



BANK OF ENGLAND

Speech

Cyborg supervision – the application of advanced analytics in prudential supervision

Speech given by

James Proudman

Executive Director, UK Deposit Takers Supervision

Workshop on research on bank supervision

Bank of England

19 November 2018

I am grateful to Sian Besley, David Bholat, Charlotte Bull, Stephen Denby, Ryan Lovelock, Clair Mills, Philip Sellar, Pete Thomas and Sam Woods for their assistance in preparing these remarks and conducting background research and analysis.

I. INTRODUCTION

Recognising faces comes instinctively to humans. Until fairly recently, however, it proved beyond the ability of computers. Advances in artificial intelligence (AI) - the use of a machine to simulate human behaviour – and its subset, machine learning (ML) – in which a machine teaches itself to perform tasks – are now making facial recognition software much more widely available. You might even use it to access your bank account.

Because it is so easy for us but so hard for computers, facial recognition is a good illustration of the challenges faced in developing AI. Enabling a machine to teach itself to recognise a face requires sophisticated algorithms that can learn from data. Advances in computational power and algorithmic techniques are helping machines become more human and super-human like. ML also requires lots of data from which to learn: data are the fuel that powers it – the more data used to train the algorithms, the more accurate their predictions typically become. Hence advances in AI are often associated with Big Data and the recent huge advances in the volume and variety of data available (see Figure 1).

As the sophistication of algorithms and volume of data rise, the uses of AI in every-day life are expanding. Finance is no exception. In this speech I want to explore the impact of AI and advanced analytics more broadly, on the safety and soundness of the firms we supervise at the PRA, and how we are starting to apply such technology to the supervision of firms. In particular, I want to explore the seeming tension between the PRA's supervisory regime that is firmly centred on human judgment, and our increasing interest and investment in automation, machine learning and artificial intelligence.

II. CHANGING THE NATURE OF THE RISKS WE SUPERVISE

Like many other firms, banks are looking to harness the power and speed of AI. If you were to take some parts of the media at face value you might be tempted to conclude that a revolution is underway. There are plenty of examples of innovation to point to – from the use of ML-driven financial-market trading algorithms; to the introduction of online banking platforms that generate alerts to customers on trends and irregularities in their spending habits; to new apps that suggest switching utility providers to the cheapest provider.¹

On closer inspection, however, the situation seems rather less revolutionary and more evolutionary. No hard data on industry-wide uptake are available but intelligence from supervisors is that the scale of adoption of advanced analytics across the industry so far is relatively slow. There is clearly, however, the potential for usage to accelerate. At the macroeconomic level, changes in technology, including AI, could, over time, profoundly affect the nature of the financial services consumed and may result in changes to the structure of the financial services industry. This set of issues is being explored at the Bank of England by Huw Van Steenis in his review of the future of the financial system. What matters to

¹ McWaters, J (2018)

us as prudential supervisors is the extent to which the development of advanced analytics changes the risks to the safety and soundness of the firms we supervise.

Increasing levels of automation, machine learning and AI could improve the safety and soundness of firms in some ways. For example, until recently, most firms were using a rules-based approach to anti-money laundering monitoring. But this is changing and firms are introducing ML software that produces more accurate results, more efficiently², by bringing together customer data with publicly available information on customers from the internet to detect anomalous flows of funds.³

ML may also improve the quality of credit risk assessments, particularly for high-volume retail lending, for which plenty of data are available and can be used for training machine learning models. Recent research, for example, analysed more than 120 million mortgages in the US written between 1995 and 2014 and identified significant non-linear relationships between risk factors and mortgages becoming non-performing. These 'jumps' in the chance that a loan defaults – sometimes with just a small change in circumstances – are precisely the kind of non-linear relationships for which machine learning models are well suited.⁴

ML is also starting to influence how wholesale loans are arranged. In contrast to retail lending, the idiosyncratic risks and limited data available for corporate lending make typical automated underwriting more difficult. But ML can still be used to improve the quality of underwriting by making use of non-traditional data. For example, natural language processing of annual reports and social media can give firms useful information on the quality of the credit.⁵

