherent drawbacks, since, by construction, they are centralized, represent single points of failure, and an irresistible attraction for both regulators and hackers (for different reasons, of course). In many instances, particularly, in the case of Theter, there is no transparency regarding underlying reserves, which makes it difficult for users to trust these coins. Unless very carefully structured, fiat-backed coins have to rely on third parties, most notably banks, to keep the corresponding collateral and execute transactions on their behalf. As a result, they suffer from high costs of carry of the collateral. Inherent low profitability is a major danger in its own right, which cannot be overestimated. In order to overcome it, stable coins require a large scale to become profitable.

4.3. Coins Collateralized with Assets

When economic environment in a country is stable, it makes sense to collateralize crypto coins with the corresponding fiat. However, in some parts of the world it is not a viable option since the local fiat itself is not stable versus the USD and other major currencies. In such cases, coins collateralized with real assets can be developed, see, e.g., [8], [9] where the so-called digital trade coins (DTC) are introduced. These coins have significant utility value in stable environments as well, for example, for supply-chain financing and, such as oil trading or cross-border financing. In addition, asset-backed coins also useful by limiting the freedom of central banks to manipulate their currencies and serving as a much-needed counterbalance to the USD, in its role of the world reserve currency. The USD domination causes serious trade imbalances, triggers trade wars and exacerbates the international frictions. To alleviate these ills, several central banks have suggested that it makes sense to complement fiat currencies with a supranational currency. By combining the ideas of narrow banking (for stability), and asset-backed digital currencies (for efficiency and transparency, we can dramatically improve the global financial system. This is an exciting possibility of designing a digital supranational currency backed by a diverse and widely held assets, which could combine the best features of historical currencies, including finality of settlement, partial anonymity, and usability on the web.

It is obvious that, if so desired, a basket of fiat currencies can be used as a collateral for a DTC. This idea was recently popularized by the Libra Association, [4], without giving any credit to the similar ideas introduced in [8], [9].

4.4. KYC and AML Considerations

The role of KYC and AML requirements for regulatory compliance cannot be overestimated. KYC procedures have to be performed when new tokens are issued, or existing ones redeemed for fiat. Given that stable tokens are going to circulate on a public blockchain, which keep an immutable record of all transactions, AML can be performed indirectly by analyzing the corresponding transaction social graph. In general, stable token transactions are more transparent than conventional physical cash transactions and are more or less on a par with traditional banking transactions. It is rather telling that when Bitcoin and Ethereum just got introduced, criminal elements were extremely excited about their illicit usage. However, eventually they realized that immutable record of transactions opens new avenues for the law enforcement agencies to monitor and prosecute their activities. Hence, at present, they are leaning back to cash transactions.

4.5. Conclusions

Although the idea of distributed ledgers is not new, modern technology gives it a new lease of life. Potentially, distributed ledgers have numerous applications in finance. Cryptocurrencies are the best known but not the only ones. While conventional cryptocurrencies are interesting, they are not very useful for real economy. Digital cash holds
a great promise and can change the whole financial ecosystem to the better. Asset-backed cryptocurrencies can serve as a much-needed counterpoint for fiat currencies, and a way forward toward ensuring world-wide financial stability and inclusion. It can also limit the impact of trade wars which disrupt existing supply chain.

5. REFERENCES

1. McLeay, M., Radia, A., and Thomas, R., 2014. Money Creation in the Modern Economy. Bank of England Quarterly Bulletin 2014 Q1. SSRN: https://ssrn.com/abstract=2416234

2. Ethereum, 2018, A Next-Generation Smart Contract and Decentralized Application Platform. White Paper. https://ethereum.org/en/whitepaper/

3 Gupta, M., 2018. Blockchain for Dummies. 2nd IBM Limited Edition. John Wiley and Sons, Inc. http://gunkelweb.com/coms465/texts/ibm_blockchain.pdf

4. The Libra Association, 2019. An introduction to Libra. White Paper, pp. 1-12.

5. Lipton, A., 2016. Modern monetary circuit theory, stability of interconnected banking network, and balance sheet optimization for individual banks. International Journal of Theoretical and Applied Finance, 19(06), p.1650034.

6. Lipton, A., 2018. Blockchains and distributed ledgers in retrospective and perspective. The Journal of Risk Finance, 19(1), pp. 4-25.

7. Lipton, A., 2019. 'Towards a Stable Tokenized Medium of Exchange', in Brummer, C. (ed.) Cryptoassets. Legal, Regulatory and Monetary Perspectives. Oxford: OUP, pp. 89- 116.

8. Lipton, A., Hardjono, T. and Pentland, A., 2018. Digital trade coin: towards a more stable digital currency. Royal Society Open Science, 5(7), p.180155.

9. Lipton, A. and Pentland, A., 2018. Breaking the bank. Scientific American, 318(1), pp. 26- 31.

10. Nakamoto, S., 2008. Bitcoin: A peer-to-peer electronic cash system. https://bitcoin.org/bitcoin.pdf

11. Sila, 2018. Sila: A Stable Fiat-Backed Digital Tokenized Means of Exchange. White Paper.

DATA DRIVEN DECISION MAKING IN TRADING AND INVESTING

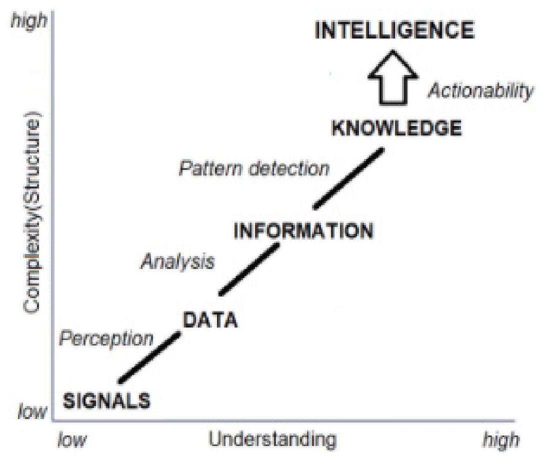
Cris DOLOC,
Quantitative & Computational technologist

1. A NEW PARADIGM

We are currently witnessing the onset of a powerful tidal wave that could have a long-lasting impact on our civilization for generations to come. Jim Gray, one of the greatest computer scientists of the 20th century and recipient of the Turing award, has predicted the advent of this new era more than a decade ago. He has called it the "Data exploration" era [1], or a period where theory, experimentation and simulation will come together to solve some of the most important problems of our time. This paradigm is driven by the availability of extremely large amounts of data generated as a result of the latest technological innovations. This is the dawn of a new era where the process of decision-making is going to be driven by data and powered by algorithms.

The concept of data is central to this new paradigm. Data represents the modern vehicle used to encode the surrounding reality, while being an orthogonal dimension to the algorithm concept which is the complementary device used to decode the same reality. The impact of Big Data on the financial industry nowadays has the proportions of a tsunami. Solving problems of increasing complexity demands more high-quality data and better algorithms for processing.

The topic of Financial Machine Learning has attracted a lot of interest recently, specifically because it represents one of the most challenging problem spaces for the applicability of Machine Learning. The applicability of Machine Learning techniques has increased dramatically in the field of financial trading, especially because the need to speed-up and automate tasks that humans do not have the capacity to handle. Financial trading and investment rely on accurate inputs being fed into business decision-making models, which traditionally were handled by humans who made judgements based on these inputs. Nowadays this functionality is provided by informatics systems that can compute at a massive scale and extract Intelligence from a multitude of sources to reach better decisions practically instantaneously. However, one should emphasize that the utility of data can only be measured by our ability to extract actionable Intelligence from it. This transformation process is quite opaque, often misunderstood and very domain specific.



While the ability to extract actionable insights from data is central to scientific discovery and the innovation process in general, it is absolutely critical in the decision-making process in trading and investing. The last decade has seen the emergence of a new interdisciplinary field called Data Science that provides an extensive toolset that covers a large array of domains, from Mathematics and Statistics, to Computer Science as well as more specialized scientific methods. The extraction of actionable insights from data is a very complex transformation that involves distilling Data into meaningful Information, encoding it into Knowledge to eventually achieve the desired outcome – actionable Intelligence (see figure above).

2. THE CURRENT AI HYPE

John von Neumann, one of the most prolific polymaths of all times, said once: "If people do not believe that mathematics is simple, it is only because they do not realize how complicated life is". Dealing with complexity is a hard-core requirement for the survival of biological entities, especially for humans. And dealing with complexity requires the use of properly defined terminology. The use of consistent and proper terminology is paramount to Science and Engineering, and it is especially essential when it comes to the use of novel technologies. As such it is important to define an adequate level of engineering and scientific clarity when it comes to using the term Artificial Intelligence, specifically as it relates to the financial industry.

The term AI has become the mantra of our time as this label is used more and more frequently as an intellectual wildcard by academicians and technologists alike. The AI label is particularly abused by media pundits, analysts and venture capitalists. The liberal use of terms like AI disruption or AI revolution is the manifestation of a systemic failure to understand the technical complexity of this topic. The zeitgeist of the AI pop culture envisions a future defined by a hyper intelligent singularity. In this worldview, digital AI villains will coexist with humans, as abstracted, ethically deprived and omnipotent overlords that will control our very existence. We are led to believe in a future world of superintelligence that has moved beyond our comprehension, and therefore we as humans are running the risk of abdicating our greatest gift and responsibility – the human rationality!

The utopian vision of general AI (as opposed to its limited pattern-matching current incarnation) is based on the extrapolation of the past exponential growth of computer power. This growth extrapolation is core to the belief that "singularity is near". The phenomenally complex problem of General AI is a not numerical one. According to leading experts like Prof. Leslie Valiant [2], it is plainly wrong to assume that human intelligence is a computationally bound problem. According to Immanuel Kant being human means having consciousness and virtuous behaviour. If one believes in the thesis that humanity requires consciousness, the myth of virtual humans could be exposed for what it is, i.e. an irrational construct that negates the very nature of humans. To embed human intelligence in an agent would require programming consciousness into it. This is not a computational problem that will be solved as Moore's Law continues playing out over time. There is no level of computer power where one may have any indication that a computer would magically have consciousness. This is not a processing or memory limit bound problem, and finding a solution is not a natural consequence of just increased computational power.

Since the beginning of the AI era, the goal of emulating human intelligence was central to this new field of science. The "human-imitative" view has later evolved towards the development of a more applicable engineering discipline in which algorithms and data are brought together to solve a variety of pattern recognition, learning, and decision-making problems. More and more, the AI research started to intersect with other engineering and scientific disciplines. A different approach was adopted, where the "human-imitative" perspective was replaced by a more practical one – "intelligence-augmentation". The systems needed not to be intelligent themselves, but to reveal patterns that humans could use. Examples are search engines, recommender systems of natural language translation.

While solving the challenge of understanding General Intelligence will be quintessential to the development of Artificial Intelligence it may also represent the foundation of a new branch of engineering. Like many other classic engineering disciplines that have emerged in the past (i.e. Civil, Electrical or Chemical), this new engineering discipline is going to be built on already mature concepts such as information, data, algorithm, computing and optimization. Some people call this new discipline Data Science. No matter the label employed, this new field will be focused on leveraging large amounts of data to enhance human life, so its development will require perspectives from a variety of other disciplines; from quantitative sciences like Mathematics and Statistics to Computational, Business and Social sciences.

3.THE CURRENT STATUS

The accelerated pace of technological progress of the past decades has contributed to the generation of vast amounts of data, and the birth of what one calls the “Big Data” age. Recent advances in High Performance Computing (HPC) and hardware acceleration (GPUs and FPGAs), coupled with new discoveries in algorithmic processing has created the conditions to train and utilize very complex Machine Learning algorithms at an unprecedented speed. All these new technological developments are going to have a revolutionary impact on many industries, and as usual, the financial industry will most likely be at the forefront of it. A new concept is already making its way in today’s financial world: “Data-Driven Decision Making”. Nowadays both trading and investing are more and more driven by large-scale data analysis, and the new concept of “alternative” data is becoming more ubiquitous.

The most important development of the last decade, i.e. the advent of Deep Learning, came as a surprise to many researchers in the field of AI. It was the result of combining the availability of massive amounts of data with great computational power, and it was assisted by new innovations in the field of hardware acceleration. The use of this novel approach has yielded surprising outcomes, particularly in speech and image recognition, as well as for most classification tasks.

Another exciting new field of research for the ML community is related to the problem of learning to learn. The goal of Meta-Learning is to train a model on a variety of learning tasks, such that it can solve new learning tasks using only a small number of training samples. Fast learning is a hallmark of human intelligence. Recognizing objects from a limited number of observations or rapidly learning new skills after a limited practice is something that a human could do “by default”. The next generation of “artificial” agents should be able to learn and adapt quickly after being exposed to only a few examples and continuing to adapt as more data becomes available. As an example, the Alpha Go program that beat the Go world’s champion cannot hold a simple conversation about the game of Chess for which it was not trained. By contrast, a human can adapt and act “intelligently” to a wide range of new, unseen situations. Enabling the “artificial” agents to acquire such versatility is will be the next big revolution in Computational Intelligence.

The ML algorithms have an implicit problem space where they could be extremely effective – i.e. speech and image recognition. One of the most important questions that need to be addressed is whether ML algorithms could be successfully applied to financial problems. This exercise will require a solid understanding of the nature of the problem space. What is really special about the financial markets? How are the characteristics of this particular problem space going to impact the applicability of general-purpose ML algorithms? In order to answer all these important questions, one needs to carefully consider the nature of several important mechanisms that are responsible for the uniqueness of financial market data, such as:
The Data generation process – is the process that generated the data stationary such as to ensure that training and testing could be performed on the same “kind” of data?

The Data quality – is the financial market data associated with “manageable” noise-to-signal ratio such that overfitting will be kept under control?
The Dimensionality of the feature space – could one engineer a feature space of enough dimensions such that the “measurable” properties of the financial markets could be properly learned by the ML algorithms?
Moving from the current paradigm of “Representation Learning” (or learning by “imitation” from data) to a paradigm of “Machine Consciousness” (where the agent will be able to make decisions based on independent “thinking” by using an approach that is more akin to human behaviour) is still a distant milestone.

Until these disruptive scientific and technological breakthroughs will occur, the best course of action is to continue to make steady progress on automating some of the most complex tasks at hand and in finding new ways of learning from data by using the great computational tools already available. At the same time, we would need to redirect the focus of the education process form mastering ever more complex tools and frameworks, to the development of solid and scalable “problem-solving” skills. One should hope that this quantitative accumulation of knowledge will one day mutate into the great discovery that will trigger the “singularity” point.

CONCLUSIONS

Recognizing Data as the principal contemporary artifact used to “encode” the surrounding reality, will shed more light into the importance of deeply understanding the generative processes associated with the observed data. The idea that the availability of more data is enough to solve any problem is simply incorrect. The sheer amount of available data or the availability of the most efficient algorithms and fastest compute platforms are not going to magically make up for the lack of understanding of the process that generates the data. Data by itself is nothing more than an opaque artifact that needs to be captured, filtered, analyzed and modelled in order to decipher from it the true meaning of the process that generates it. Identifying and understanding the nature of the process that generates the Data should be the ultimate goal of Data Science. The goal of this paper was to create some awareness about both the promises and the formidable challenges that the new era of Data-Driven Decision-Making is bringing forth for the financial industry. This also could be seen as a message of confidence in the possibilities offered by the new era of Data Exploration. At the same time, practitioners need to have realistic expectations of these possibilities, since moving from the current paradigm of “Representation Learning to a paradigm of “Machine Consciousness” is still a distant milestone. Until research in Cognitive sciences makes more progress in the understanding of how human intelligence works and how it can be emulated, financial quantitative practitioners will have to concentrate their efforts on applying novel Computational Intelligence methods [2] to their specific problem domain and focus on improving the automation aspects of the process that transforms financial Data into actionable market Intelligence.

I will conclude by making a prediction about the emergence of a new engineering discipline, one that I would label as d2ice - data-driven intelligence and computational engineering. This new engineering discipline will be at the same time Data-driven, Computationally intensive, Biology-inspired and Human-centered. All of these components will need to coexist and to be nourished with up-to-date knowledge in order to properly develop the new breed of engineer of the 21st century. This new profession will represent a significant evolution in respect to the quant of the 20th century, as it will address the needs of the modern economy by making itself available to a variety of industries, way beyond the realm of the Finance. Fields like Medicine, Healthcare, Education and the Internet-of-Things are going to be big consumers of this novel specialty. Since the modern society is so immersed and dependent on data and the methods to extract actionable information from it, this new breed of engineers will position themselves at the core of the system that drives the most important business decisions. This new discipline will fulfill a very important societal role and will have a critical contribution to the success of next industrial revolution.

REFERENCES

[1] Jim Grey, "The Fourth Paradigm: Data-Intensive Scientific Discovery", Microsoft Research, 2009. Content available at: http://itre.cis.upenn.edu/myl/JimGrayOnE-Science.pdf

[2] Leslie Valiant, "Probably Approximately Correct", published by Basic Books, NY, 2013

FaIR NETWORK CONTRIBUTORS

Lukasz SPRUCH

Lukasz is a Reader (Associate Professor) at the School of Mathematics, University of Edinburgh. He is also the director of Finance and Economics programme at The Alan Turing Institute, London. Previously Lukasz was a Nomura Junior Research Fellow at the Institute of Mathematics, University of Oxford, and a member of the Oxford-Man Institute for Quantitative Finance.
Lukasz has a broad research interest in Quantitive Finance, Machine and Reinforcement Learning, Statistics and Mean-Field Game Theory.

Blanka HORVATH

Dr Blanka Horvath is a Lecturer in Financial Mathematics at King's College London as well as a Honorary Lecturer at Imperial College London and a researcher at The Alan Turing Institute, where she is co-lead of the Machine Learning in Finance theme. Blanka holds a PhD in Financial Mathematics from ETH Zurich, a postgraduate Diplom in pure Mathematics from the University of Bonn and an MSc in Economics from the University of Hong Kong. In her latest research she focusses on non-Markovian models of financial markets such as Rough Volatility models as well as modern DNN-based market generators. Prior to her position at King's College, Blanka worked at JP Morgan on the refinements of the Deep Hedging programme. Her work on DNN-based calibration of Rough Volatility models was awarded the Rising Star Award 2020 of Risk magazine.

Michel CROUHY

Dr. Michel Crouhy is Senior Advisor to the "Direction Generale" of Natixis, and was until recently Head of Research & Development at NATIXIS Corporate and Investment Bank, a subsidiary of Groupe BPCE (Banques Populaires – Caisses d'Epargne). He is currently in charge of implementing a "Scenario-Based Stress Testing Framework" which is fully integrated into the capital and liquidity planning process of the bank. He is also the founder and chairman of the scientific commitee of the NATIXIS Foundation for Research and Innovation, which promotes and supports academic research and world-class events in the area of mathematical finance and data science.

