# AI Ethics

## A Call for Action at a Global Level

Global Forum / Shaping the Future

Muscat, Oman

17 & 18 October 2022

October 2022

*"Our future is a race between the growing power of technology and the wisdom with which we use it."*

Stephen Hawking

# Acknowledgements

Föreningen Global Forum/Shaping the Future (the Global Forum/Shaping the Future Association) sets out to bring together multiple stakeholders from around the world, inviting them to apply their collective knowledge, experience and opinions for the purpose of defining which critical ethical issues must be addressed in order to prepare, design, develop and deploy trustworthy Artificial Intelligence (AI) systems.

## Artificial Intelligence and its main challenges

AI is a disruptive shift, which is probably the one of the deepest and the fastest humanity has ever experienced.

AI is already able to perform tasks commonly associated with intelligent beings– diagnosing disease, translating languages, identifying individuals, providing customer and employee service, and so forth. In the coming years, AI application areas will continue to grow exponentially. Machines might, in a not-distant future, achieve what some experts call 'Artificial General Intelligence' (AGI), i.e., the ability to undertake most cognitive tasks at least as well as humans.

This incredibly fast development of AI applications inevitably raises fundamental questions about the relationship between such intelligent machines and humans. Diagnostic errors, false positives in facial recognition, disclosure of personal data without consent and the perpetuation of structural and institutional prejudices are too often the consequences of the biases inherent in AI applications. In these circumstances, the value of AI applications is diminished by failing to protect data and by violating basic privacy principles.

Unwanted outcomes can by countered by thoughtful application of ethical principles governing the creation and application of AI. Besides, it is to be feared that traditional law covers less and less of the digital world. New rules of law must be invented and implemented. Indeed, we are witnessing an acceleration of scientific and technological change, which may make laws rapidly obsolete vis-à-vis the state of art of technology and the society's expectations. The challenge is therefore to innovate innovation in the legal and regulatory domain in order to adapt law- and policy-making to the Internet age.

Public and private organizations around the world have begun to define ethical frameworks for the design and implementation of AI – each one based on its own values, motivations, and principles. To take one example, the EU is working on an "AI Act" – a transformative legal framework which would, if implemented, have profound and far-reaching consequences for organizations that develop or use technologies incorporating AI. Ethical constructs and how they are applied vary, however, reflecting the particular political, cultural and socio-economic contexts in which they arise. In diverse locations, many people express concern with overreach and misuse of technology in the hands of both governments and powerful corporate giants. In any ethical framework governing the design and implementation of AI, based on rules that reflect a recognized set of critical issues and risks, the hope is that the AI deployments are trustworthy.

The Global Forum stands on that line.

## Our ambition

**Raising critical questions from varying perspectives reflecting divergent and convergent views of industry, academia and government when using AI**

We believe that organizations and communities working towards some ethical framework can benefit from engagement in an inclusive review that raises questions about which policies and processes are essential for ensuring sound governance and trust in the deployment of AI.

Although the word "ethics" implies a judgement on values – which unquestionably are not always shared or interpreted in the same way across global countries – and could be replaced, for instance, with the phrase "code of conduct", the Global Forum has decided to keep it in order to comply with the most common terminology in the world.

As a unique organization that can acknowledge the international dimension of AI ethics, not just for technologists, but for the impact upon society, the Global Forum is able to address cross-boundary issues, such as manufacturing, regulation, health, education, and so forth. It has an interesting set of people involvement that brings together divergent and convergent views of industry, academia, government and civil society. A role for the Global Forum, at least to begin with, is to raise the critical questions that come from these varying perspectives and threads of thinking.

## Our vision: Questions in their consideration of the use, development and deployment of AI

The combined knowledge and vision of participants enacted that far will form the basis of a concise document that can be widely disseminated, taking full advantage of the independent and extensive network of the Global Forum, with the goal of influencing companies and governments to include the questions in their consideration of the use, development and deployment of AI.

We are aware that a lot of what we fear is the unknown, as we are still at an early stage in the roll-out of AI. Our hope is that ethical principles will be developed, adopted and enforced to keep humanity as a whole on a trajectory leading towards a future world that we and our children, and their children, will wish to live in.

