COUNCIL OF EUROPE

CONSEIL DE L'EUROPE

Strasbourg, 23 June 2021

CAHAI(2021)07

AD HOC COMMITTEE ON ARTIFICIAL INTELLIGENCE (CAHAI)

Analysis of the Multi-Stakeholder Consultation

As approved by the Consultation and Outreach Group (CAHAI-COG)
at its 3rd meeting

www.coe.int/cahai

Table of contents

INTRODUCTION

In 2020, the Ad hoc Committee on Artificial Intelligence (CAHAI) carried out a feasibility study which carefully examined the reasons why today it is necessary to have an adequate legal framework to protect human rights, democracy and the rule of law in light of the new challenges posed by artificial intelligence (AI) systems, which are being increasingly used in our daily lives and societies. In 2021, the CAHAI began a reflection on the main elements of such a framework, which will be based on the Council of Europe's standards on human rights, democracy and the rule of law.

The CAHAI decided that a multi-stakeholder consultation would take place in 2021, to gather the views of representative institutional actors (not individuals) on certain key issues arising in the development of the-above-mentioned elements.

The consultation should help the CAHAI Legal Frameworks Group, which is in charge of preparing the main elements of this future legal framework, to inform its choices based on the feedback that has been collected during the consultation. In effect, the elements are expected to address key issues such as the values and principles on which the design, development and application of AI should be based, the areas where more safeguards are needed, and the kind of policies and solutions that need to be adopted for AI systems to respect the Council of Europe's values. On these and other issues, it was considered important that the debate would be as broad as possible and would encompass the points of view of different actors: government representatives and public administrations, international organisations, business, civil society, academia and the technical community.

The consultation was open from 30 March 2021 to 9 May 2021 and was based on a questionnaire approved at the 4th CAHAI plenary meeting. The questionnaire appears as Appendix I to this document. It includes, apart from the pre-screening questions, 8 closed questions, 8 multiple choice questions, 22 Likert scale questions, 11 open questions and 8 sub open questions mostly representing the option “other”. In this timeframe, the CAHAI received 260 responses to the multi-stakeholder consultation and 1 written contribution.

The CAHAI had previously decided that transparency would be an essential principle of the consultation. It had also agreed that the various responses received during the consultation would be included in a compilation of responses and would provide the basis for the development of this specific report, which would be discussed first by the CAHAI Consultations and Outreach Group (CAHAI-COG) and then by CAHAI itself.

Both documents should be published on the CAHAI's website. The present analytical report, once reviewed by the CAHAI at its 6th plenary meeting on 5-7 July, would then be presented to the CAHAI Legal Frameworks Group as a non-binding support tool, to be used in its work of preparation of the elements of the legal framework.

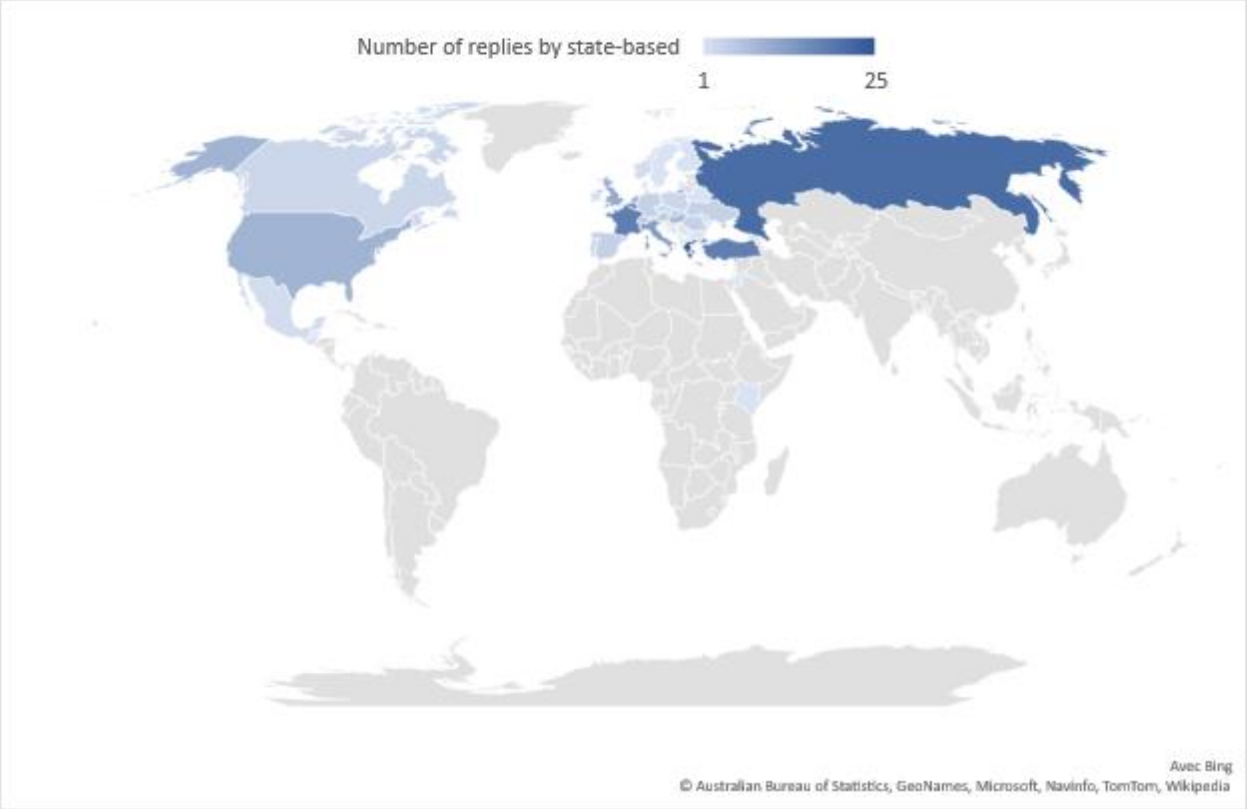
The analysis presents, in the form of graphs and response percentages, the different positions expressed by the contributors on each question of the questionnaire. This quantitative analysis is supplemented by a qualitative analysis of the answers given to open questions. Also, for this type of analysis, the approach taken has been to present as neutrally and comprehensively as possible the diversity and numerical consistency of the various positions expressed.

The analysis was prepared by the secretariat of the CAHAI with the support of Professor Marc-Antoine Dilhac, scientific expert (France) and reviewed by the CAHAI-COG at its 3rd meeting on 22-23 June 2021.

Preliminary section: Pre-screening questions result

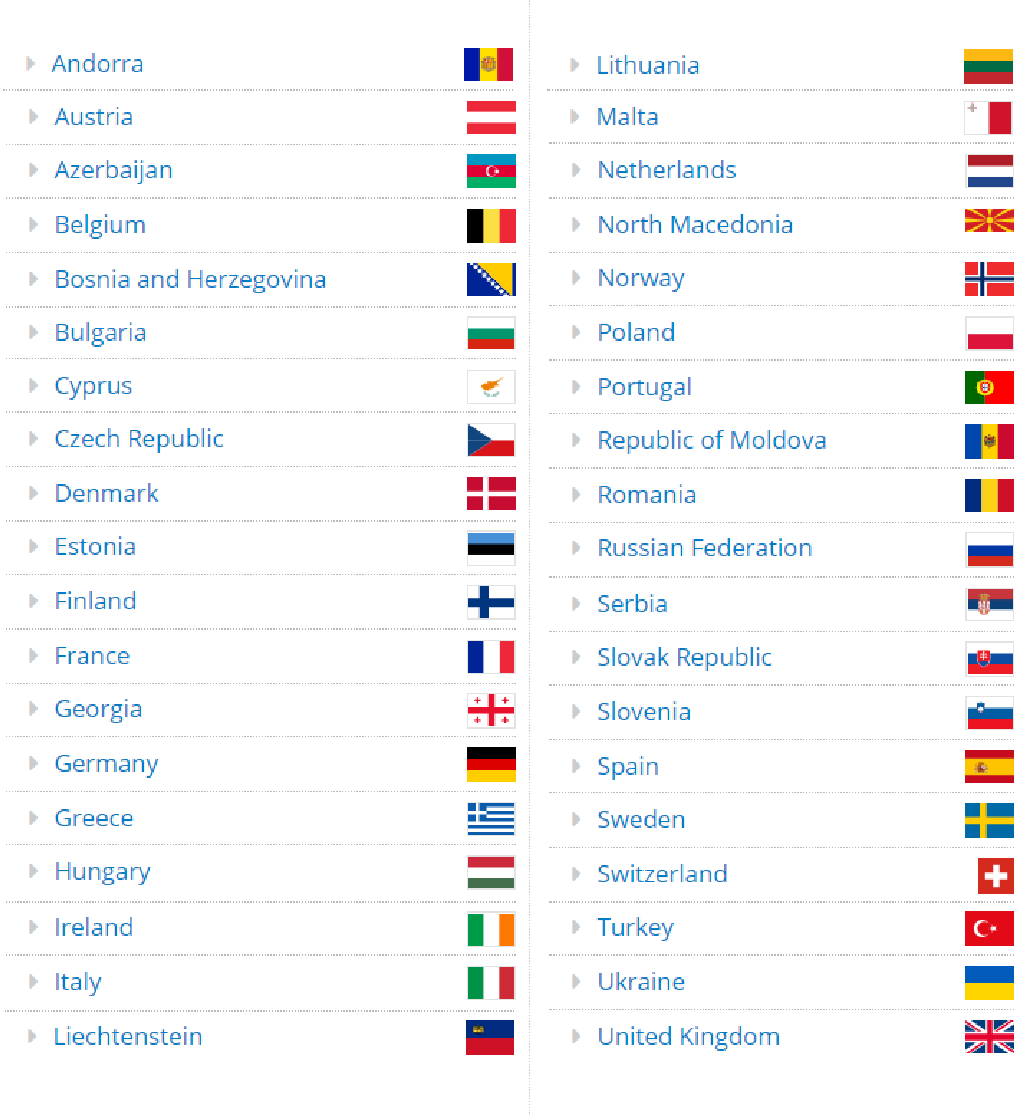
1. Pre-screening questions

Geographic representation (state-based institution/organisation)


Number of replies by state-based
1
25
Avec Bing
© Australian Bureau of Statistics, GeoNames, Microsoft, Navinfo, TomTom, Wikipedia

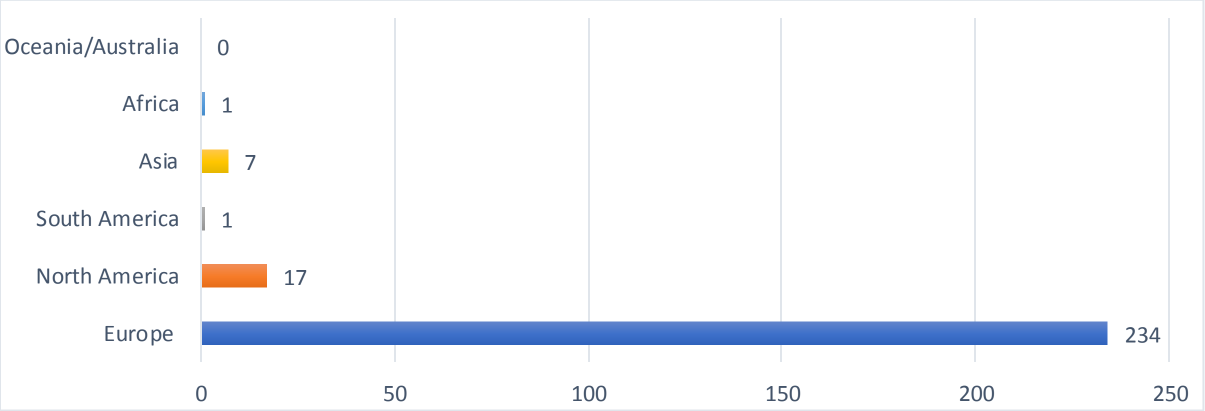
As mentioned earlier, the total number of responses received in the framework of the consultation is 260. 49 state-based respondents (in Europe and elsewhere) have contributed to the consultation. The European Commission addressed to the secretariat a letter describing its regulatory proposal on AI.

234 respondents are representing institutions or organisations based in Europe, in particular from the countries listed below.



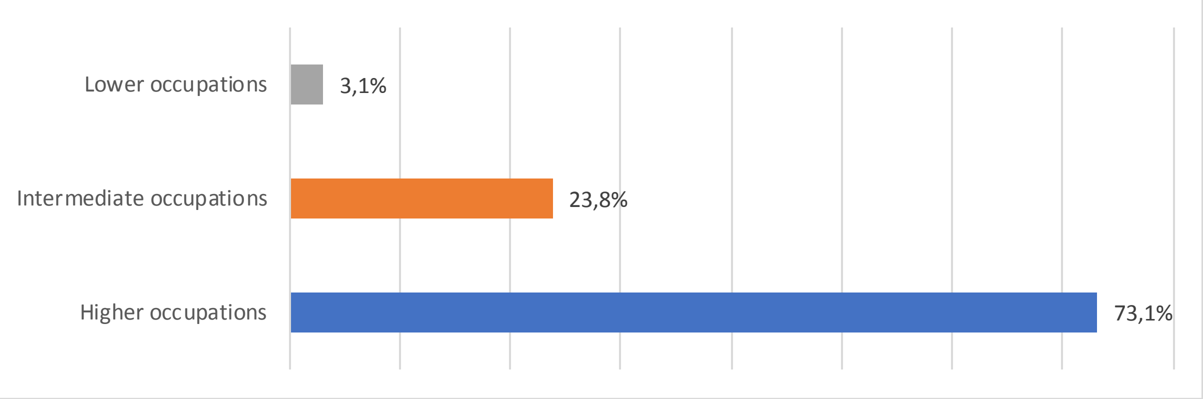
Andorra
Austria
Azerbaijan
Belgium
Bosnia and Herzegovina
Bulgaria
Cyprus
Czech Republic
Denmark
Estonia
Finland
France
Georgia
Germany
Greece
Hungary
Ireland
Italy
Liechtenstein
Lithuania
Malta
Netherlands
North Macedonia
Norway
Poland
Portugal
Republic of Moldova
Romania
Russian Federation
Serbia
Slovak Republic
Slovenia
Spain
Sweden
Switzerland
Turkey
Ukraine
United Kingdom

The consultation received submissions from respondents representing institutions or organisations from several non-European states such as Kenya, Mexico, Chile, Canada, United States of America, Guatemala, Israel, Jordan, Nicaragua and Hong Kong.



It should be specified that international inter-governmental organisations represent a wide number of states beyond European borders which are not represented in the above graph, as only state-based affiliation has been take into consideration.

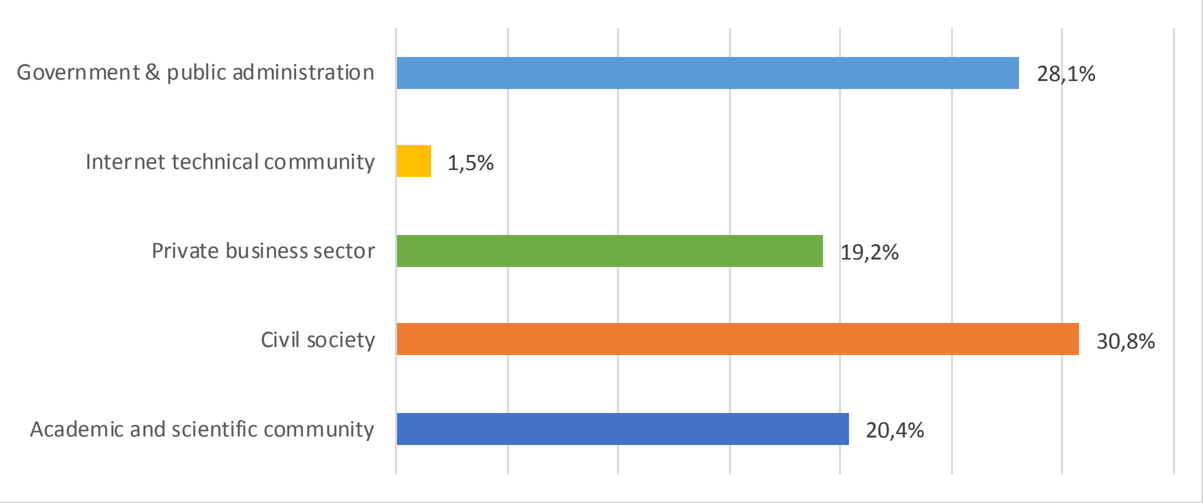
Socio-professional category representation



Respondents mostly hold a high occupation (73%) and, to a lesser extent, an intermediate occupation (24%).

3% of respondents hold a lower occupation.

Stakeholder groups representation

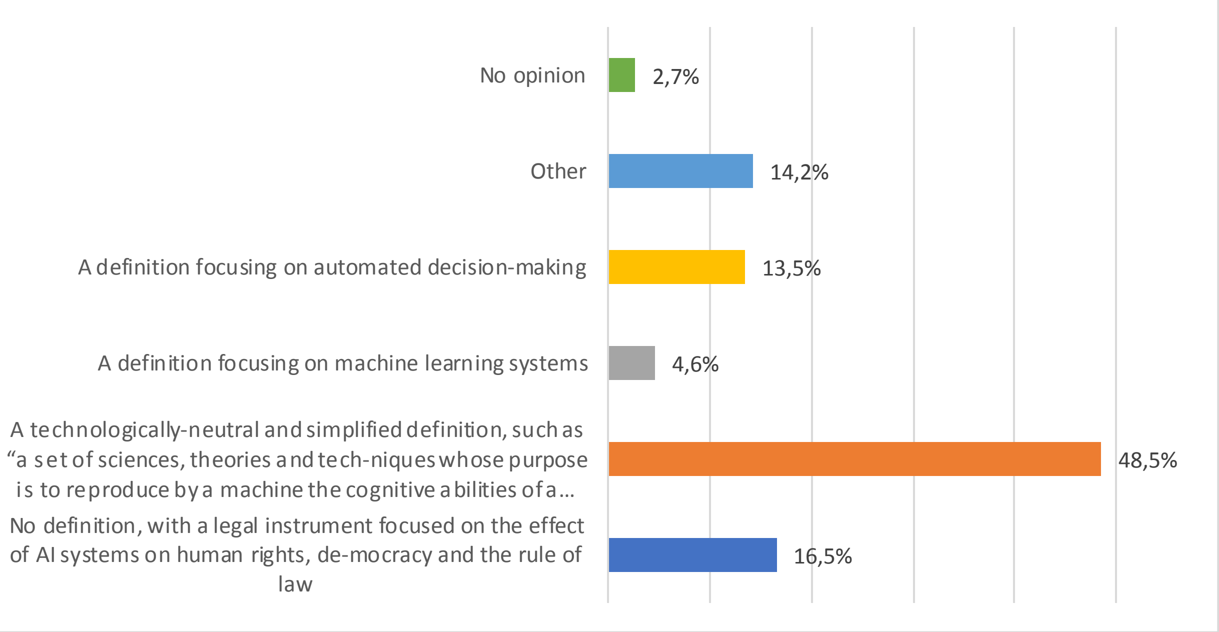


Respondents are representing rather evenly government & public administration (28%), private business sector (19%), civil society (31%), academic and scientific community (20%).

A clarification must be made regarding the representatives of the internet technical community who are apparently less represented (1,5%). However, it should be noted that the number of the most significant players in this community is quite restricted[1] and that the number of responses received for this consultation is therefore sufficiently representative of this community for the purpose of the present consultation.

Section 1: Definition of AI Systems

2. **In view of the elaboration of a legal framework on the design, development and application of AI, based on the standards of the Council of Europe on human rights, democracy and the rule of law, what kind of definition of artificial intelligence (AI) should be considered by the CAHAI**



This question is a closed question and one answer among the 5 proposals was expected.

The question was about the types of definitions of AI, not about the content of the definitions themselves. Three options were available:

[1] As regards the internet technical community, information is available at the following links: https://giswatch.org/en/economic-social-and-cultural-rights-escrs/economic-social-and-cultural-rights-and-internet-techni; Internet Technical Community Coordination — RIPE Network Coordination Centre

- A technologically neutral and simplified definition, such as “a set of sciences, theories and techniques whose purpose is to reproduce by a machine the cognitive abilities of a human being”.
- A definition focusing on machine learning systems.
- A definition focusing on automated decision-making.

Respondents also had the option to add an option ("Other") or to recommend not using a definition of AI at all ("No definition").

