# Big Data and Financial Ethics: The Significant Capabilities of Artificial Intelligence Necessitate Human Guidance and Input

**Billie Trinder**

**Abstract:** Innovations in artificial intelligence will revolutionise the financial industry and present new risks and ethical concerns. Consequently, financial institutions should self-regulate. There are two reasons to do so. Firstly, customers cannot be expected to regulate the use of their own data in instances where this precludes accessing a service. Secondly, external regulation lags behind technological advances. The increasing complexity of algorithms and subsequent ethical questions are explored, along with algorithms' potential for bias. Ensuring human participation throughout the algorithmic decision-making processes helps to mitigate associated risks and provides an avenue for implementation of an ethical framework.

## Introduction

Adoption of AI innovations has accelerated enormously in the financial sector. For traditional banks, the stakes are high – as of 2017, more than 80% of executives in the financial services industry believed that business was at risk from financial technology firms (PricewaterhouseCoopers 2017, p2). But as banks recognise the threat posed by FinTech, it is also evident big data lies at the heart of this digital revolution, and financial institutions have unique access to the data. At the 2017 Google Cloud Next Conference held in San Francisco, Chief Information Officer of HSBC Bank, Darrel West, placed AI at the frontiers of finance, stating that "apart from our $2.4 trillion dollars of assets on our balance sheet, we have at the core of the company a massive asset in the form of our data" (FinTech Innovation 2017).

The actual and potential uses for such data are wide ranging and significant – possibilities have been demonstrated in key areas including credit risk assessment and fraud detection. Yet, these opportunities come with significant risks and unique ethical issues. Given that risk affects the proper operation of financial institutions and markets – something we observed all too well in 2008 – the risks associated with big data seems to be undertheorized (Cockcroft 2018, p327).

**Complexity and Transparency**

Debate and regulations focus on the right to privacy at the moment but use of big data analytics in the financial industry gives rise to more material risks. Consequences of 'biased algorithms' are an especially significant concern when it comes to using big data analytics for credit rating. Non-traditional data can be useful for assessing creditworthiness. This was evident as early as 2002, when JP Martin, an executive at Canadian Tyre, collated information his company had collected from credit card transactions over the course of a year. From the data Martin was able to predict a person's likelihood to default on payments through analysis of purchases (Duhigg 2009). These purchases were categorised in terms of "riskiness". For example, those who bought premium birdseed were in the bottom 1% of risk (and thus some of the least likely to default on a payment), and those who bought chrome scull accessories in the riskiest 1%.

From the masses of information collected from a person's internet activity every day, analytics can make incredibly accurate predictions about traits and patterns of behaviour. However, unlike Martin's correlations between purchasing and risk, these predictions can be drawn from data with no obvious relevance to a person's financial activity. In fact, predictions can appear to have no logical relationship to the material on which they are based. A study of over nine million Facebook 'likes' conducted by Cambridge Psychometrics Centre revealed that 'liking' curly fries or Morgan Freeman's voice was a strong indicator of high intelligence, and people who 'like' the page "That Spider is More Scared Than U Are" were likely to be non-smokers (2013).

This ability to find correlations between seemingly unrelated pieces of information becomes a danger when the logic upon which the algorithm works is lost. Machine learning means that some software applications are able to reassess and alter their operations based on input of new data sets. Additionally, large institutions employ multiple people for the development of a single program. As a result of both these factors, algorithms can become so complex that in retrospect their working and logic may be rendered opaque even to their creators. Most obviously, this is an issue of transparency loss. By relying on systems, which are incomprehensible even to the institutions using them, effective auditing becomes a virtual impossibility. When it comes to banks using big data analytics for key tasks such as credit rating, transparency issues become institutional problems. These "black box algorithms" grow to greater problems when we mistake their mathematical, apparently independent modus operandi for objectivity and infallibility. Whilst such programs process data at a speed and scale far greater than can be done by any human brain, the fact the algorithm foundations are often built on data sets collected from human subjects makes results susceptible to error and bias.

One often cited example of biased algorithmic decision-making is in the criminal justice system. Across the United States, courtrooms are increasingly using algorithms to predict risk of reoffence. The allocated score informs vital decisions such as assigning bond amounts, and is even used by judges for reference in criminal sentencing. A 2016 study (Angwin et al. 2016) revealed that these scores were only 61% accurate in predicting criminal activity over the subsequent two years, and 20% accurate in predicting violent crimes. Even more concerning, in retrospect, it became evident that the algorithms exhibited significant racial bias, falsely flagging black defendants as future reoffenders almost twice as often as white defendants.

Examples like these are plentiful, from gender affecting job related ads, to Amazon's same day delivery service being unavailable in black neighbourhoods. Such bias could just as easily be unconsciously built into an algorithm used to determine credit scores. Additionally, the fact that data mining algorithms can process an incredible number of factors for these types of "risk assessment" tasks means that discrimination could emerge through correlations between a protected class and other attributes. For example, association between credit score and race from using a person's address. This type of algorithmic "discrimination by proxy" (Datta et al. 2017, p1) could be harder to identify, and thus correct. Additionally, as the results of these algorithmic processes become part of a person's data footprint, the risk of 'cascading disadvantage' emerges, as algorithmic decision making reinforces itself.

**Responsibility**

As financial institutions seek to catch the big data wave, legislation is necessary but not sufficient for the foreseeable future. If anything was made clear by Mark Zuckerberg's hearing before the Senate Commerce and Judiciary Committees, it is that regulators are not just struggling to keep up with technology – they are foreigners, hopelessly lost in the technological landscape to which those like Zuckerberg are native. This somewhat unnerving dynamic was not lost on spectators, as twitter flooded with jokes that played on the Senators' ignorance: "Senator: *lowers glasses, squints at phone* how can I tell if my granddaughter blocked me?" One tweet suggested that "you should have to prove you can successfully post something to FB before getting to question zuck."

