DATA PROTECTION, FAIRNESS AND SHARING

An introduction to data protection, fairness and sharing

When considering the building blocks of a discussion regarding Data Protection, Fairness, and Sharing, it is useful to identify three key players, and how the stakes and challenges of these players intertwine.

As the player from whom the data originates, it seems natural to begin with the consumer, also referred to as the data subject. In the first stage of the data lifecycle, and in what is probably the most basic instance of sharing data, consumers provide their data to corporations in various ways. Only in more recent years have the interests of consumers, such as the protection and ownership of their data, come to the forefront of the debate on data. This, in turn, can come into conflict with the interests of the second player- the corporation that uses this data, which may be referred to as a data controller or a data processor. Ultimately this user strives to derive value from this data, for instance a financial gain or a strategic advantage.

As the data economy has grown and developed, a gross imbalance of power has arisen between consumers and corporations, and the regulator has stepped in as a key player to mitigate this. However, the regulator also recognises the benefits of leveraging data for both players. Therefore, their mandate is twofold, pledging to both protect the consumer, while enabling and encouraging the free movement of data. As such the third player plays a vital role in harmonising the interests of the other two players.

This paper explores the opportunity for financial institutions in the era of big data. It addresses the risks that are likely to arise, and the responsibilities of the regulator and of financial institutions themselves to alleviate these risks, thus allowing financial institutions to fully realise the power of their data and the technologies that can be applied to it.

A call to action

In the age of big data lies a fruitful opportunity for the financial sector. With this opportunity comes a call to action for the industry- to confront the reality that the industry is hoarding a goldmine of data that is far from reaching its potential.

While other sectors have seen enormous benefits of utilising data, the financial sector still lags behind. Meanwhile, in just over a decade, 'Data Giants' such as Facebook and Google have turned data into a commodity that is surpassing the value of oil. If we consider the intelligence that an organisation such as Amazon reaps from customer insights generated by data analytics tools, and thus contemplate the potential knowledge that financial institutions could gain about their customers from financial data, we are pressed to ask ourselves why the financial industry is not doing a lot more with their data.

Perhaps what sets the Data Giants apart is the innovative and modern mindset that they maintain towards data. They also have the advantage of being incumbents in relatively new industries where there is either a zero-tolerance policy towards legacy solutions, or where antiquated systems simply do not exist.

An unnerving element of this call to action, is that given the surge of the importance of data for the financial industry, combined with how accomplished these 'Tech Giants' have become in mastering the art of data, a threat arises that these corporations may emerge as competitors in the financial industry. This leads us to question why the financial sector has fallen so far behind, and what is it that the Tech Giants are doing that the financial industry is not?

Having access to large amounts of data is a definite strength of the financial industry. However, this advantage comes with the caveat that if this data is not refined and converted into a manageable format, then its value cannot be exploited. This is why implementing data standards is critical for the industry.

The history of poor data management for the financial industry is quite an unforgiving one. For example, if a common data format for mortgages had existed in 2007, banks could have seen growing risks in mortgage-backed securities more readily instead of just relying on top-level ratings. Unfortunately, ten years after the crisis, much of the financial data in the industry remains trapped in black box, legacy systems in either a non-existent or non-extractable format. It follows that the big data problems of financial institutions can largely be solved by releasing this data through the implementation of data standards.

The necessity of better data standards has come to light for many industry players, including the regulator. Data standards can also play an important role in addressing the various risks that arise while utilising data in conjunction with modern technologies, which both the regulator and financial institutions ought to be aware of.

Acknowledging the risks

You cannot have a conversation about data without acknowledging various concerns with respect to privacy, data protection and the ethical use of data. It is not that data itself is a double-edged sword, but how it is used undoubtedly can be. This prompts us to question

how to strike the balance between securing the colossal benefits of data and ensuring that the risks of misuse don't materialise.

The GDPR deserves recognition as a momentous advance in protecting individuals from the harmful effects of data. It has certainly brought the protection of the consumer and their ownership of their personal data into light, and since its adoption consumers have become more aware of how often they share their data.

