

## REFERENCE DOCUMENTS

### OECD AI Principles

#### **Recommendation of the Council on Artificial Intelligence**

**Adopted May 21, 2019**

THE COUNCIL,

HAVING REGARD to Article 5 b) of the Convention on the Organisation for Economic Co-operation and Development of 14 December 1960;

HAVING REGARD to the OECD Guidelines for Multinational Enterprises [[OECD/LEGAL/0144](#)]; Recommendation of the Council concerning Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data [[OECD/LEGAL/0188](#)]; Recommendation of the Council concerning Guidelines for Cryptography Policy [[OECD/LEGAL/0289](#)]; Recommendation of the Council for Enhanced Access and More Effective Use of Public Sector Information [[OECD/LEGAL/0362](#)]; Recommendation of the Council on Digital Security Risk Management for Economic and Social Prosperity [[OECD/LEGAL/0415](#)]; Recommendation of the Council on Consumer Protection in E-commerce [[OECD/LEGAL/0422](#)]; Declaration on the Digital Economy: Innovation, Growth and Social Prosperity (Cancún Declaration) [[OECD/LEGAL/0426](#)]; Declaration on Strengthening SMEs and Entrepreneurship for Productivity and Inclusive Growth [[OECD/LEGAL/0439](#)]; as well as the 2016 Ministerial Statement on Building more Resilient and Inclusive Labour Markets, adopted at the OECD Labour and Employment Ministerial Meeting;

HAVING REGARD to the Sustainable Development Goals set out in the 2030 Agenda for Sustainable Development adopted by the United Nations General Assembly (A/RES/70/1) as well as the 1948 Universal Declaration of Human Rights;

HAVING REGARD to the important work being carried out on artificial intelligence (hereafter, “AI”) in other international governmental and non-governmental fora;

## *Artificial Intelligence and Democratic Values*

RECOGNISING that AI has pervasive, far-reaching and global implications that are transforming societies, economic sectors and the world of work, and are likely to increasingly do so in the future;

RECOGNISING that AI has the potential to improve the welfare and well-being of people, to contribute to positive sustainable global economic activity, to increase innovation and productivity, and to help respond to key global challenges;

RECOGNISING that, at the same time, these transformations may have disparate effects within, and between societies and economies, notably regarding economic shifts, competition, transitions in the labour market, inequalities, and implications for democracy and human rights, privacy and data protection, and digital security;

RECOGNISING that trust is a key enabler of digital transformation; that, although the nature of future AI applications and their implications may be hard to foresee, the trustworthiness of AI systems is a key factor for the diffusion and adoption of AI; and that a well-informed whole-of-society public debate is necessary for capturing the beneficial potential of the technology, while limiting the risks associated with it;

UNDERLINING that certain existing national and international legal, regulatory and policy frameworks already have relevance to AI, including those related to human rights, consumer and personal data protection, intellectual property rights, responsible business conduct, and competition, while noting that the appropriateness of some frameworks may need to be assessed and new approaches developed;

RECOGNISING that given the rapid development and implementation of AI, there is a need for a stable policy environment that promotes a human-centric approach to trustworthy AI, that fosters research, preserves economic incentives to innovate, and that applies to all stakeholders according to their role and the context;

CONSIDERING that embracing the opportunities offered, and addressing the challenges raised, by AI applications, and empowering stakeholders to engage is essential to fostering adoption of trustworthy AI in society, and to turning AI trustworthiness into a competitive parameter in the global marketplace;

On the proposal of the Committee on Digital Economy Policy:

I. AGREES that for the purpose of this Recommendation the following terms should be understood as follows:

–*AI system*: An AI system is a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments. AI systems are designed to operate with varying levels of autonomy.

–*AI system lifecycle*: AI system lifecycle phases involve: *i)* ‘design, data and models’; which is a context-dependent sequence encompassing planning and design, data collection and processing, as well as model building; *ii)* ‘verification and validation’; *iii)* ‘deployment’; and *iv)* ‘operation and monitoring’. These phases often take place in an iterative manner and are not necessarily sequential. The decision to retire an AI system from operation may occur at any point during the operation and monitoring phase.

–*AI knowledge*: AI knowledge refers to the skills and resources, such as data, code, algorithms, models, research, know-how, training programmes, governance, processes and best practices, required to understand and participate in the AI system lifecycle.

–*AI actors*: AI actors are those who play an active role in the AI system lifecycle, including organisations and individuals that deploy or operate AI.

–*Stakeholders*: Stakeholders encompass all organisations and individuals involved in, or affected by, AI systems, directly or indirectly. AI actors are a subset of stakeholders.

### ***Section 1:***

#### ***Principles for responsible stewardship of trustworthy AI***

II. RECOMMENDS that Members and non-Members adhering to this Recommendation (hereafter the “Adherents”) promote and implement the following principles for responsible stewardship of trustworthy AI, which are relevant to all stakeholders.

III. CALLS ON all AI actors to promote and implement, according to their respective roles, the following Principles for responsible stewardship of trustworthy AI.

IV. UNDERLINES that the following principles are complementary and should be considered as a whole.

**1.1. Inclusive growth, sustainable development and well-being**

Stakeholders should proactively engage in responsible stewardship of trustworthy AI in pursuit of beneficial outcomes for people and the planet, such as augmenting human capabilities and enhancing creativity, advancing inclusion of underrepresented populations, reducing economic, social, gender and other inequalities, and protecting natural environments, thus invigorating inclusive growth, sustainable development and well-being.

**1.2. Human-centred values and fairness**

- a) AI actors should respect the rule of law, human rights and democratic values, throughout the AI system lifecycle. These include freedom, dignity and autonomy, privacy and data protection, non-discrimination and equality, diversity, fairness, social justice, and internationally recognised labour rights.
- b) To this end, AI actors should implement mechanisms and safeguards, such as capacity for human determination, that are appropriate to the context and consistent with the state of art.