Formerly, Crouhy was senior vice president in the Risk Management Division, at the Canadian Imperial Bank of Commerce (CIBC), in charge of risk analytics, model vetting and model risk management, operational risk, economic capital attribution and customer behaviour analytics. Prior to his career in the industry, Crouhy was a professor of finance at the HEC School of Management in Paris. He has been a visiting professor at the Wharton School of the University of Pennsylvania and at the University of California, Los Angeles.

Crouhy is a founding member of PRMIA, the Professional Risk Managers' International Association. He is a member of the Research Advisory Council of the Global Risk Institute in Financial Services and of the Credit Committee of the International Association of Financial Engineers (IAFE). He is a member of the Scientific Executive Board of the "Institut Louis Bachelier". He is the author and co-author of several books, the most recent ones being "Risk Management" (McGraw-Hill - 2001), "The Essentials of Risk Management" (McGraw-Hill – second edition in 2014) and "Contingent Claims Analysis in Corporate Finance" (World Scientific – 2019). Crouhy holds a PhD from the Wharton School of the University of Pennsylvania and is a Doctor Honoris Causa from the University of Montreal.

Dan GALLAI

Dr. Dan Galai is the Abe Gray Professor of Finance and Business Administration at the Hebrew University, school of business administration in Jerusalem. He was a visiting professor of finance at INSEAD and at the University of California, Los Angeles and has also taught at the University of Chicago and at the University of California, Berkeley.

Dr. Galai holds a Ph.D. from the University of Chicago and undergraduate and graduate degrees from the Hebrew University. He has served as a consultant for the Chicago Board of Options Exchange and the American Stock Exchange as well as for major banks. He has published numerous articles in leading business and finance journals, on options, risk management, financial markets and institutions, and corporate finance. He is a coauthor ofRisk Management published by McGraw- Hill, July 2000.

Zvi WEINER

Professor Zvi Wiener is the head of the Finance Department at the Business School of the Hebrew University of Jerusalem. His areas of expertise include risk management, financial engineering and the valuation of complex financial products. Zvi Wiener is one of the founders of PRMIA (Professional Risk Managers' International Association, see www.prmia.org), has served as a co-chair of the worldwide Education and Standards Committee of PRMIA and is currently a co-director of PRMIA Israel. His research was published in financial journals including the Journal of Finance, the Review of Financial Studies, Journal of Banking and Finance, Journal of Derivatives, Journal of Corporate Finance and many others and can be found at his website: http://pluto.mscc.huji.ac.il/~mswiener. In 2014 he was awarded the Teva Prize named after Dan Suesskind for research on dividend policy, in 2012 he received PRMIA award for Outstanding Service and Leadership, in 1997 he received the Alon Fellowship and in 1994 the Rothschild Fellowship. In the years 1996, 2005, 2013 and 2015 he received research grants from the Israel Academy of Sciences. Zvi Wiener has a rich consulting experience. He consulted the Ministry of Finance, the Ministry of Defense, the Bank of Israel, the Israel Securities Authority, the Tel Aviv Stock Exchange, credit rating companies, many banks, insurance firms and pension funds, as well as leading law firms on various financial problems.

FaIR NETWORK CONTRIBUTIONS

MODEL BASED MACHINE LEARNING

Lukasz SPRUCH,
Reader (Associate Professor) at the School of Mathematics, University of Edinburgh and Director of Finance and Economics programme at The Alan Turing Institute, London.

1. INTRODUCTION

Mathematical and Statistical modelling is ubiquitous in the financial industry and drives key decision processes. Broadly, one can distinguish between two approaches to mathematical modelling [Lehman, 1990, Cox et al., 1990, Hand, 2019]. The first approach that goes under the name phenomenological, handcrafted or iconic modelling is based on the formulation of a simplified representation of phenomena we aim to model. The second approach, known as empirical or data driven, seeks to summarise the relationships in the data is some convenient way, and the model, typically, does not have any theoretical bases. The former approach is dominant in the world of quantitive finance, where one often seeks parsimonious models that can be well-calibrated to data, the latter is at the heart of modern machine learning which is extremely effective in extracting stat istical relationships in seemingly large dimensional data sets but their black-box natures means that the individual parameters do not have a meaningful interpretation.

The recent advances in machine learning bring new opportunities to quantitive, and with greater computational power, more complex models can now be used. This is due to the fact that argu ably the most complicated and computationally expensive step of calibration has been addressed. Indeed, in the seminal paper [Hernandez, 2016] used neural networks to learn the calibration map from market data directly to model parameters. Subsequently, many papers followed [Liu et al., 2019, Ruf and Wang, 2019, Ruf and Wang, 2019, Benth et al., 2020, Gambara and Teichmann, 2020, Sardroudi, 2019, Horvath et al., 2019, Bayer et al., 2019, Bayer and Stemper, 2018, Vidales et al., 2018, Cuchiero et al., 2020]. These approaches focused on the calibration of the fixed parametric model but did not address perhaps even the more important issue, which is model selection and model uncertainty. The question of how seamlessly integrated modern machine learning methods with well-understood risk models that are currently used in the finance industry is perhaps the most critical. While these new techniques bring many new opportunities, they also carry new risks. Many of the machine learning tools produce opaque models that do stand on the solid mathematical foundation and do not back by decades of research in the area of quantitive finance.

In recent work [Perez Arribas et al., 2020, Gierjatowicz et al., 2020] authors showed that it possible to combine both approaches achieving the best of both worlds. By augmenting classical risk models with modern machine learning approaches, we are able to benefit from expressibility of neural networks or signature of the path, while staying within the realm of classical models, well understood by traders, risk managers and regulators. This mitigates, to some extent, the concerns that regulators have around the use of black-box solutions to manage financial risk.

These models, called Neural-SDEs or Sig-SDEs, depending on the choice of feature space can be used to find robust bounds for prices of derivatives and the corresponding hedging strategies while incorporating relevant market data. These models are an instantiation of generative modelling and are closely linked with the theory of causal optimal transport. Neural- and Sig- SDEs allow consistent calibration under both the risk-neutral and the real-world measures. Thus the model can be used to simulate market scenarios needed for assessing risk profiles and hedging strategies.

In Neural- and Sig-SDE setting we let the data dictate the model, while still keeping a strong prior on the model form. This is achieved by using SDEs for the model dynamics but instead of choosing a fixed parametrization for the model SDEs we allow the drift and diffusion to be given by an overparametrized neural networks or linear combination of the (log-)signature terms. Neural and Sig-SDE are shown to not only provide a systematic framework for model selection, but also, quite remarkably, to produce robust estimates on the derivative prices. Here, the calibration and model selection are done simultaneously. In this sense, model selection is data-driven. For the neural SDE model is overparametrised, and there is a large pool of possible models and the training algorithm selects a model. For Sig-SDEs one can benefit from linear structure and decide whether it is more appropriate to work with over or under parametrised model. To simplify the presentation from now on, we focus on neural-SDEs and refer the reader for related ideas presented in signature feature space to [Perez Arribas et al., 2020]. We believe that it is helpful to view these new approaches to modelling in finance through the lens of robust finance paradigm.

2. ROBUST FINANCE PARADIGM

Let us now consider a probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ and a random variable $\Psi \in L^2(\mathcal{F}_T)$ that represents the discounted payoff of a illiquid (path-dependent) derivative. The problem of calculating a market consistent price of a financial derivative can be seen as equivalent to finding a map that takes market data (e.g. prices of underlying assets, interest rates, prices of liquid options) and returns the no-arbitrage price of the derivative. Typically an Ito process $(X_t^\theta)_{t\in[0,T]}$, with parameters $\theta \in \mathbb{R}^p$ has been the main component used in constructing such pricing function. Such parametric model induces a martingale probability measure, denoted by $\mathbb{Q}(\theta)$, which is then used to compute no-arbitrage price of derivatives. The market data (input data) here is represented by payoffs $\{\Phi_i\}_{i=1}^M$ of liquid derivatives, and their corresponding market prices $\{\mathfrak{p}(\Phi_i)\}_{i=1}^M$. We will assume throughout that this price set is free of arbitrage. To make the model $\mathbb{Q}(\theta)$ consistent with market prices, one seeks parameters θ^* such that the difference between $\mathfrak{p}(\Phi_i)$ and $\mathbb{E}^{\mathbb{Q}(\theta^*)}[\Phi_i]$ is minimized for all $i = 1, \dots, M$ (w.r.t. some metric). If for all $i = 1, \dots, M$ we have $\mathfrak{p}(\Phi_i) = \mathbb{E}^{\mathbb{Q}(\theta^*)}[\Phi_i]$ then we will say the model is consistent with market data (perfectly cal ibrated). There may be infinitely many models that are consistent with the market. This is called Knightian uncertainty [Knight, 1971, Cohen et al., 2018].

Let $\mathcal{M}$ be the set of all martingale measures / models that are perfectly calibrated to market inputs. In the robust finance paradigm, see [Hobson, 1998, Cox and Obloj, 2011], one takes con servative approach and instead of computing a single price (that corresponds to a model from $\mathcal{M}$) one computes the price interval $(\inf_{\mathbb{Q}\in\mathcal{M}} \mathbb{E}^{\mathbb{Q}}[\Psi], \sup_{\mathbb{Q}\in\mathcal{M}} \mathbb{E}^{\mathbb{Q}}[\Psi])$. The bounds can be computed using tools from martingale optimal transport which also, through dual representation, yields corresponding super- and sub- hedging strategies, [Beiglbock et al., 2013]. Without imposing further constrains, the class of all calibrated models $\mathcal{M}$ might be too large and consequently the corres ponding bounds too wide to be of practical use [Eckstein et al., 2019]. See however an effort to incorporate further market information to tighten the pricing interval, [Nadtochiy and Obloj, 2017, Aksamit et al., 2020]. Another shortcoming of working with the entire class of calibrated models $\mathcal{M}$ is that, in general, it is not clear how to obtain a practical/explicit model out of the measures that yields price bounds. For example, such explicit models are useful when one wants consistently calibrate under pricing measure $\mathbb{Q}$ and real-world measure $\mathbb{P}$ as needed for risk es timation and stress testing, [Broadie et al., 2011, Pelsser and Schweizer, 2016] or learn hedging strategies in the presence of transactional cost and an illiquidity constrains [Buehler et al., 2019].

3. NEURAL SDES

Fix $T > 0$ and for simplicity assume constant interest rate $r \in \mathbb{R}$. Consider parameter space $\Theta = \Theta^b \times \Theta^\sigma \subseteq \mathbb{R}^p$ and parametric functions $b : \mathbb{R}^d \times \Theta^b \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \times \Theta^\sigma \to \mathbb{R}^{d \times n}$.

Let $(W_t)_{t \in [0,T]}$ be a n-dimensional Brownian motion supported on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{Q})$ so that $\mathbb{Q}$ is the Wiener measure and $\Omega = C([0,T];\mathbb{R}^n)$. We consider the following parametric SDE

$$dX_t^\theta = b(t, X_t^\theta, \theta)dt + \sigma(t, X_t^\theta, \theta)dW_t. \tag{1}$$

We split X^θ which is the entire stochastic model into traded assets and non-tradable components. Let $X^\theta = (S^\theta, V^\theta)$, where S are the traded assets and V are the components that are not traded. We will assume that for all $t \in [0,T]$, $x = (s,v) \in \mathbb{R}^d$ we will assume that $b(t,(s,v),\theta) = \big(rs, b^V(t,(s,v),\theta)\big)$ and $\sigma(t,(s,v),\theta) = \big(\sigma^S(t,(s,v),\theta), \sigma^V(t,(s,v),\theta)\big)$. Then we can write (1) as

$$\begin{aligned} dS_t^\theta &= rS_t^\theta\, dt + \sigma^S(t, X_t^\theta, \theta)\, dW_t\,, \\ dV_t^\theta &= b^V(t, X_t^\theta, \theta)\, dt + \sigma^V(t, X_t^\theta, \theta)\, dW_t\,, \\ X_t^\theta &= (S_t^\theta, V_t^\theta)\,. \end{aligned} \tag{2}$$

Observe that σ^S and σ^V encode arbitrary correlation structures between the traded assets and the non-tradable components. Moreover, we immediately see that $(e^{-rt}S_t)_{t\in[0,T]}$ is a (local) martin gale and thus the model is free of arbitrage.

In a situation when (b, σ) are defined to be neural networks, we denote by $\mathcal{M}^{\text{nsde}}(\theta)$ the class of all solutions to (1). Note that due to universal approximation property of neural net works, see [Hornik, 1991, Sontag and Sussmann, 1997, Cuchiero et al., 2019], $\mathcal{M}^{\text{nsde}}(\theta)$ contains large class of SDEs solutions. Furthermore, neural networks can be efficiently trained with the stochastic gradient decent methods and hence one can easily seek calibrated models in $\mathcal{M}^{\text{nsde}}(\theta)$. Finally, neural SDE integrate black-box neural network type models with the known and well studied SDE models. One consequence of that is that one can: a) consistently calibrate these under the risk neutral measure as well as the real-world measure; b) easily integrate additional market information e.g constrains on realised variance; c) verify martingale property. Given a loss function $\ell : \mathbb{R} \times \mathbb{R} \to \mathbb{R}^+$, the search for calibrated model can be written as

$$\theta^* \in \arg\min_{\theta\in\Theta} \sum_{i=1}^{M} \ell(\mathbb{E}^{\mathbb{Q}(\theta)}[\Phi_i], \mathfrak{p}(\Phi_i))\,, \text{ where } \mathbb{E}^{\mathbb{Q}(\theta)}[\Phi] = \int_{C([0,T],\mathbb{R}^d)} \Phi(\omega)\mathcal{L}(X^\theta)(d\omega)\,.$$

Let us connect neural SDEs to the concept of generative modelling, see [Goodfellow et al., 2014, Kingma and Welling, 2013]. Let $\mathbb{Q}^{\text{market}} \in \mathcal{M}$ be the true martingale measure (so by definition all liquid derivatives are perfectly calibrated under this measure i.e. $\mathbb{E}^{\mathbb{Q}^{\text{market}}}[\Phi_i] = \mathfrak{p}(\Phi_i)$ for all $i = 1, \dots, M$). We know that when (1) admits a strong solution then for any $\theta \in \mathbb{R}^p$ there exists a measurable map $G^\theta : \mathbb{R}^d \times C([0,T];\mathbb{R}^n) \to C([0,T];\mathbb{R}^d)$ such that $X^\theta = G^\theta(\zeta, W)$. Hence, one can view (1) as a generative model that maps μ, the joint distribution of X_0 on $\mathbb{R}^d$ and the Wiener measure on $C([0,T];\mathbb{R}^n)$ into $\mathbb{Q}^\theta = (G_t^\theta)_{\#}\mu$. We see that by construction G is a causal transport map i.e transport map that is adapted to filtration $(\mathcal{F}_t)_{t\in[0,T]}$, see also [Acciaio et al., 2019, Lassalle, 2013].

One then seeks θ^* such that $G_{\#}^{\theta^*}\mu$ is a good approximation of $\mathbb{Q}^{\text{market}}$ with respect to user specified metric. In this paper we work with

$$D(G_{\#}^{\theta}\mu, \mathbb{Q}^{\text{market}}) := \sum_{i=1}^{M} \ell\left(\int_{C([0,T],\mathbb{R}^d)} \Phi_i(\omega)(G_{\#}^{\theta}\mu)(d\omega), \int_{C([0,T],\mathbb{R}^d)} \Phi_i(\omega)\mathbb{Q}^{\text{market}}(d\omega)\right)$$

There are many Neural SDE models that can be calibrated well to market data and that produce significantly different prices for derivatives that were not part of the calibration data. In practice these would be illiquid derivatives where we require model to obtain prices. Therefore, we com pute price intervals for illiquid derivatives within the class of calibrated neural SDEs models. To be more precise we compute

$$\inf_{\theta}\left\{\mathbb{E}^{\mathbb{Q}(\theta)}[\Psi] \,:\, D(G(\theta)_{\#}\mu_0, \mathbb{Q}^{\text{market}}) = 0\right\}, \quad \sup_{\theta}\left\{\mathbb{E}^{\mathbb{Q}(\theta)}[\Psi] \,:\, D(G(\theta)_{\#}\mu_0, \mathbb{Q}^{\text{market}}) = 0\right\}.$$

4. CONCLUSIONS

i. Neural SDEs provide a systematic framework for model selection and produce robust estimates on the derivative prices. The calibration and model selection are done simultaneously and the thus the model selection is data-driven.
ii. With neural SDEs, the modelling choices one makes are: networks architectures, structure of neural SDE (e.g. traded and non-traded assets), training methods and data. For classical hand crafted models the choice of the algorithm for calibrating parameters has not been considered as part of modelling choice, but for machine learning this is one of the key components. In it has been shown in [Gierjatowicz et al., 2020] that the change in initialisation of stochastic gradient method used for training leads to different prices of illiquid options, thus providing one way of obtaining price bounds. Furthermore even for basic local volatilty model that is unique for continuum of strikes and maturities, produces ranges of prices of illiquid derivat ives when calibrated to finite data sets.
iii. The above optimisation problem is not convex. Nonetheless, empirical experiments in [Gier jatowicz et al., 2020] demonstrate that the stochastic gradient decent methods used to minimise the loss functional D converges to the set of parameters for which calibrated error is of order 10^{-5} to 10^{-4} for the square loss function. Theoretical framework for analysing such algorithms is being developed in [Siska and Szpruch, 2020]. ˇ

REFERENCES

[Acciaio et al., 2019] Acciaio, B., Backhoff-Veraguas, J., and Zalashko, A. (2019). Causal optimal transport and its links to enlargement of filtrations and continuous-time stochastic optimization. Stochastic Processes and their Applications.

[Aksamit et al., 2020] Aksamit, A., Hou, Z., and Obloj, J. (2020). Robust framework for quantifying the value of information in pricing and hedging. SIAM Journal on Financial Mathematics, 11(1):27–59.

[Bayer et al., 2019] Bayer, C., Horvath, B., Muguruza, A., Stemper, B., and Tomas, M. (2019). On deep calibration of (rough) stochastic volatility models.

[Bayer and Stemper, 2018] Bayer, C. and Stemper, B. (2018). Deep calibration of rough stochastic volatility models.

[Beiglbock et al., 2013] Beiglbock, M., Henry-Labordere, P., and Penkner, F. (2013).Model-independent bounds for option prices—a mass transport approach. Finance and Stochastics, 17(3):477–501.

[Benth et al., 2020] Benth, F. E., Detering, N., and Lavagnini, S. (2020). Accuracy of deep learning in calibrating HJM forward curves. arXiv preprint arXiv:2006.01911.

[Broadie et al., 2011] Broadie, M., Du, Y., and Moallemi, C. C. (2011). Efficient risk estimation via nested sequential simulation. Management Science, 57(6):1172–1194.

[Buehler et al., 2019] Buehler, H., Gonon, L., Teichmann, J., and Wood, B. (2019). Deep hedging. Quantitative Finance, pages 1–21.