Even in jurisdictions with advanced data protection legislation, the law inevitably lags behind industry. With implementation of the General Data Protection Regulation (GDPR) on the 25$^{th}$ of May 2018, the European Union led the charge with the most significant and comprehensive data protection legislation to date. The new regulation effectively addresses misuse of data as was present in the case of the Facebook-Cambridge Analytica scandal, including the requirement for purpose limitation (data not be used for a purpose other than that for which it was collected). The GDPR also legislates all relevant information be presented to the subject in clear, appropriate language.

While this regulation marks a significant step forward, even the GDPR is inadequate when it comes to managing the relationship between data subjects and the financial industry. The regulation recognises that lawful data processing requires the processor to obtain the freely given consent of the data subject. But, as we increasingly rely on financial institutions for the execution of basic daily functions including deposit and transfer of funds, exchanging currencies and taking out loans, an individual cannot be reasonably expected to have data privacy at the expense of not having access to basic financial services. As Abraham points out, true ability to control one's own data comes at the price of "technological and financial hermitage".

In a context where allowing access to personal data is the norm, restricting personal data puts a person at a disadvantage. For instance, where big data analytics is being used to determine a person's credit score, the choice to restrict this information could result in an unsuccessful application, as this lack of data could be seen to indicate comparative risk or a desire to hide unsuitability. As such methods and processes are typically categorized as trade secrets, and financial institutions are especially secretive with regard to such matters, it becomes difficult to challenge a negative decision.

**Recommendations**

As individuals and their governments wield little power when it comes to ensuring they are treated fairly where algorithmic decision making is concerned, financial institutions themselves must develop an ethical framework and code of conduct. It is in their own interests. Market researcher Gartner suggests that in 2018, approximately half of all business ethics violations could be attributed to improper use of big data (Noyes 2015). Issues including biased data mining algorithms open banks up to anti-discrimination lawsuits, and loss of trust and reputation from their customers. Conversely, as people become more aware of issues of artificial intelligence in the financial industry, an opportunity emerges for firms who retain consumer trust through effective management of data.

As big data is changing the financial industry at an incredible rate, any solution to the risks associated with big data and artificial intelligence must also look to safeguard against future issues. Aside from change, the only evident constant is ever-increasing complexity. A host of concerns emerge when understanding of the working logic of an algorithm is lost. As algorithms can be inaccurate or biased due to their basis in imperfect data sets, the human role in algorithmic decision making must be made purposeful, active and consistent. One way consistent purpose can be achieved is through the use of subject based searches. Subject based searches require that an analyst determines the variables or patterns for which an algorithm will search. Conversely, pattern based searches do not require this – the algorithm itself can search for and find patterns without direction (Dahashner 2016, p4).

Using subject based searches helps to ensure that recommendations are properly assessed, and that an analyst can be held accountable for outcomes due to her active participation at all vital points of the process. This is also helpful as a way of ensuring that data mining systems can always be reduced to a 'human language explanation', and therefore, prevents the algorithm from evolving to the point at which it is no longer 'auditable'. Ensuring deliberate human participation in these processes would also offer an avenue for implementation of a code of ethics. Such codes help ensure organisations avoid unwanted outcomes of AI processes. Finally, this mode of analysis lends itself to greater understanding of variables. The analyst can thereby make more insightful assessments of causation, rather than just correlation, helping to avoid random or even discriminatory findings (Zarsky 2011, p291).

## Conclusion

The ethical issues and risks associated with use of artificial intelligence in the financial industry does not justify rejection of these new technologies. Rather, the opportunities artificial intelligence and big data offer are massive and if used properly, could herald positive change for ethics in finance. To mitigate risks, financial institutions must take it upon themselves to properly and honestly self-regulate by ensuring humans retain a deliberate and informed role in making important decisions.

-x-

## Bibliography

*Analytics and Machine Learning at the Frontiers of Finance* (2017) Fintech Innovations. Accessed: https://www.enterpriseinnovation.net/article/analytics-and-machine-learning-frontiers-finance-1246944296

Angwin, J. et al. (2016*) Machine bias*. ProPublica. Accessed: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

Cockcroft, S. (2018) *Big Data Opportunities for Accounting and Finance Practice and Research*. Australian Accounting Review. 28(3).

Danaher, J. (2016) *The Threat of Algocracy: Reality, Resistance and Accommodation*. In: Philosophy and Technology. 29(3)

*Digital records could expose intimate details and personality traits of millions* (2013) Cambridge Psychometrics Centre. Accessed: https://www.cam.ac.uk/research/news/digital-records-could-expose-intimate-details-and-personality-traits-of-millions

Datta, A. et al. (2017) *Proxy Non-Discrimination in Data-Driven Systems: Theory and Experiments with Machine Learnt Programs*. Cornell University. Accessed: https://arxiv.org/abs/1707.08120

Duhigg, C. (2009) *What does your Credit-Card Company Know About You*. The New York Times. Accessed: https://www.nytimes.com/2009/05/17/magazine/17credit-t.html

Noyes, K. (2015) Mind your analytics, or you could soon be in ethical hot water. Computer World. Accessed: https://www.computerworld.com/article/2993922/big-data/mind-your-analytics-or-you-could-soon-be-in-ethical-hot-water.html

*Redrawing the lines: FinTech's growing influence on Financial Services* (2017) Global FinTech Report 2017. PricewaterhouseCoopers. Accessed: https://www.pwc.com.au/financial-services/fintech-redrawing-the-lines-global-report-2017.pdf

Zarsky, T. (2011) *Governmental Data Mining and its Alternatives*. Penn State Law Review. 116(2).