**1.3. Transparency and explainability**

AI Actors should commit to transparency and responsible disclosure regarding AI systems. To this end, they should provide meaningful information, appropriate to the context, and consistent with the state of art:

- i. to foster a general understanding of AI systems,
- ii. to make stakeholders aware of their interactions with AI systems, including in the workplace,
- iii. to enable those affected by an AI system to understand the outcome, and,
- iv. to enable those adversely affected by an AI system to challenge its outcome based on plain and easy-to-understand information on

the factors, and the logic that served as the basis for the prediction, recommendation or decision.

#### **1.4. Robustness, security and safety**

a) AI systems should be robust, secure and safe throughout their entire lifecycle so that, in conditions of normal use, foreseeable use or misuse, or other adverse conditions, they function appropriately and do not pose unreasonable safety risk.

b) To this end, AI actors should ensure traceability, including in relation to datasets, processes and decisions made during the AI system lifecycle, to enable analysis of the AI system's outcomes and responses to inquiry, appropriate to the context and consistent with the state of art.

c) AI actors should, based on their roles, the context, and their ability to act, apply a systematic risk management approach to each phase of the AI system lifecycle on a continuous basis to address risks related to AI systems, including privacy, digital security, safety and bias.

#### **1.5. Accountability**

AI actors should be accountable for the proper functioning of AI systems and for the respect of the above principles, based on their roles, the context, and consistent with the state of art.

#### ***Section 2:***

#### ***National policies and international co-operation for trustworthy AI***

V.RECOMMENDS that Adherents implement the following recommendations, consistent with the principles in section 1, in their national policies and international co-operation, with special attention to small and medium-sized enterprises (SMEs).

#### **2.1. Investing in AI research and development**

a) Governments should consider long-term public investment, and encourage private investment, in research and development, including interdisciplinary efforts, to spur innovation in trustworthy

## *Artificial Intelligence and Democratic Values*

AI that focus on challenging technical issues and on AI-related social, legal and ethical implications and policy issues.

b) Governments should also consider public investment and encourage private investment in open datasets that are representative and respect privacy and data protection to support an environment for AI research and development that is free of inappropriate bias and to improve interoperability and use of standards.

### **2.2. Fostering a digital ecosystem for AI**

Governments should foster the development of, and access to, a digital ecosystem for trustworthy AI. Such an ecosystem includes in particular digital technologies and infrastructure, and mechanisms for sharing AI knowledge, as appropriate. In this regard, governments should consider promoting mechanisms, such as data trusts, to support the safe, fair, legal and ethical sharing of data.

### **2.3. Shaping an enabling policy environment for AI**

a) Governments should promote a policy environment that supports an agile transition from the research and development stage to the deployment and operation stage for trustworthy AI systems. To this effect, they should consider using experimentation to provide a controlled environment in which AI systems can be tested, and scaled-up, as appropriate.

b) Governments should review and adapt, as appropriate, their policy and regulatory frameworks and assessment mechanisms as they apply to AI systems to encourage innovation and competition for trustworthy AI.

### **2.4. Building human capacity and preparing for labour market transformation**

a) Governments should work closely with stakeholders to prepare for the transformation of the world of work and of society. They should empower people to effectively use and interact with AI systems across the breadth of applications, including by equipping them with the necessary skills.

- b) Governments should take steps, including through social dialogue, to ensure a fair transition for workers as AI is deployed, such as through training programmes along the working life, support for those affected by displacement, and access to new opportunities in the labour market.
- c) Governments should also work closely with stakeholders to promote the responsible use of AI at work, to enhance the safety of workers and the quality of jobs, to foster entrepreneurship and productivity, and aim to ensure that the benefits from AI are broadly and fairly shared.

### **2.5. International co-operation for trustworthy AI**

- a) Governments, including developing countries and with stakeholders, should actively co-operate to advance these principles and to progress on responsible stewardship of trustworthy AI.
- b) Governments should work together in the OECD and other global and regional fora to foster the sharing of AI knowledge, as appropriate. They should encourage international, cross-sectoral and open multi-stakeholder initiatives to garner long-term expertise on AI.
- c) Governments should promote the development of multi-stakeholder, consensus-driven global technical standards for interoperable and trustworthy AI.
- d) Governments should also encourage the development, and their own use, of internationally comparable metrics to measure AI research, development and deployment, and gather the evidence base to assess progress in the implementation of these principles.

VI. INVITES the Secretary-General and Adherents to disseminate this Recommendation.

VII. INVITES non-Adherents to take due account of, and adhere to, this Recommendation.

VIII. INSTRUCES the Committee on Digital Economy Policy:

- a) to continue its important work on artificial intelligence building on this Recommendation and taking into account work in other

*Artificial Intelligence and Democratic Values*

international fora, and to further develop the measurement framework for evidence-based AI policies;

b) to develop and iterate further practical guidance on the implementation of this Recommendation, and to report to the Council on progress made no later than end December 2019;

c) to provide a forum for exchanging information on AI policy and activities including experience with the implementation of this Recommendation, and to foster multi-stakeholder and interdisciplinary dialogue to promote trust in and adoption of AI; and

d) to monitor, in consultation with other relevant Committees, the implementation of this Recommendation and report thereon to the Council no later than five years following its adoption and regularly thereafter.

