Feature

KEY POINTS

- Multinationals (including financial services firms) rely increasingly on big data and artificial intelligence as they strive to find efficiencies and to innovate, with multiple touchpoints across each organisation.
- While regulators are now focusing on this topic, the pace of innovation and development means that use of big data and artificial intelligence has outstripped legislation and regulation.
- Use of big data and artificial intelligence carries both risk and opportunity, from a human rights perspective. Careful consideration should be given to the quality of data being deployed, and the potential impacts on stakeholders' privacy and right to non-discrimination. Equally, effective use of big data could, for example, serve to combat financial crime, help oversee good conduct, and support financial institutions' human rights due diligence.
- Thoughtful governance of data and artificial intelligence is therefore key. A set of core principles concerning data ethics, together with centralised senior management oversight, may be appropriate for many organisations. A well-designed, implemented and documented process for overseeing all aspects of data handling will both mitigate risks and be the best defence in the event of errors or challenge.

Authors Susy Bullock, Matthew Nunan, Michelle Kirschner and James Cox

Big data, ethics and financial services: risks, controls and opportunities

With use of big data growing exponentially over the past ten years, how are legislators and regulators addressing big data and artificial intelligence, and what are the key considerations for financial services firms at this time? We explore these themes below.

INTRODUCTION

The use of big data by multinationals, including financial institutions, has grown exponentially over recent years: from two zettabytes globally in 2010, to an anticipated 175 zettabytes by 2025.[1] With one zettabyte equivalent to approximately 250 billion DVDs (according to UC Berkeley), the sheer scale of big data usage is breath taking.

But what is "big data"? Gartner's *Glossary for Information Technology* defines it as "high volume, high-velocity and/or high-variety information assets that demand cost effective, innovative forms of information processing that enable enhanced insight, decision making and process automation". In layman's terms (and drawing on the Oxford English Dictionary definition), it means very large amounts of data, typically so voluminous and complex that its manipulation and management present significant logistical challenges. For that reason, management of big data often goes hand-in-hand with deployment of artificial intelligence (AI), which is the use of a non-human system to learn from experience and imitate human intelligent behaviour.

There are multiple touchpoints for big data within financial institutions: from supporting key risk management and control functions through due diligence and financial crime analysis, to facilitating business analysis and marketing strategies, and targeted credit and lending decisions by the business. More recently, financial institutions have identified the value of big data as a freestanding, tradeable commodity – with some institutions now monetising big data accrued through their own business services or research.

Yet, while the advantages of big data have been hailed by business leaders, policy makers, scientists and educators, the pace of growth of both big data and AI has outstripped the pace of regulation. A dearth of hard law and/or uniform regulatory guidance and enforcement to tackle these global developments compounds emerging and deep-rooted concerns around ethical use of data, with perceived serious risks to rights to privacy and rights to freedom from discrimination. As the 2017 European Economic and Social Committee Study on the Ethics of Big Data observed:

> "Being exposed to the influence of data analytics can be a lifelong experience for individuals that, nonetheless, still have little awareness of how their data are used to predict their behaviour and shape their virtual identity. This knowledge asymmetry makes individuals vulnerable, with limited resources to fully exercise their fundamental rights and freedoms."

In this article, we explore four key observations for financial institutions in this challenging environment. Before doing so, however, we consider the current legal and regulatory position for big data and AI.

LEGAL AND REGULATORY LANDSCAPE

Discussion of the use of big data and AI is often synonymous with discussion of the protection of fundamental human rights such as data privacy and freedom of expression: as captured in the Universal Declaration of Human Rights 1948 and reiterated in various international treaties and covenants, such as the International Covenant on Civil and Political Rights 1976 and the International Covenant on Economic, Social and Cultural Rights 1976.

However, at both the international and domestic level, there is little law directly addressing these concerns in the specific context of big data or AI.

Biog box
Susy Bullock is a partner at Gibson Dunn specialising in commercial and financial services litigation and business and human rights. Susy is one of the leaders of the firm's Environmental Social Governance Practice. In her previous role, Susy was EMEA Head of Litigation at UBS. Email: sbullock@gibsondunn.com

The OECD Principles on Artificial Intelligence and the G20 AI Principles offer some guidance, albeit not hard law. Both sets of principles promote innovative and trustworthy AI, establishing that AI should:

- benefit people and the planet through inclusive growth and sustainable development;
- be designed to respect the rule of law, human rights and democratic values;
- be transparent and explainable; and
- be robust, secure and safe, with risks continually assessed and managed.

