INTERGOVERNMENTAL MEETING OF EXPERTS (CATEGORY II) RELATED TO A DRAFT RECOMMENDATION ON THE ETHICS OF ARTIFICIAL INTELLIGENCE

Introductory Statement by the Secretariat
Timeline of elaboration process of the Recommendation

- **Mar 2020**: Appointment of 24-member AHEG for preparation of the first draft of the Recommendation on the Ethics of AI
- **May 2020**: First version of a draft text of the Recommendation produced by AHEG
- **Jun – Aug 2020**: Extensive multi-stakeholder consultation process: public online consultation, 11 regional and subregional consultations, citizen deliberation workshops, consultations with UN system
- **Sep 2020**: First draft of the Recommendation produced by AHEG and transmitted to Member States
- **Sep – Dec 2020**: Comments received from 48 Member States and 2 Observers
- **Jan – Mar 2021**: Secretariat, led by SHS and involving FPWG, integrated comments and revised draft text of the Recommendation
- **Mar 2021**: Final report containing the draft text of the Recommendation submitted to Member States
Overview of the Recommendation on Ethics of AI
Aims of the Recommendation

To guide the development and use of AI in a way that benefits all of humanity and promotes sustainable development and peace.

To assist Member States in facing the changes stemming from AI technologies, and preparing the response.
Values and Principles of the Recommendation

4 VALUES

• Respect, protection and promotion of human dignity, human rights and fundamental freedoms
• Environment and ecosystem flourishing
• Ensuring diversity and inclusiveness
• Living in peaceful, just and interconnected societies

10 PRINCIPLES

• Proportionality and do no harm
• Safety and Security
• Fairness and non-discrimination
• Sustainability
• Privacy and data protection
• Human oversight and determination
• Transparency and explainability
• Responsibility and accountability
• Awareness and literacy
• Multi-stakeholder and adaptive governance and collaboration
11 Areas of Policy Action

1. Ethical impact assessment
2. Ethical governance and stewardship
3. Data policy
4. Development and international cooperation
5. Environment and Ecosystems
6. Gender
7. Culture
8. Education and Research
9. Communication and Information (*new)
10. Economy and Labour
11. Health and Social Well-Being

2 innovative tools:
• Ethical impact assessment
• Readiness assessment methodology
Strengths of the Recommendation

- Majority of Member States expressed strong support for the Recommendation, including its structure, scope, framing and comprehensiveness
- Unique perspective of UNESCO, given its universal membership and multidisciplinary expertise
- Solid focus on ethics
- Addresses ethical implications for all AI actors in all stages of AI life cycle
Points of clarification
Interplay between ethics and human rights

*Ethics as a “dynamic basis for the normative evaluation and guidance of AI technologies, referring to human dignity, well-being and the prevention of harm as a compass and rooted in in the ethics of science and technology”.*

- AI technologies are always evolving, so inclusive and pluralistic ethical reflection is constantly required
- Ethics as a dynamic process does not refer to a change in values, but to adapt existing frameworks as required to ensure that values are always respected
- Ethical frameworks are designed to provide a response in cases where regulation does not or cannot exist
Addressees of the Recommendation

The Recommendation is specifically addressed to Member States. However, Member States cannot face all the implications of AI alone. Therefore, a multi-stakeholder approach is promoted.

Member States are encouraged to set out appropriate frameworks and ensure that other stakeholders, particularly the private sector, adhere to them.
Non-binding nature of the Recommendation

The Recommendation is non-binding.

However, we hope that Member States will use this instrument to advance their own policies and decisions regarding AI.
Regulations

• The Recommendation is not proposing more or less regulation, but more **effective** regulation

• The Recommendation does not aim to impose how this is to be achieved, leaving it up to Member States’ discretion

• The Recommendation was carefully drafted not to stifle innovation
Based on Member States’ comments, it was decided that:

• The Recommendation will focus on civilian issues

• The reference to weaponization will be removed
Gender equality

- Some Member States applauded the strong emphasis on gender equality and asked to further expand this.
- Others were concerned that the significant focus on gender comes at the expense of other concepts.
- Gender is addressed in a separate policy area because it is one of UNESCO’s two global priorities.
- The Recommendation also includes policy actions to increase diversity and combat stereotyping of all forms.
Some Member States demanded stronger protection for intellectual property rights, while others asked for stronger wording to ensure that intellectual property rights do not interfere with other principles.

The Recommendation refers to international law, which includes intellectual property law.

The existing provisions aim to ensure that intellectual property is not used to obscure other human rights and fundamental freedoms.
Human oversight and determination

- Some Member States supported the stance of the Recommendation on keeping control always in the hands of humans.

- Others believed that human involvement might not be necessary in some situations as the algorithm could reach more accurate decisions on its own.

- Keeping the control with the human comes hand in hand with strengthening human capacity to understand the decision of the algorithm and to challenge/change it if needed.
The concept of data sovereignty was retained due to its importance of safeguarding against the abuse of data that is being processed outside of the country it belongs to.

The Recommendation aims to support a right balance between the concept of data sovereignty & data sharing and the free flow of data.
Some Member States asked to differentiate between AI systems based on the level of risk they pose.

The Recommendation has not adopted the risk-based approach, recognizing that ethical considerations should be taken into account for all AI systems regardless of their level of risk.
Thank you!

More on the UNESCO Recommendation on the Ethics of AI:

https://en.unesco.org/artificial-intelligence/ethics