[Cohen et al., 2018] Cohen, S. N. et al. (2018). Data and uncertainty in extreme risks-a nonlinear expectations approach. World Scientific Book Chapters, pages 135–162.

[Cox and Obloj, 2011] Cox, A. M. and Obloj, J. (2011). Robust hedging of double touch barrier options. SIAM Journal on Financial Mathematics, 2(1):141–182.

[Cox et al., 1990] Cox, D. R. et al. (1990). Role of models in statistical analysis. Statistical Science, 5(2):169–174.

[Cuchiero et al., 2020] Cuchiero, C., Khosrawi, W., and Teichmann, J. (2020). A generative adversarial network approach to calibration of local stochastic volatility models. arXiv preprint arXiv:2005.02505.

[Cuchiero et al., 2019] Cuchiero, C., Larsson, M., and Teichmann, J. (2019). Deep neural networks, generic universal interpolation, and controlled odes. arXiv preprint arXiv:1908.07838. [Eckstein et al., 2019] Eckstein, S., Guo, G., Lim, T., and Obloj, J. (2019). Robust pricing and hedging of options on multiple assets and its numerics. arXiv preprint arXiv:1909.03870.

[Gambara and Teichmann, 2020] Gambara, M. and Teichmann, J. (2020). Consistent recalibration models and deep calibration. arXiv preprint arXiv:2006.09455.

[Gierjatowicz et al., 2020] Gierjatowicz, P., Sabate-Vidales, M., Siˇ ska, D., Szpruch, L., and ˇ Zuri ˇ c,ˇ Z. (2020). Robust ˇ pricing and hedging via neural sdes. arXiv preprint arXiv:2007.04154.

[Goodfellow et al., 2014] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. (2014). Generative adversarial nets. In Advances in neural information processing systems, pages 2672–2680.

[Hand, 2019] Hand, D. (2019). What is the purpose of statistical modelling? Harvard Data Science Review, 1(1). [Hernandez, 2016] Hernandez, A. (2016). Model calibration with neural networks. Risk.

[Hobson, 1998] Hobson, D. G. (1998). Robust hedging of the lookback option. Finance and Stochastics, 2(4):329–347.

[Hornik, 1991] Hornik, K. (1991). Approximation capabilities of multilayer feedforward networks. Neural Networks, 4(2):251–257.

[Horvath et al., 2019] Horvath, B., Muguruza, A., and Tomas, M. (2019). Deep learning volatility. [Kingma and Welling, 2013] Kingma, D. P. and Welling, M. (2013). Auto-encoding variational bayes. arXiv preprint arXiv:1312.6114.

[Knight, 1971] Knight, F. H. (1971). Risk, uncertainty and profit, 1921. Library of Economics and Liberty. [Lassalle, 2013] Lassalle, R. (2013). Causal transference plans and their monge-kantorovich problems. arXiv preprint arXiv:1303.6925.6 L. SZPRUCH

[Lehman, 1990] Lehman, E. L. (1990). Model specification: The views of fisher and neyman, and later observations. Statistical Science, 5:160–8.

[Liu et al., 2019] Liu, S., Borovykh, A., Grzelak, L. A., and Oosterlee, C. W. (2019). A neural network-based framework for financial model calibration. Journal of Mathematics in Industry, 9(1):9.

[Nadtochiy and Obloj, 2017] Nadtochiy, S. and Obloj, J. (2017). Robust trading of implied skew. International Journal of Theoretical and Applied Finance, 20(02):1750008.

[Pelsser and Schweizer, 2016] Pelsser, A. and Schweizer, J. (2016). The difference between lsmc and replicating portfolio in insurance liability modeling. European actuarial journal, 6(2):441–494.

[Perez Arribas et al., 2020] Perez Arribas, I., Salvi, C., and Szpruch, L. (2020). Sig-sdes model for quantitative finance. arXiv, pages arXiv–2006.

[Ruf and Wang, 2019] Ruf, J. and Wang, W. (2019). Neural networks for option pricing and hedging: a literature review. Available at SSRN 3486363.

[Sardroudi, 2019] Sardroudi, W. K. (2019). Polynomial semimartingales and a deep learning approach to local stochastic volatility calibration.

ON THE TRANSFORMATIVE ROLE OF NEW TECHNOLOGIES ON MARKET MODELS IN QUANTITATIVE MODELLING

Blanka HORVATH,
Lecturer in Financial Mathematics at King's College London as well as a Honorary Lecturer at Imperial College London and a researcher at The Alan Turing Institute

THE RELATIONSHIP OF CLASSICAL QUANTITATIVE FINANCE MODELS AND DEEP LEARNING

Deep learning has undoubtedly had a major impact on the possibilities of mathematical modelling in finance. The application of deep neural networks in this field has helped push the boundaries of the achievable further in terms of:

• providing numerical alternatives to established means of forecasting, pricing and hedging,
• speeding up calculations beyond what was possible with traditional numerical methods, and modelling higher dimensional scenarios than previously possible,
• finding (approximatively) optimal strategies under consideration of market frictions, beyond the regimes that were analysed and understood by classical methods,
• but also in terms of generalisations of the very concept of a financial market model to what we now call market generators.

This is by far no exhaustive list of the many avenues through which financial modelling is aided today by deep learning. With the above examples we are only scratching the surface of a wealth of ongoing developments.

But perhaps conceptually more import than the direct impacts listed above is an emerging symbiosis of algorithms and the data. The influence and impacts in this symbiosis goes both ways:

Data used in training influences the algorithms and the chosen architecture has implications on the structure of data on which it can be applied reliably. The former comes as no surprise: Deep learning models being so flexible, the data used for training influences the type of output of the application. The latter in turn calls for rethinking risk management routines profoundly. Effects of the latter influence are visible for example in [2, 12, 13], where a suitable network architecture has decisive implications on network performance. Risk management should therefore include ongoing monitoring on the structure of the data and the suitability of the network architecture to that data as well as regular updates of the training data that algorithms are exposed to.

THE CASE FOR MORE FLEXIBLE DATA-DRIVEN MARKET MODELS: The technological advances and tools developed in recent years create the desire for market models that are directly data-driven and closely reflect evolving market reality. The implications of the influence of data on algorithms becomes quite visible in [4]: When training data is numerically generated by the Black-Scholes

model, the trained network (the hedging engine) approximates the corresponding delta as an opti-al hedging strategy. Analogous effects can be seen when the training data is generated by Heston paths. In fact, the approach in [4] is inherently model agnostic: For an arbitrary1 collection of sample paths provided to the network in the training phase, the hedging engine generates response strategies that are indeed approximatively optimal in market regimes where path distributions resemble the ones presented during the training phase. The more realistic the market paths presented to the network during the training phase, the better performance can be expected on real data. Classical models are however known to suffer from a relative

inflexibility. This and similar model-agnostic neural network based financial applications therefore drove interest towards such market scenario generators (or simply market generators) that are capable of closely reflecting real market dynamics, in a model-free and directly data driven way. This leads us back to neural networks as powerful approximators of functions and distributions to build generative models for financial markets. Such generative models are based on the idea of transforming random samples of latent variables to distributions observed in data samples via differentiable functions, which are approximated by a neural network by backpropagation. The application of such frameworks to financial settings is what we refer to as market generators. The case for classical models: Market generation is currently still in its infancy. The flexibility of these models, the great variety in which they appear and their inherent synergy with the data are all factors that make them difficult to risk manage. Classical models on the other hand are well understood and a wealth of research and analysis of their properties is available, and so is decades worth of experience in their risk management. Furthermore, classical models have also been developed with the aim to reflect (despite their relative inflexibility) a selection of stylized facts of financial markets [8]. Some model families [10] have the impressive ability to exhibit these with a remarkable precision and provide an overarching theoretical basis for their behaviour across markets and applications. Most importantly, the distribution and dynamical properties of classical models can be carefully controlled by a handful set of parameters at each point in time. They can therefore conveniently provide training sets for deep learning algorithms (see [13]) where certain properties of the training set are present or absent to test the robustness of a deep learning algorithm under those market conditions. Synergies: Through the available analytical and numerical solutions classical models also provide much needed benchmarks–in the cases where these are applicable–for deep learning applications to facilitate risk management of the latter. On the other hand, approximative solutions obtained by neural networks can facilitate finding analytical solutions for classical models under market conditions where these hadn't been available.

THE TRANSFORMATIVE POTENTIALS OF MARKET GENERATION

In several situations it is advantageous to rely on simulated data rather than working with data directly. Market generators–in contrast to classical models–are so flexible that they are capable of generating data samples that are statistically indistinguishable from a given original dataset. This flexibility opens the door to applications that were so far not typically a core part of financial modelling in the classical setting.

Applications of Market Generators: Sailing on new waters of modelling.
(i) Data anonymisation: When the available data is confidential, it is desirable to generate anonymised datasets that are representative of the true underlying distribution of the data but cannot be traced back to their origin. Financial data and medical data are often proprietary, or confidential. When testing investment strategies or the effectiveness of a treatment it is imperative not to be able to trace back the datasets to the individual client or patient.
Evaluating whether the produced data is representative of the distribution that a dataset stems from, depends on the distributional properties (evaluation metrics) that we control for. Thus the question of adequate performance evaluation metrics is a central matter for research on market generators, see [5]. Also the level of anonymity achieved by such generative procedures is a highly interesting question on its own. The study presented in [16] is devoted to understanding the latter question in more detail.
(ii) Small original training datasets: Though big data as a concept is ubiquitous today, in more situations than not, the amount of data available for training a neural network is small rather than large. When there are natural restrictions on the number of available original samples (constraints on the number of experiments, restrictions on the access to data), the available data may not be sufficient to train the neural network application at hand (e.g. hedging engine). Clearly, the more complex the application, the more data samples are needed to train it.

Once such a generative network is available, more complex neural network applications can also
be trained, using the market generator that produces the necessary amount of training samples for the latter. Generative models for sparse data environments therefore need to be relatively parsimonious and trainable on a low number of data samples. Further practical applications of market generators include (but are not limited to) the following use cases:

(iii) Outlier detection: Once the distribution of a dataset can be statistically identified and reproduced (even if the data does not follow any known parametric distribution) it be comes possible to identify typical regimes–that is, datapoints that are typical for the distribution—as well as outliers or atypical events. Detecting outliers and atypical events enables us to identify occurrences of regime switching in a market, gives the basis for fraud detection and for the identification of human-machine interfaces in automation: that is, such events that alert an automated machine to hand over the handling of an atypical process to a human with appropriate responsibilities.

(iv) Backtesting: When developing a trading strategy, carrying out a backtest to measure how the strategy would perform in a realistic environment is of crucial importance. However, using historical data may result in overfitting of the trading strategy. Having a market simulator capable of generating realistic, independent samples of market paths would allow a more robust backtest, less prone to overfitting.

(v) Risk management of portfolios – be it of financial derivatives or trading strategies – is of utmost importance. A realistic market simulator can be used to generate synthetic paths to estimate various risk metrics, such as Value at Risk (VaR).

These applications are to date fairly unexplored, however they are gaining more and more relevance in a landscape where data increasingly assumes a central role in quantitative applications. And so, market generators have the very real potential of creating a whole new era of financial modelling.

REFERENCES

[1] D. Bartl, S. Drapeau, J. Ob l´oj and J. Wiesel. Data driven robustness and uncertainty sensitivity analysis. In preparation, 2020.

[2] C. Bayer, B. Horvath, A. Muguruza, B. Stemper,and M. Tomas. On deep pricing and calibration of (rough) stochastic volatility models. Preprint, arXiv:1908.08806, 2019.

[3] P. Bonnier, P. Kidger, I. Perez Arribas, C. Salvi and T. Lyons. Deep Signature Transforms. ArXiv:1905.08494, 2019.

[4] H. B¨uhler, L. Gonon, J. Teichmann, and B.Wood. Deep hedging. Quantitative Finance, 19(8) p. 1271-1291, 2019.

[5] H. B¨uhler, B. Horvath, T. Lyons, I. Perez Arribaz, B. Wood. Generating financial markets with signatures. Preprint SSRN:3657366, 2020.

[6] C. Cuchiero, W. Khosrawi, J. Teichmann. A generative adversarial network approach to calibration of local stochastic volatility models. Preprint, arxiv.org:2005.02505, 2020.

[7] I. Chevyrev and H. Oberhauser. Signature moments to characterize laws of stochastic processes. Preprint, arxiv.org:1810.10971, 2018.

[8] R. Cont. Empirical properties of asset returns: stylized facts and statistical issues. Quantitative Finance, (1): 223-236, 2001.

[9] P. Friz, M. Hairer. A Course on Rough Paths: With an Introduction to Regularity Structures. Cham: Springer International Publishing, 2014.

[10] J. Gatheral, T. Jaisson, M. Rosenbaum. Volatility is rough, Quantitative Finance, 18:6, 933-949, 2018.

[11] P. Henry-Labordere. Generative models for financial data. Preprint, SSRN 3408007, 2019.

[12] B. Horvath, A. Muguruza, M. Tomas. Deep learning volatility. To appear in Quantitative Finance, 2020.

[13] B. Horvath, J. Teichmann, Z. Zuri`c. Deep Hedging under Rough Volatility. Working Paper, 2020.

[14] Patrick Kidger and Terry Lyons. Signatory: differentiable computations of the signature and logsignature transforms, on both CPU and GPU. Preprint, arXiv:2001.00706, 2020.

[15] A. Kondratyev, C. Schwarz. The market generator. Preprint, SSRN:3384948, 2019.

[16] A. Kondratyev, C. Schwarz, B. Horvath, Data anonymisation, outlier detection and fighting overfitting with Restricted Boltzmann Machines. Preprint, SSRN:3526436, 2020.

[17] T. Lyons. Rough paths, signatures and the modelling of functions on streams, Preprint, arXiv:1307.7244, 2014.

[18] H. Oberhauser and F. Kir´aly. Kernels for sequentially ordered data. Journal of Machine Learning Research, 20(31):1-45, 2019.

THE IMPACT OF FINTECHS ON FINANCIAL INTERMEDIATION: A FUNCTIONAL APPROACH

Michel CROUHY, Dan GALAI and Zvi WIENER

EXTENDED SUMMARY

Fintech refers to the technological start-ups that are emerging to rival traditional banking and financial players. It covers a wide array of services from P2P and crowdfunding platforms, mobile payment solutions, international money transfers to online portfolio management tools, service comparators/aggregators (APIs), ICOs (Initial Coins Offerings), and more.

Recent technological innovations are disruptive for many industries, especially the financial industry. They have made possible the new financial intermediation offered by FinTech. Without the easy access to large digitalized dataset, the massive increase in storage capacity and processing capacities, especially with the "cloud", FinTech would not have emerged. It's been a while since we've seen a game-changer, but FinTech may be the one.

Bigtech is a subset of the broader fintech sector. Bigtech companies are defined as "...technology companies with established presence in the market for digital services". In the U.S., bigtechs include the so-called "GAFAM" firms: Google, Amazon, Facebook, Apple and Microsoft. "BATX" refers to the four largest Chinese technology companies: Baidu, Alibaba, Tencent and Xiaomi. Bigtechs often started their finance activity in the payments sector as a means to facilitate their core business, e.g., e-commerce. They subsequently leveraged their platform to expand into vast ecosystems, which extend far beyond payments into other operations, such as lending, insurance, savings and investment products, either independently or in association with established financial institutions.

FinTech is not a natural evolution in financial intermediation but a true revolution. One major consequence of these technological improvements is disintermediation. The financial system being essentially an intermediary is first in line to be disintermediated.

The growth opportunities are huge. Venture capital funds were quick to recognize these opportunities and provide a significant amount of funding.

Previous academic review papers analyze the various areas of operations in which fintechs engage. We take a different approach and map fintechs based on the six functions of financial intermediation rather than on their main activities. We follow Merton's functional approach in which the six core functions performed by a financial system are analyzed. According to Merton, a financial system provides:
A. A payment system for the exchange of goods and services;
B. A mechanism for the pooling of funds to undertake large-scale indivisible investments in an enterprise;
C. A way to transfer economic resources through time and across geographic regions and industries;
D. A way to manage uncertainty and control risk;
E. price information that helps coordinate decentralized decision-making in various sectors of the economy;
F. A way to deal with asymmetric-information and incentive problems when one party to a financial transaction has information that the other party does not.

In banks all these core functions are fully integrated. Technology allows for the unbundling and fragmentation of financial services among specialized fintech companies. Fintechs deliver these services in a cheaper and friendlier way, and to new clients, not served by traditional institutions.

Banks offer basic services such as low-cost checking, and so-called "sticky customer relationships" used to allow them to earn attractive margins in other areas such as asset management, credit card fees, or foreign-exchange transactions. Banks have unique expertise in credit underwriting. They are heavily regulated and offer protection to depositors through deposit insurance. Historically, consumers have been very loyal customers and slow to change banking services providers. Banks still enjoy greater trust than tech companies, although it is changing.

Fintechs are moving beyond strictly addressing a customer's financial needs (product centric) to offering a wider range of services (platforms and ecosystems).

Let's now address each of the six core functions of a financial system.

A. A PAYMENT SYSTEM FOR THE EXCHANGE OF GOODS AND SERVICES

The first core function of a financial system, i.e., "a payment system for the exchange of goods and services" is currently the area with the many new entrants with innovative products such as digital currencies also referred to as crypto-currencies or e-money, and mobile wallets.

Payments is today a great business: revenues totaling up to $1.9 trillion globally in 2018. Due to the pandemic, it is expected that the revenues from payments will decline by 7% in 2020 due to the reduction in the volume of transactions in the retail sector.

Mobile wallets have been widely adopted by consumers worldwide: Apple Pay, Ali Pay, WeChat Pay, Samsung Pay and Android Pay. Digital wallets are rapidly replacing cash in executing both customer-to-customer (C2C) and customer-to-business (C2B) transactions..

The adoption of digital wallets worldwide has been facilitated by the ubiquity of smartphones, particularly in developing countries where few have access to the banking system. A good example is Kenya, where M-Pesa® allows users to deposit money into an account linked to a cell phone, to send payments to sellers of goods and services, and to withdraw deposits for cash. It offers many of the essential services provided by banks without requiring customers to have a bank account. This feature of fintechs is termed «inclusion» of new unbanked customers.

There is a new wave of technological innovations that allows the transactional use of so-called "crypto-assets", i.e., digital tokens recorded in distributed ledgers and held in electronic safety "wallets" stored in the cloud. In crypto-currency systems, encryption techniques control the generation of currency units with the use of blockchain technology or DLT (Distributed Ledger Technology). This technology allows person-to-person transactions, independently of the banking system, with transactions authenticated by many computers (minors), around the world without the need for a trusted financial intermediary like a bank to verify the transaction.