# 1. Basic issues in considering the use of AI applications

There are fundamental concerns about the future of societies in the light of the rapid expansion of the use of AI. These issues are above and beyond what any particular application or algorithm may address, and more significantly deal with the broader context of the value that AI can bring to society as well as the potential for harm. As the use of AI proliferates, thoughtful organizations will examine these questions in detail:

## 1.1. Digital literacy and Education

- How can AI empower people through education and give them tools that are easy to use to enforce the boundaries set by new policies?

- How can AI support learners, for example through cycles of mastery, greater depth of understanding and experiential learning, so as to increase educational outcomes?

- How can AI allow educational institutions to keep up to the changing needs of society in a VUCA (Volatility, Uncertainty, Complexity, Ambiguity) world with minimum investment?

- How can AI systems be used to educate children or students who have no access to schools or universities?

## 1.2. Employment

- How can AI systems be used to enhance skills?

- Are there easier areas of training/re-training that could be identified and become available on shared systems to increase access for workers?

- Will AI enable a sensible division of labor between technology and humanity?

- How to promote access to new kind of jobs created by the development of AI systems?

- As AI displaces some of the work that might be deemed more repetitive or labor-intensive, how are we going to provide jobs to people in a growing and ageing world population?

## 1.3. Environment

The Global Forum considers that any conversation about AI ethics should address the situation humanity is facing right now, and therefore the point on the global environment should be emphasized as a universal value.

- Are AI applications impacting energy consumption?

- Is AI adding to climate change?

- Is it possible for AI applications to promote a carbon-neutral environment?

- How can more AI be deployed for cataloguing and defending biological diversity?

- How can AI help mitigate climate change, reduce biodiversity loss and reduce pollution?

## 1.4. Bioscience, genetics and health

- Are AI systems used to give access to health care to isolated people in the world?

- Can AI be put to use in setting off a public health revolution?

## 1.5. Fake news- media- challenges

- Can AI strengthen and reward the development and diffusion of methods to disclose and neutralize fake news?

- How can professional journalism and quality in news flourish in the face of AI?

- Can AI systems serve to maintain and respect the plurality of values and life choices of individuals?

## 1.6. Governance

- What framework and value system are best used to assess the impact of AI?

- How can we build ethical AI regulations that are scalable and inter-connected globally?

## 1.7. Specific issues for western-world countries

- How do we work with countries that do not share the same values or interests?

- What do countries do to maintain and advance their competitive and innovative capabilities in AI?

- Is open source an issue at stake? Will AI favor proprietary technology or help spur open source technology allowing users to engage in development work and share ideas with others (will new generations of platforms such as CrowdGauge, Crowdmap, Map Server, and Shareabouts be enabled?)

## 2. Questions that will help illuminate potential ethical principles

Once there is a potential application proposed for AI, the process of applying ethical principles or requirements in the design, development and implementation must be considered. In order to develop an ethics policy applicable to all such projects, organizations and individuals could start considering a set of questions that would describe such a policy.

Although the analysis of the global landscape of published AI ethics guidelines shows that there is currently no single principle that is common everywhere, there are themes – or "values" – which are referred to very often, in particular fairness, non-discrimination, transparency, justice, non-maleficence, responsibility, privacy, respect for human autonomy and dignity, explainability, beneficence, trust, accountability, safety and security, human control of technology, and promotion of human values.

The convergence of guidelines on similar principles, but also the differences in how each stakeholder defines these same principles generates inconsistency and confusion in global discussions among and between governments and companies. Companies are tempted to do "AI ethics shopping" where they mix and match ethical principles, guidelines, codes of conduct, etc., for the kind of ethics that is best retrofitted to justify their actual behaviors, rather than revising these behaviors by benchmarking them against shared public, ethical standards. Such attitudes will of course do nothing to eliminate the distrust that exists between enterprises, which in majority are in favor of self-regulation, and governments, which appear to be increasingly reluctant to let business regulate AI by itself and hence, are inclined to consider lawmaking. Yet, in order to move towards a mutually beneficial relationship, both enterprises and governments need to work together to bridge the trust gap.

