Data Ethics

Overview

Data ethics is becoming an important area within financial services. This is hardly surprising.

As data analytics, including intelligent and autonomous systems, play an increasingly important role in our world – transforming many different areas of our lives in profound and positive ways - these new technologies and applications pose complex ethical and economic questions that need to be addressed.

The FCA delayed their Call for Input on the access and use of data in wholesale markets in September this year. However, there are growing regulatory questions around how regulated firms are using data. This extends beyond the implications of GDPR and the use of personal data.

The FCA has outlined a number of priority areas to better understand how the use of data and machine learning could shape products and services and the potential implications for consumers and the functioning of markets. These include exploring whether they, as a regulator, should put in place policy frameworks for how firms collect and use data, to protect consumers and enhance market integrity.

Separately, the Centre for Data Ethics and Innovation ("CDEI"), a UK Government body, was established in 2018. Responsibilities for the CDEI include:

- Analysing and anticipating risks and opportunities.
- Agreeing and articulating best practice, including identifying best practice for the responsible use of data and AI.
- Advising on the need for action, enabling safe and ethical innovation in the use of data and AI.

The UK Government updated its Data Ethics Framework in 2018. Given the direction of travel of other public sector requirements (such as IR35), we think it is only a matter of time before elements are transposed into the private sector, and in particular financial services given the FCA's expected focus in this area.

Why is Data Ethics important?

Understanding the moral dilemmas that underpin data ethics can help organisations to understand and mitigate reputational risks. This is also an important tool to drive stakeholder trust – with both clients and employees.

As the use of regulatory analytics increases, the rights of the individual need to be balanced with increasing regulatory-driven and business-driven opportunities for using data. For example, deploying facial recognition unlocking systems instead of passwords might enhance security, however it is increasingly intrusive to individuals.

We have set out the key ethical challenges firms are facing, along with the principles from the UK Government's 'Data Ethics Framework', which could form the basis of financial services regulatory requirements.

Data Ethics

Challenges

Data analytics provides an opportunity to generate enormous positive outcomes, but also to create significant challenges. The first step in developing a more ethical approach to data is therefore to identify the threats that data analytics could pose. We think that the starting point to embedding data ethics in your organisation is by managing such challenges.

Privacy and transparency

Data can use large and sometimes sensitive datasets. This raises the question around privacy and the use of personal data. Thus, it is important to solidify the transparency in relation to personal data processing, ensuring better information is provided to individuals on the way in which their data is processed – this goes beyond making a generic GDPR disclosure to individuals stating their personal data may be processed.

Human agency and oversight:

AI systems should enable equitable societies by supporting human agency and fundamental rights, and not decrease, limit, or misguide human autonomy. Firms need to consider the impacts the broader application of AI may have in their organization, and what the long-term impacts could be, particularly their desired human to machine ratio, in order to ensure human interaction is not completely lost.

Governance and accountability

The creation and use of powerful new technologies requires effective governance and regulation, ensuring they are used safely and with accountability.

In the case of AI, new standards or institutions may be needed to oversee its use by financial institutions, particularly where the use case involves data sets that contain individual data.

Sustainability

To keep pace with evolving digital strategies, the evolution of sustainable data ethics codes must go beyond check-the-box compliance and enforcement of the rules. New data ethics codes must objectively consider the effects new technology and data-uses have on people. Banks must remember that not everything, that is legally compliant and technically feasible, is ethically and morally sustainable.

Managing the AI black box

AI applications often operate like “black boxes” for decision making. Ethics should be a key component in how algorithms are used. Firms should consider potential ethical implications when deploying an AI solution or use case and whether deploying certain use cases will in themselves produce attributable patterns on individuals or their behaviours.

Bias awareness

Erroneous decisions made by algorithms trained on biased data can lead to poor outcomes. Relying on these outcomes can lead to mis-judgements. Firms need to be more transparent, not just with clients but also with employees on how their data is being used.

Data Ethics

The future of regulation

The UK Government has established a data ethics framework for the public sector, which covers seven principles. We have provided an overview of these principles, as we think some of these could be applied across financial services. Firms should consider driving an industry solution, before there is regulatory intervention.

Principle 1: Be aware of relevant legislation and codes of practice

Individuals must understand the relevant laws and codes of practice that relate to the use of data:

- Legislation: individuals should be aware of existing regulation and legislation.
- Data protection by design: individuals should be aware of existing requirements around GDPR and how data usage is impacted by GDPR.
- Data minimisation: GDPR states that personal data usage should be limited to what is necessary.
- Information governance: personal data should be collected, stored, shared, processed and deleted as covered by the GDPR and DPA 2018.

Principle 2: Use data that is proportionate to the user need

Use of data must be proportionate. Individuals must not start any internal project if the use of data is not proportionate to the user need:

- Personal data can be de-identified, also known as pseudonymising or anonymizing.
- Data sources (how and where the data is from).
- Repurposed operational data.
- Repurposed third party data.
- Statistics: covered by the statutory Code of Practice for Statistics and subject to independent regulation.

Principle 3: Understand the limitations of the data

Though legal and proportionate, there may be limitations to the data that make the proposed approach inappropriate, unreliable or misleading.

Things to consider when deciding if a source of data is suitable include:

- provenance (for example how and why the data was collected).
- errors in the data.
- bias.
- if metadata and field names are ambiguous.

Principle 4: Use robust practices and work with the right skillsets

Individuals must ensure they employ robust and consistent practices.

This involves:

- ensuring accountability of algorithms.
- avoiding outputs of analysis which could result in unfair decision making.
- designing for reproducibility.
- testing models under a range of conditions.
- defining acceptable model performance: false negatives and false positives.

Principle 5: Make the development work transparent and be accountable

All work must be accountable, which is only possible if people are aware of and can understand the work being performed. How is that possible?

- Documenting work clearly (e.g. the construct of algorithms) is an essential part of working in an open and accountable manner.
- By giving transparency and interpretability to algorithms.

Principle 6: Embed data use responsibly

To embed data ethics responsibly, this last principle encourages individuals to put in place appropriate long-term processes to monitor policies by determining:

- the implementation plan, including ongoing agile iteration of your live service.
- sustainable and ongoing evaluation methods.
- the method for feedback into the data model, including when it is necessary to retrain using newly collected data.