While a sizable number of respondents preferred that AI not be defined, the vast majority of respondents opted for a technologically neutral and simplified definition of AI. Many respondents commented on the definition presented as an example of a neutral and simplified definition of AI and many alternative definitions were suggested. Although this was not the focus of the question, these comments strengthen the consensus for the use of a technologically neutral and simplified definition of AI, even if the precise content of the definition continues to be a topic for discussion and remains to be determined.

Respondents did not really offer additional kinds of definition to consider, but a few who checked “Other” suggested using a very broad and encompassing definition to cover a wide range of AI systems, such as “computational techniques”.

3. What are the reasons for your preference?

Behind the wide variety of arguments made in favour of one kind of definition or rejection of any definition of AI, there is a very broad consensus on four normative criteria of a relevant definition of AI for a legal framework:

1. A definition of AI must allow the legal framework to apply to real-world use cases;

2. A definition of AI should not unduly narrow the scope of application of the legal framework and should ensure that it covers a wide range of harms and risks caused by computer technologies;

3. A definition of AI should not be too restrictive and should allow for a variety of computing technologies and systems to be included;

4. A definition of AI should be future-proof and not make the legal framework obsolete when the definition is outdated;

Respondents who argue that no definition of AI can meet these normative criteria, consistently chose the "No definition" option. Their argument is that AI is a fast-changing set of technologies and that a definition of AI would quickly make the legal framework obsolete (criterion 4). The remaining respondents consider that a definition of AI can satisfy these criteria and that a definition of AI that identifies some computing technologies among others is a precondition for a legal framework applicable to AI uses.

These respondents are divided about the kind of definition needed. For some respondents, a definition of AI that is too broad (e.g., the CAHAI definition) makes the legal framework vague and inapplicable (criterion 2). They then choose the definition of AI that focuses on automated decision-making. While these respondents are aware of the limitations of this definition, they argue that it allows to target the systems that have the most significant impact on human rights today. These systems, even if AI technologies evolve, are here to stay. However, a massive objection has been raised against this view: a definition focusing on automated decision-making is misleading as an AI system does not make an actual decision.

In any case for the majority of other respondents, a definition that is too specific, unduly limits the scope of the legal framework and fails to cover the risks posed by various AI technologies. A broader definition would include, among other things, automated decision-making. These respondents favoured the proposal of a technologically neutral and simplified definition. It is important to note that a technology-neutral definition does not mean that the technology is neutral – such a definition is agnostic about the neutrality of the technology. However, the reference to neutrality is confusing and polarizing and could be replaced by a less divisive term. Indeed, most respondents settled instead on the simplicity of the definition of AI that CAHAI should adopt, rather than its neutrality, the main argument being that a definition of AI should be accessible to as many people as possible, both experts and laymen.

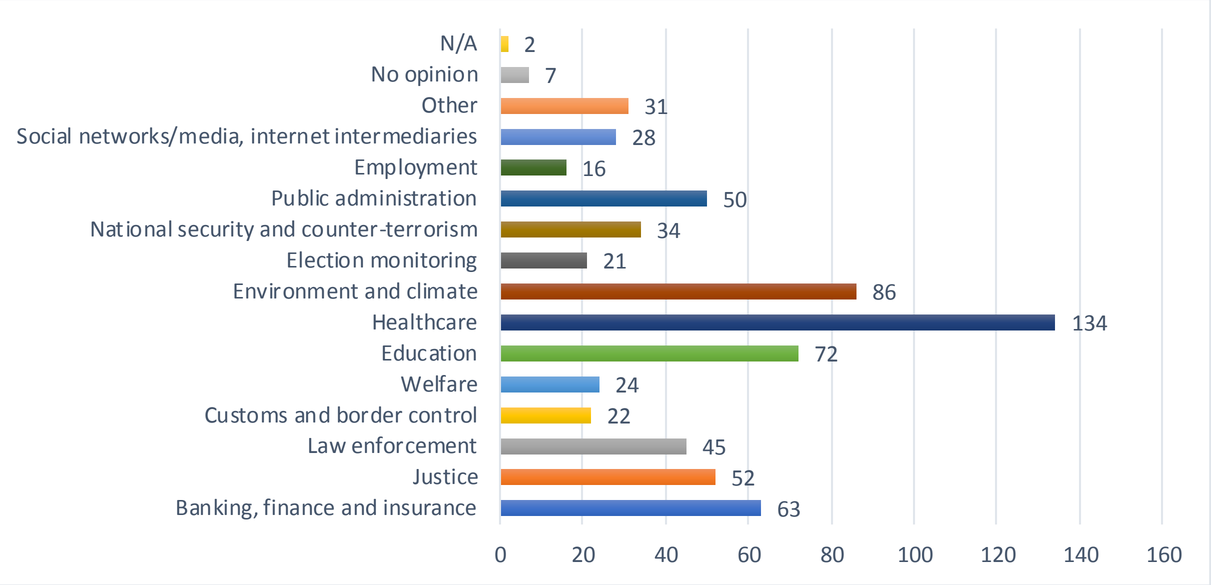
Finally, several respondents urge CAHAI to avoid anthropomorphism in the description of AI technologies. In particular, it was argued that AI does not reproduce human cognitive abilities, but at most can simulate them.

Respondents refer to four other definitions of AI that could substitute or improve on the CAHAI definition: 1. the definition proposed by the European Commission's High Level Expert Group on AI (Ethics Guidelines for Trustworthy AI, 2019); 2. the definition developed by the non-profit organisation AlgorithmWatch; 3. the definition elaborated by UNESCO's World Commission on the Ethics of Scientific Knowledge and Technology (COMEST); and finally, 4. the definition offered by the OECD Expert Group on Artificial Intelligence (AIGO).

Section 2: Opportunities and Risks arising from AI Systems

Opportunities arising from AI systems

4. Please select the areas in which AI systems offer the most promising opportunities for the protection of human rights, democracy and the rule of law



A maximum of 3 selected options on 15 options were allowed per respondent for this question.

The 3 most selected proposal areas in which AI systems offer the most promising opportunities for the protection of human rights, democracy and the rule of law are Healthcare (134), Environment and climate (86) and Education (72).

It is closely followed by Banking, finance and insurance (63), Justice (52) and Public administration (50).

Employment is the less selected by respondents (16) as an area in which AI systems offer the most promising opportunities for the protection of human rights, democracy and the rule of law.

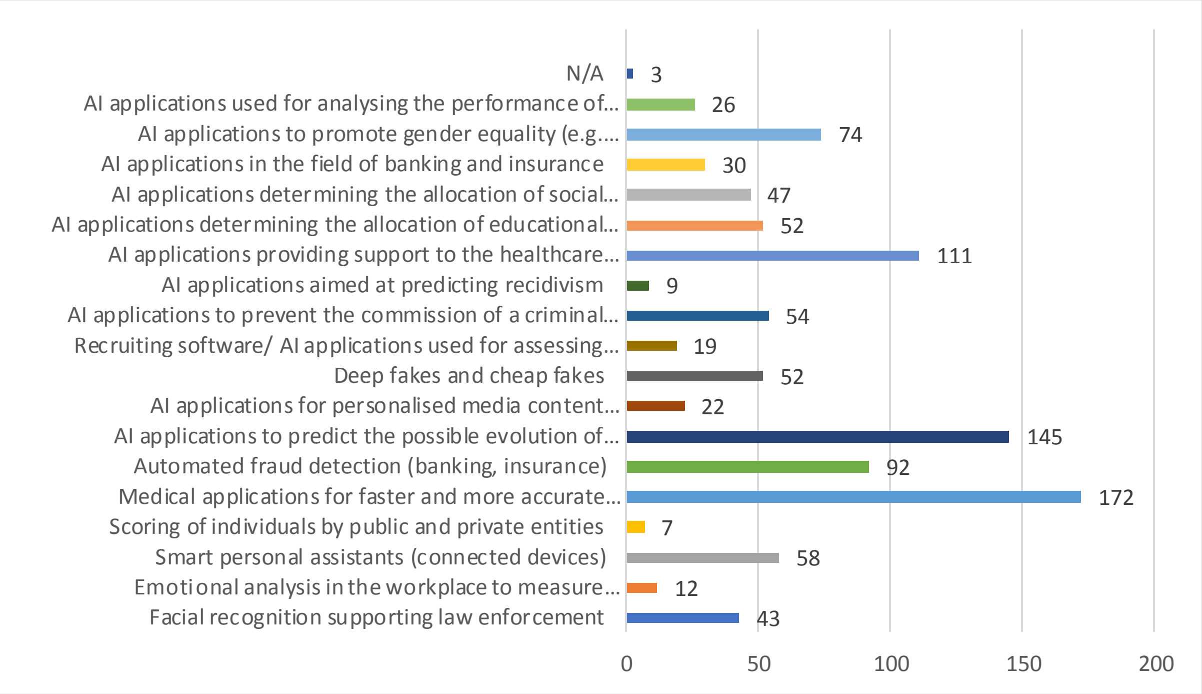
7 respondents have selected the option "no opinion" and 31 have selected the option "other" and provided another proposal and explained why.

Many respondents argued that this question is difficult to answer as the same AI system can have a positive or negative impact regardless of the application area; and in all areas, there is a wide variety of AI systems, some producing beneficial results, others producing results detrimental to human rights. For instance, in the field of law enforcement, AI systems can help prevent crimes by predicting when and where they are more likely to occur, but it can also reinforce discrimination against people living in the neighbourhood where these crimes are more likely to occur. It all depends on the goals being pursued, the application context, and the ability to use the technology properly.

With this caveat in mind, some areas seem more promising than others, and regardless of the sector, there is broad agreement among respondents that AI development offers more opportunities or less risks for the protection of human rights, democracy and the rule of law in the areas of healthcare, environment and education. However, representatives of state entities tend to choose "public administration" or "finance" (banking, insurance) over "environment". There are reasons for this choice, which will made clear in the response to the next question.

Some respondents suggested adding scientific discovery and journalism (or information) to the list.

5. **Please indicate which of the following AI system applications in your view have the greatest potential to enhance/protect human rights, democracy and the rule of law?**



A maximum of 5 selected options on 18 options were allowed per respondent for this question.

The 5 most selected AI system applications which have the greatest potential to enhance/protect human rights, democracy and the rule of law are:

1. Medical applications for faster and more accurate diagnoses (172);
2. AI applications to predict the possible evolution of climate change and/or natural disasters (145);
3. AI applications providing support to the healthcare system (triage, treatment delivery) (111);
4. Automated fraud detection (banking, insurance) (92);
5. AI applications to promote gender equality (e.g. analytical tools) (74).

The 5 less selected AI system applications which have the greatest potential to enhance/protect human rights, democracy and the rule of law are:

1. Scoring of individuals by public and private entities (7);
2. AI applications aimed at predicting recidivism (9);
3. Emotional analysis in the workplace to measure employees' level of engagement (12);
4. Recruiting software/ AI applications used for assessing work performance (19);

5. AI applications for personalised media content (recommender systems) (22).

Finally, one respondent has provided a different answer which does not appear on the list of choices.

Consistent with their response to the previous question, respondents favour applications in areas of healthcare, environment and finance (banking and insurance). If education does not rank high, this is probably due to the fact that the only AI system listed in education was more of a public administration tool than an educational tool.

6. Please briefly explain how such applications would benefit human rights, democracy and the rule of law.

It should be noted at the outset that while many respondents identified AI in health, the environment or finance as beneficial, few showed how these AI applications specifically promote human rights, democracy and the rule of law. This seemingly anecdotal finding could reveal a difficulty amongst the respondents, to distinguish between AI for good (or AI for humanity, promoted in many ethical charters which, until very recently, have steered the debate on AI governance) and AI for human rights, democracy and the rule of law. For example, an application that can predict an imminent natural disaster and evacuate threatened populations promotes the good, but the connection with human rights, democracy and the rule of law is neither direct nor obvious in itself.

From the large number of arguments put forward, it appears that AI applications can contribute to the protection of human rights, democracy and the rule of law in two ways:

1. By combating practices that undermine the exercise of procedural and civil rights such as the right to fair trial and non-discriminatory treatment, or the right to security and freedom of speech;

2. By improving access to public services that condition the enjoyment of substantive and social rights such as the right to education.

AI applications in health and education more often fall into the second category of applications that improve access to essential public services.

Medical applications that enable faster and more accurate diagnosis were identified by respondents as being among the most beneficial to human rights, democracy and the rule of law. Indeed, in addition to the strictly medical capabilities of AI, many respondents stressed that the use of AI will facilitate access to healthcare for the greatest number of people by reducing the costs of diagnosis, therapeutic treatment and medical tracing, and by freeing up human resources to improve the care of patients and treat them with dignity. This would better protect the fundamental right to health and the right to a healthy life, a right recognised by various charters and declarations such as the Universal Declaration of Human Rights, the Constitution of the World Health Organisation, or the International Covenant on Economic, Social and Cultural Rights and the European Social Charter.

AI applications can also help to promote the right to education by personalising learning and thus combating the phenomena of drop-out and attrition, and by making educational services accessible to vulnerable, marginalised or isolated populations (rural communities, for example). This right to education is recognised by the European Convention on Human Rights.

Respondents also argued that AI systems, in particular smart personal assistants, can improve access to administrative services by making information more available, facilitating and speeding up administrative procedures.

These AI applications also help fight discriminatory practices as they improve accessibility to public services and social good for everyone. However, this can only be the case, many participants noted, if these algorithmic systems are overseen and regulated and if the data that fuels them are unbiased. On the assumption that AI would, or could, be less biased than humans, most AI applications on the list then support human rights, democracy and the rule of law. Crime prevention applications are seen by many respondents as beneficial provided they are not biased and are used correctly. However, answers to question 6 are in conflict with the answers to the previous questions, since these applications belong to application areas (law enforcement and justice) that were considered the least promising for human rights, democracy and the rule of law - the answers to the subsequent questions will show that they are also deemed the most threatening ones.

To summarise the answers to this question, the beneficial use of AI for human rights, democracy and the rule of law hinges on two conditions:

- Computational methods to reduce biases of AI systems or ensure that they do not replicate human biases;
- A legal framework to prevent misuse of AI.

7. What other applications might contribute significantly to strengthening human rights, democracy and the rule of law?

Respondents mainly commented on their previous arguments about the benefits of AI and many reiterated their interest in the analytical and detection capabilities of AI, including tracking fraud, scams, cyber-attacks, copyright infringements, hate speech, fake news and emotional manipulation.

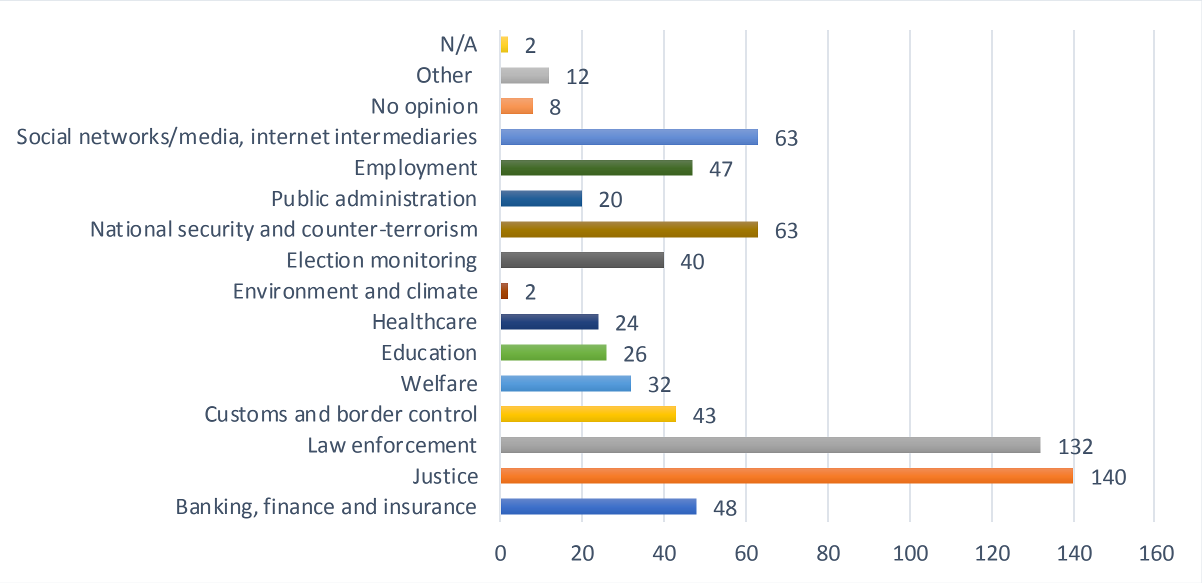
In addition, most of the proposals duplicate the applications listed in the previous question or do not demonstrate the benefits for human rights, democracy and the rule of law. Among the relevant proposals, we can highlight the following:

- Application to help disabled people to gain autonomy;
- Advanced analytics for investigative journalism;
- Automated translation enabling speakers of minority languages to participate more actively in public debate, deliberation or decision making.
- Advanced analytics for historical and cultural investigations. Such AI applications contribute to the protection of human rights in two ways: they enhance mutual understanding and promote peaceful relations between cultural communities; they help people to reconnect with or better understand their own social identity, which is a basic condition for self-esteem and social autonomy.

The last two proposals reveal a blind spot which should be covered by the legal framework: the area of culture.

Impact on human rights, democracy and the rule of law

8. Please select the areas in which the deployment of AI systems poses the highest risk of violating human rights, democracy and the rule of law



A maximum of 3 selected options on 15 options were allowed per respondent for this question.

The 3 most selected proposal areas in which the deployment of AI systems poses the **highest risk** of violating human rights, democracy and the rule of law are Justice (140), Law enforcement (132), National security and counter-terrorism (63) and Social networks/media, internet intermediaries (63).

It is followed by, Banking, finance and Insurance (48) and Employment (47).

Environment and climate option is the less selected one by respondents (2) as an area in which the deployment of AI systems poses the highest risk of violating human rights, democracy and the rule of law.

Finally, 8 respondents have selected the option “no opinion” and 12 have selected the option “other” and provided another proposal.

While “health care”, “environment” and “education” were the most promising areas for the protection of human rights, democracy and the rule of law, they are now, unsurprisingly, among the areas of least concern. On the contrary, "justice" (20%), "law enforcement" (19%) and "national security and counter-terrorism" along with "social networks" (9%) are considered very risky for human rights, democracy and the rule of law.

9. Please briefly explain how such applications might violate human rights, democracy and the rule of law.

Unsurprisingly, this question received the most comments, as it is very relevant for the development of a future legal framework on AI based on human rights, democracy and the rule of law standards. While some respondents reiterate that AI can be both harmful and beneficial depending on its use in all areas of the list, it is clear to the vast majority of respondents that some sectors present greater risks for the protection of human rights, democracy and the rule of law than others: justice, police, national security. The identification of these areas is guided by clear and sound reasoning:

1. The first three areas belong to the reserved domain of the state; these are the areas in which the state exercises coercion to protect rights, preserve the conditions of democracy and ensure the rule of law for all. As some respondents note, these are also the areas where the power imbalance is such that citizens are most vulnerable and subject to the power of a third party; those who deploy AI applications in these areas have increased power over those who are governed by these applications.

2. Furthermore, although AI applications can be misused and cause harm in all the areas under consideration, the risks and nature of the harm are not the same: in the administration of people and society, the risks of error and bias are considered to be much greater than in the case of predicting a climate disaster or the evolution of a tumour. Moreover, the harms

caused by AI applications are not of the same nature, even though their effects may be more dramatic in healthcare or the environment. In the case of a wrong diagnosis, it cannot be argued that the right to health has been directly violated. On the other hand, in the case of an application that may result in a person being sentenced to person to prison and which fails to meet transparency requirements or to provide an explanation, his or her right to a fair trial is directly violated.