The cryptocurrency "Bitcoin" was the first implementation in 2009 of the blockchain technology. It has no intrinsic value, and fulfills only the first attribute, out of three we expect from money: medium of exchange, store of value and unit of account. It is a speculative investment, not a currency. Bitcoin has been followed by the creation of more than 2,000 cryptocurrencies, the most notable at this time being Ethereum. So far, the pricing of these cryptocurrencies has been highly speculative. Several initiatives have been launched to invent so-called "stablecoins", cryptocurrencies backed by real-world assets. To summarize, the "payments" industry is undergoing rapid changes driven by a combination of increasing customer demand, the need for technology investments and potential regulatory pressure. Fintech invaders have recognized the value of payments as a gateway to customer acquisition.

B. A MECHANISM FOR THE POOLING OF FUNDS TO UNDERTAKE LARGE-SCALE INDIVISIBLE INVESTMENTS IN AN ENTERPRISE

The second core function of a financial system according to Merton is ***"a mechanism for the pooling of funds to undertake large-scale indivisible investments in an enterprise"***.

Lending has always been the core of commercial banking business: Banks collect short-term deposits and extend long-term loans to retail customers and companies (maturity intermediation).

Peer-to-peer (P2P) lending platforms were first launched in the early 2000s, initially concentrating on lending among individuals. "Crowdfunding", the funding of a private project by raising small sums from a large number of people via the Internet can be viewed as collaborative P2P. Crowdfunding has been used to fund various sorts of entrepreneurial ventures, such as artistic projects, including movies, concerts and books, as well as non-profit community-oriented social projects.

Another fintech-based innovation in capital raising is the issuance of tokens rather than shares to finance ICOs (initial coin offering): offering of tokens instead of shares to finance corporations. It can be noted that the majority of the ICOs were found to be fraudulent.

With lending innovation, alternative new credit models have emerged, including automated loan approvals. These platforms have developed data mining and machine learning techniques for credit scoring using non-traditional data sources (behavioral information, social networks like Facebook,...) and powerful data analytics to price risks and lower operating costs.

C.A WAY TO TRANSFER ECONOMIC RESOURCES THROUGH TIME AND ACROSS GEOGRAPHIC REGIONS AND INDUSTRIES

This third core function ***"A way to transfer economic resources through time and across geographic regions and industries"*** relates to the production of financial products to retail investors.

Current liquidity can be transferred to the future via bank savings accounts, or by investment in stocks and bonds. Individuals can also transfer wealth from the future to the present by borrowing for various maturities. For centuries, commercial banks, have provided basic financial products to fulfill this function, which is one of the banks' core businesses. For their part, commercial banks offer time intermediation by collecting short-term deposits and investing them in longer-term loans. Investment banks, on the other hand, specialize in trading securities and offer portfolio management and advisory services as to how to move resources to the future through securities investments.

P2P in this area allows small individual depositors to finance large-scale loans (Lending Club, Kabbage, SoFi,...) Once the loan is originated, the credit platform acts as an agent for the investors by servicing the loan. Investors have the flexibility to invest 100% of their capital in one loan, or diversify and make partial investments in multiple loans. P2P platforms have set up systems of loan selection where investors specify criteria for the loans they want to purchase. These systems then filter out these loans for the investor. P2P platforms generate revenue from

origination fees charged to the borrowers, servicing the loans, as well as additional charges such as late fees.
Neobanks are playing a growing role in this area. Neobanks are pure digital institutions with two big cost advantages relative to traditional banks: the absence of branches and their up-to-date cloud-based software. In partnership with Visa and Mastercard the neobanks also offer credit and debit cards, with no fee even when used internationally.

These neobanks have the common characteristic that they don't make money yet. Still there is uncertainty on the future of these neobanks as they need to convince clients that they are financially solid and will remain solvent

In asset management, fintech providers, including "robo-advisors", offer an interface for retail investors to trade and invest either directly or through financial intermediaries. Robo-advisors are progressively replacing human financial advisors, since the automated services are less expensive than traditional financial advisory services. Robo-advisors are computer programs that generate investment advice based on customer data. Through the use of machine learning (ML) techniques, robo-advisors offer a less costly alternative to human financial advisors, particularly for individual retail investors pursuing passive investment strategies. The cost of entry is also much lower.

D. A WAY TO MANAGE UNCERTAINTY AND CONTROL RISK

The fourth function of financial intermediaries according to Merton is ***«a way to manage uncertainty and control risk»***.
Leading risk softwares in the 90s are: Riskmetrics (JP Morgan), Creditmetrics (JP Morgan), KMV (Moody's), Kamakura, Algorithmics (Fitch, IBM, SS&C). They can be viewed as the fintech response to Basel II regulatory demands.

An important innovation in bank capital regulation emerged as a response to the GFC. Following the GFC the Fed in the US and the EBA in Europe imposed on banks to run stress tests with the purpose of ensuring that banks have sufficient capital to absorb losses under a financial downturn, while maintaining sufficient capital to keep making loans. The pandemic highlighted the need of stress tests in controlling risks in the financial markets. Fintechs are now offering banks solutions related to stress testing and the generation of scenarios.

A key industry in which technology is in the process of revolutionaizing traditional financial intermediation is insurance. Insurtechs is reinventing the insurance business through the use of new digital technologies. For example, the pricing of insurance contracts relies heavily on ML techniques, and not only on traditional approaches where clients are placed in broad categories such as age, gender, etc.

Smart contracts may radically change existing contracts, such as car insurance contracts. With a smart contract, the car insurance is embedded in the vehicle and data generated by the driver's use of it is fed continuously to the insurance contract. The terms of the contract can then be adjusted according to the transmitted data. This development is underway as insurance companies propose usage-based insurance policies, known as "pay-as-you-drive". Smart contracts will facilitate more accurate pricing of these policies.

The new technologies change the risk landscape and introduce new risks such as cyber-risk. Risk managers now consider cyber-risk as the biggest threat to the business of their enterprise. Fintech companies are emerging to help various organizations around the world to automatically 24/7 test their cybersecurity measures and identify their main risks and the most cost effective way to protect their organization against hackers.

E. PRICE INFORMATION THAT HELPS COORDINATE DECENTRALIZED DECISION-MAKING IN VARIOUS SECTORS OF THE ECONOMY

The fifth function of financial intermediaries according to Merton is providing *"**price information that helps coordinate decentralized decision-making in various sectors of the economy**"*. This function relates to processes and technologies that contribute to the efficient allocation of resources in the economy.

Major contributors to the information processing are:

- Robo-advisors,
- Algorithmic trading or Algo-trading: algorithmic trading encompasses a wide variety of trading and portfolio management strategies, such as Quant Funds, financial products such as Exchanged Traded Funds (ETFs) and High Frequency Trading (HFT).
- Personal Financial Management (PFM) emerged with Open Banking

F. A WAY TO DEAL WITH ASYMMETRIC-INFORMATION AND INCENTIVE PROBLEMS WHEN ONE PARTY TO A FINANCIAL TRANSACTION HAS INFORMATION THAT THE OTHER PARTY DOES NOT.

The sixth function of financial intermediaries according to Merton is providing "a way to deal with the asymmetric-information and incentive problems when one party to a financial transaction has information that the other party does not".

In this area we should note the critical contribution of:

- Artificial intelligence (AI) which describes algorithms in decision-making that can perform functions typically associated with human intelligence. This includes understanding speech, pattern recognition, reasoning, knowledge representation, planning, machine learning, analysis of social interactions, etc.
- Data intelligence systems which are the major source of information on any aspect of life today is the Internet. The major task is to unveil the information, verify its validity, and analyze its economic implications.
- Smart contracts.

CONCLUDING COMMENTS

In our paper, we follow Merton's functional approach to analyze the impact of fintechs on financial intermediation. This approach is the most appropriate since the six basic functions outlined by Merton are independent of the structure of the financial system.

The functions can be fully integrated, as is the case with universal banking, or can be unbundled and fragmented, allowing for the delivery of at least some financial services by non-bank fintechs. Many fintechs started as niche players developing a specific service, but have since transitioned from product-centric to customer-centric enterprises. They use data as an extremely important factor of production and lever it to provide more diverse customer services. An example among many others is Oscar, a young health insurance company, which is evolving into an enterprise offering extensive online health care monitoring and advice. This trend is expected to expand in the future.

Are fintechs going to replace the banks and insurance companies? The most probable answer is no. Banks maintain several important advantages over fintechs. First, they have the reputation and trust of many loyal customers. Second, banks have equity capital, generally in excess to what they need. Third, banks operate under a regulatory umbrella and enjoy government support. This was made evident during the 2008 financial crisis, when the Federal Reserve in the U.S. and other central banks intervened to help the banks overcome the crisis. Banks exploit these advantages, and they acquire directly, or through related companies, many of the emerging fintechs to enhance banking services. Very often traditional financial intermediaries supply the third-tier capital (in the form of long-term loans) to non-bank lending platforms.

Since banks can cross-sell many financial products on their platform, we can expect a combination of many niche firms and large banks. Since banks are also heavily regulated and slow in making changes, it can explain how small outsiders can still survive.

It can be expected that in the future, new regulation will encompass fintechs in such a way that any financial service offered on any platform will be subject to the same regulation. The position taken by commercial banks is that regulation should follow the principle of "same product – same regulation", in order for banks not to suffer adverse discrimination and to ensure a "level playing field". In principle, regulation should be based on functions, and not on tools. As we have seen in the past, regulation has a major impact on the competitive structure of the financial intermediation industry. The presence of fintechs forces regulators to rethink how to structure the financial services industry to enhance consumer protection and stability without impairing either innovation or competition.

Banks continue to adopt automation at a fast pace, and offer a growing number of online digital services. As a result, many banks close branches, shrinking the size of departments engaged in advising clients, approving loans and providing other services. Some major financial institutions have acquired online brokerage and robo-advisory fintechs, and now offer these automated services internally. Only relatively few fintechs enjoy IP protection, hence their services can be duplicated by incumbent banks and insurance companies which have large research and IT departments.

It is expected that some fintechs, those not acquired by traditional financial institutions, will likely undertake horizontal mergers and acquisitions to become multi-faceted service companies.

In the spirit of Merton's conclusion *"the functional perspective emphasizes the discovery of more efficient ways to perform one or more of the six basic functions of the financial system"*, the evidence is clear that fintechs are reimagining these services and making their delivery more efficient, at times reaching market segments that have been ignored by incumbent financial intermediaries.

ROUND TABLE SPEAKERS

Sandrine UNGARI

Sandrine is currently head of the quantitative cross-asset research team at Société Générale. Within the Cross-Asset Research group, the quantitative research team is active in risk premium strategies, derivatives and structured products, portfolio risk modeling, and provides research to investors worldwide. The group has been recognized as a market leader in quantitative research and was ranked No. 1 in the Extel survey in the Quantitative Strategies category. Sandrine joined Société Générale in 2006. Previously, she worked as a quantitative analyst at HBOS Treasury and Reech Sungard in London. She graduated from ENSTA (Paris) and holds a Master's degree in Quantitative Finance from the University of Paris VI. She is a lecturer at the University Paris Diderot.

Charles-Albert LEHALLE

Responsible for data analysis at Capital Fund Management (CFM, Paris) and Visiting Scholar at Imperial College (London), Charles-Albert studied Machine-Learning for Stochastic Control during his PhD, 20 years ago. He started his career as IA Project Manager at the Renault research center and joined the financial community in 2005 with the advent of automated trading. In 2016, Charles Albert was awarded the prize for the best article in finance by the Institut Europlace de Finance (IEF) and published more than fifty academic articles and book chapters. He is co-author of the book «Market Microstructure in Practice» (World Scientific Publisher, 2nd edition 2018), analysing the main characteristic of today's markets. He is the Scientific Director of the interdisciplinary research programme «Finance and Insurance Reloaded» at Institute Louis Bachelier.
This program explores the influence of new technologies (from the artificial intelligence blockchain) on the banking/financial/ insurance industry.

Nicole EL KAROUI

Nicole is Professor Emeritus at Sorbonne University (Paris VI), after having been a professor at the École Polytechnique for ten years. A specialist in stochastic control, she has used her ideas in financial markets, contributing to the emergence of the international field of «financial mathematics». École Polytechnique gave her the opportunity to create a research team in this field, with an international reputation. Her current research concerns the risk of longevity seen from the perspective of population dynamics and dynamic utility. Very concerned about the question of transmission to youth, she created in 1990 in Paris VI, the DEA (Master of Probability and Finance), the first in a scientific university. Through this training, she became known in financial markets around the world, even making the front page of the Wall Street Journal in 2006. French «quants» are highly sought-after in investment banks. Following the creation of the «Financial Risks» chair by Société Générale, she is actively working on the creation of the Risk Foundation in 2007 and at the Institut Louis Bachelier, convinced that this structuring of ties between academics and professionals represents a very important asset for France.

ROUND TABLES DEEP LEARNING VS. THEORY: ILLUSTRATION WITH DEEP HEDGING

Among the FaIR's round tables, this chapter will focus on a very interesting discussion that happened during the Deep Hedging round table. The focus of this round table is on the use of data implied by artificial intelligence in the financial system as well as the resulting challenges.

Over the past couple decades, artificial intelligence, as a technological innovation, has lead a important number of transformations in the financial landscape. Some transformations are still ongoing and some new opportunities will most likely come in the next couple of years. This is intrinsically linked to the availability of large amount data. For a number of reasons, including bias reduction and liability of automated algorithms, the data cannot just be fed to deep learning models making predictions, some level of explicability is desired.

In this first excerpt, Sandrine Ungari (SG) presents her view on how Deep Learning tools can be used for products valuation and risk assessments. She also stresses important challenges that might have limited deep learning adoption in the financial sector.

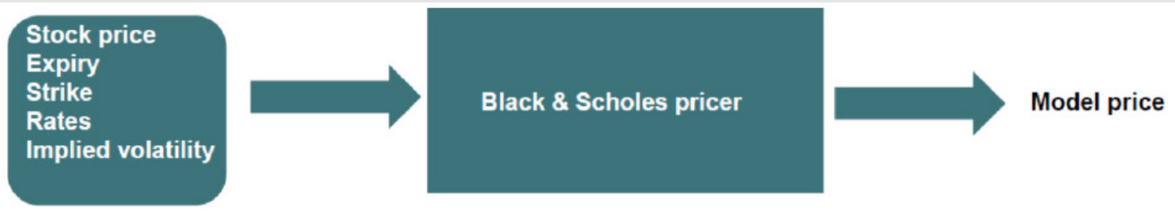
19/12/2019,
AFGAP,
ORGANIZED BY FaIR

Recently, more and more researchers have been trying to apply these approaches to the valuation and risk analysis of derivatives.

Using deep-learning for the valuation of derivatives is not a new idea. The first paper on the subject was published in 1994 in The Journal of Finance by J.M. Hutchinson, A.W. Lo and T.A. Poggio (A Nonparametric Approach to Pricing and Hedging Derivative Securities Via Learning Networks). The authors developed the idea of using a neural approach to value derivative contracts.

This approach was not pursued at the time. It is back in the spotlight today, thanks to advances in optimization and the computing power of modern computers. In a classical approach, the model takes a series of parameters as inputs: the underlying price, strike, maturity, rates, implied volatility. A formula is applied, such as the Black Scholes formula for instance, and a price is obtained.

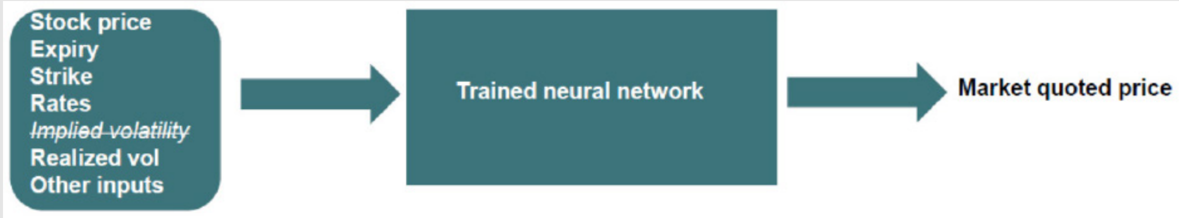
Graph 8. Valuation of derivatives in a classical model



Source : SG Cross Asset Quant Research

With neural networks, we take a maximum of input, and historical or simulated data. So the network is calibrated to minimize the gap between prices coming out of the network and market prices. The model is non-parametric and is the result of a calibration.

Graph 9. Valuation of derivatives with a neural network

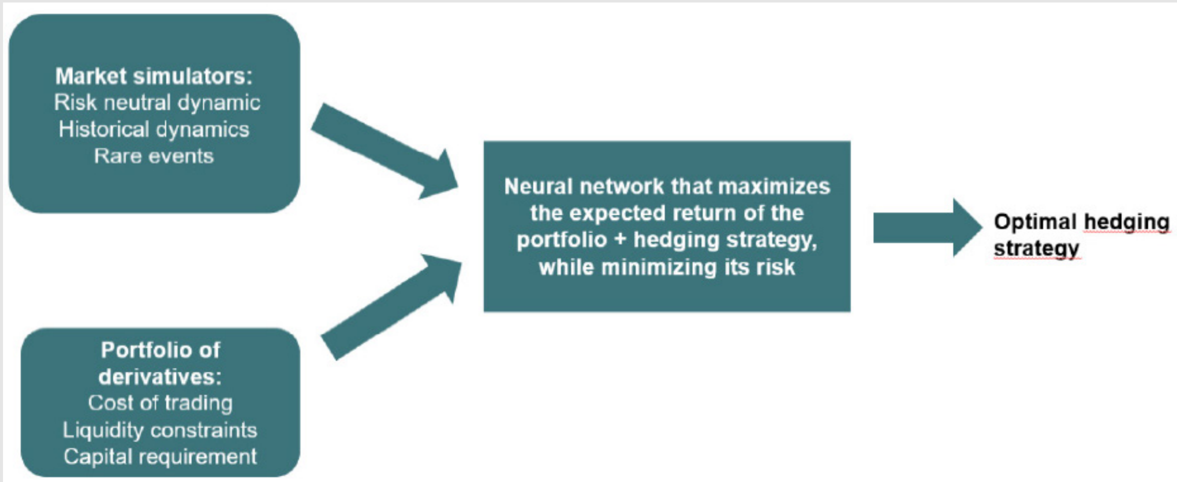


Source : SG Cross Asset Quant Research

Deep-hedging uses the same idea. It is an approach that was developed and proposed by a group of researchers at ETH Zurich and JP Morgan (Deep Hedging, the Journal of Quantitative Finance, Jan 2019, Hans Bühler, Lukas Gonon, Josef Teichmann, Ben Wood).

This approach simulates a lot of market data and optimizes the hedging strategy. Rather than calibrating the network on the price of an instrument, we calibrate the network in such a way that we minimize the portfolio variance containing the hedge instruments and let the network set the optimal hedging strategy.

