Data Ethics

July 2021

Leading industry practitioners convened to discuss the ethical use of data in the fourth session of DataTalk, building on the publication of the IIF's new Data Ethics Charter. This charter outlines a set of principles for the ethical handling of customer data, in an initiative launched at the direction of the IIF Board of Directors and developed with a core drafting group of Chief Data Officers and their deputies from 26 IIF member firms. This note provides a brief summary of the key themes from the discussion, respecting that the conversation was held under Chatham House rule, and comments are not attributed.

Data ethics is all about trust. This is instinctive for financial institutions. They safeguard their clients' money and help them fulfill their ambitions. To establish and maintain trust, institutions need to be transparent about how they collect data, store and protect it, and share it with others. Telling customers what rights they have over the use of their data and the value that the institution's data practices offers to them is also critical to building trust. Financial firms start from a very strong position: operating under strict regulatory requirements to protect client interests. A recent report by the Bank for International Settlements found that customers trust financial institutions with their data more than they do governments or technology firms.

Create an ethical culture. Institutions need to stress the importance of trustworthy data practices throughout the organization and get buy-in from all employees. That extends from teams that collect data, to the people who write algorithms, to customer-facing staff. There is potential for bias in the underlying data itself and in the models built on that data, whether in AI/ML or more 'traditional' statistical models used in underwriting and customer lifetime value management decisions, as well as those using model outputs in designing products. Institutions need to ensure trustworthy data practices across all those steps, creating a culture to ensure that anyone who wades into the data is clear about the ethical framework that applies to it.

Make your ethics operational. That starts with the data. Is it fit for purpose? Institutions want to ensure that the data is accurate and complete, and appropriate for solving the problem at hand. Ethics can get very manual in AI and machine learning applications, ironically. Technical staff have to be very hands on to validate data and the AI/ML models. They may be under pressure to bring new products to market quickly, and may want to take data from the internet to train their models on the fly, which can introduce unintended problems. Firms need clear governance principles and robust controls to avoid those risks.

Defining fairness, and achieving it. Underpinning trust and an ethical culture is a desire to ensure 'fairness' – but it's not always simple as to what exactly that might mean. Is 'fair' going to be the same for everyone? Some may expect to see universal treatments, while others might focus on ways to benefit the historically disadvantaged. Where AI/ML models may introduce new fairness questions, they could also help to remedy legacy fairness issues in traditional approaches. Financial institutions need to be transparent in how they think of fairness, and how they can pursue and deliver on that aspiration.

Ethics in a dynamic environment. Ethical principles may be fixed but their application is anything but. That's because the technological frontier keeps advancing. As algorithms get increasingly complex and institutions adopt more pre-trained models, it can become harder to ensure they don't make unintended correlations or that the data they have been trained on are appropriate. Newer techniques like Edge computing raise their own issues about whether the data being used satisfies your principles. Data ethics isn't a once-and-done business. Organizations have to continuously learn and adapt.