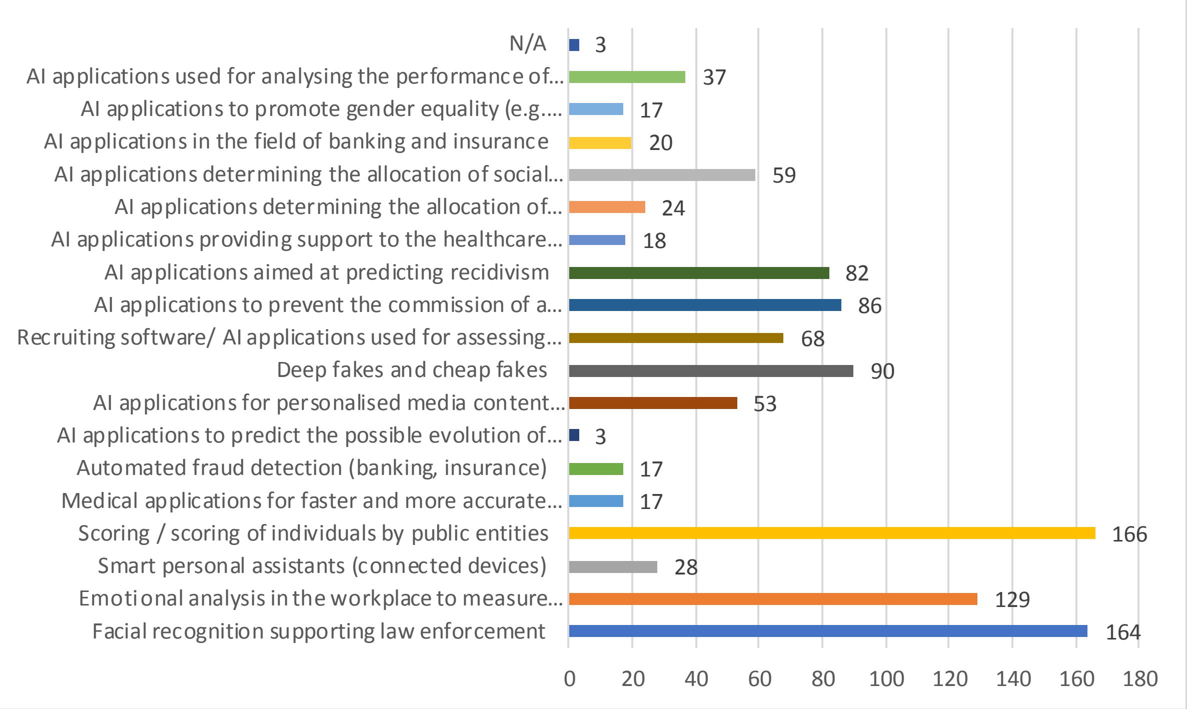
3. Finally, the deployment of AI systems in the justice, law enforcement (including border control) and national security areas are not just tools but modes of delivering justice and governing.

After stating the reasons these areas are riskier, the respondents identify the harms that people can suffer and the causes of these harms. The causes are mainly errors in algorithmic processing, its inaccuracy and bias, and poor data quality. The main harms are the following:

- Judicial uncertainty for citizens, the inability to defend their rights, and the denial of fair trial and due process
- Unequal treatment and discrimination (racial, religious, socio-economic, sexual and gender related) leading to the exclusion from social welfare;
- Violation of psychological integrity and well-being, including through surveillance, manipulation of opinions and emotions;

Most respondents anticipate the next two questions on the types of application most risky for the protection of human rights, democracy and the rule of law. We therefore defer to the analysis of question 11 to examine their arguments about the types of AI applications.

10. Please indicate the types of AI systems that represent the greatest risk to human rights, democracy and the rule of law



A maximum of 5 selected options on 18 options were allowed per respondent for this question.

The 5 most selected types of AI systems that represent the greatest risk to human rights, democracy and the rule of law are:

1. Scoring of individuals by public and private entities (166);
2. Facial recognition supporting law enforcement (164) and;
3. Emotional analysis in the workplace to measure employees’ level of engagement (129);
4. Deep fakes and cheap fakes (90);
5. AI applications to prevent the commission of a criminal offence (86).

The 4 less selected types of AI systems that represent the greatest risk to human rights, democracy and the rule of law are:

1. AI applications to predict the possible evolution of climate change and/or natural disasters (3);
2. Evenly, medical applications for faster and more accurate diagnoses (17) and AI applications to promote gender equality (e.g. analytical tools) (17) and Automated fraud detection (banking, insurance) (17)
3. AI applications providing support to the healthcare system (triage, treatment delivery) (18);

4. AI applications in the field of banking and insurance (20).

11. Please briefly explain how such applications might violate human rights, democracy and the rule of law.

The respondents' choices are generally consistent with their previous answers. AI systems in the areas of justice and law enforcement come out on top, along with applications in the media, but also in social welfare, an area that was not flagged as risky. In their comments, respondents focus on three types of applications of concern:

1. Facial recognition;

2. predictive justice applications;

3. social rating applications.

In all cases, respondents raise the cross-cutting issue of bias and discrimination: for the majority of respondents, facial recognition is inherently biased and is much less accurate in identifying the faces of darker-skinned people than, for example, white people. Both predictive justice and social scoring applications are prone to reproducing and reinforcing the discrimination that disadvantaged, stigmatised and marginalised populations already face because these applications are powered by historical data that reflect the existing biases against these populations. Finally, emotional analysis applications are not able to correctly interpret the facial expressions of people from cultures less well represented in the data; their results tend to discriminate against these people, but also against people who generally behave in ways that are out of step with dominant social norms. Beyond these four types of application, the issue of discrimination is raised for all the other applications.

Facial recognition technologies receive the most comments, which clearly indicates that they are a test case for the development of the legal framework for the use of AI. While the risks of mistakes leading to prosecutions and unjustified preventive incarceration are clearly identified, it is mass surveillance that is the main concern of respondents. As the issue of surveillance is not as cross-cutting as discrimination, the number of occurrences of the term surveillance is impressive. But it should be noted that there is a significant imbalance between stakeholders in the different sectors: surveillance is mainly a concern of civil society organisations with almost two hundred occurrences of the term, compared to barely ten for the private sector, and just under twenty for the public sector. These data highlight the wide gap between the expectations of civil society and those of the private and public sectors with regard to the legal framework.

According to respondents, AI mass surveillance systems using facial recognition infringe on the fundamental right to privacy and put everyone, but particularly people of colour who are misidentified by facial recognition systems, at risk of arbitrary arrest. But as several respondents from different sectors noted, even when surveillance does not lead to repressive intervention, it inflicts psychological harm and reduces the agency of those who feel threatened, i.e. the ability to act freely in the enjoyment of their guaranteed rights. Mass surveillance creates a "chilling effect on civil society and activism" and undermines the exercise of fundamental freedoms such as freedom of expression, association or protest.

Respondents are also concerned about the use of AI in the judicial system with applications aimed at predicting recidivism and helping judges to make decisions. Some respondents feel that while AI can reduce the processing time of traffic tickets and minor disputes, the lack of accuracy and reliability of these systems generates errors (false positives and false negatives) that are difficult for citizens to challenge. The lack of explainability and transparency, which makes these systems black boxes for citizens and judges alike, increases judicial insecurity. This seriously undermines the rights of the defence, the adversarial principle and therefore the fundamental right to a fair trial.

Social scoring applications raise the same issues:

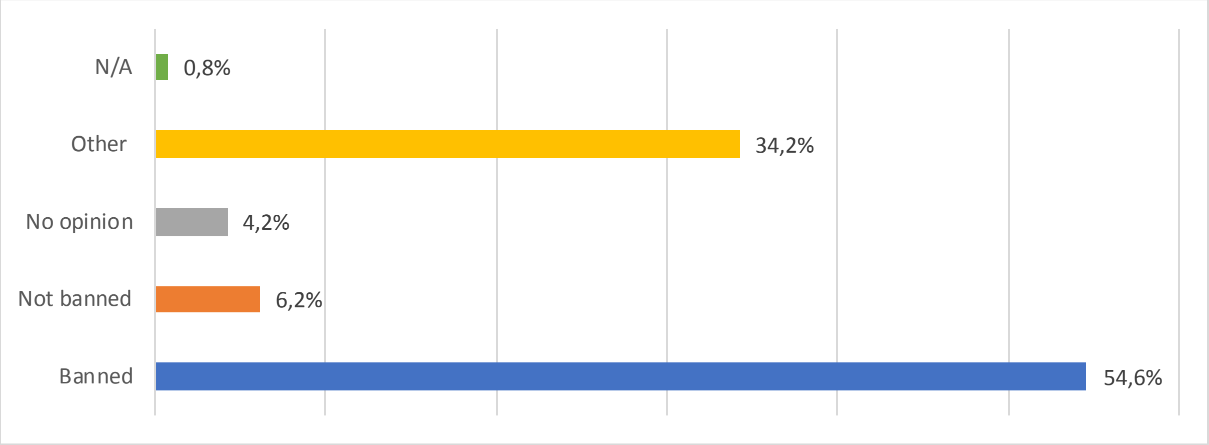
- they are not free of errors and biases
- they aggravate the situation of those who have a low score by potentially depriving them of essential services; depending on the analysis provided, the identification of the most disadvantaged differs: the young, the poor, racial minorities, etc. An intersectional analysis would show that discrimination is multifactorial and that the most disadvantaged are those who experience several discriminations as a result of belonging to several groups.
- they then increase lasting inequalities;
- and finally, by failing to provide valid explanations, they deprive people of recourse against poor evaluation and low scoring.

To complete the picture, since this overview only looks at the applications that have elicited the most comments or disapproval, we must add a word about the applications for emotional analysis. These applications come third in the selection of the most damaging applications, but curiously they are little commented on. Most of the arguments mentioned above apply to these applications: bias, error and discrimination; surveillance and breach of privacy, which must be protected even in the workplace; and, more specifically, they can harm the career advancement of employees subject to these management methods.

12. What other applications might represent a significant risk to human rights, democracy and the rule of law?

Some respondents suggest micro-targeting applications designed to influence political opinions and the course of election campaigns; others mention cyber-attack applications, but these applications are de facto unlawful. The majority of respondents suggest adding LAWS, lethal autonomous weapons, deployed by the military and potentially in use by the law enforcement. However, it should be recalled that military uses do not fall within the competence of the Council of Europe.

13. In your opinion, should the development, deployment and use of AI systems that have been proven to violate human rights or undermine democracy or the rule of law be:



Most of the respondents consider that the development, deployment and use of AI systems that have been proven to violate human rights or undermine democracy or the rule of law should be banned (55%).

Only 6% of respondents consider that it shouldn't be banned and 4% have no opinion on this question.

Finally, 34% of respondents have selected the option "other" and provided another proposal.

Although the absolute majority of responses supported a ban on AI systems that have been proven to violate human rights or undermine democracy or the rule of law, a significant proportion of respondents chose the option "other". Unfortunately, too few respondents provided an explanation for this choice or offered an alternative. Nevertheless, a robust alternative to a simple ban is proposed, based on an incremental approach:

1. AI systems designed in such a way that they infringe human rights should be banned. Their purpose invalidates them.

2. AI systems that unintentionally or indirectly violate human rights must be corrected. Precautionary measures should be taken and redress mechanisms put in place to stop the human rights abuse pending adjustment of the application.

3. Only if the developers or deployers fail to correct it (whether or not they intend to) should the system be removed.

4. Finally, exceptions to the general prohibition should be made for certain uses or compelling reasons.

Furthermore, any general ban should be justified by an intolerable and permanent violation of human rights, democracy and the rule of law. It is therefore crucial to have a well-defined list of prohibited AI applications that is not subject to interpretation.

Questions 14, 15, 16

The next three questions form a single block. They aim to assess the respondents' support for or rejection of regulatory measures, according to the levels of risk posed by AI systems. Three levels of risk are proposed for consideration:

- high risks of violation with high probability of occurrence
- low risks of violation with high probability of occurrence
- high risks of violation with low probability of occurrence

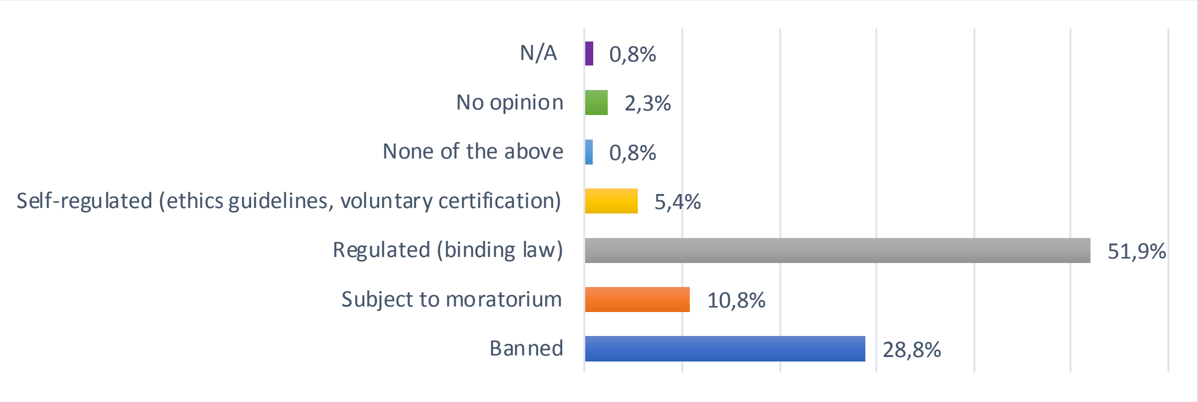
For applications presenting high risks of violating human rights, democracy and the rule of law, with a high probability of occurrence (Q14), respondents overwhelmingly (52%) support regulation by means of binding laws, and a significant proportion (29%) even support their prohibition or the adoption of a moratorium (11%). Self-regulation does not appear to be a credible option in this case (5%).

Although not majority, self-regulation turns out to be a more relevant option for 28% of respondents in the case of low-risk applications with a high probability of occurrence (Q15). We can extrapolate that this percentage would increase if the question was about low risk applications with a low probability. As this case is not problematic, the question was not asked. However, support for binding regulation remains predominant with 58% of responses.

Finally, respondents again massively support binding regulation for high-risk applications with a low probability (Q16). In this case, the choice of self-regulation drops to 14% and that of a moratorium rises to 14%.

The answers are very consistent and give clear indications for the establishment of a legal framework. The main conclusion to be drawn is that the respondents support the establishment of binding regulation for AI applications that present risks of violating human rights, democracy and the rule of law, regardless of the level of risk. Self-regulation does not appear to be a credible approach to addressing these risks.

14. In your opinion, should the development, deployment and use of AI systems that pose *high risks*[2] *with high probability*[3] to human rights, democracy and the rule of law be:



Most of the respondents consider that the development, deployment and use of AI systems that pose *high risks with high probability* to human rights, democracy and the rule of law should be regulated by binding law (52%).

29% consider that it should be banned and 11% that it should be subject to moratorium.

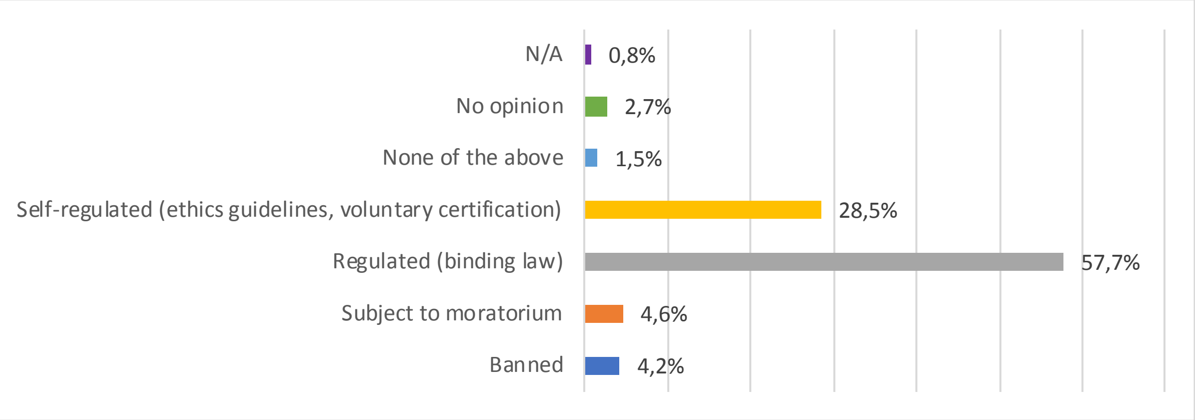
Only 5% respondents consider that it should be self-regulated by soft law instruments such as ethics guidelines or voluntary certification.

Less than 1% of respondents have selected the option “none of the above” and 2% of respondents have selected “no opinion” on this question.

[2] high negative impact on human rights, democracy and rule of law

[3] high probability of occurrence of these risks

15. In your opinion, should the development, deployment and use of AI systems that pose *low risks*[4] *with high probability*[5] to human rights, democracy and the rule of law be:



Most of the respondents consider that the development, deployment and use of AI systems that pose *low risks with high probability* to human rights, democracy and the rule of law should be regulated by binding law (58%).

28% consider that it should be self-regulated by soft law instruments such as ethics guidelines or voluntary certification.

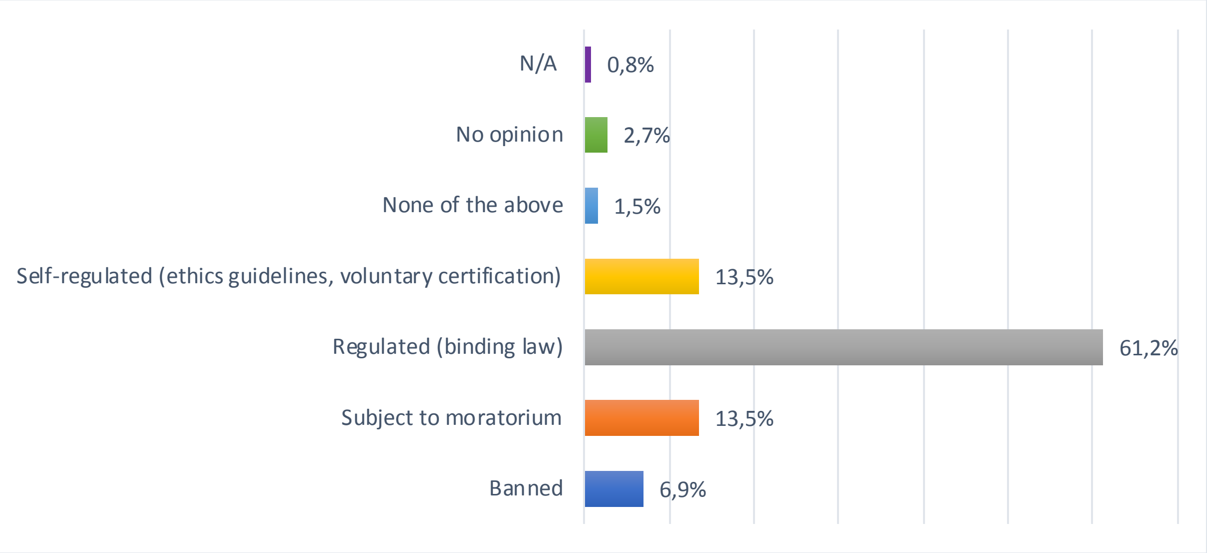
Only 5% of respondents consider that it should be subject to moratorium and 4% of respondents consider that it should be banned.

2% of respondents have selected the option “none of the above” and 3% have no opinion on this question.

[4] Low negative impact on human rights, democracy and rule of law

[5] high probability of occurrence of these risks

16. In your opinion, should the development, deployment and use of AI systems that pose *high risks*[6] *with low probability*[7] to human rights, democracy and the rule of law be:



Most of the respondents consider that the development, deployment and use of AI systems that pose high risks with low probability to human rights, democracy and the rule of law should be regulated by binding law (61%).

13% of respondents consider that it should be self-regulated by soft law instruments such as ethics guidelines or voluntary certification and the same number that it should be subject to moratorium.

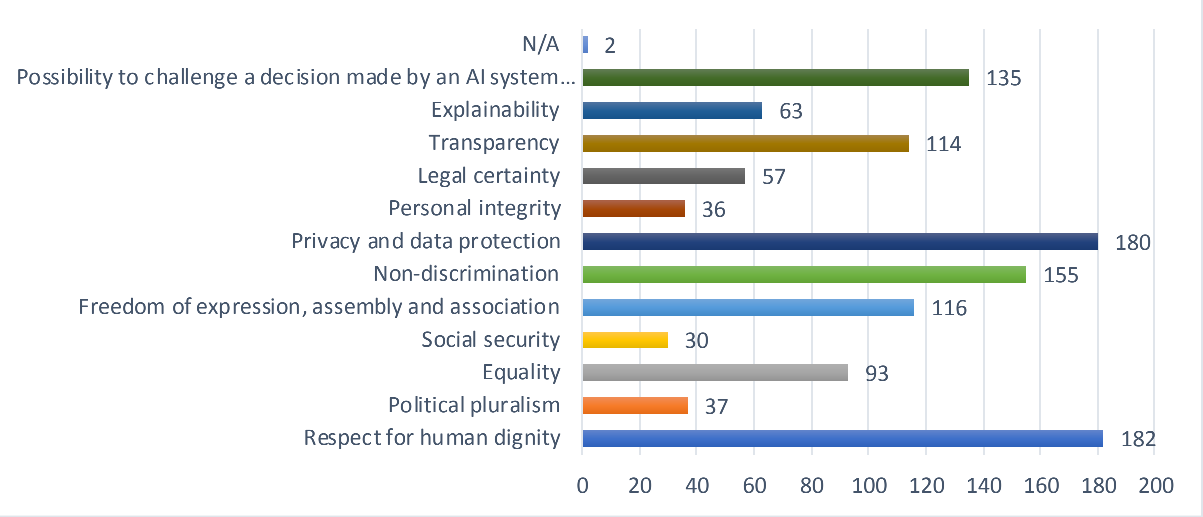
Only 7% of respondents consider that it should be banned.