Graph 10: Hedging risk management with deep pricing



Source : SG Cross Asset Quant Research

The theory and approach behind the concept is much more complicated. For the time being this type of technique is primarily based on simulations. The available market data is not enough to calibrate neural networks. Artificial data must therefore be generated. There is a whole branch of research, popular at the present time, which consists of studying how to simulate realistic data, and which is consistent with the reality of financial markets.

The challenges of deep learning : the last issue that I wanted to address is the constraints and challenges of this type of approach. In the study published by the FCA and the Bank of England, several questions emerge about the implementation of these techniques.

So the first question is the management of historically used systems. This is not a problem specific to deep learning, but one that arises as soon as one seeks to do something new in an industry such as banking or insurance. Switching from an old system to more modern technologies is always complicated to handle. More particularly in Machine Learning (ML), there is the integration of these techniques into the processes used to run the business. This issue is related to the training needed to get users to apply these new techniques. Users must be familiar with the concepts discussed at the beginning of this presentation. This is quite a cultural change that needs to be made within organizations. Another point that is often cited is the lack of explainability. This is a major problem with the deep learning approach. In a non-parametric approach, one cannot change a parameter and see how the result of the model is impacted. There is a lack of intuition on how the model reacts and on the origin of the results. For example, Deep Mind has created an algorithm capable of playing Go, called AlphaGo. In the game where AlphaGo first beat a professional Go player, the computer came up with a move that no one understood or could explain at the time. In retrospect, it was a stroke of genius.

Graph 11. Constraints and challenges of Machine Learning in practice

Source: Machine Learning in UK Financial Services – Bank of England, FCA

With deep learning, you can have completely unexpected results. As an investor, a trader or a risk manager, you don't know if it's a stroke of genius, or just a numerical problem. In practice, this kind of behavior can become troublesome. Insufficient data is also a theme that is regularly mentioned. Machine Learning is a data thirsty approach. We only have about ten years of data on most financial markets. This can be a problem in the implementation of this kind of algorithm.

Related to this, other questions arise regarding the use of alternative data: how should the use of personal data be regulated? What are the governance processes for data, how is it managed? How do we ensure its reliability? How do we ensure that data is reliable? That's a good question. You can't have an 100% guarantee that the data is reliable. On the other hand, we can set up checks to get the best possible representation of reality. For instance, if we take trading data, we can observe the prices being traded, and we have to make sure that there are no errors in the databases.

So here is a brief introduction to machine learning, deep learning, and deep hedging, which already raises many questions.

The specific case of Deep Hedging is especially relevant as one of the most direct application of deep learning in finance. In financial intermediation, replication strategies have always been an important determinant of risk sharing among financial actors. There are however a few technical considerations that should be taken into account when applying deep learning to hedging strategy replications. In the following excerpt from the Deep Hedging round table, Charles-Albert Lehalle proposes a general presentation of the so called Deep Hedging technique as well as the main challenges of using deep models for hedging. In particular, he explains how being able to reason in the risk neutral world is fundamental for hedging strategies, and also that a deep model will only minimize the mean error whereas a traditional model can minimize the local error. This tends to show that research is still needed in order to better leverage the use of deep learning models in hedging.

Charles-Albert LEHALLE

It is time to discuss the case of deep hedging. Deep hedging is a way to develop strategies for replicating the risk of an uncertain position that is based on reinforcement learning. Let's review the principle of risk replication. If I have a position that has non-linear exposures to stochastic processes, i.e. uncertain dynamics of economic variables such as volatility, prices, rates or correlations, how much should I buy or sell each day of tradable products in order to reduce this exposure? At the end of the day, when the hedged portfolio matures, if my strategy is well designed, I have exactly the amount to be delivered to my client, or I am short of the amount they are going to pay me. These replication strategies are very important for investment banks because not only do they reduce exposure to economic risks, but they also provide an estimate of the average cost of hedging over the product lifespan, and therefore allow them to price it.

In a «traditional» hedging technique, price dynamics are estimated on the basis of implicit variables.

If we consider the operating principle of «traditional» hedging, price dynamics are simulated using implicit variables. Several trajectories are generated and the combinatorics of a very large part of the possible hedging strategies are explored and compared with the simulations. Only the most efficient strategies are selected and form the trajectory components of the so-called optimal strategy. It should be noted that the implicit process amounts to not using a historical database of market prices of negotiable instruments, but to using solely the current prices of derivatives, i.e. products of the same nature as the hedged product, which also leads to the hedging strategies' implementation. It is argued that we are not using the historical measure, which is the database measure that would be used by a learning algorithm, but the risk-neutral measure.

This measure has many attractive properties. The one I prefer is that it guarantees that if, during the lifespan of the replicated product, there are other products of the same nature available for purchase or sale, then the bank will be able to «resell its risk» and will no longer have any exposure of any kind. The price that will have been calculated in this manner at the beginning of the product's life will be quite fair. This would be less the case under historical probability. By using this risk neutral measurement to generate trajectories to evaluate the performance of hedging strategies, the bank therefore perfectly fulfils its role as a financial intermediary. It places itself temporarily between several buyers and sellers of different risks and does not take «historical risk», as an investment fund might do.

With a deep neural network, one can generate fewer replication strategies, to focus on those that seem to work better.
Once we understand this principle of generating trajectories in order to evaluate the performance of potential strategies, we can use learning techniques to improve it in a preliminary way. When

we systematically cover almost all possible replication strategies, it is very costly from a computer point of view, especially when we increase the problem's dimension, i.e. when we want to cover portfolios with many assets and many nonlinearities. With a «deep neural network», one can generate fewer replication strategies, trusting the neural network parameters to quickly focus on those that seem to work better, and thus converge more quickly towards the optimal strategy, or at least that seems optimal. Sufficient comparisons have already been made by academics to convince us that it is indeed more economical in computing time and memory.

Traditional methods guarantee a maximum «local» error, whereas neural networks only guarantee an «average» error.

Obviously, once we realize this, we only want to increase the size of the problem to handle cases that are simply impossible to handle with traditional methods. Indeed, it is possible. But this raises the question, since we no longer have any elements of comparison, and since we are on a large scale, of verifying that these strategies are indeed valid. Indeed it must be understood that traditional methods today guarantee a maximum «local» error, in the sense that the error is controlled in the same way everywhere, whereas neural networks only guarantee an error «on average». For a measured level of error it is not known if the neural network makes an error of this order of magnitude everywhere or if it makes very large errors in some places and very small ones in others. And in large dimensions, we know that what is called the «curse of large dimensions» is very much misleading the calculations «on average». This implies that a method has yet to be developed that translates this «average» guarantee into a «nearly everywhere» guarantee. There are leads but nothing conclusive to my knowledge. The next step, if one is comfortable with using learning methods to develop hedging strategies, is to use similar methods to generate trajectories from data, rather than making simulations from modelled processes whose parameters are implied under riskneutral measurement. And that's still a very big change. On the one hand we can say that this is an opportunity to get closer to reality, notably by taking into account self-correlations in a non-parametric way, but on the other hand we must avoid leaving the attractive
properties of the neutral risk world. Some researchers are working on this subject but it is complex. Indeed, no one wants hedging strategies to be influenced by six recent months of bull markets; it is only a matter of time before the market becomes more stable. It is not up to intermediaries, such as investment banks, to take these kinds of bets, these kinds of risks, which are typically investment fund risks.

The success of this type of project depends on a small team of experts working together with business experts on well-defined projects.

I will summarize some key aspects for me. First of all, implementing a secondary innovation coming from AI is not simple, it requires effort and investment. The right way to foster the success of this kind of project is to have a small team of experts working together, for a short period of time, with business experts on well-defined projects. Always keep in mind that a learning algorithm is made up of three things that all three need to be tracked: external computer libraries, internal code and databases. This creates a certain complexity in versioning, deployment and monitoring. Another very important aspect comes from the data bias. We should not believe that because it is not necessary to write equations to implement learning algorithms, we do not make model choices. The choice of model is hidden in the data, and in the way the algorithm will do its best to adapt to the data. It is necessary to keep a critical, modeler's eye on the data. And when it comes to deep hedging, you have to remember all these points, and especially that of bias. We do not want a replication strategy to be influenced by recent sporadic specificities in price dynamics of tradable products. This raises some interesting issues for researchers in financial mathematics. Some of them are already tackling them, and I am convinced that they will provide answers. These will certainly not be that we should close our eyes and «trust the data».

As Sandrine Ungari pointed out, the lack of explicability can be an issue at the individual level (bad risk assessments) but also at the macro level (systemic related issues). This is especially relevant for risk evaluation as Charles-Albert Lehalle pointed it out regarding Deep Hedging. For these reasons, the regulator might need to intervene to prevent bias in order to ultimately protect investors. However, as it tends to be forgotten, bias in decisions is also an important issue in the human world. So bias should be supervised within automated algorithms but also within human-based decisions. This is essentially the point of the following discussion that Charles-Albert Lehalle had in the Deep Hedging round table.

Charles-Albert LEHALLE

The novelty of the approaches we are talking about today, deep learning methods, lies not only in their profoundly non-linear nature, which seems to be able to adapt to a wide range of relationships between variables that are not perfectly known, but it also stems from the fact that the data they gather have a great influence on the results obtained. The implementation of a learning algorithm requires:

- External» code libraries, often provided by Google, for example with TensorFlow, or Facebook, with pyTorch,
- Code developed for the application, which specifies for instance the criterion to be minimized, the choice of the optimization technique and its parameters. A database that will serve as a reference for the algorithm.

How to protect the result from a possible bias in data or even from data manipulation that will change the algorithm outcome?

A range of questions can be asked regarding the level of quality that should be required on the data. How to protect the result from a possible bias in the data or even from data manipulation that will change the outcome of the algorithm? There is the subject of cyber security of course, which is more of an industrial issue. How to secure access to the multiple interfaces that result from the use of distributed components? One can think of the cyber attack on the Mexican central bank's payment systems in May 2018. Will this type of attack be facilitated when market participants use algorithms that connect several times with many more intermediaries?

But that's not the topic for today. Let's focus on the specific issues related to the implementation of AI technologies. Difficulties can come from several factors: External libraries: how to reliably measure the impact of a change of version of one of the big learning libraries, such as TensorFlow or pyTorch? You need to be able to restart learning with the new version, on the same data as used with the older version, and to compare the «results». What exactly are the results? The parameters resulting from the training (i.e. the « mass « of the neural networks)? Or the predictions obtained by these neural networks? Probably both. But learning's random nature, because let's not forget that most of these methods rely on stochastic gradient descent, renders this comparison difficult.

- The data itself can change and be updated, when there is an inaccuracy, or we can simply increase the volume of this data. In this case, it is then necessary to start learning again and compare the results once again, using statistical criteria. It is necessary to check that the results are not «statistically different».
- Even before talking about a data change, one must consider the possible influence of «bias» in data. The classic example is that of a gender bias: if the learning base is not balanced, then the results will not be balanced either. If we think about loans' allocation, an algorithm should not be more willing to lend to men than to women, for instance. In market finance there are many other drawbacks, including the very important «contributed data» drawback, that I will talk about later.

- Obviously, the code that implements the specific functionality of an application, very often developed internally by the financial institution, and which includes a criterion to be optimized, called in statistical learning the «loss function», has a central place. It plays a very important role in the learning outcome. Which function should we choose? Particularly, is it better to choose a statistical precision criterion, or an «operational criterion», i.e. for example a gain in euros? Should a penalty be assigned to risk-taking, and how should it be combined with the gain criterion? The insurance and market finance industries are especially regulated, and rightly so, since they can have very strong negative externalities.
It is therefore necessary to find a clear framework for the use of these technologies.

Once we start thinking about the regulation of AI technologies, and to avoid embarking on the subject, which is a bit of a cliché of «algorithm ethics», it is worth taking a step back.

Why demand to control the bias of an algorithm when we do not demand control over the bias of decisions of the same nature, but made by humans.

On a personal note, I remember the debates around high-frequency trading, to which I contributed a lot. The debate was not much different from today, since high-frequency trading uses automatic trading algorithms that are calibrated with data and have relative autonomy. Once again, I have the feeling that posing the problem in an automated and technological framework has the virtue of bringing out real issues. When you study a process that involves a lot of human decisions, you can always think: «since there are humans in the loop, the description we make here is a bit approximate and never exactly respected, so we're not going to think too much about its implications». When we are faced with machine-driven processes, we ask ourselves the real questions in a much clearer and more concise way. We must therefore take advantage of this to regulate better. To regulate machines well, of course, but also to regulate humans better. What struck me in the debate on high-frequency trading is that at one point there were decisions of this type: «let's regulate machines and essentially machines». But no, the way in which humans, as intermediaries, made prices on certain illiquid products was obviously to be regulated as much as the way in which machines make prices in a systematic way, even if it is often on more liquid products.

To return to biases, we also need to supervise biases that come from human decisions

It is necessary to organize «decision making reviews», which, based on statistical criteria, will make it possible to establish whether the decisions made last month by the institution were biased, and if so, it will have to justify why. And if it is not justified, the institution will have to demonstrate that it is taking measures to prevent this from happening again. Whether these decisions are made by machines or by humans. So I'm taking a chance that automation and the use of AI technologies allow us to ask these questions. And I hope that the conclusions that will emerge from the debates will allow us to give life to regulations that will apply to processes managed by humans as well as by machines. And that's really a strength, even if it raises several problems, but I think they will be better addressed.

With the previous discussion it is clearer that artificial intelligence in risk intermediation has already proven very useful, although there are a lot of technical and regulatory challenges. However, it should be kept in mind that artificial intelligence is a generic technology. The true application of this generic technology will have to be driven by secondary innovations. More practical use cases may be created by the interaction between experts in the technology and experts in a narrow field for which the generic technology might be applied to. This is the general idea of the next excerpt from the Deep Hedging round table presented by Charles-Albert Lehalle.

Charles-Albert LEHALLE

Once the debate has been framed in terms of financial services' disintermediation, we may question the nature of what we want to call «Artificial Intelligence technologies», even if we can debate for hours about whether we should really call them that. I have to admit that everyone is using this term today, so I'm going to use it too. What is the particularity of AI when it comes to implementing it? I think it's because it belongs to what some economists call Generic Technologies.
I'll give you other examples of Generic Technologies: steam engine and electricity. Just because you've used a steam engine to drive a locomotive doesn't mean that it's easy to use a steam engine to run a loom. You have to redevelop what are called secondary innovations. Artificial intelligence works the same way. It may work very well on signal processing to recognize images or sound, but if you want to use it in a bank you have to redesign a lot of things, and readapt AI to each set of problems. This requires investment. Therefore the ecosystem needs to organize itself, in order to put in contact generic experts of this technology, and specific experts, during specific small short-term projects, to be able to manufacture secondary innovations that will respond to a particular industry or company.

The deployment of AI in banking and insurance cannot be done without investment and will require many secondary innovations.

The distinction between secondary innovations and generic technology, and especially the fact that investment is required to produce secondary innovations, creates a complicated bias. Since, on the one hand, I can read every day in the newspaper that AI is great, that it works for lots of applications. Yet, on the other hand, when I want to apply it without investing time and money, in a well thought out project, it doesn't work on the problems I'm interested in. Today's Villani report has made this clear, and it means that there will soon be areas of excellence in France, the interdisciplinary institutes of artificial intelligence (3IA), that will bring together generic experts in AI, big data, data sciences and specific needs, with their business experts. Typically in the Paris region, one of these 3IAs is called PRAIRIE. Companies will be able to make partnerships, with periods of one to three years, to try to manufacture secondary innovations. The major English universities have done this with the Turing Institute, which was created in 2015, and it seems to be working well. In the United States, each major American university has created a «small» center of this nature, such as the Center for Data Sciences at NYU, created in 2013 by Yann LeCun.

University centers of excellence are set up in France to bring together generic AI experts and business experts.

Before getting to the core subject, I wanted to stress the need to invest, no doubt in collaboration with academics, in order to obtain the secondary innovations needed to make AI work well in market finance.

In the last round table excerpt for this chapter Nicole El Karoui presents a really interesting retrospective of innovation in financial mathematics in the last 30 years. The main point is that big trends in financial intermediation, like the derivative products boom, are generally due to important technological changes. In a way similar to how financial mathematics reshaped the financial landscape and our view of it, current breakthroughs in deep learning (from computer science this time) have the potential to drastically change our approach of risk intermediation.

Nicole EL KAROUI

You could say that I have been a kind of «companion on the road» on «financial risk» markets, particularly on the issues of derivatives' hedging and the emergence of their regulation (not to mention crises) over the last thirty years. Surprisingly, it seems to me that there is a certain interest in confronting the current evolutions resulting from Big Data and its extraordinary technological tools with those of finance in the 1980s. When the futures markets (MATIF and MONEP) opened in France in 1985-86, ten years after the United States, we were in the midst of a computer revolution, corresponding to the switch from IBM mainframes to PCs. French banks took advantage of this to set up their derivatives business in a very quantitative way, notably at the Société Générale. Engineers were developing the business, focusing primarily on the possibilities offered by IT.

At the same time, at the Faculty of Science and the Ecole des Ponts et Chaussées, and later in other universities and «Grandes Ecoles», very quantitative financial mathematics courses were being set up, most of them led by women (a French exception in this highly technological world). The rapid growth of computer power (Moore's law) allowed the development of numerical probabilistic methods (Monte Carlo type) which were effective (at the time). From the experimental (start-up type) period of the 1990s, when the question of the meaning and effectiveness of dynamic hedging of derivative products was constantly questioned, we moved on to the «industrial» phase of the years 2000-2008, when financial innovation and market expansion (parallel to that of computers) boomed, despite increased regulation. All media became allowed, credit derivatives, or commodities, without any further consideration of the new risks created, since this was very profitable. The theoretical aspects were very much overlooked.

The rise of derivative products was caused, like learning machine today, by technology

The analogy with the current period concerns technology. As with credit derivatives in finance, everything seems possible and feasible. First of all, it becomes a matter of accelerating calculations and algorithms, on which we think first and foremost of technology and computer problems, and only then of analysis and interpretation. As Sandrine has pointed out, the modeller's knowledge and market intuition is no longer central to the validation process.

This reflects the 1960s, when French statisticians surrounding Benzécri were among the pioneers in developing data analysis, also known as principal component analysis etc. The idea was to geometrically represent data, and to project it on the most significant axes. One of the major innovations was to display the graphical representation of the data in the new axes. The data was no longer a statistical model, and spoke for itself. Such graphs could be found, for example, in non-specialized journals such as the «Nouvel Observateur». But this speech could be very difficult to interpret and validate, especially in medicine when confronted with « mixes « of explanatory factors. Its simplicity made it beautiful but ended up limiting its use. This question is connected with current concerns about interpretability, even if today the data dimensions are infinitely more numerous.