The Global Forum considers that as policy makers seek to develop a set of ethical principles on which to base AI developments and implementations, the themes mentioned above can suggest a series of questions that may be asked to help formulate ethics policies.

## 2.1. Human Rights

- How should fundamental human rights be protected in AI applications?

- When AI systems are aimed at guiding or supporting humans in decision making processes, making predictions, are they providing recommendations based on evidence-based practices?

- How will AI affect the emotional wellbeing of people?

## 2.2. Human Control

- When is the use of AI considered going too far?

- What are the criteria a governance structure should use for evaluating the potential value and harm of any given AI application?

- How do proponents of an AI application ensure that all stakeholder interests and concerns have been dealt with?

- How to promote engagement and human interaction in AI-rich environments?

## 2.3. Management, robustness, safety and security

### 2.3.1. Management, robustness, safety within the organization

- What management and oversight practices are required to ensure that AI applications produce accurate and evidence-based results?

- Are there ongoing oversight methods to ensure that the AI application does no harm?

- How to measure robustness and the safety of AI? Determine what is a positive or negative outcome?

- How to enforce initially ethical policies for the life cycle of the AI application?

### 2.3.2. Cybercrime and insecurity due to external attacks

- How to protect AI applications from illegitimate uses beyond the original purpose of the application?

- What security provisions must be incorporated within the AI application?

- How will attempts and successful misuse of AI applications be reported and to whom?

- What are the consequences for the misuse of AI applications?

## 2.4. Privacy and data governance

- What must AI applications do to protect the disclosure of personal data for unauthorized purposes?

- How to ensure informed consent on the use of personal data?

- How will the consent of individuals be obtained for the reuse or sharing of their personal data?[1]

- How to prohibit secondary uses of personal data?

## 2.5. Governments

- What are the limitations on government use of personal data in the design and operation of AI applications?

- How to restrict secondary dissemination of personal data for unauthorized purposes?

- How governments are enforcing ethical standards for protecting individual privacy?[2]

- How to prevent algorithmic bias?

- How could AI applications prohibit discrimination based on race, ethnic origin, gender, age or other class characteristics?

## 2.6. Companies and organizations

- Are personal data only used for the purpose for which it was collected unless consent is obtained?

- Is the secondary dissemination of personal data for unauthorized purposes prohibited?

- Does the system prohibit the secondary use of data that becomes the basis for the development of AI algorithms?

- What mechanisms can be enacted to collect less data (enforce data minimization) limiting data collection to its primary purpose?

- Is anonymization accomplished in the development of reporting structures using the AI application?

---

[1] The European GDPR (general data protection regulation) protects privacy as data sharing is based mostly on consent. Similar initiatives are followed all over the world.

[2] European laws (European Convention on Human Rights (ECHR)and Charter of Fundamental Rights of the European Union and GDPR) and European case law give some answers to solve these tensions. Any infringement to fundamental right must be in favor of general interest, necessary, proportionate, and organized by law.

- Under what circumstances should biometrics be integral to identity management?

- Is the use of synthetic data a legitimate practice for designing algorithms?

- Does data ownership influence ethical principles in AI implementations?

## 2.7. Transparency

- Are the data and processes that yield the AI system's decisions traceable?

- Are there understandable explanations for AI systems technical processes?

- Are there understandable explanations for the decisions or predictions that the AI system makes?

- Is there an open communication about the capabilities and the limitations of the AI system?

## 2.8. Social, societal and environmental well-being

### Diversity, Non-Discrimination and Fairness

- How will bias in the design or implementation of an AI application be prohibited?

- How AI systems ensure inclusivity in operations?

- How avoid AI applications to create discrimination based on race, ethnic origin, gender, age or other class characteristics?

*The present document has been elaborated by Föreningen Global Forum/Shaping the Future (the Global Forum/Shaping the Future Association). It emerged from the expertise of the members of the Föreningen Global Forum/Shaping the Future Advisory Board who worked within a think tank from June 2021 to June 2022.*