*The AI Social Contract Index 2020*

*OECD AI POLICY Adherents*

The following countries have endorsed the OECD AI Principles

*OECD Member Countries*

Australia  
Austria  
Belgium  
Canada  
Chile  
*Colombia*  
Czech Republic  
Denmark  
Estonia  
Finland  
France  
Germany  
Greece  
Hungary  
Iceland  
Ireland  
Israel  
Italy  
Japan  
Korea  
Latvia  
Lithuania  
Luxembourg  
Mexico  
Netherlands  
New Zealand  
Norway  
Poland  
Portugal  
Slovak Republic  
Slovenia  
Spain  
Sweden  
Switzerland  
Turkey  
United Kingdom

*Artificial Intelligence and Democratic Values*

United States

*OECD Non-Member Countries*

Argentina  
Brazil  
Costa Rica  
Malta  
Peru  
Romania  
Ukraine

*G-20 Countries*

China  
India  
Indonesia  
Russia  
Saudi Arabia  
South Africa

*As of December 1, 2020, 51 countries have endorsed the OECD/G20 AI Principles.*

Universal Guidelines for AI

**Universal Guidelines for Artificial Intelligence**

**23 October 2018  
Brussels, Belgium**

New developments in Artificial Intelligence are transforming the world, from science and industry to government administration and finance. The rise of AI decision-making also implicates fundamental rights of fairness, accountability, and transparency. Modern data analysis produces significant outcomes that have real life consequences for people in employment, housing, credit, commerce, and criminal sentencing. Many of these techniques are entirely opaque, leaving individuals unaware whether the decisions were accurate, fair, or even about them.

We propose these Universal Guidelines to inform and improve the design and use of AI. The Guidelines are intended to maximize the benefits of AI, to minimize the risk, and to ensure the protection of human rights. These Guidelines should be incorporated into ethical standards, adopted in national law and international agreements, and built into the design of systems. We state clearly that the primary responsibility for AI systems must reside with those institutions that fund, develop, and deploy these systems.

1. **Right to Transparency.** All individuals have the right to know the basis of an AI decision that concerns them. This includes access to the factors, the logic, and techniques that produced the outcome.
2. **Right to Human Determination.** All individuals have the right to a final determination made by a person.
3. **Identification Obligation.** The institution responsible for an AI system must be made known to the public.
4. **Fairness Obligation.** Institutions must ensure that AI systems do not reflect unfair bias or make impermissible discriminatory decisions.
5. **Assessment and Accountability Obligation.** An AI system should be deployed only after an adequate evaluation of its purpose

## *Artificial Intelligence and Democratic Values*

and objectives, its benefits, as well as its risks. Institutions must be responsible for decisions made by an AI system.

6. **Accuracy, Reliability, and Validity Obligations.** Institutions must ensure the accuracy, reliability, and validity of decisions.
7. **Data Quality Obligation.** Institutions must establish data provenance, and assure quality and relevance for the data input into algorithms.
8. **Public Safety Obligation.** Institutions must assess the public safety risks that arise from the deployment of AI systems that direct or control physical devices, and implement safety controls.
9. **Cybersecurity Obligation.** Institutions must secure AI systems against cybersecurity threats.
10. **Prohibition on Secret Profiling.** No institution shall establish or maintain a secret profiling system.
11. **Prohibition on Unitary Scoring.** No national government shall establish or maintain a general-purpose score on its citizens or residents.
12. **Termination Obligation.** An institution that has established an AI system has an affirmative obligation to terminate the system if human control of the system is no longer possible.

**Explanatory Memorandum and References  
October 2018**

**Context**

The Universal Guidelines on Artificial Intelligence (UGAI) call attention to the growing challenges of intelligent computational systems and proposes concrete recommendations that can improve and inform their design. At its core, the purpose of the UGAI is to promote transparency and accountability for these systems and to ensure that people retain control over the systems they create. Not all systems fall within the scope of these Guidelines. Our concern is with those systems that impact the rights of people. Above all else, these systems should do no harm.

The declaration is timely. Governments around the world are developing policy proposals and institutions, both public and private, are supporting research and development of “AI.” Invariably, there will be an enormous impact on the public, regardless of their participation in the design and development of these systems. And so, the UGAI reflects a public perspective on these challenges.

The UGAI were announced at the 2018 International Data Protection and Privacy Commissioners Conference, among the most significant meetings of technology leaders and data protection experts in history.

The UGAI builds on prior work by scientific societies, think tanks, NGOs, and international organizations. The UGAI incorporates elements of human rights doctrine, data protection law, and ethical guidelines. The Guidelines include several well-established principles for AI governance, and put forward new principles not previously found in similar policy frameworks.

**Terminology**

The term “Artificial Intelligence” is both broad and imprecise. It includes aspects of machine learning, rule-based decision-making, and other computational techniques. There are also disputes regarding whether Artificial Intelligence is possible. The UGAI simply acknowledges that this term, in common use, covers a wide range of related issues and adopts the term to engage the current debate. There is no attempt here to define its boundaries, other than to assume that AI requires some degree of

## *Artificial Intelligence and Democratic Values*

automated decision-making. The term “Guidelines” follows the practice of policy frameworks that speak primarily to governments and private companies.

The UGAI speaks to the obligations of “institutions” and the rights of “individuals.” This follows from the articulation of fair information practices in the data protection field. The UGAI takes the protection of the individual as a fundamental goal. Institutions, public and private, are understood to be those entities that develop and deploy AI systems. The term “institution” was chosen rather than the more familiar “organization” to underscore the permanent, ongoing nature of the obligations set out in the Guidelines. There is one principle that is addressed to “national governments.” The reason for this is discussed below.

### **Application**

These Guidelines should be incorporated into ethical standards, adopted in national law and international agreements, and built into the design of systems.

### **The Principles**

The elements of the **Transparency Principle** can be found in several modern privacy laws, including the US Privacy Act, the EU Data Protection Directive, the GDPR, and the Council of Europe Convention 108. The aim of this principle is to enable independent accountability for automated decisions, with a primary emphasis on the right of the individual to know the basis of an adverse determination. In practical terms, it may not be possible for an individual to interpret the basis of a particular decision, but this does not obviate the need to ensure that such an explanation is possible.

The **Right to a Human Determination** reaffirms that individuals and not machines are responsible for automated decision-making. In many instances, such as the operation of an autonomous vehicle, it would not be possible or practical to insert a human decision prior to an automated decision. But the aim remains to ensure accountability. Thus where an automated system fails, this principle should be understood as a requirement that a human assessment of the outcome be made.