Having a large amount of data does not eliminate uncertainty

Another bias is linked to the myth that a large number of data limits uncertainty by providing «probabilistic» information. However, this is clearly wrong if the data is not independent or stationary, hence the importance of empirical knowledge of the data panel you are working on, in relation to the problem you are trying to solve. For example, what is important in an exchange rate is its drift and trend, whether you are investing and its volatility if you are hedging. It has little to
do with it and has a significant influence on the data choice and models to be taken into account.

This bias is less important when referring to physical phenomena, which reproduce themselves in a relatively similar way over time. In the world of banking and insurance, and for much societal data in general, we are confronted with the fact that markets and societies learn quickly. Past data is therefore rarely informative for projections into the future, hence the need for increased vigilance. In fact, all three of us have raised in one form or another the topic of results' validation. The first phase is empirical, and concerns the consistency of the data produced, the «output»: outliers, comparison with the results of the previous day. The second phase is indissociable from a proper representation of problems, and this issue concerns both the data and the models.

The first thing is to have a good representation of the problems. For example, in market finance, the question of the time scale to be taken into consideration is not as trivial as it may seem. The product maturity may be very distant, but the hedging horizon is a day's time, regulatory or almost regulatory, not to mention all the intermediate maturities due to regulation. It is not easy to know how to take these elements into account simultaneously in the data vision and the algorithms that will be implemented. This is also the case in the choice of mathematical models, the only difference being that it is theoretically easier to measure the consequences, and to question the initial hypotheses if the results are not appropriate.

Mathematics requires the use of precise terminology, which helps to structure data processing

Another fundamental contribution of mathematics to the applied sciences, prior to any modeling, is the emphasis it places on vocabulary. There can be no exchanges or ambiguity.
When I started in finance, there were about ten notions under the name of volatility: market volatility, implicit, local, average, currency, and many other qualifiers that were not specified. Imagine the implications of this for data collection.

Moreover, in the matters that we are concerned about, the questions are the same: for instance what is the data for you, or for me? Raw data, re-processed, of large dimensions (which one?), reproducible, original, generated by bank algorithms, bought, sold, personal, retrieved on the web, stored on the Cloud, in your computer, obsolete, historical, etc. We are quickly getting close to a «Prevert-style» inventory. With my banking culture, I'm tempted to associate notions of risk and so this inventory is endless.

In the thirty years that I have observed (as an academic) the world of banking and insurance, whose risk is by definition time-dependent, I have never seen a sometimes weak culture of signals being put in place to warn us on possible breakdowns. This should be a necessary culture in an uncertain world, whether or not bottom-up vigilance is required. This implies increasing the diversity of views, points of view, but also training. It must be stated that although my job as an applied mathematician is to decipher the fundamentally dynamic and uncertain worlds that I study in a new light (financial markets, regulation, life insurance and longevity), I find that the paths are often blurred, and that I sometimes find it difficult to find my way around, even if the challenge is often exciting. I'll explain myself a little better, by taking an academic example. The first time I saw that to draw numbers in a perfectly random way between 0 and 1, we used a deterministic algorithm, I was more than surprised, stunned, you could say. This is the algorithm behind the random key on your computer. It's the basis of all the simulations we talked about. So there is no randomness and I had a «real soft spot» for randomness...

I was «reassured» when I discovered that Google's algorithm for finding the best page (containing a given word) consists in randomly disrupting the system to prevent it from getting trapped in a non-optimal « pit «. «Phew», I could still consider «my» probabilities to be great...

But the world is becoming very intertwined; it is very difficult to quantify the respective importance of deterministic and random aspects.

This is similar to what we were talking about. We get used to interpreting data in a world where we often don't know how and who created it.
Of course, these are only questions, but they seem important to me in our field. Besides, it is not only about interpretation but also about a form of conditioning. I've noticed that I don't think in the same way in math since I've been typing on the screen, compared to when I used to write on paper. All these effects create a Google culture, as you say, Charles-Albert, on research, on the way of seeing, on the vocabulary, which is very conditioning. You have to think about it seriously, so that you don't lose too much credibility in the way of thinking and evolving in the AI world. For instance regarding the «formatting» of algorithms, when you're the one who programmed the algorithm, you know what's inside, you get tired of doing it. You know some of its limits. Yet when you're going to use «open source» software or one of the algorithms in Python or R, there's a whole part of it that you no longer control, that you've appropriated from the outside, a black box. At a certain level, what does it give you in terms of freedom of thought, for example? Where are the singular spaces to ask yourself the proper questions?

BETTER CONNECTION TO THE REAL ECONOMY

TOWARD A MORE ACCURATE VIEW OF THE FUNDAMENTALS

MORE DATA AND BETTER INFORMATION PROCESSING

The digitalisation of the world has important consequences in terms of data production. Not only more and more data are being produced, but also new types of data are now publicly available. This creates great challenges in terms of data analysis but also considerable opportunities for many actors in the finance and insurance sector.
Historically, the finance industry has relied on a number of data sources. First, asset valuation and risk assessment was performed with financial statements and other financial data that the companies would provide to the market. Second, actors in the financial infrastructure, like exchanges and settlement houses, would also provide some very insightful data like market prices but also volumes, order-book related data and other post-trade data for derivative products. Third, national agencies would be in charge of providing aggregate data about the underlying economy. Those data sources have been used for decades in the finance industry.

In addition, there are now quite a lot of other data, coming from different sources, which are usually labelled as alternative data. Such data may include but are not limited to: satellite images, geolocation data, job posting, tweets, social network activity, or textual content in general.

Since the recent technological boom of artificial intelligence, algorithms such as Image Processing Neural Network or Natural Language Processing algorithms are now increasingly more efficient at extracting complex patterns within those alternative data.

The goal is essentially to make predictions from those new patterns. The finance industry may use alternative data to front-run the release of companies' financial statements, or agencies aggregate economic data like jobless claims or even national growth data for instance. The insurance industry may use alternative data to better estimate the damages of natural disasters or use sensors and detectors to better estimate the probability of a car accident or other disasters. All of this is referred to as Nowcasting which is the idea of better predicting the present before the release of official reports which can take from days to months to be publicly available. In other words, powerful data analysis tools, like machine learning, can be used to have a better, more accurate, representation of the underlying economy.

However, with those opportunities come some important challenges. On the one hand, a general concern with machine learning algorithms is their lack of explainability. As those algorithms heavily rely on data and lack explainability, it can be quite challenging to detect a bias in the learning process which could have important consequences whether it is used for nowcasting or other purposes. Building a safe environment for learning and executing machine learning algorithms is therefore an important objective, but this is highly non-trivial. On the other hand, there is also an important education challenge. In fact, running and actually writing and using complex machine learning algorithms require high computer science skills. But a strong economic intuition is also needed as the goal of those algorithms is to learn more about the underlying economic fundamentals. The link between them would usually have to be operated by mathematicians with high level understanding of the machine learning models (without necessarily the implementation details). This can be quite a challenge for teams that try to build and create a better understanding of the economy from data. As the hype surrounding data analysis is unlikely to stop, financial education programs are likely to evolve which has the potential to transform the financial and insurance industry in the long run.

The remaining of this chapter is organised as follows. First, Aimé Lachapelle, managing director at Emerton Data, presents his view on responsible AI in finance and insurance. Second, Christian Robert from the University of Paris-Dauphine explores improvements to the insurance industry made thanks to the use of new data. Lastly, a discussion based on excerpts from FaIR's round tables explores more the relation between alternative data, artificial intelligence and the finance and insurance landscape.

FaIR NETWORK CONTRIBUTORS

Aimé LACHAPELLE
Aimé is co-founder and Managing Partner of Emerton Data, a company specialized in AI and data transformation. He has over 13 years of experience in initiating and leading large data & AI projects, in multiple environments from start-ups to large groups.
Before founding Emerton Data, Aimé was successively Head of Pricing Innovation with Axa and Director Data Science with Capgemini Invent. He conducted end-to-end projects involving data and technology, with a focus on value delivery, mainly in the financial and industrial sectors. Aimé has developed a deep expertise in insurance, and functional expertise in innovation, data-driven pricing, advanced analytics and machine learning. Aimé holds a PhD in Applied Mathematics from Université Paris-Dauphine.

Christian ROBERT
Christian holds a PhD in applied mathematics from University Paris Diderot and is a former student of the Ecole Nationale de la Statistique et de l'Administration Economique (ENSAE). He is full professor in actuarial science at ENSAE, and formerly at Institut de Science Financière et d'Assurances (ISFA). He is a fellow of the (French) Institute of Actuaries.

FaIR NETWORK CONTRIBUTIONS

RESPONSIBLE AI IN FINANCE AND INSURANCE : SHIFTING FROM CONCEPT TO ACTION

Aimé LACHAPELLE,
Co-founder and Managing Partner of Emerton Data

THE POTENTIAL OF AI

Following the lifting, in the 2000s, of technological locks (GPU computing power) and information accessibility (Imagenet database), Artificial Intelligence technologies have considerably penetrated our daily lives.
If anybody is going to push prospective thinking to the theory of singularity – comparable to completely autonomous AI), it is now a reality that our machines, that always remain parameterized by the human expert, are capable of intelligences close to the cognitive processes usually performed by humans1.

The advent of such a technology, even in its early stages of maturity, is on paper an unequalled transformational promise for our society and our daily habits.
In the business world, and in particular in the financial services sector, the major transformation is the shift from expert-based to machine-based modeling and decision-making.

IN REALITY, THE THEORETICAL PROMISE IS QUICKLY PUT ON THE BRAKES

However, it must be acknowledged that the massive adoption of AI happened only for very few applications and business cases. Additionally, the many applications of AI already in place, initially in the insurance and banking industries, are in danger and hampered by the fear of impacts and the lack of human control resulting from that shift.
AI today suffers from a lively and powerful deployment, leading to actions that are the result of decision-making processes that humans and professionals no longer fully master. This typically arises for «black box» algorithms, having potentially learned from biased data sets.

In order to remove such a brake on the development of AI, many initiatives and reflections have recently emerged, in particular across Europe institutions and member states, on how to develop a powerful and trustworthy AI.

Still, many questions need to be answered to switch from these various upstream initiatives to reality and action. How do we create proper short-term incentives for organizations to consider Responsible AI? How do we set-up proper governance at organization level?

In the book “The Ethical Algorithm”, Kearns and Roth reflect on the core societal question of keeping control on human role versus machine role. The risks of algorithmic decision-making are very real: should we expect that avoiding these risks will excessively shrink the benefits? What will be the role for law and regulation, in particular in the financial sector, probably a crucial one to help identify what should remain human vs machine? How do we scale since everything cannot remain human assessment?

In the next paragraphs, after quickly reviewing Responsible AI generalities, we deep dive on specific challenges for bankers and insurers, with a concrete illustration in risk modeling regarding the explainability of machine learning algorithms, in the case of multiplicative underlying structures.

A PROPER GOVERNANCE OF AI NEEDS TO BE DESIGNED IN THE FACE OF ITS SPECIFIC APPLICATIONS’ IMPACTS AND RISKS

Among the various frameworks that can be found in the literature, three main pillars stand out regarding Responsible AI:
-Explainability of Artificial Intelligence algorithms and resulting models that need to be understood to a certain extent by various parties, depending on the end goal;
-Bias and fairness of algorithms and models: the resulting discrimination must be steered and controlled;
-Accuracy and robustness: often as a trade-off with the two formers. Most accurate models are indeed usually less explainable and less fair with respect to given sensitive segments (e.g. regarding unemployment or gender in the case of motor insurance risk). On this point it is clear that not all applications will require emphasis on the same topic: if in medicine diagnostics the accuracy is the main focus, it is less clear for credit scoring or risk pricing.

In its whitepaper published on February 2020, the European Commission highlights the ambition to create an “ecosystem of trust” for AI, making the observation that very disparate initiatives are being launched across its member states: German’s 5-risk levels framework, Denmark’s Data Ethics Seal, etc. Two outstanding elements from this whitepaper are worth mentioning.
First it is important to understand that setting up the governance of a trustworthy AI must pursue objectives as an extension of very fundamentals ones, of which the cornerstone is to protect fundamental rights, as it has done for years in the absence of AI. To illustrate, human decision making is not immune to bias and discrimination. For many applications, we have developed complex systems – law and appeal process in court, audit and regulation in banks, multiple “lines of defense” in risk management, peer review and collegial decisions in medicine, etc. - to steer and control the human decision making. Then basically the same must be done for AI-based models and decisions, which requires specific approaches and tools to control bias in the data (the “food” of AI), but also to control the discrimination outcome from the AI models.
The second key insight from the EU whitepaper, which gives a taste of forthcoming regulation from the EU, is to drive the governance of AI by the criticality of underlying applications and risks. Showing the wrong advertisement to the wrong customer does not have the same impact as making a medical prescription or as rejecting a credit loan or an insurance policy. Regulation must be adapted depending on the impacts and risks.

FINANCE AND INSURANCE SECTORS FACE SPECIFIC CHALLENGES AND ARE MOVING REGARDING RESPONSIBLE AI

Here we reach the core problematic: most mature AI methods and tools have been developed by the GAFAM ecosystem, mainly on advertisement, media and content applications. Therefore, they poorly address specific challenges of higher risk applications such as healthcare, transport, security or in particular financial services.

Making Responsible AI a reality in finance and insurance, an operational set of processes, methods and tools, relies on 3 challenges that are today very poorly addressed:
- Research should focus on the right topics, for instance fairness Machine Learning approaches have been developed only for classification, leaving aside all regression topics, with metrics such as disparate impact, that unfortunately do not compare discrimination resulting from the scoring to inherent observed discrimination, yet an important comparison for many applications such as insurance pricing or fraud investigation.
- Technological solutions are not ready, e.g. to audit models in terms of explainability and discrimination, to take into account fairness criteria during data preparation and model design, etc.

- Governance is complex because the topic is by its very nature transverse, and because uncomfortable trade-offs between accuracy and fairness or explainability must be achieved. What is the acceptable price to pay to constrain AI algorithms? It has implications from front-office functions (e.g. scoring, pricing) to back-office ones (e.g. risk management).
Many financial services companies are currently approaching the Responsible AI topic.
AXA was one of the first movers on the topic, with the creation of a dedicated R&D team within AXA Rev, its entity dedicated to data & new tech. "A few years ago, we launched a dedicated and substantial R&D effort on explainability and fairness", says Marcin Detyniecky, Head of R&D and Group Chief Data Scientist. This effort results in several international-class publications. "We are now starting to focus on setting-up the governance and embedding our research work in operational tools to be provided to the various entities of the group", Detyniecky concludes.

The French bank and insurance supervision, namely the ACPR, has recently led a consultation and working group, resulting in a working document on the "Governance of Artificial Intelligence in finance", co-authored by Laurent Dupont, Olivier Fliche, and Su Yang. This report presents practical guidelines for companies on the governance of AI and provides valuable insights regarding the governance of AI and its inclusion in the company's processes, e.g. regarding the integration to business processes, model validation processes, audit procedures, the definition of human-algorithm interaction, and last but not least the compliance with security.

ILLUSTRATION: EXPLAINABILITY OF AI IN THE FINANCIAL SECTOR

One of the major takeaways from this work recently performed by the ACPR is the sketch of a framework on explainability of machine learning, and in particular on the effort made on defining explainability expectations with regards to the end goal (the degree of understanding of models expected by a client cannot be the same as the one controlled by internal audit or by model validation teams for instance). The document highlights in particular the need to develop tools for the empirical audit of models.

Following the example of AXA, several large financial services groups have launched specific activities on the explainability of machine learning. JPMorgan Research has recently created a dedicated stream on explainable AI. "The challenges include the ability to explain your AI – what we call "AI explainability." When AI is used, the regulatory environment requires banks to justify or rationalize decisions. JPMorgan is trying to be the leader in applying explainability to financial markets", says Apoorv Saxena, global head of AI and machine-learning services for JPMorgan Chase.
To complete the picture, the reinsurance company SCOR is very active on the explainability topic, as confirmed by Antoine Ly, Head of Data Science: "We are currently including explainability of models as a core part of our model governance processes".

Explainability in risk modeling is a perfect example of the need to undertake the development of specific methods in order to make Responsible AI operational. Let's deep dive on the specific example of machine learning explainability for actuaries.

Many methods used to explain machine learning models have been developed during the last few years, with the release of the concept of surrogate models. Here the principle is to fit an explainable model (linear, few variables) on a more complex one, that is on a black box model with sometimes tens of thousands of parameters, which arises e.g. when using deep learning. Several methods and toolboxes are available, but very few of them are usable in practice since explainability remains arguable and since there is no easy consensus nor adoption. Among the most famous ones we can cite LIME and SHAP. This latest method has a nice property: the contributions of variables sum to the total, which is a key property for human expert understanding.

On the other hand, if we take the example of actuaries' risk modeling (e.g. auto claims, home flooding or robbery), as it is well-known that risks factors have a multiplicative effect, these experts are used to a certain kind of statistical models called Generalized Linear Models. Designing such models remains overall very manual and requires a strong effort. For these experts, the switch to machine learning models that are more automated, more complex, and less interpretable, is a real challenge. It has been very much discussed in the community for years, with no breakthrough.

Explainability is probably one of the main reasons for this conservatism. As far as all the research effort was focused on developing additive explainability, actuaries are not equipped with explainability methods and tools that would allow them to assess the multiplicative contribution of risk factors in machine learning models.
A recent paper has been released by the Emerton Data team, "X-SHAP : towards multiplicative explainability of machine learning", co-authored by Luisa Bouneder, Yannick Léo and Aimé Lachapelle, to specifically address this pain point for actuaries and to provide them with a tool to explain machine learning models the way they are used to understanding models.

CLOSING REMARKS

Making Responsible AI a reality relies on 3 main challenges: focus research on the right topics, develop adequate technological solutions, and set-up a governance in a complex and transverse environment.

Among the 3 challenges mentioned, only the first one seems to be already on track: better connecting the R&D effort to business reality. Indeed, this is a clear dedicated emerging activity for several bankers and insurers. The previous example of providing actuaries with specific explainability approaches that suit their current practices is part of the same goal.

Regulators and supervisors are expected to play a key role to make bankers and insurers implement Responsible AI in practice and to address the 2 remaining challenges (technological solutions and governance), by providing guidelines and rules.
One of today's major questions is: How do we leverage regulation and supervision to create short-term incentives to make AI more Responsible, facilitating the development of tools and governance in banking and insurance organizations, without overly slowing down the actual innovation trend?
E.g. supervision requirements can stimulate the development of empirical audit tools but can also slow down Fintech/Insurtech innovation.
How do we ensure consistency with other regulations, e.g. GDPR which imposes time constraints to delete individual data, when AI regulation could impose to keep record of datasets used to train the algorithms?

NEW DATA AND PROCESS IMPROVEMENTS IN THE INSURANCE INDUSTRY

Christian ROBERT,
PhD in applied mathematics from University Paris Denis Diderot and is a former student of the Ecole Nationale de la Statistique et de l'Administration Economique (ENSAE)

Insurance companies have always relied on their own claims data and a few variables obtained at the time of underwriting to price insurance contracts and offer their customers customized premiums. As one of the oldest and most traditional sectors of the economy, it has proven to be quite resistant to changes brought by new technologies until recently. But today, the sector is undergoing a profound transformation thanks to artificial intelligence and the arrival of alternative data such as images, texts, exchanges on social networks or information transmitted by the internet of things. Insurers are now using machine learning in traditional tasks to increase operational efficiency, improve customer service, detect fraud, ... but also in new tasks.

IMPROVING CLAIMS AND CLAIMS PROCESSES

The insurance industry handles thousands of claims and responds to an even greater number of requests from its customers. In some cases, claims do not require the work of human employees and it is interesting to have automatic tools to respond to these claims and to leave more time for employees to deal with more complex claims. Insurance companies are automating parts of their claims handling process today, saving a lot of time and improving the quality of service. Machine learning can improve this process and automatically move claims through the system - from the initial report to analysis and contact with the customer.

Lemonade is a US insurance company founded in 2015 that uses AI to process claims faster and provide customers with prompt payments using various applications such as chatbots. More and more insurance companies are using algorithms capable of extracting handwritten or typed forms from a database with nearly 100% accuracy, helping insurers to reduce response and claim times.

Another example is the technology provided by Tractable, a start-up company that offers software to assess damage and predict repair costs using photos of claims to speed up claims processing. This start- up is changing the way motor insurers will assess claims amounts in the future. The technology examines property damage and predicts repair costs based on real-time photos so that claims can be settled more quickly. Instead of taking days or even weeks to assess a claim after a car accident, insurers using Tractable now only need a few minutes. Previously, this image recognition technology was not advanced enough, and companies did not have access to enough data for their models to operate accurately.

New image-based technologies could transform the way non-life insurers manage their claims and could be an important part of the broader transformation that artificial intelligence brings to the insurance claims process.

INCREASINGLY SOPHISTICATED UNDERWRITING ALGORITHMS

There is a famous saying in the world of actuaries: «There are no bad risks, only bad underwriting». This means that insurance companies should be able to cope with most risks provided they find good pricing. However, many insurers continue to use traditional econometric methods combined with historical risk factors to assess risks. In this context, machine learning associated with new data can provide actuaries with new tools and methods to help them classify risks and calculate more accurate predictive pricing models that should ultimately reduce anti-selection. One example is motor insurance telematics, which is the combination of data collected by vehicles, wireless telecommunication technologies that facilitate the flow of information over large networks, and predictive algorithms. Current motor insurance models that use variables such as age, education level, marital status, vehicle type, etc. have found their limits in identifying the most dangerous and potentially life-saving driving behaviours on our roads. We are seeing the appearance of in-car devices in our vehicles that monitor our driving behavior.

PERSONALIZATION IN MARKETING

Customers expect to receive personalized services that match their needs, preferences and lifestyle. Creating personalized insurance experiences using advanced analytics and automated learning is now an option that firms can use to improve their marketing effectiveness. They can use data on individual preferences, behaviors, attitudes, lifestyle details and hobbies to create customized products, loyalty program services and recommendations. This is achieved by using machine learning algorithms applied to new data to develop suggestions tailored to specific customers through sophisticated selection and matching mechanisms.

By taking advantage of machine learning solutions, insurers can offer their clients personalized services, machine-generated insurance advice or even use chatbots to communicate directly with them. Younger consumers are prone to new technologies and the majority of them are willing to interact with machines in this context.

FRAUD DETECTION AND PREVENTION

Fraud is a major problem that costs for example the US insurance industry more than $40 billion a year. Insurance companies are looking for ways to effectively limit fraud and that could have a positive impact on their profit and loss accounts. And this is where machine learning algorithms can help. They begin to be intensively used by industry leaders to identify claims that are more likely to be fraudulent than others and subject them to further investigation by human employees. Machine learning tools allow insurance companies to take action against fraud much faster than when relying solely on human analytical capabilities.

NOWCASTING NATURAL DISASTER DAMAGES WITH SOCIAL NETWORK

Researchers have begun to use social media platforms to obtain information about natural disasters themselves. For example, it has been shown that the number of photographs uploaded to Flickr is highly correlated with the physical variables that characterize natural disasters (the atmospheric pressure during Hurricane Sandy, for example). Although it is not clear what causes this correlation - external information, network effects or direct observer effects - the correlation suggests that digital traces of a natural disaster can help measure its physical strength or impact very quickly. Based on a similar concept for tweets, other studies have tested the links between the spatial and temporal distribution of tweets and the physical extent of flooding, or the link between the prevalence of disaster-related tweets and the distribution of damage predicted by modelling in the case of Hurricane

Sandy, or some studies have explored the feeding of tweets transmitting observations from socialsensors in the field into predictive models to predict the perceived intensity of earthquakes.

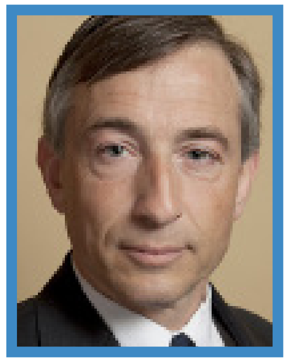
Social media messages posted in the aftermath of a natural disaster have a predictive value that goes beyond the detection of the event itself. The exploitation of these digital traces provides more accurate knowledge of the situation, helping to estimate the consequences

of the disaster on the population and infrastructure in an almost immediate manner.
To date, however, automatic damage assessment has not attracted much attention from insurers, but this should soon change.

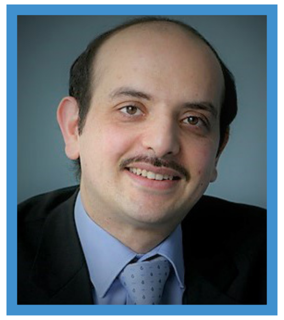
In conclusion, since insurance companies have always worked with data, it makes sense for them toride the wave of digital transformation and implement machine learning solutions that allow them to examine these data more deeply to discover new information. And these solutions are useful for everything from fraud detection to the development of underwriting algorithms that determine the best underwriting strategies. Machine learning is on the rise, so we are bound to see these applications mature and new ones appear on the insurance scene to accelerate the digital transformation of the industry.

ROUND TABLE SPEAKERS

Charles-Albert LEHALLE
Responsible for data analysis at Capital Fund Management (CFM, Paris) and Visiting Scholar at Imperial College (London), Charles-Albert studied Machine-Learning for Stochastic Control during his PhD, 20 years ago. He started his career as IA Project Manager at the Renault research center and joined the financial community in 2005 with the advent of automated trading. In 2016, Charles Albert was awarded the prize for the best article in finance by the Institut Europlace de Finance (IEF) and published more than fifty academic articles and book chapters. He is co-author of the book «Market Microstructure in Practice» (World Scientific Publisher, 2nd edition 2018), analysing the main characteristic of today's markets. He is the Scientific Director of the interdisciplinary research programme «Finance and Insurance Reloaded» at Institute Louis Bachelier.
This program explores the influence of new technologies (from the artificial intelligence blockchain) on the banking/financial/ insurance industry.



Helman LE PAS DE SÉCHEVAL
Executive Vice President and General Counsel of Veolia. He is a veteran executive with a proven track record in both the finance sector and the environmental industry. He has extensive experience in capital markets, insurance, mergers and acquisitions, financial markets and private equity. He is a graduate of the École Normale Supérieure and the École Nationale Supérieure des Mines in Paris.



Oussama CHERIF
Oussama is currently the Director of Innovation for the Smart Automation Solutions division of the Fives Group. During his 12-year career at Fives, he has also created and led the implementation of AI-based digital tools in industrial sectors such as the production of Aluminum, Steel or Machine Tools and more recently for the 3D metal printing industry. Before joining Fives' innovation department in 2008, Oussama was one of the founders and later technical director at Miriad Technologies, a machine learning start-up. Oussama Cherif Idrissi El Ganouni is an alumnus of the Ecole Polytechnique and ENSTA. He also received a PhD from ENS Cachan in Applied Mathematics on Image Segmentation and Parallel Computing.

Bruno DURAND

Bruno Durand is currently referent for Decision and Planning for Autonomous Driving for the Renault Group. During his 20 year career at Renault, he has also participated in numerous embedded software development projects as well as creating in 2016 an algorithmic development team dedicated to the data fusion of sensors used for ADAS (Advanced Driver Assistance). Before joining Renault, Bruno was responsible for several innovative projects at Miriad Technologies, a machine learning start-up. Bruno Durand holds a PhD from the Ecole Normale Supérieure de Cachan on «Multi-signal Neural Classification and Temporal Distortions» and its application to the recognition of decreased driver alertness.

Anne-Sophie TAILLANDIER

Anne-Sophie Taillandier graduated from Supélec (CentraleSupelec) in 1992 and received a thesis in applied mathematics (machine learning) in 1998 from ENS Paris Saclay. She has extensive professional experience with various software publishers. She began her career at Dassault Systèmes in 1998 where she held various positions for 10 years. Before joining ITM, she was the CTO of LTU Technologies, a company specialized in image recognition. She also joined Cap Digital as an expert member of the knowledge commission. Since June 2015, she has been Director of TeraLab a Big Data and AI platform at the Innovation Department of the Institut Mines Telecom. TeraLab drives research and innovation by providing human and technological means as well as secure, independent and neutral infrastructures, enabling data providers, innovative companies and researchers to overcome technical and scientific barriers based on real-life use cases and data.

Stéphane HERBIN

received an engineering degree from the Ecole Supérieure d'Electricité (Supélec), the M. Sc. degree in Electrical Engineering from the University of Illinois at Urbana-Champaign, and the PhD degree in applied mathematics from the Ecole Normale Supérieure de Cachan under the supervision of Robert Azencott. He was employed by Aérospatiale Matra Missiles (now MBDA) from 1998 to 2000. He joined ONERA in 2000, and has been working since then in the Information Processing and Modelling De-partment. His research adresses mainly the design of models and algorithms for data interpretation with a focus on images and videos.

Alain TROUVÉ

Head of Department of Ecole normale supérieure de Cachan, Cachan (ENS Cachan)

Nicolas VAYATIS

Nicolas Vayatis is a university professor and director of the Borelli Center (formerly known as CMLA) at the Ecole Normale Supérieure Paris-Saclay. He is a specialist in machine learning and mathematical modeling and leads a research group working in predictive modeling, network science, and industrial and biomedical mathematics. Nicolas is also co-directing the Master M2 Research Mathematical-Vision-Apprenticeship (MVA) that trains more than 200 students per year in AI research.

ROUND TABLES: EXCERPT 1, 2, 3, 4
ALTERNATIVE DATA : NEW TOWARDS NEW CORPORATE GOVERNANCE AND PRICE FORMATION PARADIGM ?

This chapter will present excerpts from two of the FaIR's round tables that are the alternative data round table and the IA, Math and Industry round table.

- Alternative Data : new towards new corporate governance and price formation paradigm ?, 14/01/2020, As part of the Euronext annual conference, accelerating growth in europe
- Artificial Intelligence, Mathematics and Industry, 14/05/2019, ENS Paris-Saclay, Cachan, Elements of Mathematics for AI

Artificial intelligence is being more and more popularized, made available to the general public and therefore to specialised actors in the finance and insurance industry. This has been achieved through two main channels. First, high grade machine learning tools have been open-sourced for everyone to contribute and access. Second, the world is producing more data than ever, which can be analysed precisely with those open-source tools. However, the finance and the insurance industry have been using data for a long time, so what's new? In the first excerpt from the Alternative Data round table, Charles-Albert Lehalle shares his view on the general definition and classification of data to point out what alternative data refers to.

14/01/2020,
PAVILLON CAMBON, PARIS,
AS PART OF THE EURONEXT ANNUAL CONFERENCE, ACCELERATING GROWTH IN EUROPE

Charles-Albert LEHALLE

Alternative data is an original name because it doesn't define it at all and in fact we started to say alternative because it's non-financial data, i.e. in the financial world we are used to dealing with data generated by the markets, the most important data being used the price, which results from the confrontation between supply and demand and which summarizes a lot of information. Next, we looked at flow information, at the volumes of data exchanged, which is financial data generated by the financial world, then fundamental data on companies: balance sheets, which are also financial data generated by accounting, and finally by accounting through companies' activity, and more recently many data sources have been invited into all this available information in the world around us. So to give you a few easy to understand examples: satellite images, geolocalization on our cell phones, credit card use, company job postings, therefore company job advertisements, transcripts of speeches by company managers, etc.
So all of this is data that is certainly not data generated by the financial world, but which provides a great deal of information about the economic reality that surrounds us. If we want to try to understand a little bit and classify this data more, we can classify it by «originators»: who is the originator of this data. The first source of alternative data is you and me, individuals, through our activity we produce data. So when you drive a car, there is geolocation, when you enter a parking lot there is a count of people in the parking lot.
When we consume with our credit cards, data is also generated. Our own activity, as individuals, when we have a watch that is connected, well, this watch can connect to the internet, put our itineraries, if we go jogging, on the internet every morning, the itinerary is stored, etc. So our activity as human beings and as part of the economic world is available and becomes data. Then there is data that is generated by companies, so I'm not talking about financial data, but just by their activity, so purchases and sales of raw materials, transportation for all the data that makes transportation, all the transportation data, by plane, by sea, by train, so all this activity of companies, web traffic, so B2B companies that connect, individuals - I didn't name them those - who connect to websites and so that makes a second category of alternative data available. The third category is the one that is over, so in fact these are data that have existed for a long time but that were not electronic, it's the data generated by agencies, so we can think of the INSEE in France, which publishes a lot of data on the state of the economy on a quarterly basis. Now there are a lot of independent agencies, there are news agencies, of course, which also provide figures on various objects in the world around us, and then there are various government agencies. All these agencies make up figures that are available. Three years ago now there was an Open Data initiative of the Banque de France. If you're an academic, in particular, you can be eligible for this Open Data initiative and use data that is preserved and kept by the Banque de France.

The increasing use of artificial intelligence creates opportunities to look for patterns in this large amount of data with the goal of getting a better understanding of the underlying economy. By being more connected to the real economy, financial institutions and regulatory agencies hope to be able to improve their reactivity and overall understanding of fundamentals, whether it is driven by profit in the case of private institutions or by policy improvements in the case of regulatory agencies. In the following, there are two excerpts from the Alternative Data round table. First Helman le Pas de Sécheval and then Charles-Albert Lehalle both present some important use cases for alternative data and the questions that would be raised by those use cases.

Helman LE PAS DE SÉCHEVAL

"So to complete what Charles-Albert has just said, I would like to take two examples and try to see how this data differs from more traditional data.
So, for the first example, you know what has a big impact on markets, is the publications of famous agencies, e.g. GDP growth, economic growth which is strong or weak or slowing down or accelerating, it usually has an immediate impact on markets, even before we release economic data, or employment data. Obviously everybody is looking forward to when US employment comes out, quarterly, and the markets usually react to this good or bad news.

So there's a researcher in the United States, whose name I have, whose name is Apoorv Jain, and he decided to try to get the information faster than official agencies were releasing it. He simply analyzed 1.2 billion tweets from 230,000 people who were posting messages on Twitter and selecting of course «I lost my job» or «I found a job». And so clearly with the help of that analysis he managed to get, not the number obviously, not the precise figure, but clearly the trend. He was able to negotiate, trade as we say in the jargon on this market information. Let me take a second example, which is even closer to a manager job.

It's a hedge fund called Point72 Asset Management which was cross-checked by three sources of data: firstly, classic online research, secondly, credit card transactions, then of course aggregated, anonymized, large volumes of credit card transactions, and finally, just like on Twitter, reviews of messages posted on social networks. So with the help of all this, it concluded that it was likely that Weight Watchers, a producer of low-fat meals, was losing ground to its competitors. So, it didn't have the amount, nothing precise. You would say it got the news ahead of time somehow and he was making money. So, I think what's interesting actually is to try to characterize these data. These alternative data are well named: alternative data are to traditional data what alternative management is to management. So we can see that it is not official information, it is not financial statements, nor analyst presentations. They are not press releases from press issued by listed companies or agencies in the case of my first example. Yet it is clear that this is data that requires not only formatting but also analysis, a thorough analysis. You must extract the signal from the noise and then you have to go and do the right analysis. So as Charles- Albert said, there are four essential data sources : Firstly, there is everything that involves satellite images and drones, secondly, there is everything that comes out of the payment means that have been digitized over the past twenty years. So all the bank card transaction data. There is the classic skimming of what can be found on the internet which is more and more abundant, and then finally there are these social networks in which everyone, finally, when you look at the number of individual terminals, the progression of the number of individual terminals and the use of these individual terminals by people, you can see that the data explosion has been exponential and that as a result it has become Finally, when you look at the number of individual terminals, the progression of the number of individual terminals and the use of these individual terminals by people, you can see that the explosion of data has been exponential and that as a result it has become an extremely interesting mass of data. If you know how to work with it, you can find things."

Charles-Albert LEHALLE

"So maybe before I give the floor to business, I'll talk about central banks that have a similar problem, since a central bank's prerogative is to make governance decisions at the country level, and ultimately the only weapon a central bank has is surprise, i.e. it's managing the surprise effect. If information is available on the state of the economy in real time, in an extremely relevant way, then the academic discipline that is developing on this is called now casting, by analogy with forecasting, which is to guess the future, then we are only trying to guess the present. Therefore, if we have the state of the economy in real time, what is the capacity to surprise a central bank?

So, that's a real challenge and there's a real transformation today of central banks that have several hundred economists who usually publish reports at the rate of a human being by digesting information, they're also specializing and having labs, data labs, using a machine to try to understand how to make sensors on the state of the economy that allow them to «compete» in quotes with investors and market participants.

So, the problem already exists at the country level. Obviously, at the company level I think similar things are happening, where there is a kind of competition. Traditionally, when you think about market actors in the old economic theory, market players and then company players, the latter have the private information and the ability to decide, and that's what makes it possible for them to make money, to earn money, to have a successful business. On the side of market participants, it is rather information on the flows that are available. In an investment bank, what they have is information on flows, so they will give a price to the liquidity, which depends on whether or not everyone wants to buy. So, all I see are flows, and I learn from this that the price will go up, so I won't give the same price for liquidity to people who want to buy as to people who want to sell, and that's how these two forces are regulated. Intermediary actors, market participants who

have information about financial flows and little information about economic and fundamental reality, and fundamental companies and investors who lease capital by following longer-term and medium-term information. Nowadays we can ask ourselves if the balance is not changing. It has already happened, we have already had the debate on the frequency trader when it was pointed out, on the frequency trader, when they had access to fundamental information in any case, e.g. they had computers, they could read the news, they could read the dispatches. So, they were more on equal footing with a traditional market maker who only sees the flows and who already has a lot of effort to assimilate all the financial flows that they have to respond to, the requests to buy and sell, well, they're on an equal footing because they know all the flows but no more than the investor who knows the information, who can understand what's going on in the news, read the newspaper and so on.