But the increased use of ML and AI may also increase some risks to the safety and soundness of firms. Implementing ML and AI at scale is likely to require considerable investments by firms in their data and technology capabilities. While in the long-run these investments could increase revenue, in the short-term they are likely to increase costs. They will also amplify execution and operational risks. And even if firms eventually are successful in embedding new tools and techniques, these may make their businesses more complex and difficult to manage. For example, while ML models could alter banks' trading and retail businesses – enabling them to make better decisions more quickly – the opacity, however, of these models may also make them more difficult for humans to understand. Boards, senior management and staff in firms may consequently need different skills to operate an effective oversight, risk and control environment.

² Breslow, S., Hagstroem, M., Mikkelsen, D., Robu, K (2017)

³ Arnold, M (2018)

⁴ Sirignano, J., Sadwhani, A., and Giesecke, K (2018)

⁵ Institute of International Finance (2018)

III. CHANGING THE METHODS BY WHICH WE SUPERVISE

Advanced analytics are also likely over time to lead to changes to the way we do our jobs as supervisors. To see how, it is perhaps easiest to go back to the basics of what prudential supervision actually is.

Our approach to promoting safety and soundness is based upon forward-looking judgement-based supervision, in which we identify the key risks facing firms and set supervisory strategies to mitigate them. Described as a business process, it can be broken down into a number of simple steps: 1) rule-setting and reporting; 2) analysis and monitoring; and 3) setting and communicating a supervisory strategy to mitigate identified risks. Each of these aspects of supervision is amenable to automation, machine learning or AI to some extent.

With respect to rule setting, for example, a project is underway to use advanced analytics to understand the complexity of the PRA rulebook.⁶ We hope to use the results to identify ways to simplify our rules to make them easier to comply with.

The PRA Rulebook contains 638,000 words –77,000 words longer than *War and Peace* in English translation. The complexity of the language used can make the text difficult to read. Another layer of complexity is added because of cross-references and links between different parts of the Rulebook, requiring the reader to refer backwards and forwards, disrupting reader flow.

Figure 2 is a visualisation of the Rulebook. Each node is a part the PRA Rulebook. Each line between the nodes is a cross-reference in the text. When parts of the rulebook are linked together, tweaking one part can have unintended consequences for others. We can quantify the interconnectedness of different parts of the rulebook using the PageRank algorithm, the same algorithm used by Google’s search engine. A higher score implies greater connectivity of a particular part to other parts. Happily, most parts of the rulebook are self-contained and “structurally simple”. Looking further into the future, a bigger win might be to automate the rulebook entirely.

Regulated banks are required to submit large quantities of data to regulators. The cost of collecting and reporting data to meet regulatory requirements is a significant burden to both regulators and regulated firms. Regulatory data collections also have significant time lag s, normally 4 to 6 weeks.

One solution is to make the data reporting process better tailored to the needs of supervisors. Digital regulatory reporting (DRR) is the automation of regulatory data collection, and could potentially lead to significant improvements in both the cost and timeliness of data. The idea is based on machine readable reporting requirements that firms’ systems could automatically interpret and satisfy via a secure regulator-firm digital link. This would allow regulators to collect data on an ad hoc basis from firms as required, in close to real time without any manual intervention at either end. That would enable

⁶ Amadjarif, Z, Brookes, J., Garbarino, N., Patel, R., and Walczak, E (n.d) [forthcoming]

supervisors to specify the data they needed to solve a particular puzzle – exposures to a particular country, for example – and transmit that data request to firms in a machine readable form. The data would then be ‘grabbed’ directly from firms’ systems and sent back to supervisors automatically. The FCA and Bank of England are currently undertaking a DRR pilot with participants from a number of regulated firms. It is too early to say what the outcome of this early pilot will be, but initial findings suggest it is feasible. There remain significant technical challenges to be overcome. And regulators would need to guard against the moral hazard that could arise if firms perceived that responsibility for the accuracy and congruence of data had transferred from regulated entities to regulators.