2% of respondents have selected the option “none of the above” and 3% of have no opinion on this question.

[6] high negative impact on human rights, democracy and rule of law

[7] Low probability of occurrence of these risks

17. What are the most important legal principles, rights and interests that need to be addressed and therefore justify regulating the development, deployment and use of AI systems?



A maximum of 5 selected options on 12 options were allowed per respondent for this question.

The 5 most selected legal principles, rights and interests that need to be addressed and therefore justify regulating the development, deployment and use of AI systems are:

1. Respect for human dignity (182)
2. Privacy and data protection (180)
3. Non-discrimination (155)
4. Possibility to challenge a decision made by an AI system and access to an effective remedy (135)
5. Transparency (114)

The 3 less selected legal principles, rights and interests that need to be addressed and therefore justify regulating the development, deployment and use of AI systems are:

1. Social security (30)
2. Personal integrity (36)
3. Political pluralism (37)

Among the principles proposed, some are general ethical, political and legal principles, others apply more specifically to data and AI governance. The principle of respect for human dignity is the most popular, with 182 respondents choosing it. This overarching principle of international and European law is enshrined in the preamble and Article 1 of the 1948 Universal Declaration of Human Rights and in the preamble of the European Convention on

Human Rights. Although subject to differences in interpretation, the highest number of replies suggests that this principle should take a central place in a new legal framework on AI based on human rights, rule of law and democracy.

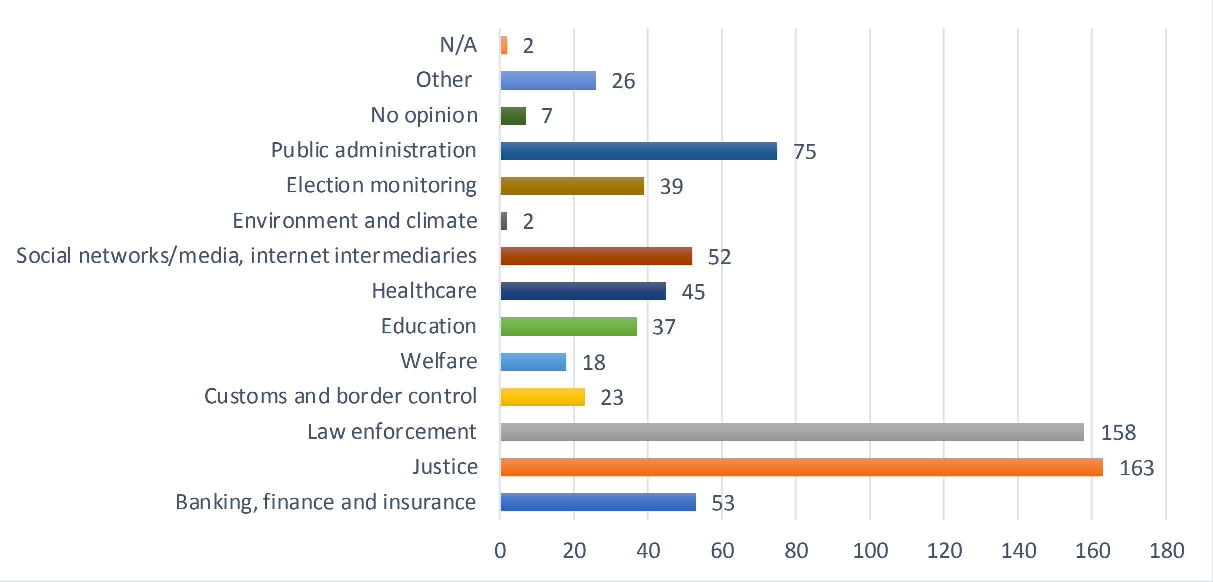
Furthermore, the recognition of equal human dignity grounds the principle of equal treatment or non-discrimination that 155 respondents chose. This result would probably have been more important if the principle of non-discrimination had been combined with that of equality.

Recognition of equal dignity also implies respect for privacy, the second most popular principle: a notion that emerged in modern times to protect people from state interference, then in the nineteenth century from press intrusion, and now in the present day from social networking. Privacy is a condition of physical and psychological well-being and integrity.

The fourth general principle recognizes the right of individuals to challenge a decision that affects them. This procedural principle underpins the rule of law, and therefore democracy too. The recognition of this principle as a basis for an AI regulatory framework has very important implications: the right to challenge is conditional on the right to access relevant information, to obtain a meaningful explanation, to be heard by human judges, to demand a human review, to obtain compensation or redress when applicable. In the rest of the questionnaire, respondents offer answers consistent with this choice.

The fifth choice is that of transparency. Transparency is not always easy to interpret and respondents who refer to it a lot rarely explain it. Transparency is the democratic principle that requires public persons to (1) make public relevant information about the exercise of their office, (2) explain their choices or decisions since by definition they affect the public. Applied to AI, the principle means that public institutions and private actors make available information on the use of algorithms and explain the decisions they take on the basis of algorithmic recommendations.

18. In your opinion, in what sectors/areas is a binding legal instrument needed to protect human rights, democracy and the rule of law?



A maximum of 3 selected options on 13 options were allowed per respondent for this question.

The 3 most selected sectors/areas in what a binding legal instrument is needed to protect human rights, democracy and the rule of law are Justice (163), Law enforcement (158) and Public administration (75).

It is followed by Banking, finance and insurance (53), Social networks/media, internet intermediaries (52) and Healthcare (45).

Environment and climate option is the less selected one by respondents (2) as a sector/area in what a binding legal instrument is needed to protect human rights, democracy and the rule of law.

Finally, 7 respondents have selected the option “no opinion” and 26 have selected the option “other” and provided another proposal.

The results echo the previous responses, but some respondents argue that all areas should be subject to binding regulation as fundamental rights are at stake in all of them, although in some areas regulation should be more stringent because of the higher risk of human rights violations.

Section 3: Potential Gaps in Existing Binding Legal Instruments Applicable to AI

In the following section, respondents have been asked to indicate to what extent you agree or disagree with the following statements or if you have no opinion on a given issue.

Key findings

The next section aims to assess the need for a new legal instrument for the development of AI. More specifically, the question is whether binding regulation of AI is the most effective way to prevent and mitigate the risk of violations of human rights, democracy and the rule of law, or whether other regulatory approaches could be implemented with more concrete results. For instance, an approach based on self-regulation of companies could be both efficient and sufficient. Companies have the capacity to implement a variety of self-regulatory instruments such as voluntary certification or ethics guidelines, the latter being preferred by 31,5% of respondents as shown in the results of Q21. Respondents mention a few ethical frameworks: the *Ethics Guidelines for Trustworthy AI* (2019) by the High-Level Expert Group on AI (European Commission) is the most cited, together with the UNESCO *Draft Recommendation on the Ethics of AI* (2020) and the *European Ethical Charter on the use of artificial intelligence in judicial systems* elaborated by the European Commission for the efficiency of justice (CEPEJ), followed by *Unboxing artificial intelligence: 10 steps to protect human rights* (Council of Europe, 2019) and the *Montreal Declaration for a Responsible Development of AI* (2018). Two other documents are mentioned once: *Rome Call For AI Ethics: A Human-Centric Artificial Intelligence* (2020), and *Towards Trustworthy AI: Malta Ethical AI Framework for Public Consultation* (2019).

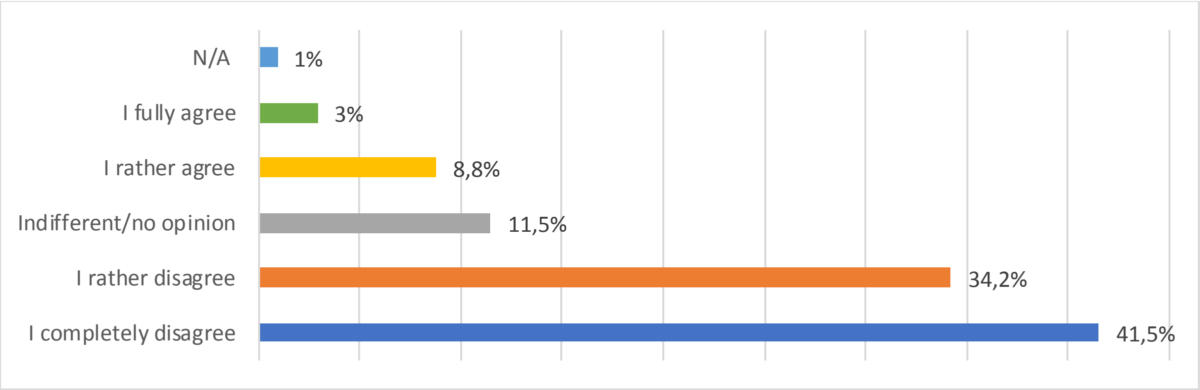
Other self-regulatory instruments are then proposed such as code of conducts (e.g. *Code of Data Ethics*, Russia), standards for the industry (e.g. IEEE P7000), assessment lists (e.g. *Assessment List for Trustworthy Artificial Intelligence* by the High-level- expert-group on AI, European Commission) and principles for good governance both public and private (e.g. the OECD *Recommendation of the Council on Artificial Intelligence*, 2019). But a majority of stakeholders in the different sectors, 75,5% and 85% respectively, believe that self-regulation is neither more efficient than government regulation (Q19), nor sufficient (Q20) to prevent and mitigate the risk of violations of human rights, democracy and the rule of law. And some respondents explicitly deplore that the European Ethical Charter on the use of artificial intelligence in judicial systems, the OECD *Recommendation of the Council on Artificial Intelligence* or *Montreal Declaration for a Responsible Development of AI* are not binding.

As a matter of fact, as shown in the results of Q22, even existing binding legal instruments are not sufficient to regulate AI systems for 74% of respondents. While many respondents from different continents and legal cultures praise the European General Data Protection Regulation (EU GDPR) as the most effective binding legal instrument and propose it as a model for the regulation of AI, they do not think the existing legal instruments either apply specifically to AI or prevent effectively violation of human rights, democracy and the rule of law (Q24). Examples of existing regulatory instruments provided in response to Q23 fall in one category or the other.

For instance, the Universal Declaration of human rights, the Charter of Fundamental Rights of the European Union or the European Convention on Human Rights are binding legal instruments but do not apply specifically to AI, if at all. In contrast, the White Paper on Artificial Intelligence (European Commission, 2020) or the Model Framework on AI Governance (Singapore, 2019) specifically address AI governance but are not binding legal instruments.

That said, the EU GDPR and Convention 108 (Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data, 1981) are proposed as the closest legal instruments to a binding regulatory framework for AI. Convention 108+ (the modernised Convention 108 with the Amending Protocol) is even closer as it incorporates considerations on algorithmic decision-making processes. Respondents also draw the attention on existing legal instruments outside Europe, such as US FDA framework as applied to algorithms (Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device), US Fair Credit Reporting Act, which applies to credit algorithmic systems, and the Directive on Automated Decision-Making of the Government of Canada (2020).

19. Self-regulation by companies is more efficient than government regulation to prevent and mitigate the risk of violations of human rights, democracy and the rule of law.

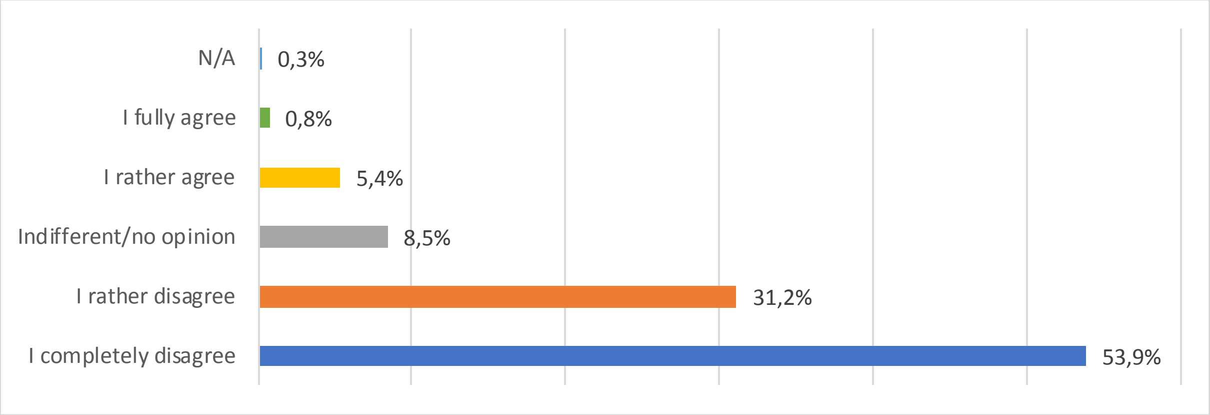


A strong majority of the respondents (75,5%) disagree with the statement that Self-regulation by companies is more efficient than government regulation to prevent and mitigate the risk of violations of human rights, democracy and the rule of law (108). 89 respondents rather disagree with it.

8 respondents fully agree and 23 rather agree.

30 respondents have no opinion or are indifferent to this statement.

20. Self-regulation by companies is sufficient to prevent and mitigate the risk of violations of human rights, democracy and the rule of law



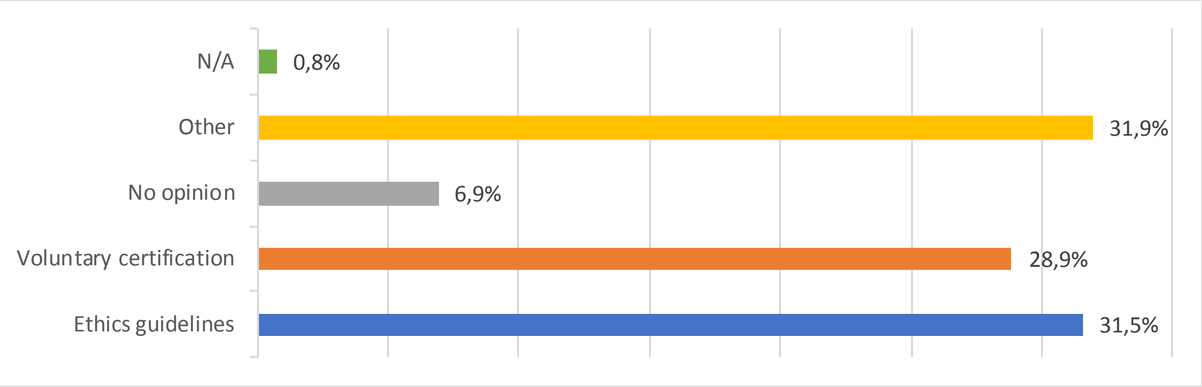
A strong majority of respondents (85%) disagree with the statement that Self-regulation by companies is sufficient to prevent and mitigate the risk of violations of human rights, democracy and the rule of law:

- 140 respondents completely disagree with the statement.
- 81 respondents rather disagree with the statement.

Only 2 respondents fully agree and 14 rather agree.

22 respondents have no opinion or are indifferent to this statement.

21. Which of the following instruments of self-regulation do you consider to be the most efficient?



The self-regulation instrument considered to be the most efficient by a short majority of respondents (who haven't selected the option "other") is ethics guidelines (31,5%).

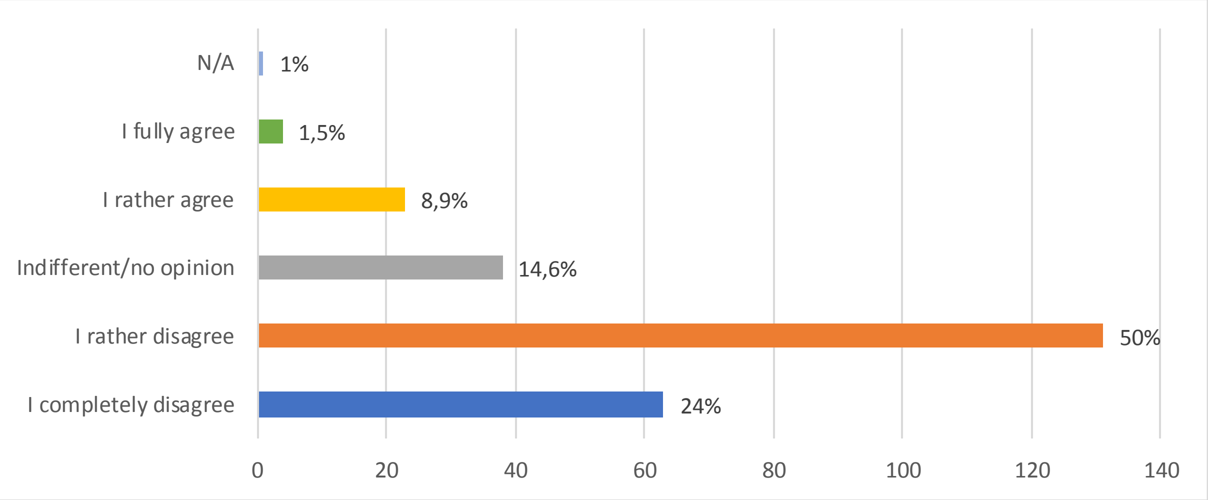
75 respondents consider voluntary certification as the most efficient instrument of self-regulation.

18 respondents have no opinion on the question.

However, most of the respondents (32%) have selected the option "other"

Other instruments include standards, assessment lists, codes of ethics and codes of conduct.

22. Existing international, regional and/or national binding and/or non-binding legal instruments are sufficient to regulate AI systems in order to ensure the protection of human rights, democracy and the rule of law.



A **strong majority** of respondents (74%) disagree with the statement that existing international, regional and/or national binding and/or non-binding legal instruments are sufficient to regulate AI systems in order to ensure the protection of human rights, democracy and the rule of law:

- 50% of respondents rather disagree with the statement.
- 24% of respondents completely disagree with the statement.

Only 1,5% of respondents fully agree and 9% rather agree.

15% of respondents have no opinion or are indifferent to this statement.

23. Please provide examples of existing international, regional and/or national (binding and/or non-binding) instruments that in your view are effective in guiding and regulating the design, development and use of AI systems to ensure compatibility with the standards for human rights, democracy and the rule of law

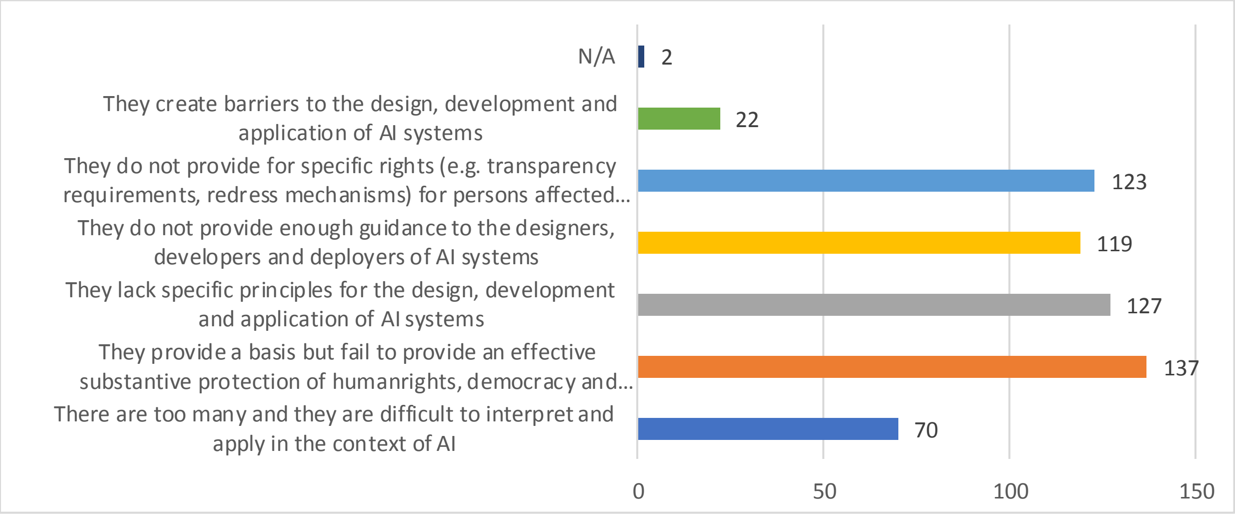
The following table bring together examples of existing international, regional or national instruments proposed by respondents.

