CAHAI FEASIBILITY REPORT ("Report") COMMENTS

Actors: AI developers, deployers and users

1. Definition: The Report prefers to use the phrase "algorithmic systems" and does not define the term of AI. The task of finding a viable definition of AI should belong to technical and standardization groups such as NIST, ISO, CEN-CENELEC, IEEE. AI is a dynamic and evolving technology. To define AI within a legal instrument would hamper it and limit its evolving dimensions to the static definition of AI given under the relevant legal instrument. NIST's Cloud Computing is one of the best examples for that. If a standardization body or a technical group defines AI, it would be easy to modify that definition based on technological developments regarding AI Technology.

2. Concerning learning models of the algorithm, we suggest to take into account "federated learning or collaborative machine learning without centralized training data" for the future work of the CAHAI.

3. Plenty of Impact Assessments: The impact of AI on human rights was explained in detail based on the rights set out by the ECHR and ESC. The Report also encompasses the effects of AI on democracy and rule of law. The Risk Based Approach is an essential tool to evaluate the AI risk on human rights, democracy and rule of law. To do that the Report mentions conducting "human rights impact assessment". However, looking at the EU Commission's work and the literature, we observe that there are plenty of impact assessments such as:

- Algorithmic impact assessment (AIA)
- Data protection impact assessment (DPIA)
- Data security impact assessment
- Environmental Impact Assessment
- Human Rights Impact Assessments (HRIA)
- Privacy Impact Assessments (PIA)
- Ethical Impact Assessments, and
- Surveillance Impact Assessments (SIA)

CoE offers human rights impact assessment. However, the EU offers ethical impact assessment and also privacy & data protection impact assessment. The multitude and variety of impact assessments could create a burden particularly on SMEs that develop AI technologies. Therefore, we suggest that an approach to consolidate different kinds of impact assessments should be adopted by the CAHAI.[1]

[1] *We would like to note at this point a remark made by Jan Kleijsen of the CoE regarding the clashes between the legal instruments to be developed by the EU and the CoE and as to how to ensure coherence in text and later developing practice: According to Mr. Kleijsen, the CoE aims to include different stakeholders in its work and inform all related parties about relevant developments in an effort to minimise such clashes. Secondly, Mr.*

4. We agree on the "green lines" and "red lines", which is quite similar to the EDPB's white and blacklists concerning DPIA. Establishing green lines and red lines would be helpful for AI developers and deployers, whether there is a need to conduct an HRIA or DPIA or AIA or not. Red lines also point to "high risk AI applications". Green and red lines are not numerus clausus, which means the CoE and/or the CAHAI could expand the list based on new and emerging AI applications.

5. Clarifying and broadening the scope of the existing rights and obligations: The GDPR's extended and comprehensive data subject rights list would be a good example in that regard.

6. Regarding 7.1.5 principle of transparency and explainability of AI systems and 7.1.7 Accountability and Responsibility: Algorithmic Impact Assessments are a crucial tool in establishing algorithmic accountability. The GDPR's version of an Algorithmic Impact Assessment serves as a central connection between its two approaches to regulating algorithms: individual rights and systemic governance. To that extent Malgieri and Kaminski suggest a model of multi-layered explanations drawn from Algorithmic Impact Assessments. Since there are several layers of algorithmic explanation required by the GDPR, the Authors recommend that data controllers disclose a *relevant summary of a system*, produced in the DPIA process, as a first layer of algorithmic explanation, to be followed by *group explanations* and more granular, *individualized explanations.*

7. Paragraphs 124-125 of the Report:

The said paragraphs concern the risk based approach and determination of high risk & low risk. The Report is not as clear in that regard as the EU Commission's White Paper on AI. As Ryan Budish eloquently elaborates in his article, *"The Commission's white paper creates two categories of risk: (1) low-risk applications that will not face any new restrictions beyond existing law; and (2) high-risk applications that will face new restrictions about training data, record keeping, transparency, accuracy, human oversight, and more. The challenge, however, is in determining exactly what "high risk" and "low risk" actually mean. To that end, the white paper offers some limited guidance in the form of two criteria. First, an application is high risk "where, given the characteristics of the activities typically undertaken, significant risks can be expected to occur." And second, an application is high risk when it is "used in such a manner that significant risks are likely to arise." What is apparent is that these two cumulative criterion do not actually define "high risk," and instead circularly assert that an application is high risk*

