



A nod to the future

Eversheds Sutherland and Smart Dubai consider the future of AI governance and how a proactive approach can create a trusted ecosystem and address ethical concerns.

Could 2019 be the year Artificial Intelligence (“AI”) takes off? Around the world, governments are launching initiatives to consider the policy, governance, socio-economical and regulatory implications of AI. In this article, we will focus on one key area – AI ethics guidelines.

In December 2018, the European Commission published draft guidelines on ethical AI for consultation (final version due in April 2019). The Singapore Personal Data Protection Commission published a model for AI governance in January 2019 and many more regulators and governments are either drafting or consulting on domestic versions of AI guidelines.

Closer to home, Smart Dubai published its AI Ethics Principles and Guidelines on January 8, 2019 (the “AI Ethics Guidelines”) following discussions with international counterparts and industry participants. We will consider the AI Ethics Guidelines in further detail below.

WHAT IS AI?

In this article, we will adopt the AI definition from the AI Ethics Guidelines, namely “*The capability of a functional unit to perform functions that are generally associated with human intelligence such as reasoning, learning and self-improvement*”.

The text further clarifies that this incorporates component parts, such as algorithms and machine learning, and does not necessitate autonomy or an ability to perform functions associated with general human intelligence. Given this definition, we can frame the discussion that follows within the context of technologies that already exist, or are soon to be developed, and with a decennial, rather than generational, timeframe in mind.

HOW TO GOVERN AI?

Some argue that any step toward regulation is premature – that AI application is growing in industry and we need to take account of actual application and not stifle

innovation. Others would counter that it is critical this area is regulated given AI's potential to adversely impact both society and individuals. Furthermore, those in favour of regulation argue that AI already impacts many vital decisions and the sphere of influence is only set to increase as societies become increasingly complex and reliant on technology. AI is currently used for many important decisions such as filtering university applications, providing credit ratings, assessing loan applications,

setting insurance premiums, screening job applications, assessing social welfare entitlements and more.

Increasingly, such algorithms are being called out for reinforcing rather than correcting biases and with such prevalent use, this raises important legal and ethical questions.

In a step towards addressing these questions, Smart Dubai has developed and published a toolkit consisting of a set of core principles and practical guidelines

AI Ethics and Dubai: A city government's attempt to govern the ungovernable

WHY AI ETHICS MATTERS?

Dubai is a city of bold statements, from architecture to ambitions for growth and technology adoption. One technology for which we have significant ambition is Artificial Intelligence ("AI").

Smart Dubai - as the city's technology and data accelerator - is clear that exploration of the innovation potential of a technology is matched by governance that mitigates against the downside risks of complex AI systems. We see ethics - being respectful to a core set of values - as core to adoption of AI and wider market development.

For Smart Dubai, the imperative stems from the fact that smart city services will put, not just data, but AI full-square into the lives of people. It is still early stages but given that much of AI's transformative impact lies within city services and operations, and cities operate best on trust, it is vital that government openly acknowledges AI ethics and brokers a discussion around them. Further, in Dubai you will find the emergence of the 'learning-by-doing' approaches needed to navigate the AI fog. Remote international and national government bodies cannot replicate this proximity and active research and development.

THE CITY AS AN (ETHICAL) LABORATORY – THREE MAIN ELEMENTS

1. **Smart Dubai's AI Lab (a partnership with IBM)** - a production line for AI city services, on which the planning, experimentation and stabilisation of AI use cases takes place. The Lab is also leading on the vital training of public sector staff.
2. **Ethical AI principles and guidelines** - a differentiating factor of our work is a self-assessment toolkit, designed to help developers and implementers of AI to judge their systems from various ethical angles. The broader evidence base that this city-level feedback loop generates allows us to further enhance assistance to Dubai's AI ecosystem.
3. **AI Ethics Advisory Board** - comprising legal, technology, and ethics experts, and experts from academia and government.

Deliberately diverse, this board will bring together an array of perspectives so it can govern and provide strategic direction for the first two elements of our AI work. Together, these three elements form a powerful proposition, based around one truth - AI ethics is a field in which no-one has all the answers. Indeed, the diverse competencies and experience needed makes collaboration essential.

FUTURE DIRECTION: COLLABORATION AND REGULATION?

The most important note to strike in ending this piece is that like everyone else, we are in the early stages of AI adoption and in this sense there is no clear path to hard(er) regulation. By publishing our ethical AI principles, guidelines and our voluntary self-assessment tool though, Smart Dubai is signalling, first, the importance of ethics to broader adoption, and second, the leadership and partnership working that is needed, if managing the upside and downside risks of AI is to be normalised.

We will iterate the guidelines and tool - a process that may involve admitting where they cannot have a positive impact, because of the lack of clarity around the potential development path and applications for this technology. We will examine their usefulness to small businesses, larger corporations, other parts of government, and the public. We will also look at the feasibility of AI audits and advisory services, and the capabilities required.

This move towards adaptive and agile governance that enhances, rather than limits, innovation is how we will deliver ethical AI today, and in the longer term, build trust in the technologies driving the digital economy of tomorrow.



ANDREW COLLINGE,
advisor, Smart Dubai Data

to inform ethical requirements in AI design and use. They have also published a self-assessment tool to assist developers in evaluating their designs.

It is important to note that the framework is intended to evolve over time and is not legally-binding.

Core principles

The core principles are Ethics; Security; Humanity and Inclusiveness. A high-level summary is set out below.

To be ethical, AI systems should be fair, transparent, accountable and understandable. This means that organisations should take care when using datasets to ensure that biases are considered and mitigated. Those that design, develop and deploy the system are responsible and accountability can never be attributed to the system itself. Proper governance is, therefore, a crucial aspect. Additionally, AI systems should be as explainable as possible, and individuals should be able to request explanations of any decision that might affect them.

Security requires that AI is safe, secure and controllable. Designated individuals should be able to override any AI or reverse the decisions. AI should never be autonomously capable of hurting, destroying or deceiving humans.

The humanity principle is a nod towards the future and expressly requires that governance models should be developed for artificial general intelligence and superintelligence.

The inclusiveness principle stresses that AI is a global effort which requires continuous international collaboration.

The guidelines

The guidelines support the principles by offering tangible suggestions and practical examples of how stakeholders can adhere to the principles. We have set out below some key examples.

1. **External Audits** - It is interesting to note that the guidelines envision that AI could become auditable for compliance with the transparency and accountability requirements. Similar wording can be found in the European Commission's guidelines and this is one of the areas that may well become regulated in the future.
2. **Automated decisions** - the guidelines require any significant automated decision concern an AI subject to

be capable of challenge and, where appropriate, the subject should be able to opt-out of such decisions.

3. **Explaining decisions** - AI operators should consider high-level explanations (in non-technical language) of the data used by the system, the type(s) of algorithm deployed, which category the subject was placed in and, finally, the important features driving the outcome of the decision.
4. **Black box limitation** - AI operators should consider using traceability for significant decisions in order to ensure that the organisation can trace the key factors that led to a particular decision.

LEGAL ISSUES

The use of AI raises complex legal issues which are beyond the scope of this article.

Existing legislation can, to some extent, be interpreted to answer basic AI question (for example, in respect of human rights, product liability, consumer protection and more) but as the technology evolves in complexity, we expect legislation to follow and this may ultimately result in formal regulation.

As a pre-emptive measure, we therefore recommend that all organisations place AI governance on their board agenda and to adopt a proactive approach, guided by the principles, when considering the use of AI driven technology. ➡

Please note that the information provided above is for general information purposes only and should not be relied upon as a detailed legal source.



Text by:

1. **ERICA CROSLAND**, senior associate, Eversheds Sutherland
2. **CHARLOTTE WALKER-OSBOURNE**, partner and head of technology sector, Eversheds Sutherland