Non-binding instruments	**Binding Instruments**
Ethics Guidelines for Trustworthy AI (2019) - High-Level Expert Group on AI (European Commission)	European General Data Protection Regulation (2018)- EU
Ad Hoc Expert Group (AHEG) for the preparation of a draft text of a recommendation on the ethics of artificial intelligence (2020) And UPDATE : Final report on the draft text of the Recommendation on the Ethics of Artificial Intelligence (2021) - UNESCO	Universal Declaration of human rights (1948) - UN
European Ethical Charter on the use of artificial intelligence in judicial systems (2019) - European Commission for the efficiency of justice (CEPEJ),	Charter of Fundamental Rights of the European Union (2000) - EU
Unboxing artificial intelligence: 10 steps to protect human rights (2019) - Council of Europe	Convention for the Protection of Human Rights and Fundamental Freedoms (1950, revised 2010) - UE
Montreal Declaration for a Responsible Development of AI (2018) – Université de Montréal	"Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data" (convention 108 - 1981) AND "Convention 108+. Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data" (2018) - COE

A Strategy and Vision for Artificial Intelligence in Malta 2030 (2019) - Malta	SB-1121 California Consumer Privacy Act of 2018 (2018) – California Gov.
Towards Trustworthy AI: Malta Ethical AI Framework for Public Consultation (2019)	Fair Credit Reporting Act (2018) – FTC US Gov.
Ethically Aligned Design (2019)- IEEE	Directive on Automated Decision-Making (2020) – Government of Canada
Assessment List for Trustworthy Artificial Intelligence by the High-level- expert-group on AI (2020) - European Commission	Directive (EU) 2016/680 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, and repealing Council Framework Decision 2008/977/JHA (2016) - UE
OECD Recommendation of the Council on Artificial Intelligence (2019) - OECD	Council Directive 2000/43/EC of 29 June 2000 implementing the principle of equal treatment between persons irrespective of racial or ethnic origin (2000) – UE
White Paper on Artificial Intelligence. A European approach to excellence and trust (2020) - European Commission	Council Directive 2000/78/EC of 27 November 2000 establishing a general framework for equal treatment in employment and occupation (2000) - UE
Model Framework on AI Governance (2019) - Singapore	Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices (2017) - UE
Toronto Declaration - Protecting the right to equality and non -discrimination in machine learning systems (2018) – Human Rights Watch and other groups	Declaration on Fundamental Principles and Rights at Work and its Follow-up (1998, revised 2010) - ILO
Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning (AI/ML) - Based Software as a Medical Device (2020) – FDA US Gov.	Tripartite Declaration of Principles concerning Multinational Enterprises and Social Policy (2017) - ILO
U.S. LEADERSHIP IN AI: A Plan for Federal Engagement in Developing Technical Standards and Related Tools (2019) - NIST	Regulation (EU) 2019/881 of the European Parliament and of the Council of 17 April 2019 on ENISA (the European Union Agency for Cybersecurity) and on information and communications technology cybersecurity certification and repealing Regulation (EU) No 526/2013 (Cybersecurity Act) (Text with EEA relevance) (2019) - UE
Guiding Principles on Business and Human Rights (2011)- UN	Directive on Automated Decision-Making (2020) - Canada
Due Diligence Guidance for Responsible Business Conduct (2018) – OECD	Code of ethics of data (2019) - Council for Data Ethics Russia
Guidelines on Multinational Enterprises (2011) - OECD	Loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés (1978) - France

The 2020 European Social Partners Framework Agreement on digitalization (2020)– Confederation of European Business or Business Europe	LOI n° 2016-1321 du 7 octobre 2016 pour une République numérique (2016) – France
Cybersecurity Guidelines (2018) - IBA	Data Protection Act 2018 (2018) - UK
Rome Call For AI Ethics; A Human-Centric Artificial Intelligence (2020) – Rome Call	
Artificial Intelligence Mission Austria 2030 (2018) - Austria	
Responsible Use of Artificial Intelligence (2016-2021) - Canada	
Algorithmic Impact Assessment Tool (2021) - Canada	
ExplAIn (2019) - ICO and Alan Turing Institute	
Model AI Governance Framework (2019, updated 2020) - Personal Data Protection Commission	

24. If you responded disagree/completely disagree to question 22, please indicate why existing international, regional and/or national (binding and/or non-binding) legal instruments are not sufficient to regulate AI systems (select all you agree with):



There was no limitation of selected number of answers for this question.

The 4 most selected options, often together, are:

- They provide a basis but fail to provide an effective substantive protection of human rights, democracy and the rule of law against the risks posed by AI systems (137).

- They lack specific principles for the design, development and application of AI systems (127).
- They do not provide enough guidance to the designers, developers and deployers of AI systems (119).
- They do not provide for specific rights (e.g. transparency requirements, redress mechanisms) for persons affected by AI (123).

The less selected option is "They create barriers to the design, development and application of AI systems" (22).
One respondent has provided a different answer which does not appear on the list of choices.

25. Please indicate other *specific* legal gaps that in your view need to be addressed at the level of the Council of Europe

Respondents' contributions vary widely in their interpretation of the term 'legal gaps'. They can be classified as follows:

a. Regulatory context:

Some focus on the regulatory context and underscore the fact that regulatory instruments are too numerous, heterogeneous, and most often non-binding. This leaves too much room for interpretation by the Member states, creating a risk of fragmentation of the (European) internal market. In this context of regulatory competition and silo development of legislation, the priority is to establish a common legal framework that brings together the different initiatives at European and international level. This also requires the Council of Europe to cooperate with other intergovernmental organisations such as the OECD.

b. Legal methods:

Several contributions develop recommendations on the methods of law such as the attribution of the burden of proof. A reversal of the burden of proof is called for, especially in cases of discrimination, so that it is for the entity that develops or deploys an AI system to demonstrate that it does not have a negative impact (discrimination, unfairness) on individuals. The authors of this request also want the responsibility to rest on the developers and deployers rather than on the victims who would have to defend themselves or on the AI systems which cannot be held responsible.

Two further methodological recommendations are made: firstly, the legal framework should impose a method of assessing fairness that is based on the (ex post) comparison of the results of AI systems across groups that are affected by these results (predictions, decisions); secondly, the Council of Europe should adopt a risk-based approach.

c. Rights :

Respondents also suggest adding rights that are not adequately covered at this point. This would include the right to mental integrity and safety. This is particularly important for younger people, and some respondents urge the Council of Europe to recognise specific rights for children within its legal framework and to acknowledge the greater vulnerability of children. Finally, in certain circumstances, the right to refuse to be subjected to an AI system (right to opt out) should be guaranteed and alternative options should be offered without disadvantages to those who refuse.

d. Mechanisms

The majority of responses identify the "legal gaps" as missing mechanisms to ensure the effective regulation of AI and to promote human rights, democracy and the rule of law. The first type of mechanism is that of citizen participation in AI governance and the creation of a platform for authorities and companies to engage with civil society organisations and marginalised groups. Another crucial mechanism to ensure democratic transparency is the creation of public registers of AI systems used by public institutions. An AI Observatory could be established to monitor legislation and public policies on AI and share good practices.

Some complain that the lack of regulation of certification processes leads to a race to the bottom in the provision of auditing services, if not to conflicts of interest. They urge the Council of Europe to provide legal requirements for the certification of AI systems. Finally, the legal framework should include provisions on redress and remedy mechanisms in case of harm.

Section 4: Elements of a Legal Framework on AI Systems

In relation to some AI systems, we can reasonably foresee a significant risk to human rights, democracy and the rule of law. Bearing this in mind, in the following section, it has been asked to respondents to indicate to what extent respondents agree or disagree with the following statements or if they have no opinion on a given issue.

Key findings

In this section, the questions aim to identify consensus on the elements that the legal framework on AI systems should include. Five options are offered for each statement in order to measure the support or rejection of the proposals submitted. Overall, there is a high level of consensus among respondents, with most proposals receiving support of over 80% (I tend to agree and I completely agree). Two proposals receive a support rate of 67% (Q36) and 73% (Q37), which makes them relatively consensual proposals but which should be discussed further. Finally, one proposal (Q40) divides the participants with a 53% support rate and a 25% disapproval rate.

The respondents massively agree that individuals should always be informed when they interact with an AI system in any circumstances (Q26), when a decision which affects them personally is made by an AI system (Q27) and when an AI system is used in a decision-making process which affects them personally (Q28). Moreover, they support the right to a meaningful explanation of decisions based on algorithms (Q29) and the right to have algorithmic decisions or decisions based on algorithmic results checked and reviewed by humans. (Q 30, Q 31 and Q32). More precisely, they agree that individuals should always have the right that any decision taken by an AI system in the framework of judicial proceedings are reviewed by a “human” judge (93% of support) and a right to demand the review of an algorithmic based decision by a human being (85% of support). To this end, there should always be a person responsible for reviewing algorithmic based decisions in the public sector and private companies.

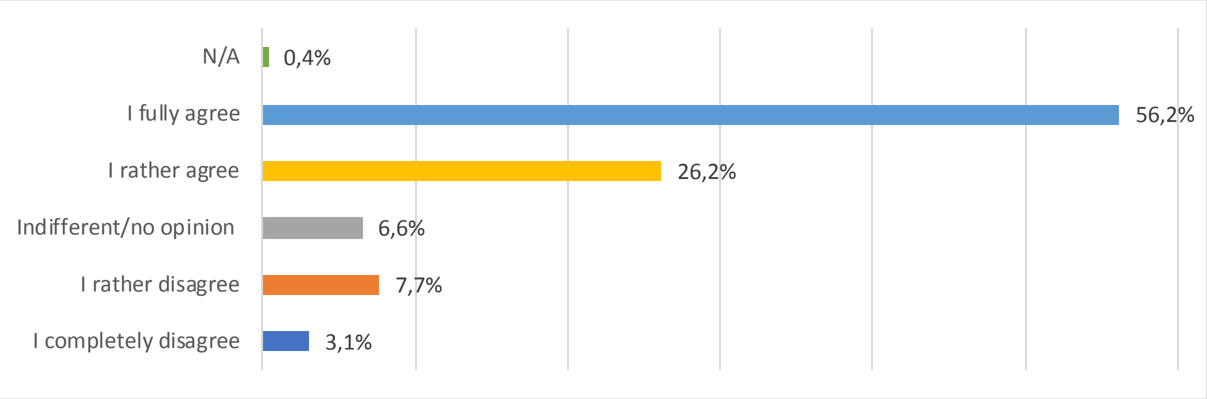
Public institutions have the duty not to use AI systems to promote or discredit a particular way of life or opinion (e.g. “social scoring) and to design, develop and apply sustainable AI systems that respect applicable environmental protection standards, according to a strong majority of respondents in Q 33 and 34.

Respondents consistently converge about oversight mechanisms that should be put in place. For 85% of respondents, Member States should establish public oversight mechanisms for AI systems that may breach legally binding norms in the sphere of human rights, democracy and the rule of law (Q38), and 94% of respondent support reporting mechanism: errors and flaws discovered in AI systems which have led or could lead to the violation of human rights, democracy and the rule of law must be reported to the competent authorities (Q39). However, while 82% of respondents agree that the code behind AI systems should always be accessible to the competent public authorities for the purposes of external audit (Q35), they are less supportive (67%) of the difference in transparency standards applicable to public entities and private entities (Q36). The difference of legal requirement between sectors tends to be less approved as confirmed by the results of question Q37, where 73% of respondents think there should be higher standards for access to an effective remedy for individuals in relation to decisions informed and made by an AI system in the field of justice than in the field of consumer protection.

It is clear from the previous responses that respondents are very concerned about AI systems for facial recognition and many of them call for a ban. But the proposal to ban the use of facial recognition in public (Q40) is more controversial and polarising: while 53% of respondents support this perspective, 25% oppose it and 22% have no definite opinion on the question. However, when asked if the information obtained through the use of facial recognition systems should always be reviewed by a human being before being used for purposes that have an impact on individual freedom (Q41), the disapproval rate drops to 9%.

The results for these two questions are in line with those for questions Q13 and Q14. A majority of respondents believe that the use of facial recognition in public places or for surveillance purposes infringes on human rights and should be prohibited. Framed as a high-risk technology, a greater proportion of respondents back regulation to ensure the proper use of AI. In the same vein, respondents agree that the use of AI systems in democratic processes (e.g. elections) should be strictly regulated.

26. Individuals should always be informed when they interact with an AI system in any circumstances.



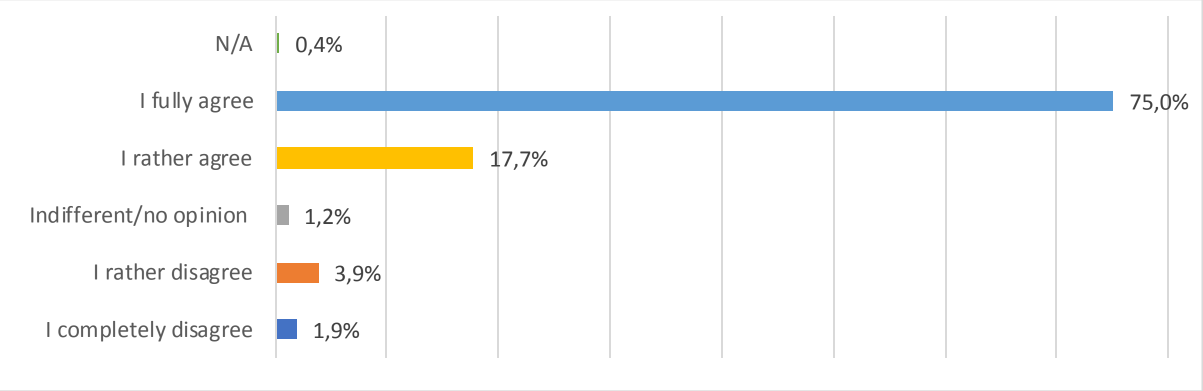
A **strong majority** of respondents agree with the statement that individuals should always be informed when they interact with an AI system in any circumstances:

- 146 respondents fully agree with the statement;
- 68 respondents rather agree with the statement.

Only 8 respondents completely disagree and 20 rather disagree.

17 respondents have no opinion or are indifferent to this statement.

27. Individuals should always be informed when a decision which affects them personally is made by an AI system.



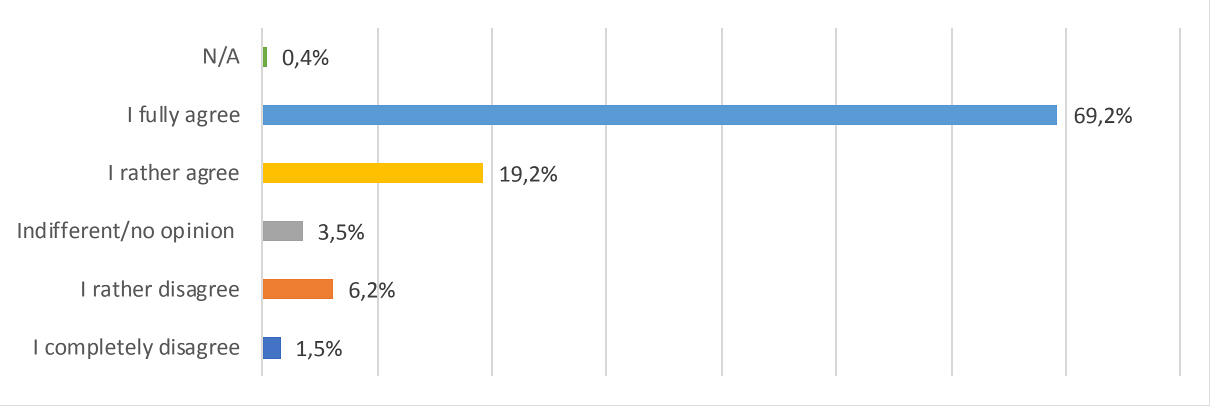
A **strong majority** of respondents agree with the statement that Individuals should always be informed when a decision which affects them personally is made by an AI system:

- 195 respondents fully agree with the statement.
- 46 respondents rather agree with the statement.

Only 5 respondents completely disagree and 10 rather disagree.

3 respondents have no opinion or are indifferent to this statement.

28. Individuals should always be informed when an AI system is used in a decision-making process which affects them personally.



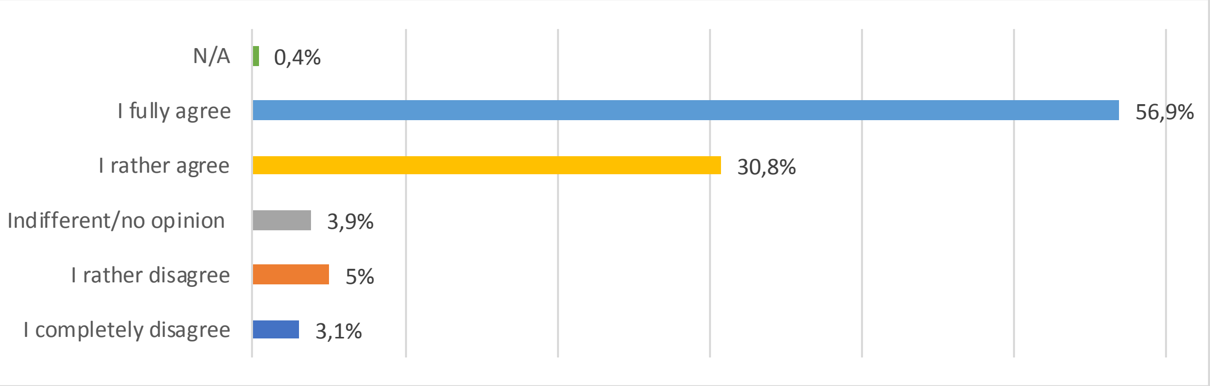
A **strong majority** of respondents agree with the statement that individuals should always be informed when a decision which affects them personally is made by an AI system:

- 180 respondents fully agree with the statement;
- 50 respondents rather agree with the statement.

Only 4 respondents completely disagree and 16 rather disagree.

9 respondents have no opinion or are indifferent to this statement.

29. Individuals should have a right to a meaningful explanation of algorithmic based decisions, in particular how the algorithm reached its output.



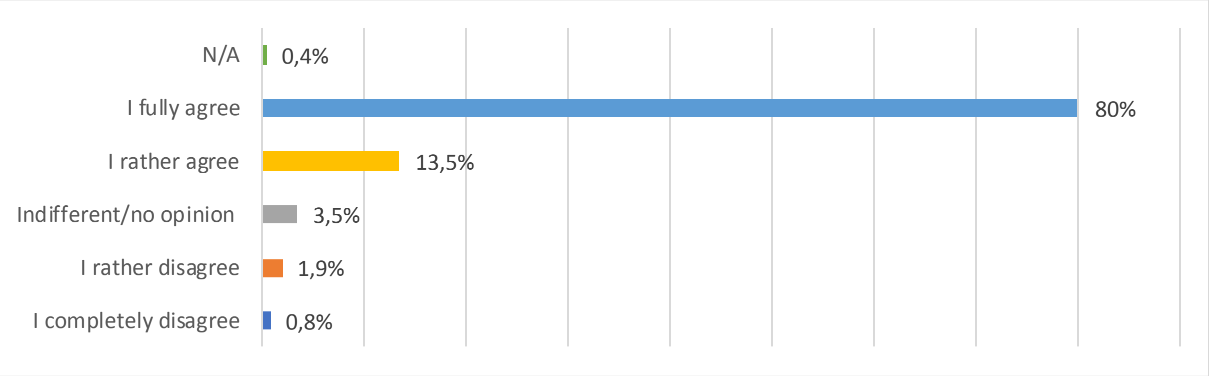
A **wide majority** of respondents agree with the statement that Individuals should have a right to a meaningful explanation of algorithmic based decisions, in particular how the algorithm reached its output:

- 148 respondents fully agree with the statement;
- 80 respondents rather agree with the statement.

Only 8 respondents completely disagree and 13 rather disagree.

10 respondents have no opinion or are indifferent to this statement.

30. Individuals should always have the right that any decision taken by an AI system in the framework of judicial proceedings are reviewed by a “human” judge.