**Identification Obligation.** This principle seeks to address the identification asymmetry that arises in the interaction between individuals

and AI systems. An AI system typically knows a great deal about an individual; the individual may not even know the operator of the AI system. The **Identification Obligation** establishes the foundation of AI accountability which is to make clear the identity of an AI system and the institution responsible.

The **Fairness Obligation** recognizes that all automated systems make decisions that reflect bias and discrimination, but such decisions should not be normatively unfair. There is no simple answer to the question as to what is unfair or impermissible. The evaluation often depends on context. But the Fairness Obligation makes clear that an assessment of objective outcomes alone is not sufficient to evaluate an AI system. Normative consequences must be assessed, including those that preexist or may be amplified by an AI system.

The **Assessment and Accountability Obligation** speaks to the obligation to assess an AI system prior to and during deployment. Regarding assessment, it should be understood that a central purpose of this obligation is to determine whether an AI system should be established. If an assessment reveals substantial risks, such as those suggested by principles concerning Public Safety and Cybersecurity, then the project should not move forward.

The **Accuracy, Reliability, and Validity Obligations** set out key responsibilities associated with the outcome of automated decisions. The terms are intended to be interpreted both independently and jointly.

The **Data Quality Principle** follows from the preceding obligation.

The **Public Safety Obligation** recognizes that AI systems control devices in the physical world. For this reason, institutions must both assess risks and take precautionary measures as appropriate.

The **Cybersecurity Obligation** follows from the Public Safety Obligation and underscores the risk that even well-designed systems may be the target of hostile actors. Those who develop and deploy AI systems must take these risks into account.

The **Prohibition on Secret Profiling** follows from the earlier Identification Obligation. The aim is to avoid the information asymmetry

## *Artificial Intelligence and Democratic Values*

that arises increasingly with AI systems and to ensure the possibility of independent accountability.

The **Prohibition on Unitary Scoring** speaks directly to the risk of a single, multi-purpose number assigned by a government to an individual. In data protection law, universal identifiers that enable the profiling of individuals across are disfavored. These identifiers are often regulated and in some instances prohibited. The concern with universal scoring, described here as “unitary scoring,” is even greater. A unitary score reflects not only a unitary profile but also a predetermined outcome across multiple domains of human activity. There is some risk that unitary scores will also emerge in the private sector. Conceivably, such systems could be subject to market competition and government regulations. But there is not even the possibility of counterbalance with unitary scores assigned by government, and therefore they should be prohibited.

The **Termination Obligation** is the ultimate statement of accountability for an AI system. The obligation presumes that systems must remain within human control. If that is no longer possible, the system should be terminated.

## **REFERENCES**

*Asilomar AI Principles* (2017)

*Aspen Institute Roundtable on Artificial Intelligence* (2016)

Association for Computing Machinery, U.S. Public Policy Counsel, *Statement on Algorithmic Transparency and Accountability* (Jan. 2017)

Council of Europe, *Convention 108* (1981)

*Council of Europe and Artificial Intelligence* (2018)

Data and Society, *Governing Artificial Intelligence* (2018)

European Commission, *High Level Expert Group on Artificial Intelligence* (2018)

*EU General Data Protection Regulation* (2018)

IEEE, *Ethically Aligned Design* (2016)

*The AI Social Contract Index 2020*

Japan, Ministry of Internal Affairs and Communications, *AI R&D Guidelines* (2016)

Garry Kasparov, *Deep Thinking: Where Machine Intelligence Ends and Human Creativity Begins* (2017)

*Madrid Privacy Declaration* (2009)

OECD, *Artificial Intelligence* (2018)

OECD, *Privacy Guidelines* (1980)

Cathy O'Neil, *Weapons of Math Destruction* (2016)

Frank Pasquale, *The Black Box Society: The Secret Algorithms That Control Money and Information* (2015)

Privacy International, *Artificial Intelligence* (2018)

*US Privacy Act* (1974)

*Toronto Declaration* (2018)

Joseph Weizenbaum, *Computer Power and Human Reason* (1976)

*Universal Declaration of Human Rights* (1948)

Social Contract for the Age of AI

***A New Social Contract for the Age of Artificial Intelligence***

*Original Signatories*

Governor Michael Dukakis, Boston Global Forum and  
President Vaira Vīķe-Freiberga, Latvia and World Leadership Alliance-  
Club de Madrid

*Additional Signatories*

Vint Cerf, Father of Internet, Google, Nazli Choucri, MIT,  
Prime Minister Zlatko Lagumdžija, Bosnia and Herzegovina,  
Tuan Anh Nguyen, Boston Global Forum, Thomas Patterson, Harvard  
University, Alex Pentland, MIT, Marc Rotenberg, CAIDP, David  
Silbersweig, Harvard University

The term “artificial intelligence” refers to the development of computer systems able to perform tasks that normally require human intelligence, such as visual perception, speech recognition, decision-making, language translation, and self-driving cars. Advances in AI have already altered conventional ways of seeing the world around us. This is creating new realities for everyone – as well as new possibilities.

These advances in AI are powerful in many ways. They have created a new global ecology and yet remain opaque and need to be better understood. Advances in AI raise policy issues that must be assessed. We must now focus through dialogue, tolerance, learning and understanding on key principles and practices for an agreement among members of society for shared social benefit that we call the Social Contract for the AI Age.

The expansion of Artificial Intelligence is widely recognized and could change our lives in ways yet unimagined. At the same time, without guidelines or directives, the undisciplined use of AI poses risks to the wellbeing of individuals and creates possibilities for economic, political, social, and criminal exploitation.

The international community recognizes the challenges and opportunities, as well as the problems and perils, of AI. Many countries have announced national strategies to promote the proper use and development of AI. These strategies set out common goals such as:

## *The AI Social Contract Index 2020*

- Scientific research, funding and culture,
- Sustainable development, and inclusive growth
- Skills, education, and talent development
- Public and private sector adoption,
- Fairness, transparency, and accountability
- Ethics, values and inclusion,
- Reliability, security and privacy,
- Science-policy links,
- Standards, human control and regulations
- Data and digital infrastructure.