They apply to all those who play an active role in the AI system lifecycle, including organisations and individuals that deploy or operate AI.

Similarly, civil society has begun to mobilise in this area. The Toronto Declaration 2018 (a declaration endorsed by multiple civil society groups and focused on ethical principles to guide the development and use of AI) calls on governments and companies to ensure that machine learning systems respect the principles of equality and non-discrimination, and that those who believe their rights have been violated have a meaningful avenue to redress.

The majority of relevant law concerns data protection and data privacy – which, within Europe, is derived primarily from the General Data Protection Regulation 2016/679 (GDPR). It is worth noting, however, that this looks set to change. In the EU, for example, the European Commission published (and also launched a consultation on) its White Paper on Artificial Intelligence in February 2020. This set out policy proposals to encourage the development and uptake of AI in the EU and plans for a new regulatory framework to address the broader risks presented by AI. A legislative proposal is expected in the first quarter of 2021.

Data privacy

Described as a "game changer for everyone" by the UK's Information Commissioner, the GDPR placed a spotlight on data protection and privacy.

The GDPR regulates the collection and use of information of identified or identifiable individuals and covers the development and deployment of AI and big data. With an extensive reach, it applies where data processing activities are:

- conducted by organisations (controllers or processors) established in the EU; or
- related to offering goods or services to data subjects situated in the EU (not restricted to EU citizens); or
- where such activities involve the monitoring of the behaviour of individuals situated in the EU.

Detailed scrutiny of the GDPR goes beyond the scope of this article, but it is relevant that the Regulation recognises that the right to personal data protection must be balanced against other fundamental rights and freedoms, including freedom to conduct a business within Recital 4. For example, Art 22 contains a general restriction on automated decision-making and profiling, which is only permissible where it is:

- necessary for the entry into or performance of a contract;
- authorised by law; or
- based on the individual's explicit consent.

The GDPR places a burden on companies to carefully consider the impact of their use of AI through impact assessments (Art 35(3)(a)); and businesses should stand ready to explain (Arts 13-15) to data subjects any algorithmic decisions that are made about them.

Following Brexit, the GDPR has been enshrined into UK law as the UK GDPR by the Data Protection, Privacy and Electronic Communications (Amendments etc) (EU Exit) Regulations 2019. However, the UK GDPR does retain some important differences. For example, the UK GDPR allows for:

- profiling whenever there are legitimate grounds for doing so and appropriate safeguards are in place; and
- data subjects' rights to be set aside if compliance with these rights would seriously impact an organisation's ability to carry out their functions when processing data for scientific, historical, statistical and archiving purposes.

The Information Commissioner's Office (ICO) is tasked with upholding information rights and enforcing the UK GDPR. Its (non-binding) guidance on AI contains recommendations on best practice and technical measures that organisations can use to mitigate risks caused or exacerbated by the use of AI. The "headline takeaway, [from the ICO's guidance] is the value of considering data protection at an early stage … This guidance should accompany that early engagement with compliance, in a way that ultimately benefits the people whose data AI approaches rely on".

Firms developing and deploying AI systems will also need to consider whether their systems are compliant with the Equality Act 2010 (EA). Importantly, the requirement to demonstrate that AI systems are not unlawfully discriminatory under the EA is separate to the requirement to show non-discrimination under the UK GDPR – compliance with one does not necessarily guarantee compliance with the other.

Of course, Europe and the UK are not unique in having data privacy legislation. In the USA, California's new privacy law, the Consumer Privacy Rights Act (CPRA), addresses "profiling" which it defines broadly as "any form of automated processing of personal information". As written the law does not say much about what companies must do with regard to "profiling", but it calls upon the California Attorney to "[i]ssu[e] regulations governing access and opt-out rights with respect to businesses' use of automated decision making technology". Given California's influence in the United States, this legislation may be a harbinger for other states.

Financial regulatory considerations

The challenges and opportunities posed by the adoption of big data and AI by financial institutions has received much attention in recent years, including from international standard-setting bodies, regulators and trade bodies.

The International Organization of Securities Commissions (IOSCO) published a consultation on proposed guidance for securities regulators in June 2020 on the development of appropriate regulatory frameworks for the supervision of financial institutions that utilise AI and machine learning. In the UK, the Financial Conduct

Feature

Biog box

Matthew Nunan is a partner at Gibson Dunn specialising in financial services regulation and enforcement, investigations and white-collar defence. Prior to joining the firm Matthew was Head of Conduct Risk for EMEA at Morgan Stanley, and responsible for the development and use of the firm's Conduct Risk metrics programme.

