FIVE PILLARS FINANCIAL INSTITUTIONS SHOULD LEVERAGE FOR ETHICAL DATA MANAGEMENT

Simon Campbell, Principal Consultant

As financial institutions weigh both the business benefits and potential consequences of having access to vast amounts of consumer data, FIs should leverage five pillars at the top of the organization and across geographies and lines of business to ethically manage data. As such, these pillars for ethical data management are agnostic to the domain or business units to facilitate organization-wide adoption.

TRANSPARENCY

Being transparent with data means ensuring that how data is being collected, stored, and used is thoroughly documented. This documentation (in a digestible and accurate format) must be accessible to both regulatory bodies and consumers whose data the FI uses to derive insights, make predictions, or decisions.

Full transparency requires going beyond the raw data itself to include the processing and feature engineering of data, as well as the intent behind any analytics or AI model in which the data is leveraged.

REGULATIONS

Thorough documentation of how data is being collected, stored, and used is not only a pillar of good data ethics but a critical aspect of business strategy to not only improve performance, but also to avoid fines and penalties from regulatory bodies.

Regulatory standards should be considered a baseline, to which FIs should be proactive in surpassing to stay ahead of potential ethical or regulatory issues that could arise in the future.

For example, GDPR gives more control to consumers over their data and the processing of it (Directorate-General EU, 2020), while ensuring the individuals behind the data are not identifiable.

GDPR serves as a reliable benchmark. However, as we have seen, technology is evolving at such a rapid pace that often regulations have a hard time keeping up. As new technologies become available that facilitate new ways of creating, accessing or leveraging data, FIs must ensure that those applications follow stringent internal standards as regulations may not yet have been developed to the same degree.

FAIRNESS AND RELIABILITY

The proliferation of data across the world is one of the reasons FIs have begun to incorporate AI and machine learning across a wide variety of use cases to minimize human effort and augment human decision making. However, a historical lack of fairness in certain areas can cause a systemic issue in AI and machine learning models if trained on compromised data.

For example, consider the Massachusetts Institute of Technology who recently uncovered a systemic inability for AI to detect early signs of breast cancer in mammogram data from women belonging to minority groups. This stemmed from the fact that these groups historically had less access to healthcare in the United States, resulting in there being far fewer data representing these groups during the training of the AI models (McGreevey, 2018).

It is important for FIs to be acutely aware of this potential pitfall across all advanced analytics and AI use-cases, however a few key examples include; credit decisioning on loan applications, pre-approval offers, and hiring and human resource decisions.

FIs must carefully evaluate the data they are using and examine it for potential bias before training and deploying an AI model on it. This will ensure that models can make decisions without prejudice that can unintentionally discriminate in future decisions.

PRIVACY AND SECURITY

With the vast amounts of data being generated and stored, how do FIs ensure it is kept secure and that the privacy of customers, employees, and partners is maintained? The question becomes even more pressing as banks begin to face regulatory pressure to make sure data assets available to third-party providers from regulations such as PSD2.

FIs must establish data architectures that they build across their lines of business to address this challenge. Architectural design must provide both a high level of security to personal data while providing the organization with the ability to quickly access, manipulate, and join datasets to extract value from the data. FIs must also accompany architectural design with a clear data governance structure, including data stewardship and data lineage as well as robust master and metadata management.

ACCOUNTABILITY FOR ETHICAL DATA MANAGEMENT

Ensuring FIs remain accountable for ethically managing data and following the previously discussed pillars discussed is not a trivial task. It requires on-going commitment from the top of the organization. To facilitate this, we suggest FIs do three key activities:

1. Embed data ethics into use-case evaluation frameworks.

2. Establish a centralized Ethics Review Committee to review all use cases before they move to implementation.

3. Understand how data flows across their organization, and apply the pillars of ethical data management across the data lifecycle.

Embedding data ethics criteria into use-case evaluation frameworks forces business decision-makers to consider:

• How they will be transparent with their stakeholders.

• How they will ensure the data is being used and managed in a way that meets regulatory and internal standards.

• If the data to be leveraged in the use-case contains any inherent or unforeseen biases.

This evaluation will help ensure ethical decision making is engrained in use-case selection and prioritization. Conducting business reviews of use-cases before they go into production is nothing new for FIs: this is commonplace to ensure the endeavor is going to result in a positive outcome for the business. However, establishing a central Ethics Review Committee to evaluate use-cases before going to production alongside the existing business review process will ensure that the use-cases that make it to production follow an ethical framework. Additionally, it ensures the organization is applying data ethics and evaluating use-cases consistently, across lines of business and geographies.

CAPCO

THE FUTURE. NOW.