Setting regulatory standards and collecting data is only the start of the supervisory process. Working out what the data mean is a second stage. Recent research has demonstrated how machines can now outperform doctors in the diagnosis of certain forms of skin cancer: machines can be taught to recognise cell clusters more accurately than the human eye can. This does not, however, imply there is no role for doctors in the treatment of cancers; quite the opposite. By using technology to perform certain roles, doctors can free up time to focus on cancer treatment and patient care⁷. This is an example of what is sometimes referred to as human-centred automation “... which considers where humans can often do tasks or make better judgements than machines, and designs automation around these strengths”⁸. In a similar way, by introducing ML to perform complex tasks, we ought to be able to free up and focus supervisors’ time where it is most needed.

Take the case of credit unions. Of the 570 or so U.K. domiciled credit institutions, about 450 are credit unions. These are very small and simple providers of credit facilities that between them account for 0.07% of the assets of the UK deposit taking sector. Because none of these lenders is sufficiently significant to the stability of the financial system as a whole, we supervise these entities in a proportionate manner. That is to say, we only intervene intensively in the event of likely failure, to ensure that insolvency is orderly and that depositors are paid out promptly.

Recent work at the Bank investigated the predictive power of the regulatory returns for these firms.⁹ It found a significant and stable correlation between simple explanatory variables and the probability of default one, two and three years later. In most banking data sets, this structural relationship is obscured by the intervening hand of supervision - leaving few if any observable banking failures.

The research has the practical application for focusing our scarce supervisory resources in a systematic and efficient way on those credit unions where they are most likely to be needed. The tool cannot yet be classified as ML, as there is no learning involved. But it does demonstrate how more advanced analytics can be used to enhance effectiveness of supervision, and we are beginning to experiment with introducing genuine ML into this tool.

⁷ I am grateful to Hermann Hauser for first bringing this argument to my attention.

⁸ Professor Peter Gahan, University of Melbourne

⁹ Coen, J., Francis, W B., Rostom, M (2017)

The task that lies at the heart of supervision – the third step I referred to above – is setting strategies to reduce prudential risks. For each firm that the PRA deems sufficiently critical, we form an assessment of the key risks to its safety and soundness. From that, we articulate a strategy of actions by the bank to mitigate the likelihood and the consequence of those risks. The nature and intensity of the supervisory strategy for a firm – and the resources we allocate – are proportional to the scale of the risks to its safety and soundness, and to the threat the firm poses to the wider economy. We then monitor progress against the delivery of the strategy, as well as the underlying risks themselves.

This approach relies on judgement – about where the key risks lie, the supervisory strategy required to mitigate those risks, and how to respond to risks crystallising. It is a matter of debate how far and how fast AI will be able to move in the direction of making complex judgements. It seems to me to be highly unlikely in the foreseeable future. Perhaps the main contribution it will make is to improve the efficiency and productiveness of strategy-setting. A typical problem faced by supervisors, for example, is the ‘needle-in-a-haystack’ problem: if something is going wrong in a firm, it can be necessary to find out who in the firm made relevant decisions, based on what information, and why the checks and balances of the firm – the board, and second and third lines of defence – did not work.

Advanced analytics can assist. The information to investigate would likely come in many forms – spreadsheets, regulatory returns, management information, e-mails, meeting agendas and minutes. And the information sources may evolve – firms’ definitions of products, business lines, risks, committees and so on do not stay the same. So – along the same lines pursued by law firms for example – one big win is the ability to produce structured data from a range of sources, the analysis of which traditionally required significant manual effort. Over time it may be possible, for example, to train tools to recognise business lines via their numerical characteristics and patterns, and their unstructured data alongside structured regulatory returns. ML also allows documents with similar characteristics to be classified together and analysed, either within or across banks. For example, it could be used to follow the escalation trail from the most junior to the most senior committees. This sort of work is labour intensive when performed by humans: aided by machines, supervisors could in future devote time to those areas where humans have a comparative advantage.