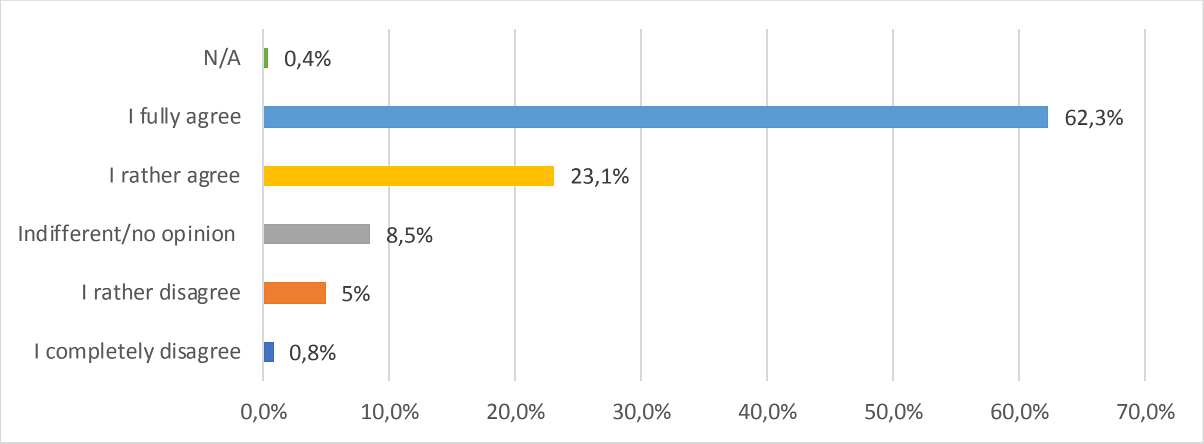
A **strong majority** of respondents agree with the statement that Individuals should always have the right that any decision taken by an AI system in the framework of judicial proceedings are reviewed by a “human” judge:

- 208 respondents fully agree with the statement;
- 35 respondents rather agree with the statement.

Only 2 respondents completely disagree and 5 rather disagree.

9 respondents have no opinion or are indifferent to this statement.

31. Individuals should have a right to demand the review of an algorithmic based decision by a human being.



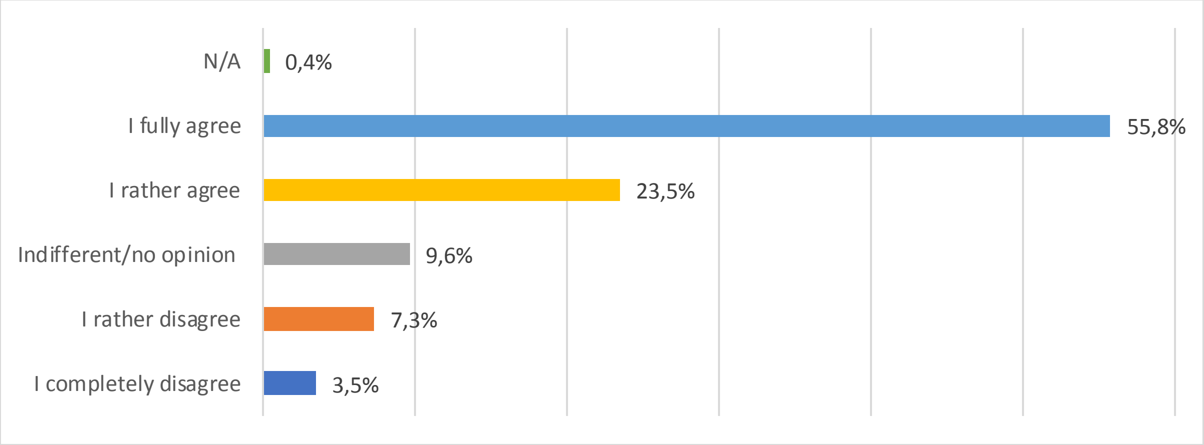
A **strong majority** of respondents agree with the statement that individuals should have a right to demand the review of an algorithmic based decision by a human being:

- 162 respondents fully agree with the statement;
- 60 respondents rather agree with the statement.

Only 2 respondents completely disagree and 13 rather disagree.

22 respondents have no opinion or are indifferent to this statement.

32. There should always be a person responsible for reviewing algorithmic based decisions in the public sector and private companies.



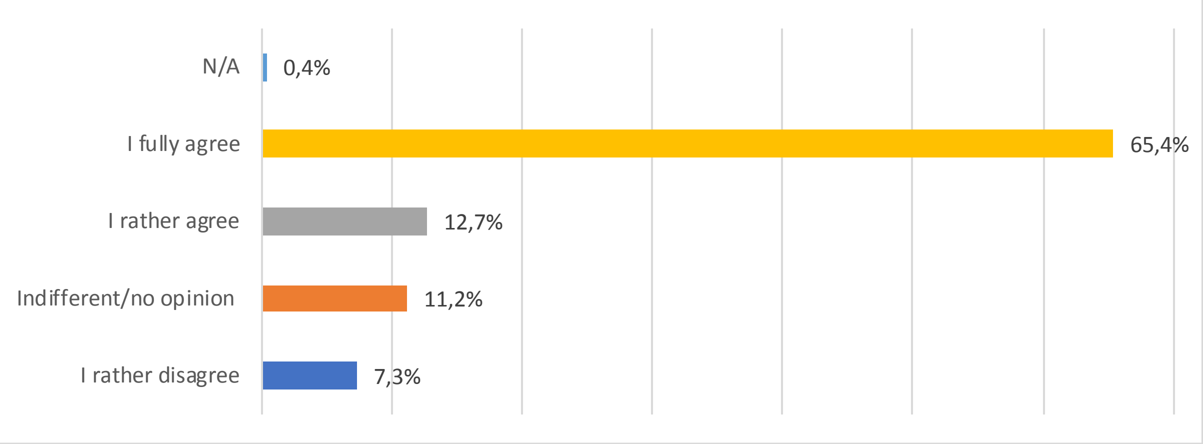
A **strong majority** of respondents agree with the statement that there should always be a person responsible for reviewing algorithmic based decisions in the public sector and private companies:

- 145 respondents fully agree with the statement;
- 61 respondents rather agree with the statement.

Only 9 respondents completely disagree and 19 rather disagree.

25 respondents have no opinion or are indifferent to this statement.

33. Public institutions should not use AI systems to promote or discredit a particular way of life or opinion (e.g. “social scoring”).



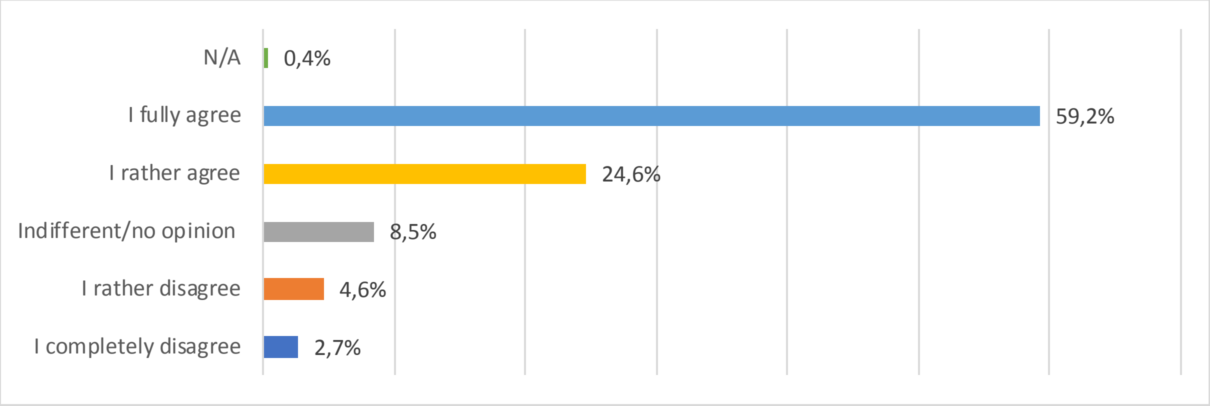
A **wide majority** of respondents agree with the statement that public institutions should not use AI systems to promote or discredit a particular way of life or opinion (e.g. “social scoring”):

- 170 respondents fully agree with the statement;
- 33 respondents rather agree with the statement.

Only 8 respondents completely disagree and 19 rather disagree.

29 respondents have no opinion or are indifferent to this statement.

34. States should be obliged to design, develop and apply sustainable AI systems that respect applicable environmental protection standards.



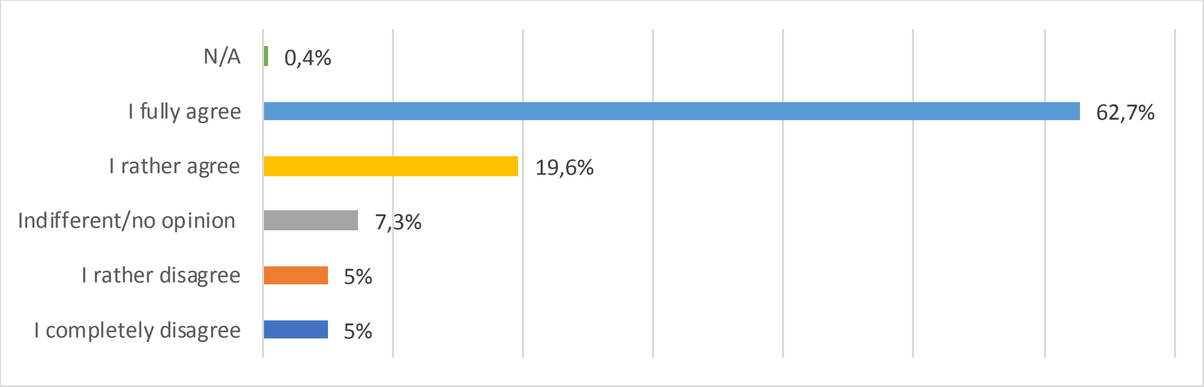
A **strong majority** of respondents agree with the statement that states should be obliged to design, develop and apply sustainable AI systems that respect applicable environmental protection standards:

- 154 respondents fully agree with the statement;
- 64 respondents rather agree with the statement.

Only 7 respondents completely disagree and 12 rather disagree.

22 respondents have no opinion or are indifferent to this statement.

35. The code behind AI systems used in the public and private sectors should always be accessible to the competent public authorities for the purposes of external audit.



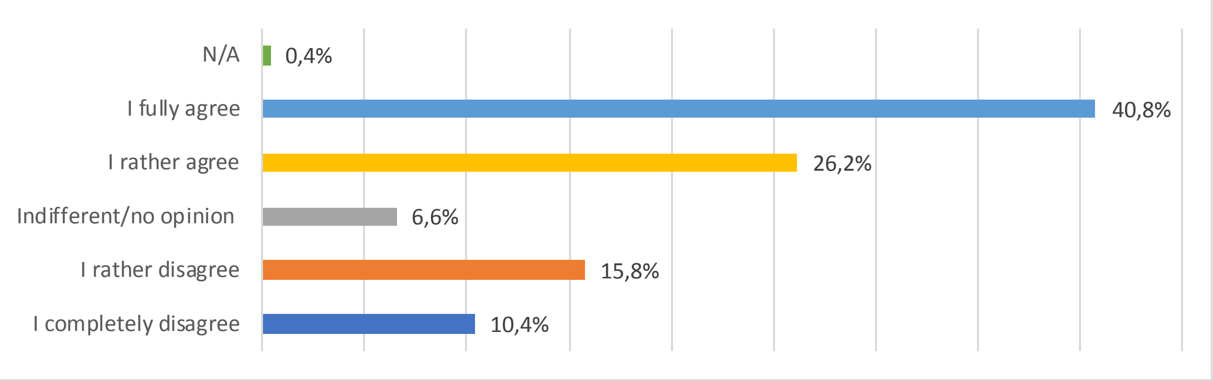
A **wide majority** of respondents agree with the statement that the code behind AI systems used in the public and private sectors should always be accessible to the competent public authorities for the purposes of external audit:

- 163 respondents fully agree with the statement.
- 51 respondents rather agree with the statement.

Only 13 respondents completely disagree and also 13 rather disagree.

19 respondents have no opinion or are indifferent to this statement.

36. There should be higher transparency standards for public entities using AI than for private entities.



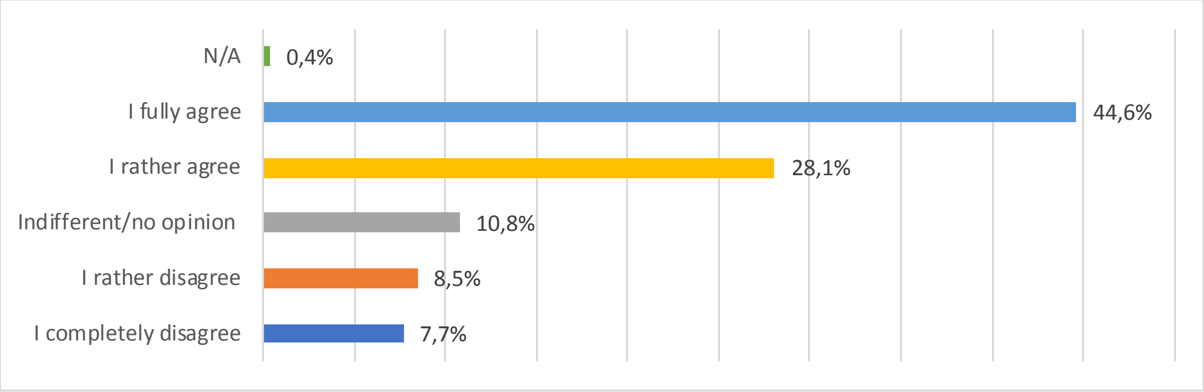
A **majority** of respondents agree with the statement that there should be higher transparency standards for public entities using AI than for private entities:

- 106 respondents fully agree with the statement;
- 68 respondents rather agree with the statement.

27 respondents completely disagree and 41 rather disagree.

17 respondents have no opinion or are indifferent to this statement.

37. There should be higher standards for access to an effective remedy for individuals in relation to decisions informed and made by an AI system in the field of justice than in the field of consumer protection.



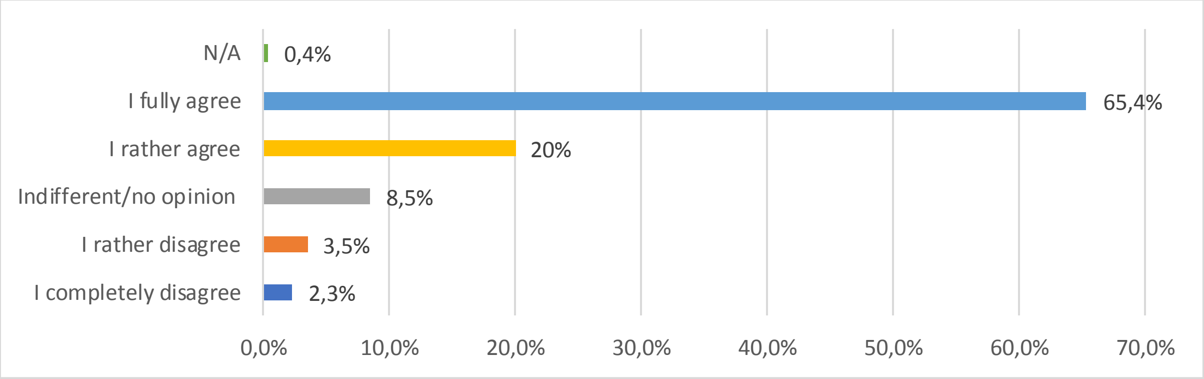
A **strong majority** of respondents agree with the statement that there should be higher standards for access to an effective remedy for individuals in relation to decisions informed and made by an AI system in the field of justice than in the field of consumer protection:

- 116 respondents fully agree with the statement.
- 73 respondents rather agree with the statement.

20 respondents completely disagree and 22 rather disagree.

28 respondents have no opinion or are indifferent to this statement.

38. Member states should establish public oversight mechanisms for AI systems that may breach legally binding norms in the sphere of human rights, democracy and the rule of law.



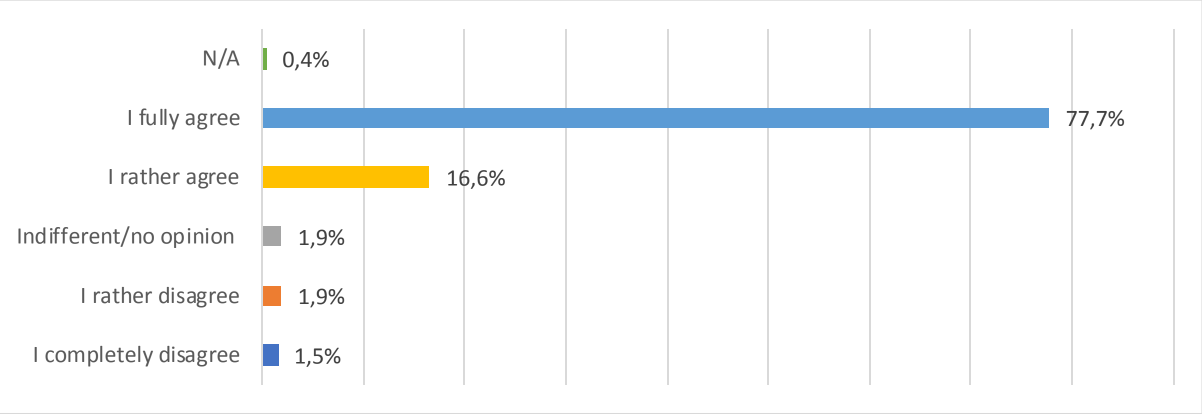
A **wide majority** of respondents agree with the statement that member states should establish public oversight mechanisms for AI systems that may breach legally binding norms in the sphere of human rights, democracy and the rule of law:

- 170 respondents fully agree with the statement;
- 52 respondents rather agree with the statement.

Only 6 respondents completely disagree and 9 rather disagree.

22 respondents have no opinion or are indifferent to this statement.

39. Errors and flaws discovered in AI systems which have led or could lead to the violation of human rights, democracy and the rule of law must be reported to the competent authorities.



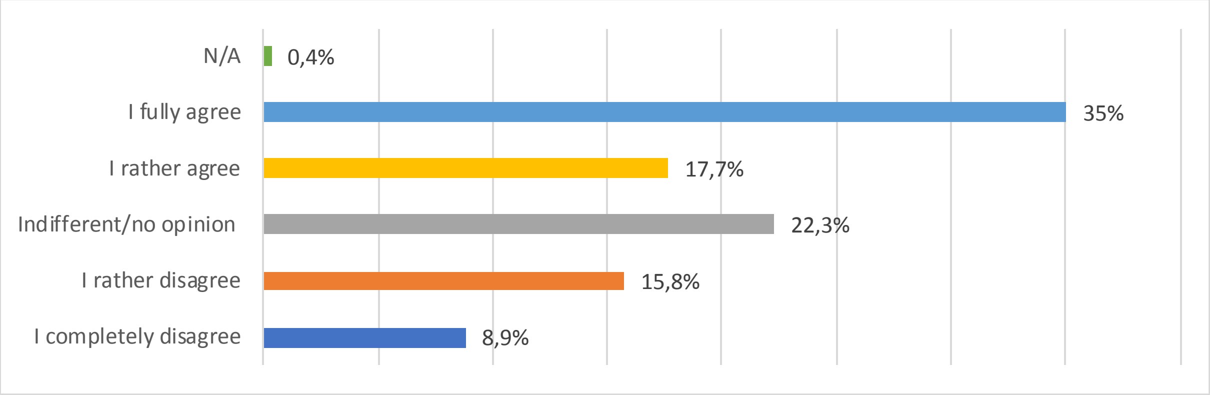
A **strong majority** of respondents agree with the statement that errors and flaws discovered in AI systems which have led or could lead to the violation of human rights, democracy and the rule of law must be reported to the competent authorities:

- 202 respondents fully agree with the statement;
- 43 respondents rather agree with the statement.

Only 4 respondents completely disagree and 5 rather disagree.

5 respondents have no opinion or are indifferent to this statement.

40. The use of facial recognition in public spaces should be prohibited.



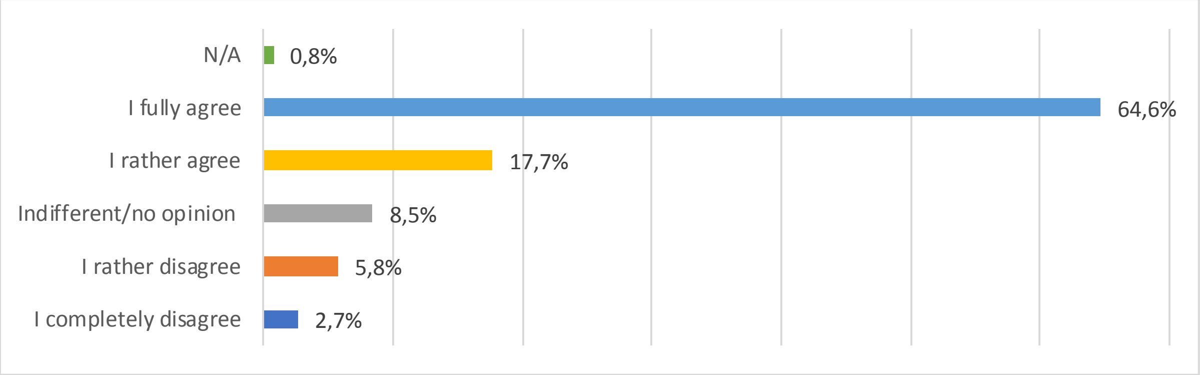
A majority of respondents agree with the statement that the use of facial recognition in public spaces should be prohibited:

- 91 respondents fully agree with the statement;
- 46 respondents rather agree with the statement.

23 respondents completely disagree and 41 rather disagree.

58 respondents have no opinion or are indifferent to the statement.

41. The information obtained through the use of facial recognition systems should always be reviewed by a human being before being used for purposes that have an impact on individual freedom, such as in relation to a person boarding an airplane, upon police arrest or in the framework of judicial proceedings.



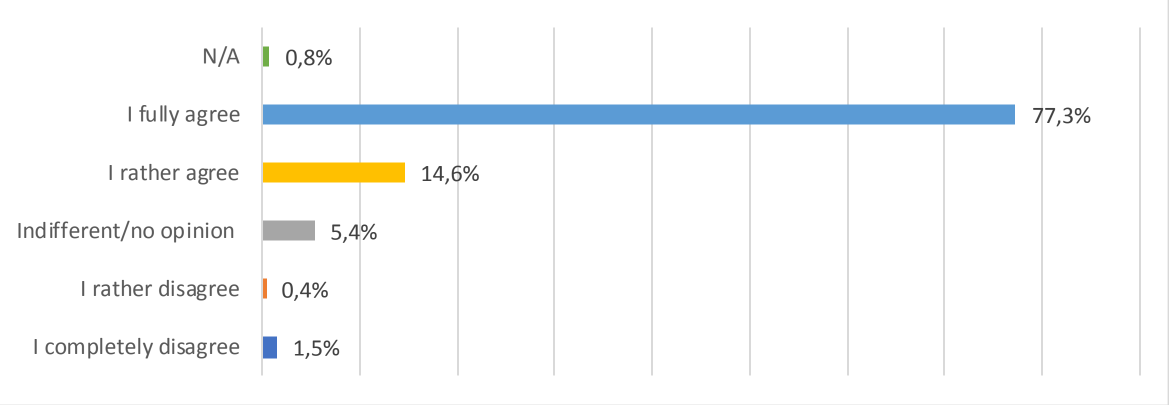
A **wide majority** of respondents agree with the statement that The information obtained through the use of facial recognition systems should always be reviewed by a human being before being used for purposes that have an impact on individual freedom, such as in relation to a person boarding an airplane, upon police arrest or in the framework of judicial proceedings:

- 168 respondents fully agree with the statement;
- 46 respondents rather agree with the statement.

Only 7 respondents completely disagree and 15 rather disagree.

22 respondents have no opinion or are indifferent to the statement.

42. The use of AI systems in democratic processes (e.g. elections) should be strictly regulated.



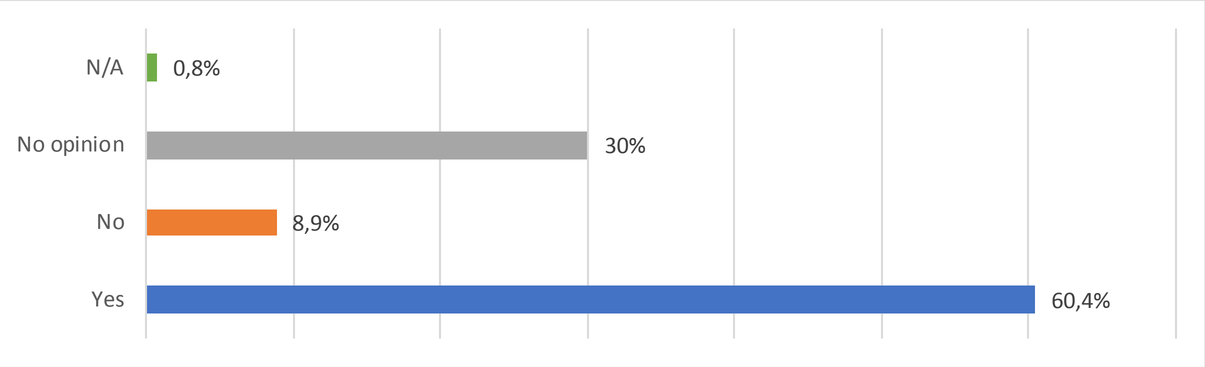
A **very strong majority** of respondents agree with the statement that the use of AI systems in democratic processes (e.g. elections) should be strictly regulated:

- 201 respondents fully agree with the statement;
- 38 respondents rather agree with the statement.

Only 4 respondents completely disagree and 1 rather disagree.

14 respondents have no opinion or are indifferent to the statement.

43. Should a future legal framework at Council of Europe level include a specific liability regime in relation to AI applications?



The majority of respondents (60%) agree that a future legal framework at Council of Europe level should include a specific liability regime in relation to AI applications.

44. If yes, what aspects should be covered?

This question received few comments compared to questions in the first and second sections. Although it was directed at those who answered "yes" to the previous question, several respondents expressed doubts about the need to create a specific liability regime as the use of AI is cross-cutting and responsibilities are not or should not be the same across different areas of use. For some, it is more appropriate to integrate responsibility for problematic uses into existing liability regimes.

For those who believe that a new liability regime is needed, the first challenge is to identify those who have control over AI systems and who can be held accountable if harm is suffered by an affected person. One of the specific problems with liability in AI, especially with deep learning, is that damage can be caused by an algorithmic decision without any defect in the design of the algorithm, in its operation and in the databases used.

Some suggest that we should keep together strict liability (with reconsidered defences and statutory exceptions) and liability based on fault. This would help to avoid "gaps in liability". Others specify that for systems that feed human decisions, it is the human decision-makers who should be held liable; and for systems that make their own decisions or act on their own analyses and predictions, it is the humans who selected and deployed the system who should be held liable. And finally, it must include entities that issue compliance certifications as well as end-users who may misuse AI systems.

The issue of burden of proof comes up here too. Not only should humans be held accountable, but the burden of proof should rest with the designers and deployers of AI systems to demonstrate that their systems are not causing harm.

Finally, according to many respondents across sectors, an adequate liability regime should include mechanisms for challenge, redress, compensation and even restoration. One private sector respondent suggested that a European fund should be set up to compensate victims quickly.

Section 5: Policies and Measures for Development

In this section, the questions 45-49 are focusing on the usefulness for respondents of compliance (5 options) and follow up (4 options) mechanisms. The last question (Q 50) is an open question aimed to offer space for respondents who wish to bring other issues to the attention of the CAHAI.

Most of respondents considers the following compliance mechanisms as highly useful: Human rights, democracy and rule of law impact assessments, audits and intersectional audits, continuous automated monitoring (Q45).

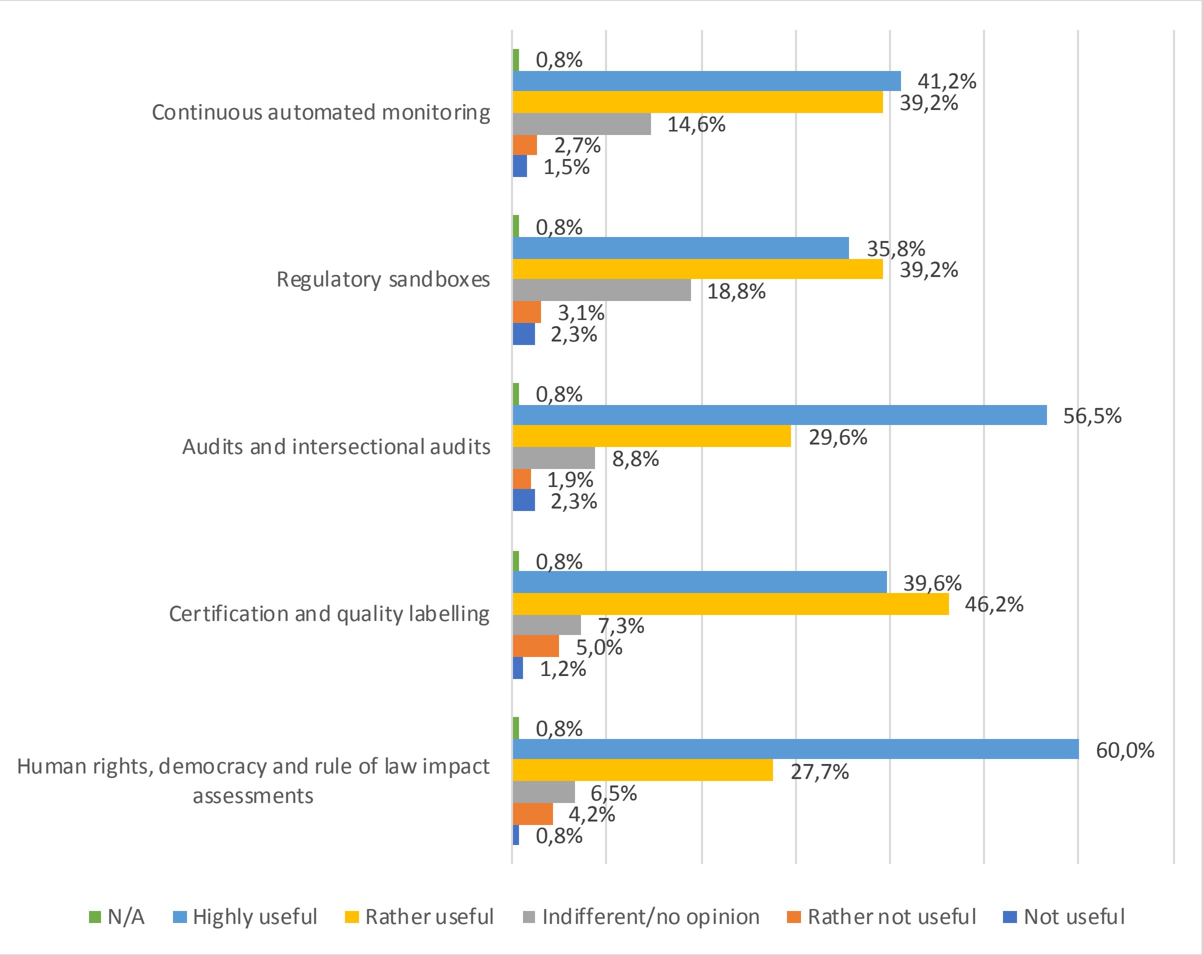
For most of the respondents, compliance mechanisms should be part of a biding instrument (Q 47). However, this preference is less evident for Regulatory sandboxes where 40% of respondents consider that Regulatory sandboxes should be part of a binding instrument against 38% in favour of a non-binding instrument. In addition, respondents who have selected the option “other” to this question, proposed several mechanisms that could complement the proposed list.

The 4 follow up mechanisms proposed in the question 48: Monitoring of AI legislation and policies in member states, Capacity building on Council of Europe instruments, including assistance to facilitate ratification and implementation of relevant Council of Europe instruments, AI Observatory for sharing good practices and exchanging information on legal, policy and technological developments related to AI systems, and Establishing a centre of expertise on AI and human rights, are consider as useful activities to be implemented by the Council of Europe in the view of the majority of respondents.

At the question “What other mechanisms, if any, should be considered?” (Q 49), many respondents recalled proposals that have been made in previous questions. However, three proposals stand out: Dispute resolution mechanism, Cooperation mechanism and Public participation.

Finally, the responses to the question 50 were very rich and informative. Respondents urge the Council of Europe to be clear, consistent, pedagogical and forceful in taking binding measures, even if they can or should co-exist with soft and non-binding regulations.

45. In your opinion, how useful would the following compliance mechanisms be in preventing and mitigating the risks to human rights, democracy and the rule of law arising from the design, development and application of AI?

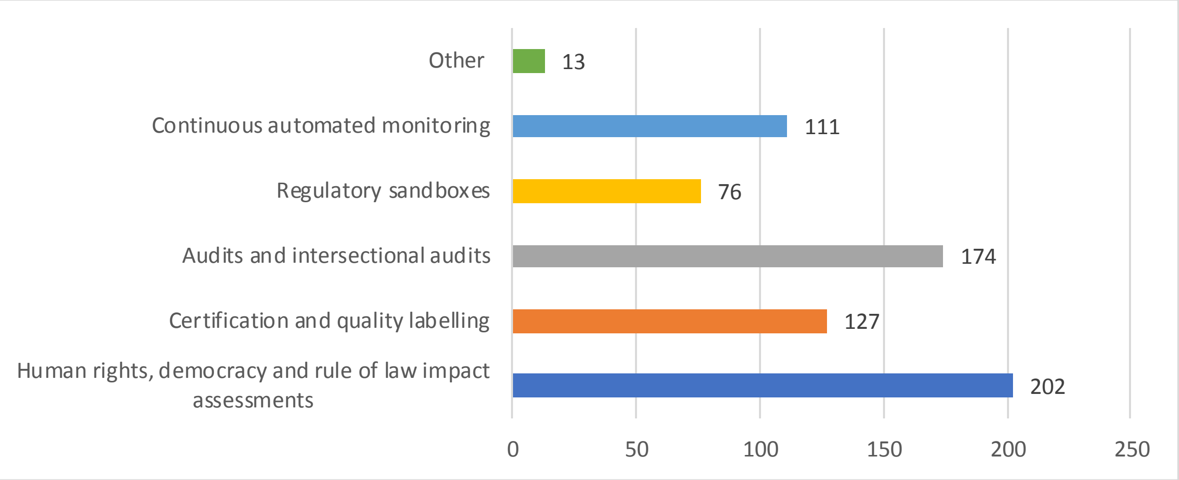


All the following compliance mechanisms proposed have been considered by most of the respondents either highly useful or rather useful.

The 3 compliance mechanisms most considered by the majority of respondent as very useful are Human rights, democracy and rule of law impact assessments (60%), Audits and intersectional audits (56,5%) and Continuous automated monitoring (41%).

Certification and quality labelling (46%) and Regulatory sandboxes (39%) have been considered by most of respondents as rather useful.

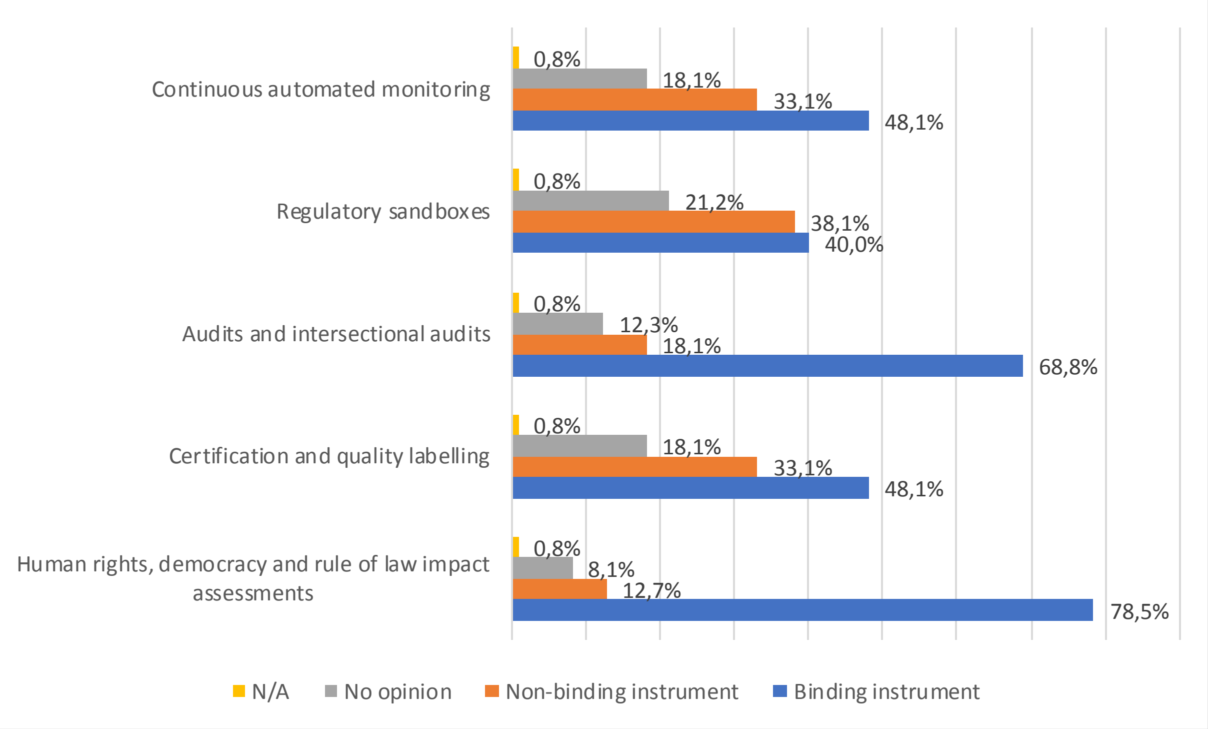
46. Please indicate what combination of mechanisms should be preferred to efficiently protect human rights, democracy and the rule of law



A maximum of 3 selected options were allowed per respondent for this question.

The most selected combination brings together Human rights, democracy and rule of law impact assessments (202), Audits and intersectional audits (174) and Certification and quality labelling (127) mechanisms.

47. Please select which mechanism(s) should be part of either a binding instrument or a non-binding instrument to best protect human rights, democracy and the rule of law.



The majority of respondents consider that a Human rights, democracy and rule of law impact assessments should be part of a binding instrument (79%). It is also the case for Audits and intersectional audits (69%).

48% of respondents are in favour of a binding instrument for Certification and quality labelling and for Continuous automated monitoring

40% of respondents consider that Regulatory sandboxes should be part of a binding instrument against 38% in favour of a non-binding instrument.

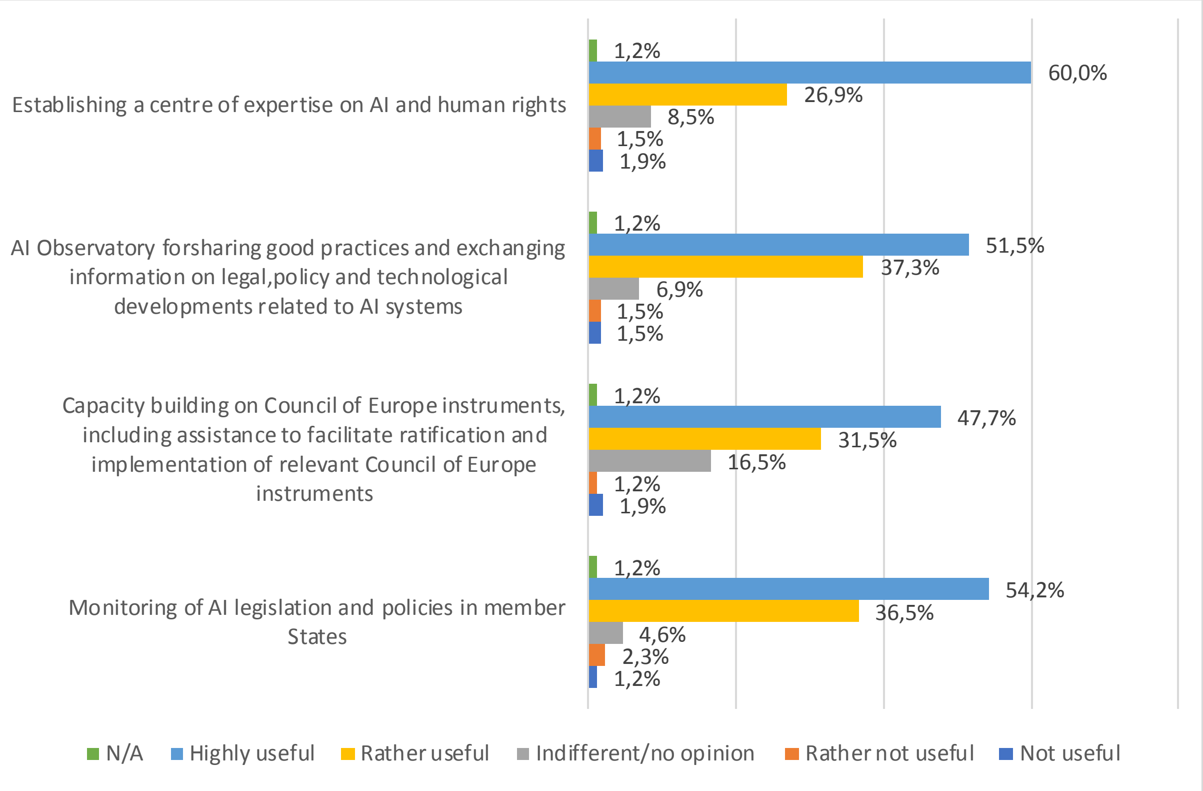
47bis. Other

According to the respondents, several mechanisms could complement the proposed list, such as codes of ethics and codes of good conduct mentioned in question 21, or a mechanism for citizen participation in the development of AI systems from the design stage; proponents of the latter proposal underline the importance of including the most vulnerable people. This would allow developers to assess the social relevance of the AI system they are designing. Some respondents thus propose a mechanism for monitoring the usefulness of AI before deciding to replace current procedures with AI-based processes.