Technological actors today can process fundamental information, or a first filtering of fundamental information, and that changes the balance, and that is a world that all actors have to be prepared to enter. How do I make decisions that are increasingly public. And so I find it harder to preserve some sort of advantage from operating on information that only concerns my company and me?"

Although it might be the case in the future, It is important to note that none of these use cases work perfectly as of 2020. There are indeed a number of limitations that prevent the ideal goal of being able to perfectly anticipate everything. First, alternative data are usually not very precise. Although regular data can also be imprecise because of aggregation, alternative data are imprecise by nature. In general, publicly available sources of data concerns flows while the internal state of inventory will have to be estimated for instance. Another important limitation relates more directly to the tools used. Indeed, machine learning algorithms, in general, lack explainability which makes it quite challenging to detect bad modelling or biased learning for instance. The following two excerpts from FaIR's round tables detail those ideas. First, in the Alternative Data round table, Helman le Pas de Sécheval and Charles-Albert Lehalle, present their views on imprecision in alternative data. The second excerpt is a discussion between Charles-Albert Lehalle, Oussama Cherif, Bruno Durand, Anne-Sophie Taillandier and Stéphane Herbin in the IA, Math and Industry round table. The second discussion is not directed specifically toward the financial system, but as a general AI concerns, the lack of safe learning environments they mention can be directly extended to the financial context.

Helman LE PAS DE SÉCHEVAL,

So the difference, to get back to it, is that it is at the same time less precise. We manage to have clues, trends, to have leading indicators as we say in economics, but not precise data. We can see that it is less reliable.If there is a place where fakes news is spreading it is first and foremost on social networks. So all these tweets that we are analyzing may be part of «wishful thinking» like we thought, people who are acting rather than corresponding to a specific reality, when we compare this to the process of production and verification of information that I called traditional, i.e. the information disseminated by listed companies or by agencies that publish statistics. But as it comes out ahead of the phase, in advance, it is obviously data that can be processed, that can be traded on the markets.

Charles-Albert LEHALLE,

Yes, in order to give an idea of the level of ease, of the noise level of the information that can be extracted from this data, I think that we have to consider that we are observing all this information related to a company in activity, and therefore often this information represents flows, because what we are observing are incoming and outgoing flows, or customers, or the supply of raw materials, or things that happen in parking lots, in warehouses, etc. And so we see flows, which are noisy because we don't have all the flows, and it's true that there are companies, there are activities that are essentially activities of flow and transformation with just-in-time flows. So, if this is the case, this extra-financial information will give a rather precise vision of what is happening in the company's balance sheet. If now we have companies that manage stocks and therefore have private information that remains inside the company and therefore is difficult to capture by external means regarding the state of the stocks, then understanding the level of noise to comprehend the state of the company's good health or not is more complicated. An anecdote about this is the use of satellite images to understand oil stocks, which were much more used in the last ten years. So it's quite easy because where the oil is stored is in large tanks, they have a removable roof to prevent the oil from coming into contact with oxygen, and so in fact there is a drop shadow, and with satellite images, by measuring the drop shadow you can see whether the tanks are full or empty. It's a technique that has been used for a very long time, i.e. imagery to understand the state of oil stocks. Which is not enough because you also need to know the state of the supply and what's there now. But if we look at the stocks, well, a few years ago the oil companies understood that this is normally private information about the way they managed their stocks and they put covers on top of all these containers, and in fact now we don't have any more information on the state of the stocks with satellite images. So this frontier, we can see that these images that are available outside the companies, it's more like flow information, it allows us to understand the state of the balance sheet and the good health or not of the company. If it's a company that essentially does just-in-time transformations, since when we see what comes in and what goes out, we actually see what is gained or not gained. So, there is private information that is preserved, that can be preserved, and this is what makes the noise level on some activities higher than on others.

ROUND TABLES: EXCERPT 5, 6, 7 ARTIFICIAL INTELLIGENCE, MATHEMATICS AND INDUSTRY

14/05/2019,
ENS PARIS-SACLAY, CACHAN,

Charles-Albert LEHALLE

This is very true. The last sequence of questions is in line with what you're saying. How is a system designed, from the beginning, with an understanding of what will happen when difficulties will arise? We can talk a bit about the man-machine interface. We all know that these systems are not 100% reliable. We hope that the person who ran them in production has put, over it, a monitoring system which will know, at some point, that it is not no longer in its comfort zone and that it is therefore going to have to hand it back to the person. The question is not to give the human the take-over by saying: it doesn't work anymore and where are the controls? you have to give back to the human in an enlightened way, by giving them very quickly enough information to put them in the right mind frame to make decisions to play their role as a human being. In addition, we give them back power when it is more complicated, since all the simple cases have been managed by the machine. I think this is a last round of questions. How is it that we can help humans take control at the appropriate moment? What do we need to do? That has to be managed with design.

Oussama CHERIF

Yes indeed. There is a first issue, that Robert raised, It seems to me that artificial intelligence makes people lazy. Handing over to them once they get used to being passive is going to be a bit complicated. Maybe we should rethink uses differently. Here, I'm going to tease Bruno a little bit on the autonomous car. Do we really need an autonomous car or is the only interesting aspect of the autonomous car to actually replace a valet? I go home or I go to a restaurant, I leave the car, empty, to park by itself. In any case, it will drive smoothly and since there are no passengers inside, it actually solves a number of problems: a primary objective, not to injure pedestrians, and a secondary one, to avoid damaging the car. I'm exaggerating a little bit, but there may be similar things to think about.

Bruno DURAND

Even if we test them, validate them, and revalidate them, our systems are indeed not perfect. What we always try to do is plan how we're going to hand over to the driver. For instance, on the service that consists of triggering emergency braking if the user forgets to brake on an obstacle, when this time becomes less than 2.5s, we switch on a light on the dashboard, which is therefore a first reminder that we send to the driver to take back the hand. If the driver does nothing and there is a collision that is getting closer, a buzzer will sound, which will also incite the driver to regain control.

This is a first example. On more autonomous examples of vehicles, which we expect to see within the next two years, the usage case is as follows: I'm on the highway or on the outskirts in a dense traffic situation. I'm driving at less than 70 km/hr. In this case, what the system will do is to position you in the lane in which you are and regulate your distance from the vehicle in front of you, managing the lane entries of people coming into the lane, motorcycles coming up from the lanes, and so on. What is the driver asked to do? The driver must be seated in the driver's seat. However, they are not required to have their feet on the pedals, nor their hands on the steering wheel, nor even to look. They can read, knit or eat. However, they are warned that they may be asked to take over. In this case, they have 10 seconds to do so. During those ten seconds, we must continue to ensure a safe trajectory.

Generally speaking, we try to assure the driver that the car is in control of the scene in which it is driving. This objective is achieved, for example, via the HMI, in which we present what the system has perceived, what obstacles and vehicles the system has perceived and what lines and lanes the system has detected.

Anne-Sophie TAILLANDIER

On research projects that I could see or follow from a distance, I found that it was very interesting to have an involvement from the social sciences and humanities,i.e. that there is a mathematician and a scientist, on the AI and algo side, and that there is also a team in the social sciences and humanities. It seems to me that AI results in collaborations with social sciences, which I did not expect and which is very interesting. This is for instance the case in the health field. You can imagine a lot of things. It is rather a quantitative approach that will be brought by the very scientific side and a perhaps more qualitative approach that will be brought by the social sciences. I find that the results are very valuable and innovative.

Stéphane HERBIN

The question of monitoring troubles a lot of people at this time, especially in all steering systems, for example. There are the car piloting systems, but the piloting systems of big planes that crash from time to time. The last accident may be due to a defect in the relationship between a system and a human. It is true that these are things that have an impact.

As for doing something interdisciplinary, yes, I think there is also neuroscience that has a big role to play in this, since we are doing intelligence.

Charles-Albert LEHALLE

That's good, we are concluding the round table by addressing the initial topics of interdisciplinarity that you claimed to be something that goes along with the artificial intelligence. Finally, even if we are not necessarily all promoters of this term, artificial intelligence, there is still something behind it that draws other disciplines than hard sciences to interact on rather social and societal issues. Perhaps that's why it deserves to be called intelligence.

The increasing need for data analysis with AI algorithms creates a shift in the skills that were used to be required by traditional actors in the finance and insurance sector. The ideal skill set would be : i) technical skills to implement a particular algorithm for alternative data collection or to deploy large machine learning models, ii) sector specific skills which are needed to create good predictive models, and iii) strong mathematical skills because in the end a machine learning model is nothing more than a mathematical model. Although some education programs are already shifting to incorporate all those skills, it will probably be quite challenging to find all of them in a single profile. This creates an important HR issue because teams with individuals with very different skills will have to efficiently work together in order to allow for this better connection to the real economy. To end this chapter, here are two discussions that happened during the IA, Math and Industry round table. Those two discussions extend the idea of the wide range of needed skills as well as a more precise discussion regarding the role of academia in this challenging context. Although those discussions concern any industry, the financial and insurance industry are not specific and will have to deal with the same issues as more general industries.

Alain TROUVÉ

The first question I'm going to ask you is the following: what should we think about the role of balances and interactions between three types of profiles? A profile that is data oriented and about hardcore learning, in which you need to know PyTorch, Tensorflow. You have to know Hadoop, Python and also perhaps calculation structure problems . A profile that is more of a mathematician, focused on models' manipulation and their properties and algo construction. Finally, the business profile, or for its academic version, the expert profile for questions

and domain, who knows the main paradigms and the prerequisites of the domain. There are therefore three types. This may be too simple a classification, but what can be said about these three elements? What are their roles? What are the balances and interactions? Does it change? Is it an obsolete vision?

Oussama CHERIF

I think the third profile is the most serious one, i. e. the one who knows the trade. It seems to me that the question to raise is the following: what is the purpose of AI? Why do we do it? Is it just for fun? In the industry, we lack experts and we especially lack the possibility of being able to transmit expertise from one area to another, from one generation to another. We want to be able to build, improve an expertise and/or be able to transmit it through tools. Do we need someone who only masters tools, like TensorFlow, etc.? All you have to do is place an offer and there will be plenty of people who will come forward. Personally, I have a little sympathy towards the mathematician profile. Moreover very often in our teams , I try to play this mathematician role, especially to allow people to think differently, because they have a very engineer-like or down-to-earth approach. At a given moment, you have to take a step back and you need to have an abstraction background to be able to do that. You have to have a mathematician background, you don't have to be a proven mathematician, but you need that kind of profile.

Anne-Sophie TAILLANDIER

I believe that AI, even if it's a bit of a buzzword right now - we used to call it big data three or four years ago -, what it's really changed in companies… I feel it strongly. Among all the projects we follow on Teralab - we have more than sixty of them, in all fields, agriculture, industry, health, etc. - we've been working on a lot of projects. Companies need to open up to the outside world. I think that what Oussama has just pointed out is very important. There is a need to be able to pass on this knowledge, especially in the car industry. There are automotive engineers who were there and who must also pass on their knowledge and be able to discuss with data scientists, data engineers, who will prepare the data, and so on. It's something that is quite cross-disciplinary. It can also shake up habits in companies. There can be internal power struggles. There may be departments that say: it's AI, it's my place. There is a question of change management, which is still very important and is an important obstacle. And that's true in all economic sectors. The data scientist, that sort of five-legged sheep that everyone is looking for, who is at the same time a mathematician, a bit of a geek, knows how to use things, has to work with people in the trades, has to be able to understand what we are looking for, and has to be able to take into account the company's information system, because there is a technical architecture that has to be implemented. In the end, it's really a team effort. This kind of teamwork can also be done with outside teams. Sometimes, internal teams need a bit of inspiration - that's why they try to interact with laboratories - in order to lift the technological or scientific barriers. So I want to stress the crucial role of people who know the business and who are sometimes from another generation than data scientists… And this exchange is critical.

Bruno DURAND

Indeed, at Renault, there is a big buzz on AI, there are lots of people and structures that are being set up. Particularly an IA expertise channel that is influencing various trades: design, manufacturing, sales and after-sales.

What we are asked to do is to make a car with different services, such as a driving assistance system.

The major constraint is that we have little or no ground truth. The cost of acquiring and labelling ground truth is very high: e.g. knowing where the vehicle in front of me is, what angle it has in relation to me, how fast it is going, what class it is in, etc. We do small data collections on equipped circuits, but far away from real traffic. For these instrumented scenarios, we equip different vehicles, cyclists and pedestrians with very precise GPS markers, and after a week, we manage to have a few kilometers with a ground truth on the positions and speeds of these objects. Apart from that, we don't have any labeled data, i.e. data on which we have a very precise ground truth. Consequently, for the generation of vehicles to be marketed within five years, we didn't start with a massive learning curve, which would have required millions and millions of data. I stress that the cost of qualified, clean data, for which we are ready to swear that it tells the whole truth, is very high. Of course, precise and measured tests are carried out, such as crash tests, for instance. The service consists in avoiding a collision with a vehicle, a cyclist or a pedestrian. For these tests, the car is allowed to brake by itself on the obstacle. I reassure you, the targets used are made of plastic.

In addition, we drive 800,000 km to verify that on this performance, we will not brake, untimely, for nothing. But most of this massive driving is done without ground truth. This being said, artificial intelligence plays a role in the processing chain of our data.

In order to carry out our services (obstacle avoidance, centering in the lane and regulation, in distance from the vehicle in front of us, etc.), we have a certain number of sensors: cameras, radars, LIDAR. We buy smart-sensors from equipment manufacturers that preprocess acquired raw data and give back in return lists of obstacles and lines, with kinematic (position, speed) and static (class, size) attributes. These sensors are far from perfect.
For example, a radar is very bad at classifying an object, unlike a camera. A radar is very good at estimating radial velocity, something a camera does quite badly. Afterwards, we use the imperfections of the different sensors to have overlaps and imperfections that are not of the same nature each time on the different kinds of sensors. We do more things like tracking with Kalman filters. In short, on the whole processing chain, some components are clearly IA (detection of objects by a camera), others are less IA typed (fusion by Kalman filter).

As is often the case for industrial projects, the efficiency of the processing chain in its entirety is is essential. We develop certain components ourselves, possibly intelligent ones, but our role is above all to master the process from start to finish. We don't need big theories, but rather to implement a lot of simple and mastered ideas We don't need big theories, but rather to implement many simple and mastered things.

Stéphane HERBIN

I will go back to the relationship between modeler and computer scientist. I think that what we're experiencing now is that we're running systems that we don't understand. Typically, what we put into systems that come from learning, which is what Bruno was talking about, are cameras that provide direct object detection. They provide things, but why did they provide them, and how? We don't really know. First of all, because there are intellectual property issues, and even, fundamentally, all those who have tried a bit of deep learning, we manage to get results, but we often don't understand why. And the question, more epistemological or societal, is also: are we ready to pay the price of not understanding something, but of using it anyway? Are we capable of accepting the fact that we won't be able to go through with the modelling? On the contrary, we are going to be able to develop systems that could really be used, derived from data in general, knowing that the data cost is important. I think we're seeing a paradigm shift here. It's a real epistemological shift. This is perhaps what we are currently experiencing. Ten years ago, in vision and image processing, we were trying to find the characteristics and the best filter that would provide the right result. In vision, almost nobody does that anymore. We extract characteristics and get by with them. If we look back, we find that it's not very interesting. We're going to lose efficiency and what do we gain? That's a bit like that too. Are we ready to lose efficiency in order to gain understanding? That's a real question.

Nicolas VAYATIS

I have a quick question. Given your experience, your position and your knowledge of the research world, how do you perceive expectations that companies have today regarding laboratories? Are there things that are changing? How do you view these relationships? Do scientists need to work in a certain way to support the transfer of knowledge and technology towards companies? The feedback we get is that there is a certain difficulty in grasping the state of the art as it is presented today, because there is really a lot of work. It's quite difficult for people who don't have the right codes to find their way around. How do you view this issue?

Bruno DURAND

As far as Renault is concerned, in parallel with the short-term work carried out for vehicles that are due to come out in two or three years, we have five CIFRE PhD students who continue to connect with different state of the art research to understand what will happen on the engineering market in the upcoming years. It stays a very normal functioning where the company, in order to learn, continues to collaborate with labs.

Anne-Sophie TAILLANDIER

Of course there is the classic model, with theses. It will continue to exist. It is in line with what you were saying, Charles-Albert, about continuing education. We can see that all this is evolving very quickly. There is a need for feedback from research done in labs that we can start to integrate into companies. Data scientists in these companies need to interact with researchers in labs. They need to be confronted with inspiring ideas and projects. Of course, companies need to get back intellectual property. They need to be able to implement what researchers are doing in labs. The Chairs model, for instance, is a model that industries continue to appreciate even if they are sometimes a little frustrated at not getting immediate return in terms of intellectual property. This Chair model may need to evolve. Chairs enable company's engineers and researchers to work together on common problems and data sharing. Being able to experiment on company data is a very important point. As AI-related technologies evolve very quickly, and there is a very strong competitive aspect between companies for applications, everyone would like to be the first to implement this type of AI or new algorithm. Companies continue to fund chairs, but are frustrated by not being able to directly implement the results internally.

Charles-Albert LEHALLE

Thanks to my position at the Institut Louis Bachelier, I am indeed seeing a change of model.

Industrial partners are asking for more than just research papers from academics with whom they collaborate. They want actionable research. It is interesting to note that there is also a similar demand within the academic community around the concept of reproducible research. When you are an editor of an academic journal, how do you set up articles, some of which are essentially proposals for algorithms? We must find a way to have a living and runnable version of an article when we publish, so that it can be really checked against real alternative datasets than those of the academic team that submits the article for publication. The journal's editors must ensure that they do not publish an algorithm that only works with the data of a particular article. This is somewhat similar industrial's concern: they require that the results of a chair or the results of a collaboration are also operable. I think this is a need that emerges in a totally cross-cutting way. In my opinion, we must try to provide the same type of response for scientific journals and industries.

LIST OF ROUND TABLES

- Robo-advisor, 16/01/2020, ACPR, Paris, Organized by FaIR (coming soon)

- Alternative Data : new towards new corporate governance and price formation paradigm ?, 14/01/2020, a part of the Euronext annual conference, accelerating growth in europe

- Deep Learning vs. Theory: Illustration with Deep Hedging, 19/12/2019, AFGAP, Organized by FaIR (coming soon)

- New Challenges in Insurance 05/09/2019, Conservatoire National des Arts et Métiers, Paris, Organized by Toulouse School of Economics (coming soon)

- Resolution of post-trades via a distributed ledger, 20/05/2019, acpr, PARIS, Organized by FaIR

- Artificial Intelligence, Mathematics and Industry, 14/05/2019, ENS Paris-Saclay, Cachan, Elements of Mathematics for AI (coming soon)

F a

I R

Finance and Insurance Reloaded

an Institut Louis Bachelier
interdisciplinary program