Setting a supervisory strategy without effective communication is pointless, as we rely on the firms to take actions to mitigate the risks. To achieve complex supervisory outcomes – which often require significant, multi-year remediation by firms – boards and senior management of firms have to understand the context and rationale for what we are trying to achieve, as well as what we would deem to be a successful outcome. So getting our communications right is key. But how clear are those communications?

Firms have developed a wide range of more-or-less polite methods for providing us with feedback on the letters we write to them. But letter writing is an art rather than a science, and evaluating objectively how clear we are does not lend itself easily to traditional forms of quantitative methods. Advances in ML,

however, are helping. We recently analysed the letters we write to firms on the key risks they face and our supervisory strategy. We quantified a number of qualitative features of these letters, for example, how blunt we are in our messaging, how personal we are in terms of to whom we address the letter, and the overall sentiment expressed by the letter. We then used an ML model called random forests to detect whether, for example, the PRA writes to firms differently than the prior regulator, the FSA. (We do.)¹⁰ On the back of that project, we have built an app that now enables supervisors to analyse their written communications. Supervisors can use the app to analyse any of their draft documents before they are sent to firms.

IV. CONCLUSION

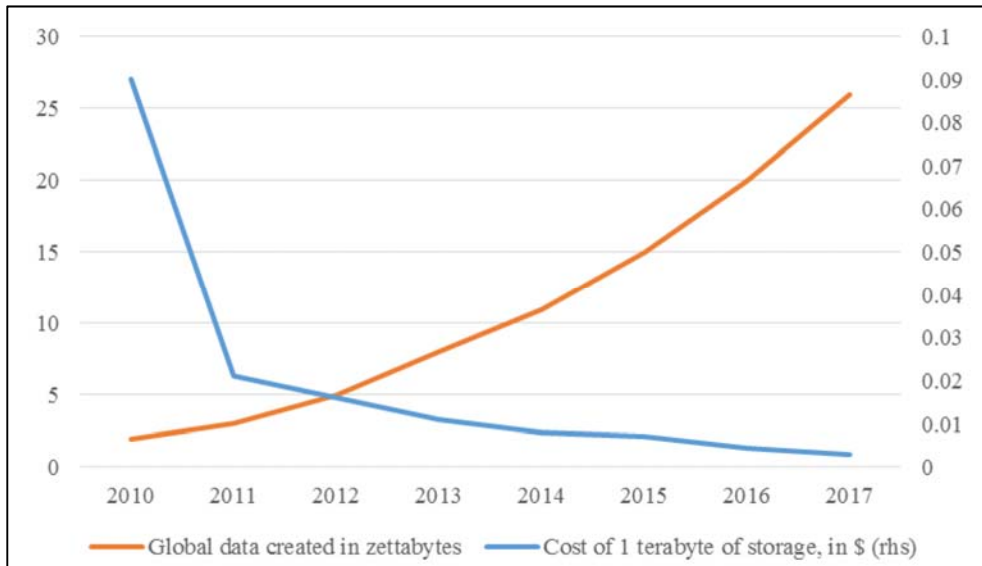
Advanced analytics, machine learning and AI seem to be everywhere now – from image and voice recognition software to driverless cars and health care. Banks too are also seeking to apply these tools and techniques to the range of their activities, many of which used to be seen as the preserve of experts: from risk assessment, to financial crime prevention and trading in the financial markets. These trends are likely to accelerate.

Banking supervisors need to adapt to technology too. Supervisors need to stay abreast of how technology is changing the risks the banks are running and how they are being controlled. And just as advanced analytics are opening an ever wider range of banks' activities to automation, so too are they creating new possibilities for us to supervise banks more efficiently and effectively. But until machines can fully replicate human cognition – a remote possibility for the foreseeable future – supervisory judgment will still have a central role to play. My central expectation is that over coming years the PRA will develop a form of 'cyborg supervision' involving humans and machines working ever more closely together and leveraging their comparative strengths.