Email: mnunan@gibsondunn.com

Authority (FCA) and the Bank of England (BoE) conducted a survey in 2019 to better understand the current use of machine learning in UK financial services. The survey report was published in October 2019. Building on this report, the Artificial Intelligence Public-Private Forum was launched in the UK to explore the means to support the safe adoption of AI and machine learning within financial services. The European Banking Authority (EBA) has also published its report on key challenges in the roll out of big data and advanced analytics in January 2020.

Regulators strive to develop regulatory regimes which are technology neutral, not least because it would be impossible for regulation to keep pace with the fast rate of technological change. In any event, from the regulatory perspective, the risks posed by the use of AI and machine learning to both consumers and markets are adequately mitigated by the application of existing regulatory principles, such as proper standards of market conduct and the fair treatment of customers.

The following high-level principles can be derived from a review of the plethora of materials published by the wide array of bodies involved in the oversight of the financial services sector in relation to regulatory priorities for the use of AI and big data by financial institutions:

- governance and oversight, including senior management accountability;
- development, testing and ongoing monitoring;
- data quality and bias;
- transparency and explainability;
- outsourcing and operational resilience; and
- ethical concerns.

KEY OBSERVATIONS

These four observations identify some of the most important risks, principles and opportunities for financial services firms.

All types of data provide opportunities for value; this can include positive human rights impacts

Financial services firms hold or generate many different types of data, each of which can be used, singly or in combination with others. That data may also be valuable to other firms – for example aggregated trading data or client databases.

Data exists in multiple places within each financial services firm and in multiple formats, such as personnel data, trading data or client information, to name just a few. For maximum value, the data has to be combined and cross referenced. It is important therefore to consider how to breakdown silos of data, so that it can be combined and interrogated collectively.

Monetising the data through use or sale is possible, although not without risk. The converse is true – failure to recognise or use data appropriately can cost a firm in fines for misconduct, poor business decisions or simply loss of commercial advantage: failing to compete with other firms who fully utilise their data.

Financial institutions face increasing expectations from regulators (and stakeholders) in terms of use of big data and AI to support their risk management, compliance and controls, for example in efforts to tackle financial crime or manage conduct risk. Further, financial institutions, like other multinationals, are facing calls to implement mandatory human rights due diligence across their operations and supply chain – with draft legislation now in the pipeline at European level, as well as some individual jurisdictions. If implemented, such legislation will impose substantive requirements on organisations, and there may well be opportunities for financial institutions to leverage existing data and strategies to meet these upcoming requirements.

Poor data quality or analytical logic may lead to misleading or specious output

The product of data analysis can only be as good as the quality of the data and the quality of the analysis. Flaws in the underlying data can, if not identified, lead to false conclusions. Similarly, erroneous logic in the analysis will lead to unsupported conclusions. Firms therefore need to be sure that both the data and the manipulation and interrogation of it are carefully vetted. This includes considering the potential for bias in any data sets and whether there are any other data points that could be gathered to check the validity of the population or the conclusion.

By way of example – a recent US study by Bartlett (2019) reported evidence of racial discrimination in face-to-face lending as well as algorithmic lending, with Latinx and African American individuals suffering higher rejection rates than everyone else (60.6% vs. 47.6%). While algorithmic lending was found to improve acceptance rate parity, it was still found to impose racial premiums for mortgage purchasers and refinance mortgages. Academics speculate that this may be due to the ability of the algorithm to discriminate on the basis of intragroup variation such as neighbourhood or shopping behaviour.

Where algorithms and AI are applied to inaccurate or poor data, the impact of any statistical biases (such as for race or gender) may be difficult to detect but will very quickly be reflected and even strengthened in the final output, undermining confidence and integrity of the relevant assessment.

The importance of data ethics: the goals and outcomes of data analysis should be considered against the relevant legal and regulatory framework and against a firm's own values

Financial service firms may have a legitimate interest in treating clients or counterparties with different characteristics differently. For example, poor credit risks will be charged more, while high volume or value clients may get preferential rates. Equally, performance related pay requires differentiation between a firm's own staff. All of these decisions will be based on data, and potentially improved by better or more detailed data analysis.

However, some of these decisions may have unintended consequences – differential pricing may have a discriminatory impact. Reward for performance based on hard data points may miss less easily measured factors. At the same time, data that is produced or delivered for one purpose may have contractual or legal bars on its use for other purposes.

