Sofia Ihsan, Trusted AI Leader for UKI, EY

Biography: Sofia Ihsan is the Trusted AI Leader for UKI and was a lead architect in the development of their Trusted AI methodology. She lead the first global pilots of the methodology and is engaged with academics, regulators and clients in trying to help shape the approach to building trustworthy AI that considers both the benefit and potential detriment to individuals, organisations, government, society and the environment. In her 21 years at EY she has undertaken a number of roles including cyber security, transaction advisory and client secondments across industries.

SUMMARY OF EVIDENCE

Whilst there is much excitement at the potential economic and social benefits of AI, our clients tell us that one of the main barriers to the adoption of AI is a lack of trust. Broad consensus on the conditions necessary to build trust in AI are that it should:

- Be Compliant with ethical and social norms, including corporate values not historically mainstream for technology such as moral behaviour, respect, fairness, be unbiased and transparent - Ethical
- Consider local and macro social impact, including its impact on the financial, physical and mental well-being of humans and our natural environment - Socially Responsible
- Have a clear line of accountability to an individual and clarity on how it operates, the data that it uses and the decision framework that is applied - Accountability
- Perform as intended, not just during the initial training or modelling but also throughout its ongoing “learning” and evolution whilst in use - Resilient

Further AI is probabilistic by its very nature and the predictions and insights it provides will sometimes be wrong. The key therefore in building trust to enable broader adoption is not only to implement robust controls through development to help ensure that it is fit for purpose and delivers to the required level of accuracy but also to have controls in place to help identify, challenge and correct when it’s wrong.

Early learning in using EY’s Trusted AI framework to look at the governance, control and therefore trust gaps through AI development and deployment has identified some common themes across organisations in different industries:
There is no overall ownership or governance structure over AI with a resultant lack of visibility of where AI is being used or built, how the risks arising from this are being managed to help ensure that it is fit for purpose, in line with ethical and societal norms and compliant with applicable regulation. This impacts how an organisation can ultimately demonstrate the AI’s trustworthiness and therefore accelerate adoption.

Whilst there may have been some consideration of ethics, there are often no formalised, accessible, up to date and embedded ethical design policies and standards including an AI ethical code of conduct and AI design principles. The absence of these make it difficult to demonstrate an ethical approach.

There is insufficient consideration of the level of interpretability and explainability required to demonstrate that the predictions and insights provided by the intelligent agent are based on sound logic, can be trusted, are not rooted in bias and treat users equitably.

There is inadequate consideration of monitoring requirements when the agent is deployed to confirm that it continues to operate to the required level of accuracy and in line with ethical and social norms as it dynamically learns from interactions with users and evolves its decision frameworks.

There are no formal guidelines for the procurement of AI components which require additional considerations such as confirmation that they are fit for purpose, built ethically and do not contain biases.

This is an evolving area where regulation is lagging the technology and feedback from clients starting to harness AI is that there is limited guidance as to what they should be doing. However, as the power and responsibility being entrusted to AI increases at pace, it is essential to have a set of guiding principles that enable safe and ethical innovation and build business, and consequently consumer confidence and trust to enable broader adoption. Conversely, without trust organisations and consumers may be unwilling to share the data upon which tomorrow’s innovation lies. We believe that the acceleration of this guidance must be a key priority for government.