¹⁰ Bholat, D., Brookes, J., Cai, C., Grundy, K., and Lund, J (2017)

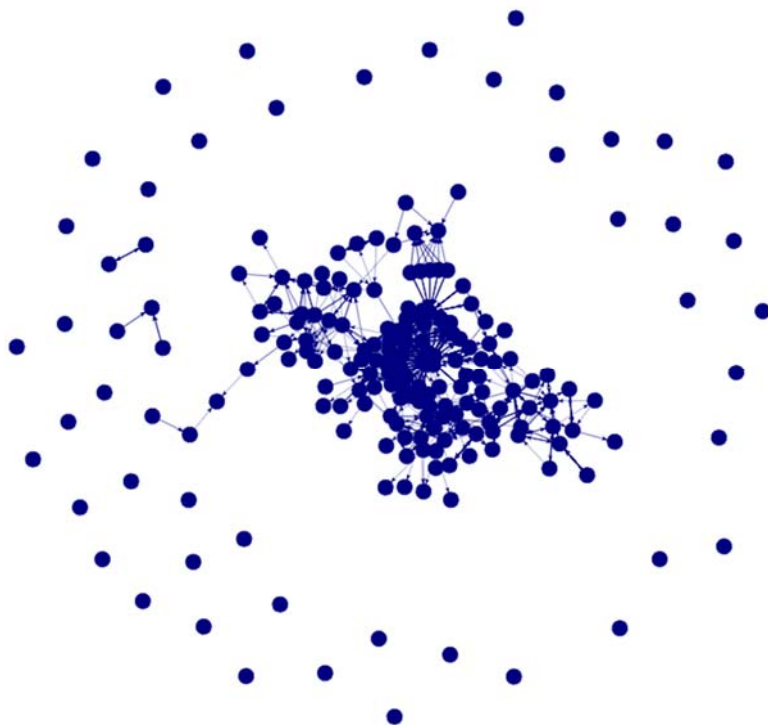
Figures

Figure 1: The quantity and cost of data



Source: Financial Stability Board (2017)

Figure 2: Textual complexity of the PRA rulebook



Source: Amadjarif et al. (n.d.)

References

Amadxarif, Z., Brookes, J., Garbarino, N., Patel, R., and Walczak, E (n.d) Textual complexity in UK prudential regulation.

Arnold, M (2018), HSBC brings in AI to help spot money laundering, *FT* available here: <https://www.ft.com/content/b9d7daa6-3983-11e8-8b98-2f31af407cc8>

Bholat, D., Brookes, J., Cai, C., Grundy, K., and Lund, J (2017), Sending firm messages: text mining letters from PRA supervisors to banks and building societies they regulate. *Bank of England Staff Working Paper* 688, available here: <https://www.bankofengland.co.uk/-/media/boe/files/working-paper/2017/sending-firm-messages-text-mining-letters-from-pra-supervisors-to-banks-and-building-societies>

Breslow, S., Hagstroem, M., Mikkelsen, D., Robu, K (2017), The New Frontier in anti-money laundering, available here: <https://www.mckinsey.com/business-functions/risk/our-insights/the-new-frontier-in-anti-money-laundering>

Coen, J., Francis, W B., Rostom, M (2017), The determinants of credit union failure in the United Kingdom: how important are the macroeconomic factors? *Bank of England Staff Working Paper* 658, available here: <https://www.bankofengland.co.uk/working-paper/2017/the-determinants-of-uk-credit-union-failure>

Financial Stability Board (2017), Artificial intelligence and machine learning in financial services, available here: <http://www.fsb.org/wp-content/uploads/P011117.pdf>

Institute of International Finance (2018), Machine learning in credit risk, available here: <https://www.iif.com/file/24725/download?token=FFv8Lh1>

McWaters, J (2018), The New Physics of Financial Services, available here: http://www3.weforum.org/docs/WEF_New_Physics_of_Financial_Services.pdf

Sirignano, J., Sadwhani, A., and Giesecke, K (2018), Deep learning for mortgage risk, available here: <https://arxiv.org/abs/1607.02470>