Given the ongoing pace of technological development, and the assertion that AI and machine learning technologies will replace many functions in the financial sector, the GDPR does not go far enough. AI and machine learning tools essentially learn from the patterns and behaviours of the underlying dataset to which they are applied, and therefore a huge risk emerges if the underlying data is biased, unfairly presented, or even incorrect or inaccurate. Biases can also arise in the interpretation of the dataset (perhaps based on how the data has been presented), and in the AI or machine learning algorithm itself.

With this in mind, controls must be in place to ensure results of using machine learning or AI technologies are explainable and auditable, meaning that they can always be tracked back to a source. A consumer must be able to depend on their bank not to deny them a mortgage on the basis of an AI algorithm that they cannot explain, or of an underlying bias in the bank's data set.

The financial services industry is highly dependent on consumers trusting the system and the institutions that they deal with, and as more customer-focused models are adopted this becomes increasingly important. As financial institutions begin to release the power of their data, they ought to have learned from recent events where misuse of data by Tech Giants has had a damaging effect on consumer trust, as well as on the reputation of these companies.

This is not to say that the financial industry doesn't have its own reputational challenges, as is evident in emerging money laundering investigations. Therefore, procuring the right culture and attitudes internally will be an important factor for financial institutions as data is used more freely. It should be acknowledged that, as financial institutions will probably find themselves eager to implement data strategies and algorithmic tools quickly in order to get ahead and catch up with the capabilities of other industries, quite easily ethical considerations could become an afterthought. This is concerning, considering that the consequences of data misuse could be very severe in the financial industry. To mitigate this, training, education and governance frameworks should be adopted. For example, this could consist of guidelines on how users of data can identify and overcome biases, or that teach users not to put blind faith in algorithms, and instead how a human input can often compliment an algorithmic decision.

The role of governments and regulators

European regulators and public authorities have the responsibility of protecting consumers from the harmful effects of misuse of data, but also to enable the European financial industry to harness the power of data and to remain competitive on a global scale. In many ways, given that post-crisis regulation is demanding financial institutions to report their data more granularly and more frequently, the regulator is encouraging institutions to optimise their data management processes in order to fulfil their regulatory obligations. This also has the intent of protecting consumers from another financial crisis.

Nonetheless, a knowledge gap exists where policy makers don't have a sufficient understanding of what technology developers in the financial sphere can do with data, making it difficult for them to regulate this and create effective guidance on data standards. Governments need to invest in hiring in-house talent with in-depth knowledge of big data, AI and Machine learning so that they can become highly educated about these topics internally.

Upon ascertaining this knowledge, regulatory policy could play an important role by adopting auditing and governance processes. This might involve a future where, in the same way that firms are audited regularly to maintain IT Security certifications such as ISO, the traceability of their algorithmic decision making, and the integrity of their data sets are audited on a frequent basis. The regulator should also have a role in publishing data standards, as well as guidelines on what it means to use data ethically.

Furthermore, naturally, larger data sets mitigate the risk of biases and the government has access to large amounts of data. This sparks the question of whether the government should lead by example by converting data into a shareable, scalable format that can be consumed by the industry and that allows for future disruption.

These are just some of the ways that governments and regulator can ensure that the ethical reality is not lost as the financial industry embraces the era of big data and the revolutionary technologies that come with it.

Embracing the technology

Embracing the latest technologies allows industry players to realise the usability of their data and effectively mine their data resources. For the private sector, at a time where margins are low and competition is high, this is extremely valuable.

Implementing data standards that are technologically digestible allows firms to put technology first. After all, standards such as basic HTML formats and others for sending and receiving data have made software and technology possible. For a standard to be technologically digestible it must be capable of being understood by a developer who knows nothing about finance, thus, allowing the industry to attract tech-talent by giving developers data that they can work with.

The fact that the financial industry is recognised as a more traditional sector should not prevent it from seizing the opportunities of a technology-led future. Without addressing its big data problems, this opportunity is likely to be missed. The winners in the age of big data will be those that invest in organising their data architecture and unlocking their data, thus realising the rich opportunities of cutting-edge technologies. ■

This paper was drafted by Sarah Murphy, In-House Counsel at Suade Labs. Suade Labs is a RegTech firm with a data driven approach to regulation and with a mission to prevent the next financial crisis by bridging the regulatory gap through modern technology.