AI is also the focus of foreign policy and international cooperation. There is a shared view that no country will be able to compete or meet the needs of its citizens without increasing its AI capacity. As well, many countries are now engaged in technology leapfrogging. It is no longer expected, nor necessary, to replicate the stages of economic development of the West—one phase at a time.

In a world as diverse as the one today, there are few mechanisms for responding to such possibilities on a global scale. Social Contract for the AI Age is designed to establish a common understanding for policy and practices, anchored in general principles to help maximise the “good” and minimise the “bad” associated with AI. Derived from the 18th century concept of a social contract—an agreement among the members of society to cooperate for social benefits—Social Contract for the AI Age is focuses on the conditions of the 21st century. It is a response to artificial intelligence, big data, the Internet of Things, and high-speed computation.

### *Foundations*

Just as earlier social contracts helped shape societies for a common purpose, the Social Contract for the AI Age has a transformative vision, one that transcends the technological features of artificial intelligence and seeks to provide foundations for a new society. Consider, for example, how the Covid-19 pandemic urgently requires a new society with new structure and order, approach — new ways to share data and coordinate action, accelerated social reliance on digital service across businesses, education, and government services. The Social Contract for the AI Age would create standards for a new international system. It focuses on the conduct of each nation, relations with international business and not for profit entities, and the cooperation of nations. Just as TCP / IP is the platform for

## *Artificial Intelligence and Democratic Values*

communication among internet users, the Social Contract for AI Age is a platform for connection among governments, stakeholders, and private and public institutions.

### *Objectives*

The Social Contract for the AI Age seeks to build a multi-stakeholder, inclusive society in all aspects of life across politics, government, economics, business, and industry. The Social Contract for the AI Age values creation, innovation, philanthropy, and mutual respect. It seeks the right of freedom on, and access to, the Internet worldwide. The Social Contract for the AI Age seeks to make the world a locus of responsible interaction—a place where every person’s contribution is recognized and everyone has a right to knowledge and access to information, where no one is above the law, where money cannot be used to subvert political process, and where integrity, knowledge, creativity, honesty, and tolerance shape decisions and guide policy.

In short, the Social Contract for the AI Age seeks to build a world where all are recognized and valued, and all forms of governance adhere to a set of values and are accountable and transparent. It is a world where global challenges are met by collective action and responsibility.

### *Principles*

Extensive and appropriate AI application to politics, governments, society, and businesses can create a Smart Democracy. The Social Contract for the AI Age creates a platform for a Smart Democracy society, and a new global supply chain, named Supply Chain 2020. As a framework for society in the AI age, the Social Contract 2020 is based on balancing power among governments, businesses, civil society, individuals, and AI assistants. The Social Contract for the AI Age is a commitment to protect property, common values, and collective norms.

- AI must respect fundamental human rights such as human dignity, rule of law, and privacy protection.
- AI systems must be considered from a multi-stakeholder perspective — for the individual and for society as a whole
- The Social Contract for the AI Age is a basis to achieve sustainable and inclusive development for a global community that is fair, equitable, and prosperous. It is designed to apply

## *The AI Social Contract Index 2020*

the concept of a people-centered economy and to create a trustworthy AI, data, and Internet ecosystem for work and life.

- The Social Contract for the AI Age should be transparent and accountable, and follow standards based on policies driven by trustworthy data. The UN Sustainable Development Goals data metrics and the World Economic Forum Environmental, Social, and Governance (ESG) metrics, should provide citizens and organizations with reliable data that enables well-informed policy choices.
- Communities must have control over their data. Data is the basis of self-determination and provides the ability to measure the impact of actions and policy in the AI realm.
- Data literacy at all levels of society, together with open, trustworthy information, is the basis for an intelligent, thoughtful society.
- Commitments of Stakeholders/ Power Centers
- Individuals, Citizens, Groups:
- Everyone is entitled to basic rights and dignity that are enhanced/promoted by AI

### *Data Rights and Responsibilities:*

- Each individual has the right to privacy and is entitled to access and control over their own data. Individuals have a right to manage their data, individually or collectively, and the right to withhold their data from businesses.
- Each individual and each community must have access to a trustworthy AI, data, and Internet ecosystem to create an inclusive, fair, people-centered economy, and society.

### *Internet Rights*

- Each individual has the right to access the Internet and any website or news system without restriction.
- Freedom of expression on the Internet is guaranteed.
- Secure digital identity allows the individual to know about, and control access to their data.

### *Education and Political Participation:*

## *Artificial Intelligence and Democratic Values*

- Each individual has the right to education through best available venues.
- Each individual must have access to education/knowledge/training about the use and impact of AI.
- Each individual has the right to unimpeded political participation.
- Everyone must have access to due process and fair trial, as well as remediation for injustice.

### *Responsibility*

- Individuals is prohibited from exercising socially disruptive behaviors, such as hacking, disseminating disinformation or online hate.
- Individuals must contribute to the common good through appropriate taxes and provision of critical personal information (with appropriate data protection) as, for example, in the collection of census data and voting for public officials

### **Governments.** *Governments have a leading role in the Social Contract for the AI Age.*

- All government should behave responsibly in the management of AI for governance and for interactions with individuals and such behavior must be easily auditable.
- All governments implement AI governance policies that respect honesty, transparency, fairness, and accountability. These standards and norms apply in every area of governance and are the basis for collaboration with international communities.
- All governments create incentives for citizens to use AI in ways that benefit society (for example each person does good work for society that will be recognized as value, and this value can be exchanged with other values such as financial value, products, services, etc.).
- All governments construct a secure, stable, and trustworthy AI, data, and Internet ecosystem for work and life to support the people-centered economy.
- All governments ensure that communities are able take control of their data and use AI with their data so that they can manage their community to suit their needs and to create prosperity for themselves and their family.

## *The AI Social Contract Index 2020*

- All governments establish norms, rules and pass laws to ensure AI benefits society.
- All governments create secure safety net for citizens in a trustworthy AI, data, and Internet ecosystem for work and life.
- All governments protect intellectual property rights without undermining free access to the information commons.

### *Collaborations between governments*

- Norms, values and standards of the Social Contract for the AI Age are designed as connections among governments. To maintain the Social Contract for the AI Age, it is essential for countries to establish a Democratic Alliance for Digital Governance. All governments should work to promote the Democratic Alliance for Digital Governance.

### *United Nations and International Organizations:*

- The United Nations should extend international human rights standards for AI, and create a UN Convention on AI and establish a specialized UN Agency on AI.

### ***Business Entities.*** *Business operations and related rights come with accountability and responsibility – nationally and internationally. Business must:*

- Enable independent audits for transparency, fairness, accountability, and cybersecurity.
- Adopt common AI values, standards, norms, and data ownership rules with penalties for noncompliance.
- Companies will be incentivized to do business only with companies and countries that uphold the Social Contract for the AI Age and endorse Supply Chain 2020.
- Collaborate with governments and civil society to help create a people-centered AI, data, and Internet ecosystem, to create trustworthy and relevant data, and to use AI to share wealth with individuals and communities.

### ***Civil Society Organizations & Community.*** *Rights and responsibilities of civil society organizations include:*

- Monitor governments and firms with respect to common values.

## *Artificial Intelligence and Democratic Values*

- Promote common values, norms, standards, and laws.
- Support AI users and assist them to serve the broad interests of society.
- Collaborate with governments, business entities, and others to create a people-centered AI, data, and Internet ecosystem for work and life, enabling all people can create meaningful data, value, and create value-based wealth for their community.
- Enable data cooperatives—the voluntary collaborative pooling by individuals of their personal data for the benefit of the group or community.
- Participate in the making of AI rules and norms.
- Enhance civil society become an intelligent, thoughtful civil society based on knowledge, smart data, critical thinking and social responsibility; and through the achievement of data literacy, to become a trusted open-data system, with validated, predictive AI tools that communities to plan their future.

***AI Assistants.*** *AI assistants provide an interface to facilitate compliance with established standards.*

- Support AI users and assist them to serve the broad interests of society.
- Engage with other power centers for mutual support and supervision.
- Community control: AI assistants should be governed by communities of users.

*The Social Contract for the AI Age will be implemented as follows:*

- The promulgation of a Code of Ethics for AI Developers and AI Users.
- The creation of a system to monitor and evaluate governments, companies, and individuals (based on their contribution to maintaining norms, standards, common values, and international laws for honesty, transparency, accountability, and responsibility).
- The recognition of the Social Contract for the AI Age by the United Nations, governments, companies, civil society and the international AI community.
- The establishment of a United Nations Convention on Artificial Intelligence to obligate governments and others to

*The AI Social Contract Index 2020*

comply with international rules and norms to protect rights in AI age.

- The development of the Democratic Alliance for Digital Governance as the global authority to enforce the Social Contract for the AI Age.
- The creation of the “AIWS City”—an all-digital virtual city based on the standards and norms of “the Social Contract for the AI Age”, “People Centered Economy”, “Trustworthy Economy”, “AI-Government”, and “Intellectual Society-Thoughtful Civil Society”.

## *Artificial Intelligence and Democratic Values*

### Universal Declaration of Human Rights

#### **Preamble**

Whereas recognition of the inherent dignity and of the equal and inalienable rights of all members of the human family is the foundation of freedom, justice and peace in the world,

Whereas disregard and contempt for human rights have resulted in barbarous acts which have outraged the conscience of mankind, and the advent of a world in which human beings shall enjoy freedom of speech and belief and freedom from fear and want has been proclaimed as the highest aspiration of the common people,

Whereas it is essential, if man is not to be compelled to have recourse, as a last resort, to rebellion against tyranny and oppression, that human rights should be protected by the rule of law,

Whereas it is essential to promote the development of friendly relations between nations,

Whereas the peoples of the United Nations have in the Charter reaffirmed their faith in fundamental human rights, in the dignity and worth of the human person and in the equal rights of men and women and have determined to promote social progress and better standards of life in larger freedom,

Whereas Member States have pledged themselves to achieve, in co-operation with the United Nations, the promotion of universal respect for and observance of human rights and fundamental freedoms,

Whereas a common understanding of these rights and freedoms is of the greatest importance for the full realization of this pledge,

Now, Therefore THE GENERAL ASSEMBLY proclaims THIS UNIVERSAL DECLARATION OF HUMAN RIGHTS as a common standard of achievement for all peoples and all nations, to the end that every individual and every organ of society, keeping this Declaration constantly in mind, shall strive by teaching and education to promote respect for these rights and freedoms and by progressive measures, national and international, to secure their universal and effective recognition and

## *The AI Social Contract Index 2020*

observance, both among the peoples of Member States themselves and among the peoples of territories under their jurisdiction.

### **Article 1**

All human beings are born free and equal in dignity and rights. They are endowed with reason and conscience and should act towards one another in a spirit of brotherhood.

### **Article 2**

Everyone is entitled to all the rights and freedoms set forth in this Declaration, without distinction of any kind, such as race, colour, sex, language, religion, political or other opinion, national or social origin, property, birth or other status. Furthermore, no distinction shall be made on the basis of the political, jurisdictional or international status of the country or territory to which a person belongs, whether it be independent, trust, non-self-governing or under any other limitation of sovereignty.

### **Article 3**

Everyone has the right to life, liberty and security of person.

### **Article 4**

No one shall be held in slavery or servitude; slavery and the slave trade shall be prohibited in all their forms.

### **Article 5**

No one shall be subjected to torture or to cruel, inhuman or degrading treatment or punishment.

### **Article 6**

Everyone has the right to recognition everywhere as a person before the law.

### **Article 7**

All are equal before the law and are entitled without any discrimination to equal protection of the law. All are entitled to equal

## *Artificial Intelligence and Democratic Values*

protection against any discrimination in violation of this Declaration and against any incitement to such discrimination.

### **Article 8**

Everyone has the right to an effective remedy by the competent national tribunals for acts violating the fundamental rights granted him by the constitution or by law.

### **Article 9**

No one shall be subjected to arbitrary arrest, detention or exile.

### **Article 10**

Everyone is entitled in full equality to a fair and public hearing by an independent and impartial tribunal, in the determination of his rights and obligations and of any criminal charge against him.

### **Article 11**

(1) Everyone charged with a penal offence has the right to be presumed innocent until proved guilty according to law in a public trial at which he has had all the guarantees necessary for his defence.  
(2) No one shall be held guilty of any penal offence on account of any act or omission which did not constitute a penal offence, under national or international law, at the time when it was committed. Nor shall a heavier penalty be imposed than the one that was applicable at the time the penal offence was committed.

### **Article 12**

No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honour and reputation. Everyone has the right to the protection of the law against such interference or attacks.

### **Article 13**

(1) Everyone has the right to freedom of movement and residence within the borders of each state.

- (2) Everyone has the right to leave any country, including his own, and to return to his country.

#### **Article 14**

- (1) Everyone has the right to seek and to enjoy in other countries asylum from persecution.
- (2) This right may not be invoked in the case of prosecutions genuinely arising from non-political crimes or from acts contrary to the purposes and principles of the United Nations.

#### **Article 15**

- (1) Everyone has the right to a nationality.
- (2) No one shall be arbitrarily deprived of his nationality nor denied the right to change his nationality.

#### **Article 16**

- (1) Men and women of full age, without any limitation due to race, nationality or religion, have the right to marry and to found a family. They are entitled to equal rights as to marriage, during marriage and at its dissolution.
- (2) Marriage shall be entered into only with the free and full consent of the intending spouses.
- (3) The family is the natural and fundamental group unit of society and is entitled to protection by society and the State.

#### **Article 17**

- (1) Everyone has the right to own property alone as well as in association with others.
- (2) No one shall be arbitrarily deprived of his property.

#### **Article 18**

Everyone has the right to freedom of thought, conscience and religion; this right includes freedom to change his religion or belief, and freedom, either alone or in community with others and in public or private, to manifest his religion or belief in teaching, practice, worship and observance.

## *Artificial Intelligence and Democratic Values*

### **Article 19**

Everyone has the right to freedom of opinion and expression; this right includes freedom to hold opinions without interference and to seek, receive and impart information and ideas through any media and regardless of frontiers.

### **Article 20**

- (1) Everyone has the right to freedom of peaceful assembly and association.
- (2) No one may be compelled to belong to an association.

### **Article 21**

- (1) Everyone has the right to take part in the government of his country, directly or through freely chosen representatives.
- (2) Everyone has the right of equal access to public service in his country.
- (3) The will of the people shall be the basis of the authority of government; this will shall be expressed in periodic and genuine elections which shall be by universal and equal suffrage and shall be held by secret vote or by equivalent free voting procedures.

### **Article 22**

Everyone, as a member of society, has the right to social security and is entitled to realization, through national effort and international co-operation and in accordance with the organization and resources of each State, of the economic, social and cultural rights indispensable for his dignity and the free development of his personality.

### **Article 23**

- (1) Everyone has the right to work, to free choice of employment, to just and favourable conditions of work and to protection against unemployment.
- (2) Everyone, without any discrimination, has the right to equal pay for equal work.
- (3) Everyone who works has the right to just and favourable remuneration ensuring for himself and his family an existence worthy of human dignity, and supplemented, if necessary, by other means of social protection.
- (4) Everyone has the right to form and to join trade unions for the protection of his interests.

**Article 24**

Everyone has the right to rest and leisure, including reasonable limitation of working hours and periodic holidays with pay.

**Article 25**

(1) Everyone has the right to a standard of living adequate for the health and well-being of himself and of his family, including food, clothing, housing and medical care and necessary social services, and the right to security in the event of unemployment, sickness, disability, widowhood, old age or other lack of livelihood in circumstances beyond his control.  
(2) Motherhood and childhood are entitled to special care and assistance. All children, whether born in or out of wedlock, shall enjoy the same social protection.

**Article 26**

(1) Everyone has the right to education. Education shall be free, at least in the elementary and fundamental stages. Elementary education shall be compulsory. Technical and professional education shall be made generally available and higher education shall be equally accessible to all on the basis of merit.  
(2) Education shall be directed to the full development of the human personality and to the strengthening of respect for human rights and fundamental freedoms. It shall promote understanding, tolerance and friendship among all nations, racial or religious groups, and shall further the activities of the United Nations for the maintenance of peace.  
(3) Parents have a prior right to choose the kind of education that shall be given to their children.

**Article 27**

(1) Everyone has the right freely to participate in the cultural life of the community, to enjoy the arts and to share in scientific advancement and its benefits.  
(2) Everyone has the right to the protection of the moral and material interests resulting from any scientific, literary or artistic production of which he is the author.

**Article 28**

## *Artificial Intelligence and Democratic Values*

Everyone is entitled to a social and international order in which the rights and freedoms set forth in this Declaration can be fully realized.

### **Article 29**

- (1) Everyone has duties to the community in which alone the free and full development of his personality is possible.
- (2) In the exercise of his rights and freedoms, everyone shall be subject only to such limitations as are determined by law solely for the purpose of securing due recognition and respect for the rights and freedoms of others and of meeting the just requirements of morality, public order and the general welfare in a democratic society.
- (3) These rights and freedoms may in no case be exercised contrary to the purposes and principles of the United Nations.

### **Article 30**

Nothing in this Declaration may be interpreted as implying for any State, group or person any right to engage in any activity or to perform any act aimed at the destruction of any of the rights and freedoms set forth herein.

GPA Resolution on AI and Accountability

**RESOLUTION ON ACCOUNTABILITY  
IN THE DEVELOPMENT AND USE OF ARTIFICIAL  
INTELLIGENCE**

**Global Privacy Assembly  
October 2020**

**Sponsors**

Privacy Commissioner for Personal Data, Hong Kong, China

Superintendence of Industry and Commerce, Colombia

Federal Commissioner for Data Protection and Freedom of  
Information, Germany

Information Commissioner's Office, United Kingdom

**Co-Sponsors**

Agencia de Acceso a la Información Pública, Argentina

Office of the Privacy Commissioner of Canada

Information Access Commission, Quebec, Canada

European Data Protection Supervisor, European Union

Data Protection Commission, Italy

National Institute for Transparency, Access to Information and  
Personal Data Protection, Mexico

Office of the Privacy Commissioner, New Zealand

National Privacy Commission, Philippines

Personal Data Protection Office, Poland

National Data Protection Commission, Portugal

Data Protection Authority, Republic of San Marino

## *Artificial Intelligence and Democratic Values*

National Commission for Informatics and Liberties, Burkina Faso

Office of the Information and Privacy Commissioner, Ontario,  
Canada

### **The 2020 GLOBAL PRIVACY ASSEMBLY:**

**Recalling** the Declaration on Ethics and Data Protection in Artificial Intelligence made by the 40th International Conference of the Data Protection and Privacy Commissioners on 23 October 2018, which endorsed *inter alia* the principle of accountability of all relevant stakeholders to individuals, supervisory authorities and other third parties, and which established a permanent Working Group (**AI WG**) to address the challenges of development of artificial intelligence (**AI**), and promote understanding of and respect for the principles of the Declaration,

**Highlighting** that the Work Programme of the AI WG includes an action to prepare a statement on the essential need for accountability and liability of human actors for AI systems,

**Taking into account** the results of a survey conducted by the AI WG in May and June 2020, to gather the views of the members of the Global Privacy Assembly on accountability for AI systems, as detailed in the Explanatory Note,

**Noting** that international organisations (including the United Nations, the Organisation for Economic Co-operation and Development, the Council of Europe and the European Commission), governments, civil society bodies, and technology companies have produced and continue to produce guidelines and recommendation on the legal and ethical development of AI, and that the need for accountability and a human-centric approach are common themes within these guidelines,

**Noting** that accountability is to be understood as the compliance and demonstration of compliance with personal data protection and privacy regulations, in particular through the adoption and implementation of appropriate, practicable, systematic and effective measures,

**Affirming** that the responsibility for the operation and effects of AI systems remains with human actors,

**Taking the view** that in order to be effective, accountability obligations should be assessed against clearly defined principles and frameworks, and

extend to both organisations that develop AI systems and organisations that use them,

**Emphasising** that the principle of accountability encompasses accountability to the people affected by the decisions made by or with AI systems, as well as to supervisory authorities and, where appropriate, to other third parties, and that beyond the compliance element, accountability should also be demonstrated in order to build trust with the stakeholders,

**Recognising** that AI systems may affect human rights in different ways, the application of specific obligations should take into account the risks for human rights as well as the importance of the principle of human accountability,

**Asserting that** in order to support the trustworthiness of organisations developing and using AI systems, these organisations should work closely with policy-makers, individuals and other stakeholders (e.g. non-government organisations, public authorities and academia) to resolve concerns and rectify adverse impacts on human rights.

**The 2020 GLOBAL PRIVACY ASSEMBLY therefore resolves to:**

1. Urge organisations that develop or use AI systems to consider implementing the following accountability measures:

- (1) Assess the potential impact to human rights (including data protection and privacy rights) before the development and/or use of AI;
- (2) Test the robustness, reliability, accuracy and data security of AI before putting it into use, including identifying and addressing bias in the systems and the data they use that may lead to unfair outcomes;
- (3) Keep records of impact assessment, design, development, testing and use of AI;
- (4) Disclose the results of the data protection, privacy and human rights impact assessment of AI;
- (5) Ensure transparency and openness by disclosing the use of AI, the data being used and the logic involved in the AI;

## *Artificial Intelligence and Democratic Values*

- (6) Ensure an accountable human actor is identified (a) with whom concerns related to automated decisions can be raised and rights can be exercised, and (b) who can trigger evaluation of the decision process and human intervention;
  - (7) Provide explanations in clear and understandable language for the automated decisions made by AI upon request;
  - (8) Make human intervention on the automated decision made by AI upon request;
  - (9) Continuously monitor and evaluate the performance and impacts of AI by human beings, and act promptly and firmly to address identified issues;
  - (10) Implement whistleblowing / reporting mechanisms about non-compliance or significant risk in the use of AI;
  - (11) Ensure the auditability of AI systems and be prepared to demonstrate accountability to data protection authorities on request; and
  - (12) Engage in multi-stakeholder discussions (including with non-governmental organisations, public authorities and academia) to identify and address the wider socio- economic impact of AI and to ensure algorithmic vigilance.
2. Urge organisations that develop or use AI systems to implement accountability measures which are appropriate regarding the risks of interference with human rights.
  3. Call upon all members of the Global Privacy Assembly to work with organisations that develop or use AI systems in their jurisdictions and globally to promote the principles adopted in its 2018 resolution, and accountability in the development and use of AI, and the adoption of accountability measures;
  4. Encourage governments to consider the need to make legislative changes in personal data protection laws, to make clear the legal obligations regarding accountability in the development and use of AI, where such provisions are not already in place; and

*The AI Social Contract Index 2020*

5. Encourage governments, public authorities, standardisation bodies, organisations developing or using AI systems and all other relevant stakeholders to work with data protection authorities in establishing principles, standards, and accountability mechanisms, such as certification, for the purpose of demonstrating legal compliance, accountability and ethics in the development and use of AI systems.

[An Explanatory Note accompanies the Resolution. The Explanatory Note summarizes the opinions of the members of the Global Privacy Assembly on the measures for demonstrating accountability in the development and use of AI.]