Biog box

London partner Michelle Kirschner leads the London financial regulatory practice of Gibson Dunn, handling both corporate advisory work and contentious regulatory matters. Email: mkirschner@gibsondunn.com

James Cox is a partner at Gibson Dunn in London specialising in UK employment law and data protection. Email: jcox@gibsondunn.com

Firms should therefore give careful consideration to the intended goals of data analysis exercises, and reflect on whether those goals in themselves are consistent with legal, regulatory and ethical principles. When exercises are complete, thought should again be given to the actions that will be based on the analysis, to consider whether they are still desirable and whether there might be any unintended undesirable consequences. With this in mind, it is notable that a number of global banks (including HSBC and UBS) have already published data ethics principles, or incorporated data ethics in their codes of conduct.

Dearth of regulation/lack of uniformity of approach between jurisdictions, mean that financial institutions must be proactive in adopting good governance and data management practices

Closely connected with observation three: the lack of regulation in respect of big data and AI, matched with growing calls for production of data to support various regulatory or other initiatives, means that proactivity and responsible conduct by financial institutions is even more important. Many of these data gathering or publication initiatives are driven by regulators, for example in the spheres of market activity, measuring sustainability or diversity statistics such as the reporting requirements established by the Equality Act 2010 (Gender Pay Gap Information) Regulations 2017.

At the same time, there is increasing concern about the growing intrusion into individuals' private lives and various privacy regulations, which often vary across different jurisdictions. In relation to the gender pay gap reporting, Government guidance notes:

> "It is important for employers to be sensitive to how an employee chooses to self-identify in terms of their gender and the requirement to report shouldn't result in employees being singled out."

Firms therefore must be aware that there may be a fine line to walk in considering how best to handle, store and supply data. The simplest method to meet firm obligations may offend against competing principles and a one-size-fits-all approach may not meet expectations in all the jurisdictions in which the firm operates.

For all of the reasons set out above, thoughtful governance processes are vital to avoid unintended consequences. In a regulated environment where personal accountability is increasingly the norm, a well-designed, implemented and documented process for overseeing all aspects of data handling will be the best defence in the event of errors or challenge. Revisiting processes regularly – to assess effectiveness and salience – will also be essential when trying to stay on top of developments.

Real consideration should also be given to the degree of transparency in the process. One of the key takeaways from recent years is that regulators expect data suppliers (clients, staff etc) to be at least offered an opportunity to understand the use that is made of their data. Two important questions are: "Do my [clients/staff] know I am using their data in this way?" and "Do I think they would feel happy if they knew?". If the answer to either is "No" firms should look again at their processes and governance.

Centralised oversight may be challenging, given the many different disciplines using big data and AI across each institution; however, a set of core principles, with centralised senior management oversight, may be appropriate. Firms will need to match the expertise and manpower available to them, to their ambitions in data gathering and analysis.

For firms intending to push boundaries, both in terms of innovative use of data or in terms of the sophistication of analytical methods, key hires may be critical and oversight mechanisms should be carefully considered.

CONCLUSION

Financial services is, as always, a rapidly developing sector, where innovation often outpaces guidance, and terminology evolves in an organic fashion. While the tendency and temptation can be to see data as a new business opportunity and to jump straight in, the potential pitfalls involved in mishandling data are sufficiently serious that every step should be carefully considered.

Regulators and stakeholders expect that big data and AI will be leveraged where appropriate to ensure effective operations and business services, but firms must also carefully consider the risks to fundamental human rights which may prevail, depending on the nature of collection, processing and deployment of big data. At the same time, the regulatory, governmental and public discussions on the future direction for big data continue and firms will have to pay attention to future pronouncements in this area, seek appropriate guidance and be prepared to revisit their approach to big data and AI on a regular basis.

In the meantime, boards of financial institutions would be well advised to address data ethics proactively: setting the tone from the top down with a holistic approach to good data governance and the implementation of centralised data ethics standards. As with so many things in the financial services sector, good planning, oversight and governance are key not only to maximising the chances of success but also to defending the position in the event of later challenge. ■

1 Alice Gast, 'Why we need to talk about big data', *World Economic Forum*, 9 January 2020, available at https://www.weforum.org/agenda/2020/01/privacy-in-a-world-of-ai-and-big-data/

Further Reading:

- Big enough? Understanding and complying with the law through computational analysis of legal data (2021) 1 JIBFL 3.
- The next generation of financial conglomerates: Big Tech and beyond (2020) 10 JIBFL 689.
- LexisPSL: News Analysis: The Big Data dilemma – fintech.