Kleijsen notes that the CoE's angle is specifically focused on democracy, fundamental rights and the rule of law. In light of this, the CoE is currently aiming to establish general principles and we will be seeing whether the member states would like to proceed with this approach and turn these general principles into a treaty. These general principles have been laid out in several ethical charters, also by the HLEG. Therefore, the two mechanisms (that of the EU and of the CoE) are expected to be of a complementary nature. The CoE aims to set a global benchmark on a number of general principles, and they are confident that the EU laws will be based on the same principles (such as transparency, do no harm, human oversight, some form of independent control, etc). The CoE's aim seems to be to turn the voluntary engagements represented by these principles into binding obligations.

if it is used in a space that is high risk and in a way that is high risk. Such a definition only defers and displaces the determination of risk."[2]

"*In order for the European Commission to responsibly deal with the risk of AI, the Commission should learn from the debates that have taken place in a range of other scientific fields as they have grappled for decades with the role of scientific certainty and quantification of risk. Through that experience, risk governance experts and scholars have developed new frameworks that continue to value scientific data but alongside other more qualitative measures of risk. Unlike the traditional approaches to risk assessment/risk management and the precautionary principle, these more expansive risk governance frameworks embrace data and methodologies that are inherently messy, uncertain, and ambiguous. In particular, these risk governance frameworks have three important features:*

(1) they focus on broadening participation in the risk governance process, including a range of key stakeholders;

(2) they value qualitative data and policy analysis; and

(3) they use deliberative, multistakeholder processes.

Collectively, the above-mentioned three features are particularly important when addressing the risks of new technologies, which often frustrate attempts to quantify the risks. Before committing their strategy to the vague categories of "high-risk" and "low-risk", the European Commission should consider the lessons learned from past risk governance debates and ensure that they are building a risk governance framework that embraces a holistic view of risk, including more qualitative measures"[3].

Before committing our strategy to the vague categories of "high-risk" and "low-risk", we should consider the lessons learned from past risk governance debates and ensure that we are building a risk governance framework that embraces a holistic view of risk, including more qualitative measures.

8. Extraterritorial Effect of National Laws & Regulations and the Need More Harmonized Rules on AI: Extraterritorial effect is a rising trend of national laws and regulations, which could create overlaps between national laws & regulations and relevant international instruments. For example; an AI developer or deployer not established in the EU, based on the criteria of Article 3 might comply with the GDPR. Article 3 of the GDPR sets forth criteria on its territorial scope. If an AI developer or deployer's country is a member state of a relevant international convention, the AI developer or deployer should have to fulfill obligations of its national law which transpose international conventions provisions into national law. If the AI developer and deployer's country is a part of the WTO agreements, this

[2] Ryan Budish: AI & the European Commission's Risky Business https://medium.com/berkman-klein-center/ai-the-european-commissions-risky-business-a6b84f3acee0

[3] Ryan Budish: AI & the European Commission's Risky Business https://medium.com/berkman-klein-center/ai-the-european-commissions-risky-business-a6b84f3acee0

could be another layer in terms of obligations to fulfill. If we were to add soft law requirements (i.e. ethical guidelines), standards and certifications obligations as another layer, that could be more complex and burdensome for AI developers and deployers to be compliant with all of these legal, ethical and technical requirements.

9. Liability for damage caused by artificial intelligence: There are new and varied approaches regarding civil liability of AI systems. However, it is important to underline that the new liability approaches find their roots in fundamental principles of civil law and commercial law:

- (objective) good faith

- prudent merchant

The liability discussions generally rotate around the above mentioned principles. The recent Artificial Intelligence and Civil Liability Report has elaborated on the questions regarding AI and liability by referring to civil liability to solidify the ambiguous approaches regarding AI.[4]

10. Certification and Quality Labelling (9.3.1 of the report): The European Commission for the efficiency of justice (CEPEJ) considers a specific certification for AI systems in the legal sector such as law enforcement, courts and judiciary. The EU has already established different types of certifications and labellings such as CE, Ecolabel, EU Trustmark, Privacy Labelling and Certification under GDPR etc. All of them have different requirements in terms of obtaining relevant certification and labelling. The CAHAI should take that point into consideration in order to align its own criteria with other certification and labelling schemes.

As a last remark concerning the audit mechanism, it should be mentioned that regardless of the name of the impact assessment mechanism (AI impact assessment or human rights impact assessment, ethical impact assessment, algorithmic impact assessment), the impact assessment is definitely helpful in order to determine the level and feature of the risk and the assessment enhances the audit mechanism.

11. As to the regulatory efforts on AI based on the CoE's standards on human rights, democracy and rule of law the CAHAI Secretariat published a very comprehensive report called *"TOWARDS REGULATION OF AI SYSTEMS Global perspectives on the development of a legal framework on Artificial Intelligence systems based on the Council of Europe's standards on human rights, democracy and the rule of law"*, Compilation of contributions prepared by the CAHAI Secretariat, December 2020. They develop 7 criteria in order to oversee the AI applications.

