Recommendation of the Council on Artificial Intelligence

Contents

Vision ................................................................. 3
Mission .............................................................. 3
Values ................................................................. 3
1. Investment ........................................................ 5
   1.1. R&D .......................................................... 5
   1.1.1. Technical Issues ....................................... 5
   1.1.2. Society, Legality & Ethics ......................... 5
   1.1.3. Policy Issues ........................................... 5
   1.2. Datasets .................................................... 6
   1.2.1. Bias ....................................................... 6
   1.2.2. Standards & Interoperability ...................... 6
2. Digital Ecosystem .............................................. 7
   2.1. Knowledge ............................................... 7
   2.2. Data ........................................................ 7
3. Policy ............................................................... 8
   3.1. Deployment ............................................... 8
   3.1.1. Experimentation ...................................... 8
   3.2. Assessment & Regulation ........................... 8
4. Workforce ........................................................ 9
   4.1. Transformation .......................................... 9
   4.1.1. Empowerment ........................................ 9
   4.2. Fairness .................................................... 9
   4.2.1. Training ............................................... 9
   4.2.2. Displacement ......................................... 9
   4.2.3. Labour Market ....................................... 9
   4.3. Collaboration ............................................ 10
   4.3.1. Workplaces .......................................... 10
   4.3.2. Jobs ..................................................... 10
   4.3.3. Entrepreneurship & Productivity ............... 10
   4.3.4. Benefits ............................................... 10
5. Co-Operation .................................................. 11
   5.1. Principles & Progress .................................. 11
   5.2. Knowledge ............................................... 11
   5.2.1. Initiatives .............................................. 11
   5.3. Standards ................................................ 11
   5.4. Metrics ...................................................... 12
   5.4.1. Evidence & Assessment ......................... 12
Administrative Information ..................................... 12
Organisation for Economic Co-operation and Development (OECD)

Stakeholder(s):
Council on Artificial Intelligence

Vision
Trustworthy AI

Mission
To promote principles for responsible stewardship of trustworthy AI

Values

**Principles**: the following principles are complementary and should be considered as a whole.

**Inclusion**: 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.

**Growth**
**Sustainability**
**Well-Being**

**Human-Centricity**: Human-centred values and fairness

**Rule of Law**: AI actors should respect the rule of law, human rights and democratic values, throughout the AI system lifecycle.

**Human Rights**
**Democracy**

**Freedom**: These include freedom, dignity and autonomy, privacy and data protection, non-discrimination and equality, diversity, fairness, social justice, and internationally recognised labour rights.

**Dignity**
**Autonomy**
**Privacy**
**Data Protection**
**Non-Discrimination**
**Equality**
Diversity
Fairness
Social Justice
Labour Rights

**Mechanisms:** 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.

**Safeguards**

**Human Determination**

**State of Art**

**Transparency:** 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:

**Explainability**

**Understanding:** to foster a general understanding of AI systems

**Awareness:** to make stakeholders aware of their interactions with AI systems, including in the workplace

**Outcomes:** to enable those affected by an AI system to understand the outcome

**Corrections:** 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

**Robustness:** 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.

**Security**

**Safety**

**Traceability:** 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.

**Risk Management:** 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.

**Privacy**

**Fairness**

**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.
1. Investment

Invest in AI

1.1. R&D

Invest in AI research and development

Stakeholder(s):

Governments:
Governments should consider long-term public investment, and encourage private investment, in research and development, including interdisciplinary efforts, to spur innovation in trustworthy AI that focus on challenging technical issues and on AI-related social, legal and ethical implications and policy issues.

1.1.1. Technical Issues

Focus on challenging technical issues

1.1.2. Society, Legality & Ethics

Focus on AI-related social, legal and ethical issues

1.1.3. Policy Issues

Focus on AI-related policy issues

1.2. Datasets

Invest in open datasets

Stakeholder(s):

Governments:
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.

1.2.1. Bias

Ensure that AI research and development that is free of inappropriate bias
1.2.2. Standards & Interoperability

*Use standards and improve interoperability*
2. Digital Ecosystem

_Foster a digital ecosystem for AI_

**Stakeholder(s)**

**Governments**: 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.

2.1. Knowledge

_Sharing AI knowledge_

2.2. Data

_Promote the safe, fair, legal and ethical sharing of data_
3. Policy

Shape an enabling policy environment for AI

3.1. Deployment

Support transition from the research and development stage to the deployment and operation stage for trustworthy AI systems

**Stakeholder(s):**

**Governments:**
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.

3.1.1. Experimentation

Provide a controlled environment in which AI systems can be tested and scaled-up

3.2. Assessment & Regulation

Adapt policy and regulatory frameworks and assessment mechanisms to encourage innovation and competition for trustworthy AI

**Stakeholder(s):**

**Governments:**
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.
4. Workforce

*Build human capacity and prepare for labour market transformation*

4.1. Transformation

*Prepare for the transformation of the world of work and of society*

**Stakeholder(s):**

**Governments:**
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.

4.1.1. Empowerment

*Equip people with skills to use and interact with AI systems*

4.2. Fairness

*Ensure a fair transition for workers as AI is deployed*

**Stakeholder(s):**

**Governments:**
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.

**Workers**

4.2.1. Training

*Offer training programmes along the working life*

4.2.2. Displacement

*Support for those affected by displacement*

4.2.3. Labour Market

*Provide access to new opportunities in the labour market*
4.3. Collaboration

Work closely with stakeholders

**Stakeholder(s):**

**Governments:**
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.

4.3.1. Workplaces

Promote the responsible use of AI at work

4.3.2. Jobs

Enhance the safety of workers and the quality of jobs

**Stakeholder(s):**
Workers

4.3.3. Entrepreneurship & Productivity

Roster entrepreneurship and productivity

**Stakeholder(s):**
Entrepreneurs

4.3.4. Benefits

Ensure the benefits from AI are broadly and fairly shared
5. Co-Operation

*Co-operate internationally for trustworthy AI*

5.1. Principles & Progress

*Co-operate to advance these principles and progress trustworthy AI*

**Stakeholder(s):**

**Governments:**
Governments, including developing countries and with stakeholders, should actively co-operate to advance these principles and to progress on responsible stewardship of trustworthy AI.

**Developing Countries**

5.2. Knowledge

*Foster sharing of AI knowledge*

**Stakeholder(s):**

**Governments:**
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.

5.2.1. Initiatives

*Encourage international, cross-sectoral and open multi-stakeholder initiatives to garner long-term expertise on AI*

5.3. Standards

*Promote the development of technical standards for interoperable and trustworthy AI*

**Stakeholder(s):**

**Governments:**
Governments should promote the development of multi-stakeholder, consensus-driven global technical standards for interoperable and trustworthy AI.
5.4. Metrics

Develop and use internationally comparable metrics to measure AI research, development and deployment

Stakeholder(s):

Governments:
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.

5.4.1. Evidence & Assessment

Gather evidence to assess progress in the implementation of these principles

Administrative Information

Start Date: 2019-05-21
End Date:
Publication Date: 2020-06-18

Submitter:
Given Name: Owen
Surname: Ambur
Email: Owen.Ambur@verizon.net
Phone: