Submission to the Consultation on the "White Paper on Artificial Intelligence – a European approach to excellence and trust"

Contribution by the Secretariat of the Council of Europe
19 June 2020

Note: This submission is a contribution from the Secretariat of two directorates general of the Council of Europe (DG1 – Human Rights and Rule of Law; DG2 – Democracy), and has been coordinated by the Secretariat of the Ad hoc Committee on Artificial Intelligence (CAHAI)

Introduction

As the continent's oldest pan-European intergovernmental organisation, the Council of Europe brings together 47 member States and 830 million citizens around one idea: the peoples of Europe can only build a common future on the basis of what unites them rather than what divides them. The Organisation has been built around three pillars (human rights, rule of law and democracy) and defends fundamental and universal values through more than 220 Conventions, including the European Convention for the Protection of Human Rights and Fundamental Freedoms. At the same time, the Council of Europe has also built up an unprecedented set of implementation, monitoring and assistance mechanisms for its member States in fields as varied as the functioning of democratic institutions, the efficiency and independence of judicial systems, social rights, culture, education, youth and sport, and has become over time the true democratic conscience of Europe.

In the field of the control of technological developments, the Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data (ETS No. 108), known as "Convention 108", was the first international instrument in the field, entering into force on 1 October 1985. Convention 108 has been ratified by the 47 member States of the Council of Europe and 6 non-member States (Cape Verde, Mauritius, Mexico, Senegal, Tunisia, Uruguay). Its modernised version known as "Convention 108+" (CETS No. 223) has already been signed by 35 member States of the Council of Europe and 3 non-member State (Argentina, Tunisia and Uruguay)[1]. The Budapest Convention on combating cybercrime (ETS No. 185) is also another important tool used widely (65 ratifications, including 21 non-member States of the Council of Europe)[2].The ban on human cloning is also the result of a protocol ETS No. 168 to the so-called "Oviedo Convention" enacted in 1998, which came into force just two years after the cloning of the sheep Dolly. The Convention was the first binding legal instrument in the field, striking a balance between freedom of research and the protection of individuals. Many of other non-binding instruments of the Council of Europe address the impact of digital technologies in their provisions (e.g. Recommendation of the Committee of Ministers to member States on guidelines to respect, protect and fulfil the rights of the child in the digital environment, Recommendation of the Committee of Ministers to member States on preventing and combating sexism, Recommendation of the Committee of Minister to member States on Human rights and business). Amongst digital technologies, increased attention has been paid to artificial intelligence (AI) and its effects on human rights, democracy and the rule of law.

In this regard, while increasing numbers of Europeans reap the benefits of AI every day - as unprecedented progress is made in a wide range of fields, including industrial productivity, health care, transportation and logistics - there is growing concern amongst the public about the broader implications of the use, and possible abuse, of automated data processing and

---

[1] For the full list, see Council of Europe's Treaty website: https://www.coe.int/en/web/conventions/full-list/-/conventions/treaty/223

[2] For the full list, see Council of Europe's Treaty website: https://www.coe.int/en/web/conventions/full-list/-/conventions/treaty/185/signatures?p_auth=ckwvNVy9

mathematical modelling for individuals, for communities, and for society at large. Can computational data analytics replace the reasoning of a trained judge when applying the law to a specific context? How does algorithmic decision-making affect the delivery of essential public services and our recruitment and employment conditions? Can individuals remain visible as independent agents in societies that are shaped by optimisation processes? And finally: how does the increasing reliance on mainly privately developed and run technology square with the rule of law and the fundamental principle of democratic societies that all power must be accountable before the law?

If economic growth remains an important public policy objective and innovation one of its key components, Council of Europe member States are bound by the obligation to ensure that the basic values of human rights, democracy and rule of law remain effectively embedded in appropriate legislative frameworks and enforced throughout on-going societal and technical evolution. The Organisation helps them fulfil this objective both through a specialised (vertical) approach, by elaborating instruments adapted to the specific challenges raising from the use of AI in each policy area (e.g. justice, democracy, autonomous vehicles, biomedicine), as well as through a cross-cutting (horizontal) approach aimed at detecting common principles and general actions to be undertaken.

With regard to the vertical approach, it is worth noting that several recent Council of Europe's non-binding instruments; such as the Recommendation 2020(1) on the human rights impact of algorithmic systems, the Recommendation of the Human Rights Commissioner on "Unboxing artificial intelligence: 10 steps to protect human rights", the Guidelines on Artificial Intelligence and Data Protection, the CEPEJ European Charter on the use of AI in judicial systems and their environment, the Recommendation of the Parliamentary Assembly of the Council of Europe about Technological convergence, artificial intelligence and human rights or the study of the MSI-AUT on responsibility and AI address issues dealt with by the White Paper. These texts confirm the importance of adopting a risk-based approach in regulating AI, propose relevant criteria for identifying high-risk applications[3] as well as steps aimed at reducing these risks (for instance, human rights impact assessments[4]) and converge on the need to define specific obligations, both at the level of member States and of private actors, to assess verify compliance with human rights requirements. It is important that risk assessment ensures both the protection of individuals and the general interests of society and that its criteria are widely discussed in open debates, set up through a democratic process and that proper citizens' oversight is ensured.

With regard to the horizontal approach, it is worth recalling that the Council of Europe's Ad hoc committee on artificial intelligence (CAHAI)[5] has been set up in order to "examine the feasibility and potential elements on the basis of broad multi-stakeholder consultations, of a legal framework for the development, design and application of artificial intelligence, based

---

[3] The Recommendation CM/Rec 2020(1) on the human rights impacts of algorithmic systems proposes a broad approach to the question of the level of the impact – positive and negative – of algorithmic systems on human rights.

[4] In the Recommendation of the Commissioner for Human Rights, "10 steps to protect human rights", it is emphasized that "Member states should establish a legal framework that sets out a procedure for public authorities to carry out human rights impact assessments (HRIAs) on AI systems acquired, developed and/or deployed by those authorities." and that is an obligation of member states to facilitate the implementation of human rights standards in the private sector.

[5] For more information on the work of the CAHAI, please see: www.coe.int/cahai

on the Council of Europe's standards on human rights, democracy and the rule of law". The European Union, through its institutions and agencies, contributes to the work of the CAHAI in an observer capacity.

Following a mapping of existing binding and non-binding frameworks, the CAHAI will engage through specific working groups and a broad multi-stakeholder consultation to develop this feasibility study and propose potential regulatory elements.

Member States of the Council of Europe represented in the CAHAI recalled the need for a coordinated approach with other international organisations contributing to building a global architecture for technology regulation, including the European Union. The present submission is therefore a further contribution to this necessary co-ordination.

Contribution to the Survey

Section 1 - An ecosystem of excellence

To build an ecosystem of excellence that can support the development and uptake of AI across the EU economy, the White Paper proposes a series of actions.

In your opinion, how important are the six actions proposed in section 4 of the White Paper on AI (1-5: 1 is not important at all, 5 is very important)?

Working with Member states	**5**
Focusing the efforts of the research and innovation community	**5**
Skills	**5**
Focus on SMEs	**4**
Partnership with the private sector	**4**
Promoting the adoption of AI by the public sector	**5**

Are there other actions that should be considered?
500 character(s) maximum

This section does not address fundamental rights commitments of member States. This ecosystem of excellence could therefore also take into account, in its foundations this essential dimension of creating a framework of trust between all stakeholders. The respect for regulatory frameworks and fundamental rights and the meaningful implementation of the relevant existing regulation, as well as a broad and effective public consultation are important prerequisites to establish such framework of trust.

More generally, it seems important that the actions are part of a global regulatory policy, not only European but also global, in line with the initiatives of the Council of Europe, the OECD and the United Nations (especially UNESCO).

Public services and agencies should also be encouraged to develop good practices for the use AI systems which could possibly serve as a model for other users. The development of operational actions and programmes around AI in sectors such as the audio-visual or cultural sectors can raise awareness of AI issues and good practices among the general public.

Revising the Coordinated Plan on AI (Action 1)

The Commission, taking into account the results of the public consultation on the White Paper, will propose to Member States a revision of the Coordinated Plan to be adopted by end 2020.

In your opinion, how important is it in each of these areas to align policies and strengthen coordination as described in section 4.A of the White Paper (1-5: 1 is not important at all, 5 is very important)?

Strengthen excellence in research	**4**
Establish world-reference testing facilities for AI	**5**
Promote the uptake of AI by business and the public sector	**4**
Increase the financing for start-ups innovating in AI	**3**
Develop skills for AI and adapt existing training programmes	**4**
Build up the European data space	**3**

Are there other areas that that should be considered?

500 character(s) maximum

While cooperation among member States is essential to prevent duplication of potentially inconsistent legislative measures and diverging implementation, Section 4A outlines further measures to ensure that the development of AI is carried out for the benefit of humans as well as societal and environmental well-being.

Innovation and technological development must serve humanity and respect human rights, rule of law and democracy. But use of digital technologies is not always useful, beneficial or mandatory: it seems important to conduct a case-by-case assessment of the benefit of these technologies, including AI, in order to prevent any form of "technological solutionism", excluding the use of technologies in full or partly in areas where human presence can bring greater benefits (care for the elderly, learning by children and youth) and encouraging human control and oversight.

Moreover, enforced and efficient cooperation of independent regulatory bodies (such as DPAs, consumer protection authorities) should not be underestimated and should be promoted as a stand-alone action.

Skills development should integrate ethical frameworks, human rights and risks assessments as a valued competence and an integral part of AI design and deployment. Efforts should be made to include under-represented groups, beyond women; this extends to section 4.F on adoption of AI in public sector. Reference should be also be made to democratic oversight and transparency in its deployment. A better case should be made as to why AI can bring greater benefits in the delivery of public services while acknowledging inherent risks in its deployment, for instance privacy concerns on health data, including access of private business to such data, data bias effects on underrepresented groups etc.

A united and strengthened research and innovation community striving for excellence

Joining forces at all levels, from basic research to deployment, will be key to overcome fragmentation and create synergies between the existing networks of excellence.

In your opinion how important are the three actions proposed in sections 4.B, 4.C and 4.E of the White Paper on AI (1-5: 1 is not important at all, 5 is very important)?

Support the establishment of a lighthouse research centre that is world class and able to attract the best minds	**4**
Network of existing AI research excellence centres	**5**
Set up a public-private partnership for industrial research	**4**

Are there any other actions to strengthen the research and innovation community that should be given a priority?

500 character(s) maximum

Firstly, it is important that research efforts focus on fundamental rights so as to promote the development of legislation, ethical standards and guidelines on AI compliant with such rights. Among the measures to be encouraged concerning the research community, particular attention should be paid to the initial training of computer professions (especially data scientists) in order to make digital humanities compulsory (e.g. applied use of technology in social sciences, technical studies on human rights in digital environments)).

Ensuring equal opportunities and encouraging a multi-disciplinary approach in common high-level research projects should be supported. with a view to encouraging the emergence of multidisciplinary/diverse project teams. Efforts must be made to attract underrepresented groups, beyond women. Conversely, training in law, human and social sciences could benefit from digital training.

The White paper refers in Section 4.C to the updated Digital Education Action Plan and claims that the Action plan will increase awareness of AI at all levels of education in order to prepare citizens for informed decisions that will be increasingly affected by AI, and it focuses on skills. However, empowering citizens by giving them the means to acquire the technical and functional skills and competences for democratic culture is no less important than their protection and safety, and enables them to tackle the challenges and risks arising from the digital environment and emerging technologies, e.g. AI, but also to benefit from the opportunities they provide.

The White paper also claims that the Plan will help make better use of data and AI-based technologies to improve education and training systems. This claim could be reconsidered in the light of the recent COVID-19 crisis. Digital technologies certainly helped sustaining education during the confinement. However, whether access to quality education is ensured, or teaching and learning are realised through digital technologies/platforms yet needs to be questioned and explored. Therefore, consideration could be given to developing new pedagogies or adapting the existing ones to the emerging technologies. This requires co-operation among education authorities, academia and developers of AI-based technologies. The Recommendation from the Committee of Ministers to the member States [CM/Rec(2018)7 on guidelines to respect, protect and fulfil the rights of the child in the digital environment](#) includes guidance on priority measures for a strategic and comprehensive approach in policy making in this field.

Education data need particular attention, as the data that will be generated by AI-based systems may affect individual learning path, thereby raising ethical considerations and issues about the collection and use of this data. This is particularly important when education is provided by or in co-operation with the private sector. While the collection and use of this data are regulated at national or international levels (GDPR, [Modernised Convention for the Protection of Individuals with Regard to the Processing of Personal Data](#) (Convention 108+), further consideration could be given to aspects specific to education in the existing or future legal frameworks.

Focusing on Small and Medium Enterprises (SMEs)

The Commission will work with Member States to ensure that at least one digital innovation hub per Member State has a high degree of specialisation on AI.

In your opinion, how important are each of these tasks of the specialised Digital Innovation Hubs mentioned in section 4.D of the White Paper in relation to SMEs (1-5: 1 is not important at all, 5 is very important)?

Help to raise SME's awareness about potential benefits of AI	**4**
Provide access to testing and reference facilities	**4**
Promote knowledge transfer and support the development of AI expertise for SMEs	**4**
Support partnerships between SMEs, larger enterprises and academia around AI projects	**4**
Provide information about equity financing for AI startups	**3**

Are there any other tasks that you consider important for specialised Digital Innovations Hubs?
500 character(s) maximum

The use of the term AI is sometimes imprecise and is a marketing term covering the use of sometimes simple algorithms. In order to create a quality offer and to distinguish innovative and virtuous SMEs, an "IA" label could be awarded to companies that really implement machine learning, deep learning, etc. This label could include a non-technical component, such as ethical commitments and readiness to undertake a regular audit based on principles such as respect for human rights and fundamental freedoms, prevention of discrimination, promotion of gender equality, inclusive economic growth, etc.

The European creative sector (culture and audiovisual) is mainly made up of small companies that do not have the capacity to develop AI knowledge and skills. Raising awareness about the impact of AI in this sector (and SMEs in general) is essential.

Section 2 - An ecosystem of trust

Chapter 5 of the White Paper sets out options for a regulatory framework for AI.

In your opinion, how important are the following concerns about AI (1-5: 1 is not important at all, 5 is very important)?

AI may endanger safety	**5**
AI may breach fundamental rights (such as human dignity, privacy, data protection, freedom of expression, workers' rights etc.)	**5**
The use of AI may lead to discriminatory outcomes	**5**
AI may take actions for which the rationale cannot be explained	**4**
AI may make it more difficult for persons having suffered harm to obtain compensation	**4**
AI is not always accurate	**5**

Do you have any other concerns about AI that are not mentioned above? Please specify:
500 character(s) maximum

The importance of ex-ante measures to prevent many of the aspects mentioned above should be stressed at this stage. Each risk must be contextualised in terms of its foreseeable wider impact on individuals or society: thus, the need for explicability seems greater in a context of decision-making and with legal effect than in a context of musical recommendations. But, at the same time, impacts on cultural choices and cultural diversity may imply discriminatory practices and limit freedom of expression, hence, affect human rights. The issue of culture, creativity, diversity and AI must not to be neglected in the debate, including about building trust.

More generally, it is important to develop the trust of the public, chiefly young people, about what AI represents (AI and Digital literacy). It is seen as central for creating an "ecosystem of trust", for the inclusion of young people and for enhancing the transformative power of technology on the well-being (not only on economics and competition).

Do you think that the concerns expressed above can be addressed by applicable EU legislation? If not, do you think that there should be specific new rules for AI systems?

Current legislation is fully sufficient	
Current legislation may have some gaps There is a need for a new	✓
There is a need for a new legislation Other	
No opinion	

Other, please specify

500 character(s) maximum

AI may already be regulated by national or European Union legislation, as well as treaties or conventions such as the European Convention on Human Rights or its data processing components by Convention 108+.

In the event of litigation, however, there is a risk that member States adopt divergent case law approaches. It seems therefore useful to clarify or adapt certain aspects of the existing regulatory framework, particularly as regards civil and criminal liability based on existing instruments and mechanisms (see Section 3 below).

There is, however, a strong need for higher level legal framework in order to identify fundamental principles common to all types of AI applications. The CAHAI (Ad Hoc Committee on Artificial Intelligence) intends to study the advisability of such a legal framework in order to provide shared legal foundations among the 47 Council of Europe Member States and lay the ground for more specific sectorial regulation in given fields.

If you think that new rules are necessary for AI system, do you agree that the introduction of new compulsory requirements should be limited to high-risk applications (where the possible harm caused by the AI system is particularly high)?

Yes	
No	
Other	✓
No opinion	

Other, please specify:

500 character(s) maximum

The idea of regulation proportional to the risks incurred has been raised in several international organisations (ONE AI at the OECD, AHEG at UNESCO in particular).

Such approach is intended to identify legal constraints which are proportionate to the risks incurred in each specific situation. However, the determination of criteria must clearly prioritise general interests, protection of human rights and fundamental freedoms. These criteria should result from open debates and set up through a democratic process. Proper citizens' oversight must be ensured. The risk-based approach must focus on the potential harmful impact both at the individual and at a wider society level and should be based on clear and transparent rules.

The Recommendation CM/Rec 2020(1) on the human rights impacts of algorithmic systems proposes a broad approach to the question of the level of the impact – positive and negative – of algorithmic systems (including AI) on human rights; this depends on the specific purpose for which algorithmic systems are used, their functionality, accuracy, complexity, their effects and the scale on which they are deployed. Their impact also depends on the broader organisational, thematic, societal and legal context in which they are used, each of which is associated with specific public and ethical values (…) The evaluation of the extent of the possible human rights impact of an algorithmic system should take account of the severity, scale and likelihood of giving rise to a human rights violation (…) In this Recommendation, the term "high risk" is applied when referring to the use of algorithmic systems in processes or decisions that can produce serious consequences for individuals or in situations where the lack of alternatives prompts a particularly high probability of infringement of human rights, including by introducing or amplifying distributive injustice.

Moreover, it should be possible to sanction the infringement of certain fundamental values, whether or not the applications are - or are not - high-risk. Thus, a strict liability regime could be envisaged in the event of a violation of human rights or a serious breach of democracy or the rule of law principles and rules (see the study Responsibility and AI, DG1(2019)05).

As AI supposes a new way of processing personal data, existing regulations need to be interpreted and consequently implemented in a way to ensure the same level of protection while facilitating free flow of data as for non-AI powered processing.

Do you agree with the approach to determine "high-risk" AI applications proposed in Section 5.B of the White Paper? [Note YM: Section 5.C]

Yes	
No	
Other	✓
No opinion	

Other, please specify:

500 character(s) maximum

See above for a general comment.

The notion of "high risk" should be defined more precisely than in the White Paper. As mentioned earlier, Council of Europe Recommendation (2020) 1 on the human rights impact of algorithmic systems *can provide guidance in this respect. If the combination of sector of application / type of application may sound very effective, the effective determination of criteria (or a system of criteria) should be the result of open debates and a democratic process, placing the general interest and human rights at the centre. The current proposal could lead to the exclusion of uses of AI that have a strong direct or indirect impact, such as content recommendation systems that contribute to the shaping of public opinion and may limit freedom of expression.*

Consideration should also be given to the addition of AI to existing systems, transforming their scope and nature. Thus, although the use of video surveillance is today strictly regulated, the addition of AI substantially transforms its functionalities and legal consequences, particularly in the public space.

If you wish, please indicate the AI application or use that is most concerning ("high-risk") from your perspective:

500 character(s) maximum

See above the two previous comments.

Many suggestions could be discussed on how to determine what is a "high risk" (such as a grid of analysis, based on criteria that are the results of an open and democratic debate).

Nevertheless, it should be stressed that this notion covers in particular public services but extends to all products and services. Private services that can be considered a public good are of particular concern (e.g. financial services, education) where discrimination/bias can have large impact. Employment and workers' rights are also particularly sensitive areas.

In general, we could consider that applications of AI which serve, intensify or otherwise create negative discriminatory outcomes based on genetic, biometric characteristics, attributes, racial or ethnic origin, political opinions, religious or other beliefs, health or sexual life would be seen as extremely high risk.

In the creative and audiovisual sector, further to the issues linked to data protection and privacy, AI regulation seems also appropriate in order to avoid unjustified and unexplained restrictions of choice to consumers or restrictions of the freedom of expression of artists.

In your opinion, how important are the following mandatory requirements of a possible future regulatory framework for AI (as section 5.D of the White Paper) (1-5: 1 is not important at all, 5 is very important)?

The quality of training data sets	**5**
The keeping of records and data	**5**
Information on the purpose and the nature of AI systems	**5**
Robustness and accuracy of AI systems	**5**
Human oversight	**5**
Clear liability and safety rules	**5**

In addition to the existing EU legislation, in particular the data protection framework, including the General Data Protection Regulation and the Law Enforcement Directive, or, where relevant, the new possibly mandatory requirements foreseen above (see question above), do you think that the use of remote biometric identification systems (e.g. face recognition) and other technologies which may be used in public spaces need to be subject to further EU-level guidelines or regulation:

No further guidelines or regulations are needed	
Biometric identification systems should be allowed in publicly accessible spaces only in certain cases or if certain conditions are fulfilled (please specify)	
Other special requirements in addition to those mentioned in the question above should be imposed (please specify)	
Use of Biometric identification systems in publicly accessible spaces, by way of exception to the current general prohibition, should not take place until a specific guideline or legislation at EU level is in place	✓
Biometric identification systems should never be allowed in publicly accessible spaces	
No opinion	

Please specify your answer:

While the definition of public space could be further substantiated (whether premises of public institutions are covered by it or not), the introduction of biometric identification system in public spaces in general is widely seen to possibly lead to a whole paradigm shift in the perception of individual's privacy.

Beside the fact that it could also result in some undesirable consequences (theft, misuse of identity, raise in cybercrime, loss in public trust, etc.) it would have a serious chilling effect on the reasonable expectations by individuals towards their privacy in public spaces. On the other hand, if accompanied by large public support, it would need to comply with art 8.2 of the ECHR and art 11 of the modernised Convention 108.

If its specific, temporary and geographically restricted usages (e.g.: fight against terrorism) is however justified and considered lawful, appropriate safeguards still ought to be put in place which width and costs should not be underestimated when deciding on the necessity and proportionality of such measure. Safeguards should include a very strict purpose limitation, prohibition of discrimination (as explained above), existence of appropriate national legislation, compliance with the principle of the proportionality and legitimacy of the processing (with special attention to the quality of consent) exigence of high quality (above all accuracy) of data and the model used; safeguards related to transparency, data security, data storage (its time limit and its infrastructure), accountability, meaningful exercise of data subject rights and effective independent oversight. Such a system would require a thorough privacy impact assessment which also include societal and ethical considerations of the use of such technique and an innovative use of the privacy-by-design principle.

Do you believe that a voluntary labelling system (Section 5.G of the White Paper) would be useful for AI systems that are not considered high-risk in addition to existing legislation?

Very much	✓
Much	
Rather not	
Not at all	
No opinion	

Do you have any further suggestion on a voluntary labelling system?

500 character(s) maximum

Even for applications that are not classified as "high-risk", a mandatory labelling of artificial intelligence applications could be imagined in certain circumstances: in the field of education, for example, or social networking platforms with regard to the prioritisation of information or content filtering. The impact on human rights should be always considered and could be assessed in specific assessments such as "human rights impact assessment" (see the Recommendation of the Commissioner for Human Rights "Unboxing AI: 10 steps to protect human rights").

Such labelling could be also essential in the audiovisual and cultural sectors where digital platforms can have a huge impact on the structuration of the offer to the European consumers and the shaping of their tastes.

What is the best way to ensure that AI is trustworthy, secure and in respect of European values and rules?

Compliance of high-risk applications with the identified requirements should be self-assessed ex-ante (prior to putting the system on the market)	
Compliance of high-risk applications should be assessed ex-ante by means of an external conformity assessment procedure	
Ex-post market surveillance after the AI-enabled high-risk product or service has been put on the market and, where needed, enforcement by relevant competent authorities	
A combination of ex-ante compliance and ex-post enforcement mechanisms	✓
Other enforcement system	
No admission	

Please specify any other enforcement system:

An intelligent regulation of artificial intelligence could be based on a combination of ex-ante and ex-post mechanisms, ranging from mandatory human rights impact assessment (see above), certification at the time of placing on the market to regular performance monitoring by an independent authority for applications assisting decision-making in public services for example. These mechanisms could, moreover, initially be experimental, sector by sector (laws of experimentation), in order to enable the right level of constraint to be gradually calibrated.

Do you have any further suggestion on the assessment of compliance?
500 character(s) maximum

These mechanisms should be established in close cooperation with the IEEE, ISO (CEN), national regulators and monitoring bodies (e.g. equalilty bodies, National Data Protection Authorities, etc). They should be encouraged/facilitated to cooperation to ensure complementarity of expertise to perform a sound impact assessment, monitoring and provide potential for redress. Moreover, other professional organisations from the concerned sectors (for instance creative and audiovisual sectors) should also be associated in the development of AI ethical guidance applicable to these sectors.

Section 3 – Safety and liability implications of AI, IoT and robotics

The overall objective of the safety and liability legal frameworks is to ensure that all products and services, including those integrating emerging digital technologies, operate safely, reliably and consistently and that damage having occurred is remedied efficiently.

The current product safety legislation already supports an extended concept of safety protecting against all kind of risks arising from the product according to its use. However, which particular risks stemming from the use of artificial intelligence do you think should be further spelled out to provide more legal certainty?

Cyber risks	✓
Personal security risks	✓
Risks related to the loss of connectivity	
Mental health risks	

In your opinion, are there any further risks to be expanded on to provide more legal certainty?

500 character(s) maximum

Among the many other risks to be considered, cognitive biases such as automation bias or anchoring bias should be taken into account at the design stage for all systems. The objective would thus be to prevent the risks, whether voluntary or not, of manipulation of individuals' decision making by providing them with all the additional elements needed to objectify their decision making. In a larger sense: perception of privacy by individuals, disconnection from a humanist perception of individuals and its unique characteristics and potentials (like in educational sector), lowering of human involvement in decision making, potential loss of jobs, etc… This is also particularly necessary for systems used for creative and audiovisual content for instance delivered through digital platforms. In such case, decision-making to offer or consume a particular cultural product is not necessarily based on objective criteria.

Automated driving is one example of how artificial intelligence could both address and create risks of physical harm in the public sphere. The trends in technological development, and the case of automated vehicles, suggest that AI and machines with autonomous functionality will become ever more present in advanced societies, and that States thoughtfully need to consider how to deal with this in their legal and regulatory framework. The Council of Europe is currently working with its member States on identifying the scope and the main elements of a new Council of Europe legal instrument on AI and criminal law, preferably a convention. Agreeing on common standards to clearly and properly allocate possible criminal responsibility and to clarify connected procedural issues as well as possible human rights implication needs to be a joint effort by public and private sector actors, so that the technology can develop successfully and in a way that respects our fundamental values.

Do you think that the safety legislative framework should consider new risk assessment procedures for products subject to important changes during their lifetime?

Yes	✓
No	
No opinion	

Do you have any further considerations regarding risk assessment procedures?

500 character(s) maximum

All risks related to the violation of fundamental rights should be included in the screening procedures. Risk assessment procedures should also (regularly) update to adapt to new product technologies/ AI methodologies and incorporate new risks identified via research and case law.

Do you think that the current EU legislative framework for liability (Product Liability Directive) should be amended to better cover the risks engendered by certain AI applications?

Yes	✓
No	
No opinion	

Do you have any further considerations regarding the question above?

500 character(s) maximum

The burden of proof regarding cases involving discrimination should be on the company/provider in line with ECRI policy standards and European best practices. Lacking this (in national legislation provisions) and other areas of concern unless there are specific applications, the burden of proof for victims of damage caused by AI should be facilitated.

Since Article 4 of the Directive requires victims to prove the defect of the product and the causal link between the damage and the product, the administration of proof could prove extremely complex in the presence of algorithmic systems composed of several subsystems created by different suppliers. The producer should be able to provide the elements necessary for any expert assessment (traces of connection, etc.) and to prove that a number of checks have been carried out before the product is placed on the market and then regularly during the operation of the system to demonstrate his commitment to prevent damage.

Do you think that the current national liability rules should be adapted for the operation of AI to better ensure proper compensation for damage and a fair allocation of liability?

Yes, for all AI applications	
Yes, for specific AI applications	✓
No	
No opinion	

Please specify the AI applications:

Serious damage, such as human rights violations or physical harm, should be sanctioned even in the absence of fault on the part of the producer of the system, according to a principle of strict liability. Insurance mechanisms could play an important role in order to allow rapid compensation of damages. For less serious damage, proportionate mechanisms should make it possible to modulate the requirements on producers.

Do you have any further considerations regarding the question above?
500 character(s) maximum

No further comments.