[4] Regarding the civil liability of AI, see also https://www.europarl.europa.eu/doceo/document/TA-9-2020-0276_EN.pdf

12. The EU Agency for Fundamental Rights (FRA) published on 14 December, 2020 a Report called "Getting the future right – Artificial intelligence and fundamental rights Artificial intelligence and big dataData protection, privacy and new technologies".

This report supports that goal by analysing fundamental rights implications when using artificial intelligence. Based on concrete 'use cases' of AI in selected areas, it focuses on the situation on the ground in terms of fundamental rights challenges and opportunities when using AI.

Main findings:

I. **Considering the full scope of fundamental rights with respect to AI**: FRA Opinion 1:
   A. "When introducing new policies and adopting new legislation on AI, the EU legislator and the Member States, acting within the scope of EU law, must ensure that respect for the full spectrum of fundamental rights, as enshrined in the Charter and the EU Treaties, is taken into account. Specific fundamental rights safeguards need to accompany relevant policies and laws."
   B. "In doing so, the EU and its Member States should rely on robust evidence concerning AI's impact on fundamental rights to ensure that any restrictions of certain fundamental rights respect the principles of **necessity and proportionality**. "
   C. "Relevant safeguards need to be provided for by law to effectively protect against arbitrary interference with fundamental rights and to give legal certainty to both AI developers and users. Voluntary schemes for observing and safeguarding fundamental rights in the development and use of AI can further help mitigate rights violations."
   Ç. "The legal definition of AI-related terms might need to be assessed on a regular basis."

II. FRA Opinion 2: "The EU legislator should consider making **mandatory impact assessments** that cover **the full spectrum of fundamental rights**. These should cover the private and public sectors, and be applied before any AI-system is used. The impact assessments should take into account the varying nature and scope of AI technologies, including the level of automation and complexity, as well as the potential harm. They should include basic screening requirements that can also serve to raise awareness of potential fundamental rights implications. The EU and Member States should consider *using existing tools, such as checklists or self-evaluation tools,* developed at European and international level. These include those developed by the EU High-Level Group on Artificial Intelligence." (emphasis added)

III. FRA Opinion 3: "The EU and Member States should ensure that **effective accountability systems** are in place to monitor and, where needed, effectively address any negative impact of AI systems on fundamental rights. They should consider, in addition to f*undamental rights impact assessments* (see FRA opinion 2), introducing specific safeguards to ensure that the accountability regime is effective. This could

include a legal requirement to make available enough information to allow for an assessment of the fundamental rights impact of AI systems. This would enable external monitoring and human rights oversight by competent bodies. The EU and Member States should also make better use of existing oversight expert structures to protect fundamental rights when using AI. These include *data protection authorities, equality bodies, national human rights institutions, ombuds institutions and consumer protection bodies*." (emphasis added)

IV. FRA Opton 4 Specific safeguards to ensure non-discrimination when using AI: "EU Member States should consider encouraging companies and public administration to assess any potentially discriminatory outcomes when using AI systems. The European Commission and Member States should consider providing funding for targeted research on potentially discriminatory impacts of the use of AI and algorithms. Such research would benefit from the adaptation of established research methodologies, from the social sciences, that are employed to identify potential discrimination in different areas – ranging from recruitment to customer profiling.This suggests a lack of in-depth assessments of such discrimination in automated decision making."

V. FRA Opinion 5 ADM: "More clarity is needed on the scope and meaning of legal provisions regarding automated decision making. The European Data Protection Board (EDPB) and the European Data Protection Supervisor (EDPS) should consider providing further guidance and support to effectively implement GDPR provisions that directly apply to the use of AI for safeguarding fundamental rights, in particular as regards the meaning of personal data and its use in AI, including in AI training datasets."

VI. FRA Opinion 6: "Effective access to justice in cases involving AI-based decisions. To effectively contest decisions based on the use of AI, people need to know that AI is used, and how and where to complain. Organisations using AI need to be able to explain their AI system and decisions based on AI. The EU legislator and Member States should ensure effective access to justice for individuals in cases involving AI-based decisions. To ensure that available remedies are accessible in practice, the EU legislator and Member States could consider introducing a legal duty for public administration and private companies using AI systems to provide those seeking redress information about the operation of their AI systems. This includes information on how these AI systems arrive at automated decisions. This obligation would help achieve equality of arms in cases of individuals seeking justice. It would also support the effectiveness of external monitoring and human rights oversight of AI systems (see FRA opinion 3). In view of the difficulty of explaining complex AI systems, the EU, jointly with the Member States, should consider developing guidelines to support transparency efforts in this area. In doing so, they should draw on the expertise of national human rights bodies and civil society organisations active in this field."