Others propose technical protocols providing access to data collection and processing practices in a standardised form (machine-readable disclosure mechanisms such as Open Ethics Transparency Protocol).

But the proposals that come up most often relate to the monitoring of algorithms. Some point out that certification only makes sense if the certified product does not change, which is difficult to ensure for AI autonomous systems. Rather than establishing certifications that are not sustainable, it is more useful to set up reporting mechanisms with regular reports on the use and impact of automated systems on human rights. Monitoring AI systems also implies to establish (and invest in) an independent AI audit ecosystem.

48. In your opinion, how useful would the following follow-up activities be if implemented by the Council of Europe?



90% of the respondents consider **Monitoring of AI legislation and policies in member states** as useful. Only 3 respondents consider it as not useful and 12 have no opinion on the question.

80% of the respondents consider **Capacity building on Council of Europe instruments, including assistance to facilitate ratification and implementation of relevant Council of Europe instruments** as useful. Only 5 respondents consider it as not useful and 43 have no opinion on the question.

89%of the respondents consider **AI Observatory for sharing good practices and exchanging information on legal, policy and technological developments related to AI systems** as useful. Only 4 respondents consider it as not useful and 18 have no opinion on the question.

87% of the respondents consider **Establishing a centre of expertise on AI and human rights** as useful. Only 5 respondents consider it as not useful and 22 have no opinion on the question.

In a nutshell, the result demonstrate that the 4 follow-up activities proposed are consider as useful activities to be implemented by the Council of Europe in the view of the majority of respondents.

49. What other mechanisms, if any, should be considered?

Many of the responses repeat proposals that have been made in previous ones. But three proposals stand out:

1. Dispute resolution mechanism

A legal instrument for the meaningful regulation of AI cannot be complete if it contains only principles and rules and no enforcement mechanisms in case of disputes in the use of AI. Therefore, the Council of Europe should develop a dispute resolution mechanism at European and international level. It would be administered by staff with expertise in AI and law. The Council of Europe should therefore facilitate the creation of special courts or mediation and arbitration centres as exist in other fields.

2. Cooperation mechanism

Several respondents from all sectors call for cooperation between European and international organisations to avoid regulatory fragmentation and to improve legal oversight and monitoring of AI systems. Some want the Council of Europe with the OECD and other organisations to work together to build a global AI policy, leverage common resources and pool the different expertise developed by these organisations. Other stakeholders advocate a multi-stakeholder format that includes companies in order to be able to adapt regulatory legal instruments to the evolving reality of AI technologies.

Although the following proposal was made in response to the next question, it complements this recommendation well, and we therefore take the liberty of mentioning it here. One respondent gives a specific example of cooperation that could inspire the work of CAHAI. It refers to the 3rd Generation Partnership Project (3GPP) as a multi-stakeholder governance mechanism in the information and communications technology (ICT) industry.

3. Public participation

Many respondents recommend not only cooperation between international organisations but the participatory inclusion of all stakeholders including the public, citizens and especially groups exposed to the negative impacts of AI. The aim is to involve and empower civil society in the context of the massive deployment of AI systems. Therefore, they propose the establishment of a platform that facilitates the sharing of good practices, the identification of trends in the development of AI and the anticipation of ethical and legal issues. This platform would also ensure the participation of the public, especially groups underrepresented in public institutions and in AI policymaking. It would enable feedback on the use of AI and the reporting of severe problems. This proposal is close to but different from that of an observatory for AI.

50. Are there any other issues with respect to the design, development and application of AI systems in the context of human rights, democracy and the rule of law that you wish to bring to the attention of the CAHAI?

The responses to this question were very rich and informative. We cannot do them justice in this summary. In essence, the respondents urge the Council of Europe to be clear, consistent, pedagogical and forceful in taking binding measures, even if they can or should co-exist with soft and non-binding regulations.

- Clear:

The contributors call for unambiguous and well-explained language, in particular so that lay people, civil society organisations that are not front-line AI stakeholders, and even professionals using AI, understand the legal framework and are able to apply it and refer to it for complaints.

- Coherent:

The Council of Europe also needs to establish its legal framework in a coherent way, i.e. to take into account other legal and policy frameworks in place. Any potential convention on AI should thus take into account the interactions with data protection frameworks developed within the Council of Europe as well as the EU Digital Services Act, the GDPR, the European framework of the Medical Devices Regulation and other European Union regulations on AI. Without a holistic approach to AI regulation, there is a serious risk of regulatory fragmentation, which was often asserted throughout the questionnaire, but also a risk of lack of trust and customer uptake of AI products and services developed by the European AI industry.

- Pedagogical:

A lot of comments were about the need of raising awareness of AI among lay people and public officials as well, and about raising ethical awareness among AI developers, designers and deployers. For some respondents, ethical awareness should be a mandatory part of training for AI professionals, which is the case for regulated professions involving a significant degree of risk to the public.

- Forceful:

While the discourse of ethics is useful and necessary to give guidance, many respondents call on the Council of Europe to protect human rights and democracy through binding legal provisions. This is very clear from the overall results of the questionnaire. The field of justice emerged as the riskiest area for the protection of people's rights and the need to regulate AI in this field seems more pressing. Some respondents, representing legal professions, submitted specific protocols to regulate the use of AI in the judicial domain. For instance, to mitigate the potential risks and impact of AI tools within court systems, the following principles should be implemented:

- The possibility to identify the use of AI (Principle of identification): all parties involved in a judicial process should always be able to identify, prior to and within a judicial decision, the elements resulting from the implementation of an AI tool.
- Non-delegation of the judge's decision-making power (Principle of non-delegation): under no circumstances should the judge delegate all or part of his/her decision-making power to an AI tool. In any case, a right to a human judge should be guaranteed at any stage of the proceedings.
- The possibility for the parties to verify the data input and reasoning of the AI tool (Principle of transparency).
- The possibility for the parties to discuss and contest AI outcomes (Principle of discussion) in an adversarial manner outside the deliberation phase and with a reasonable timeframe.
- The neutrality and objectivity of AI tools (Principle of neutrality) used by the judicial system should be guaranteed and verifiable.

Let us conclude on a paradox that appears in several contributions and that remained unnoticed by the respondents: AI should be leveraged to develop AI systems aimed to protect human rights.

This suggests the importance for the Council of Europe to elaborate policy responses and any other suitable measures which can support the development by public and private entities of AI systems in line with the requirements of human rights, rule of law and democracy.

CONCLUSIONS

The objective of the consultation was to collect feedback from a wide range of stakeholders in different sectors (government and public administration, civil society, the private sector and academia) on the main elements of a new legal framework for the development, design and application of artificial intelligence, based on the Council of Europe's standards on human rights, democracy and the rule of law. This feedback, in turn, could help measure the level of support to the different regulatory options being considered by the CAHAI-LFG and hence help this Working Group inform its regulatory proposals.

Many respondents, while contributing with important and relevant ideas to the questionnaire, complained about the rigidity of the consultation format, either because the range of options was too limited or because the open-ended questions only allowed for short responses. However, it should be borne in mind, that limiting the number of options forces careful prioritization and gives CAHAI-LFG the most useful information to identify trends and pinpoint areas of agreement or consensus, debated issues, and deadlocks. In this regard, thanks to the commitment of respondents, the consultation was a success.

A majority was found around several points, starting with the type of definition of AI that should be used in the legal framework. For 48% of participants, a technologically neutral and simplified definition of AI is preferable to narrow definitions or no definition at all. This example of definition provided by the *Feasibility Study* seems to be appropriate for establishing a legal framework.

The majority of respondents noted that AI applications can contribute to the protection of human rights, democracy and the rule of law. Different ways were identified: by combating practices that undermine the exercise of procedural and civil rights such as the right to fair trial; by improving access to public services that condition the enjoyment of substantive and social rights such as the right to education.

Furthermore, there is agreement on the areas in which AI poses the greatest risk of violating human rights, democracy and the rule of law. Two areas stand out clearly: justice and law enforcement, with national security immediately following. The other areas are less clear-cut. Obviously, while the possible future legal framework is likely to be of transversal character and hence apply to the use of AI across different areas and sectors, the feedback received indicates that a possible future legal framework should address both general and specific issues related to the deployment of AI systems in these areas.

In the event that a violation of human rights, democracy and the rule of law by an AI system is proven, a majority (55%) supports its ban. However, a response from a government provides for another option: instead of banning AI systems immediately, an incremental approach should be adopted by increasing the pressure on developers to fix flawed AI systems (the code, model or database) and increasing the penalties. This incremental approach does not apply to AI systems whose purpose by design or in the deployment environment conflicts with human rights, such as social scoring, which should be banned. It should be noted that the "moratorium" option was not given, and based on the results of the next question on AI systems with high risk of human rights infringement with high probability, we can reasonably infer that an option of strict regulation possibly including this option could have been supported by a majority of those who answered "other".

Of course, the question remains as to what criteria should be used to determine whether an AI system violates human rights by its very design. This will raise serious debates, as evidenced by the comments on the regulation of facial recognition in public places. On this point, there is no consensus. Facial recognition is an issue that deeply divides stakeholders.

The positions are sometimes very clear-cut, with no possibility of reconciliation, which shows the importance to find a fruitful space for discussion in order to elaborate a proposal that, without accommodating the most distant positions, will be supported by a solid majority.

Indeed, the responses to the questionnaire revealed a very broad consensus on the need for binding regulation of technologies that present high risks of human rights violations, regardless of the level of probability.

It is clear from the responses to the questionnaire that public authorities and states are expected to implement binding regulation and that self-regulation or soft regulation through ethical charters are not considered enough when human rights, democracy and the rule of law are at stake.

The findings of the multi-stakeholder consultation are clear indications that a legally binding instrument regulating the use of AI is considered as necessary to ensure an effective protection of human rights, rule of law and democracy, in the light of the specific issues raised by artificial intelligence that existing binding regulatory frameworks are unable to address. This option, as noted in the *Feasibility Study*, would need to gain the support of the Committee of Ministers. Based on the responses to the questionnaire, it can be stated that stakeholders from all sectors, including government organisations, overwhelmingly support this option (74%).

A possible new legal framework on the design, development and application of AI based on Council of Europe’s principles should according to the multi-stakeholder consultation give prominence to the following principles and rights: respect for human dignity, privacy and data protection, as well as non-discrimination. The possibility to challenge a decision made by an AI system and access to an effective remedy, as well as ensuring transparency of AI applications, are also considered by most respondents as important aspects to be addressed.

The consultation has also allowed to register an high level of consensus on specific regulatory choices to be made, such as for instance the duty to always inform individuals when they interact with an AI system in any circumstances, when a decision which affects them personally is made by an AI system and when an AI system is used in a decision-making process which affects them personally. Strong support is also recorded with regard to the right to a human review of algorithmic decisions, especially in sensitive contexts such as justice.

Respondents consistently converge about oversight mechanisms that should be put in place, in particular they converge on the need for human oversight when facial recognition technologies are used. For 85% of respondents, states should establish public oversight mechanisms for AI systems that may breach legally binding norms in the sphere of human rights, democracy and the rule of law, and 94% of respondent support reporting mechanism: errors and flaws discovered in AI systems which have led or could lead to the violation of human rights, democracy and the rule of law must be reported to the competent authorities and oversight mechanisms that should be put in place. Furthermore, the consultation has confirmed the perception of usefulness of compliance mechanisms to prevent and mitigate the risks to human rights, rule of law and democracy. According to a strong majority of respondents, some of these mechanisms, such as human rights, democracy and rule of law impact assessments, as well as audits and intersectional audits, should become part of a future binding instrument, whilst others can be better developed through non-binding regulation.

Appendix I. Questionnaire

Disclaimer on data protection

Personal data collected with this questionnaire are managed in accordance with the ***Secretary General's Regulation of 17 April 1989*** *instituting a system of data protection for personal data files at the Council of Europe and the DGA/DIT(2013)02 Data and Information Management Policy of the Council of Europe.*

I, in my capacity as the contact person for replies provided by my delegation, understand that any data, information or assessment, including personal data or confidential information, that I supply to the above survey will be exclusively used by the Council of Europe in the framework of the work undertaken by the CAHAI. I agree to this use being made of any information provided. I understand that, the original replies provided, containing the above personal data, would be deleted by the CAHAI secretariat by [DATE] at the latest.

I formally consent to the use of my personal data and of any other information I supplied as described above. If I submit personal data or confidential information of another person, I confirm that I have obtained the authorisation to do so from that person.

For any request relating to the exercise of your right to the protection of personal data, please contact dpo@coe.int.

For any issues, please contact secretariat.cahai@coe.int

1. **Pre-screening question of the survey:**

- Your state
- Institution: Name of the institution/body/company
- Personal capacity: Your socio-professional category (using an existing list)
- Your stakeholder groups (choice amongst government & public administration/ private business sector/ civil society/ academic and scientific community / internet technical community)

Section 1: Definition of AI Systems

2. In view of the elaboration of a legal framework on the design, development and application of AI, based on the standards of the Council of Europe on human rights, democracy and the rule of law, what kind of definition of artificial intelligence (AI) should be considered by the CAHAI (select one):

○ No definition, with a legal instrument focused on the effect of AI systems on human rights, democracy and the rule of law.

○ A technologically-neutral and simplified definition, such as “a set of sciences, theories and techniques whose purpose is to reproduce by a machine the cognitive abilities of a human being”[8].

○ A definition focusing on machine learning systems.

○ A definition focusing on automated decision-making.

○ Other (Please explain below)

Limited characters

○ No opinion

3. What are the reasons for your preference?

Limited characters

Section 2: Opportunities and Risks arising from AI Systems

Opportunities arising from AI systems

4. Please select the areas in which AI systems offer the most promising opportunities for the protection of human rights, democracy and the rule of law (select 3 maximum):

- Banking, finance and insurance
- Justice
- Law enforcement
- Customs and border control
- Welfare
- Education
- Healthcare

[8] See the CAHAI feasibility study, §5.

- Environment and climate
- Election monitoring
- National security and counter-terrorism
- Public administration
- Employment
- Social networks/media, internet intermediaries
- Other (which areas and why)

Limited characters

- No opinion

5. Please indicate which of the following AI system applications in your view have the greatest potential to enhance/protect human rights, democracy and the rule of law? (select 5 maximum):

- Facial recognition supporting law enforcement
- Emotional analysis in the workplace to measure employees' level of engagement
- Smart personal assistants (connected devices)
- Scoring of individuals by public and private entities
- Medical applications for faster and more accurate diagnoses
- Automated fraud detection (banking, insurance)
- AI applications to predict the possible evolution of climate change and/or natural disasters;
- AI applications for personalised media content (recommender systems)
- Deep fakes and cheap fakes
- Recruiting software/ AI applications used for assessing work performance
- AI applications to prevent the commission of a criminal offence (e.g. anti-money laundry AI applications)
- AI applications aimed at predicting recidivism
- AI applications providing support to the healthcare system (triage, treatment delivery)
- AI applications determining the allocation of educational services
- AI applications determining the allocation of social services
- AI applications in the field of banking and insurance
- AI applications to promote gender equality (e.g. analytical tools)
- AI applications used for analysing the performance of pupils/students in educational institutions such as schools and universities

6. Please briefly explain how such applications would benefit human rights, democracy and the rule of law.

Limited characters

7. What other applications might contribute significantly to strengthening human rights, democracy and the rule of law?

Limited characters

Impact on human rights, democracy and the rule of law

8. Please select the areas in which the deployment of AI systems poses the highest risk of violating human rights, democracy and the rule of law (select 3 maximum)

- Banking, finance and insurance
- Justice
- Law enforcement
- Customs and border control
- Welfare
- Education
- Healthcare
- Environment and climate
- Election monitoring
- National security and counter-terrorism
- Public administration
- Employment
- Social networks/media, internet intermediaries
- Other

Limited characters

- No opinion

9. Please briefly explain how such applications might violate human rights, democracy and the rule of law.

Limited characters

10. Please indicate the types of AI systems that represent the greatest risk to human rights, democracy and the rule of law (select 5 maximum):

- Facial recognition supporting law enforcement
- Emotional analysis in the workplace to measure employees' level of engagement
- Smart personal assistants (connected devices)
- Scoring / scoring of individuals by public entities
- Medical applications for faster and more accurate diagnoses
- Automated fraud detection (banking, insurance)
- AI applications to predict the possible evolution of climate change and/or natural disasters;
- AI applications for personalised media content (recommender systems)
- Deep fakes and cheap fakes
- Recruiting software/ AI applications used for assessing work performance
- AI applications to prevent the commission of a criminal offence
- AI applications aimed at predicting recidivism
- AI applications providing support to the healthcare system (triage, treatment delivery)
- AI applications determining the allocation of educational services
- AI applications determining the allocation of social services
- AI applications in the field of banking and insurance
- AI applications to promote gender equality (e.g. analytical tools)
- AI applications used for analysing the performance of pupils/students in educational institutions such as schools and universities

11. Please briefly explain how such applications might violate human rights, democracy and the rule of law.

Limited characters

12. What other applications might represent a significant risk to human rights, democracy and the rule of law?

Limited characters

13. In your opinion, should the development, deployment and use of AI systems that have been proven to violate human rights or undermine democracy or the rule of law be:

○ Banned

○ Not banned

○ Other

Limited characters

○ No opinion

14. In your opinion, should the development, deployment and use of AI systems that pose *high risks9 with high probability*[10] to human rights, democracy and the rule of law be:

?

○ Banned

○ Subject to moratorium

○ Regulated (binding law)

○ Self-regulated (ethics guidelines, voluntary certification)

○ None of the above

○ No opinion

15. In your opinion, should the development, deployment and use of AI systems that pose *low risks*[11] *with high probability*[12] to human rights, democracy and the rule of law be:

○ Banned.

○ Subject to moratorium.

○ Regulated (binding law)

○ Self-regulated (ethics guidelines, voluntary certification)

[9] high negative impact on human rights, democracy and rule of law

[10] high probability of occurrence of these risks

[11] Low negative impact on human rights, democracy and rule of law

[12] high probability of occurrence of these risks

○ None of the above

○ No opinion

16. In your opinion, should the development, deployment and use of AI systems that pose *high risks*[13] *with low probability*[14] to human rights, democracy and the rule of law be:

○ Banned

○ Subject to moratorium

○ Regulated (binding law)

○ Self-regulated (ethics guidelines, voluntary certification).

○ None of the above

○ No opinion

17. What are the most important legal principles, rights and interests that need to be addressed and therefore justify regulating the development, deployment and use of AI systems? (select 5 maximum):

- Respect for human dignity
- Political pluralism
- Equality
- Social security
- Freedom of expression, assembly and association
- Non-discrimination
- Privacy and data protection
- Personal integrity
- Legal certainty
- Transparency
- Explainability
- Possibility to challenge a decision made by an AI system and access to an effective remedy

[13] high negative impact on human rights, democracy and rule of law

[14] Low probability of occurrence of these risks

18. In your opinion, in what sectors/areas is a binding legal instrument needed to protect human rights, democracy and the rule of law? (select 3 maximum)?

- Banking, finance and insurance
- Justice
- Law enforcement
- Customs and border control
- Welfare
- Education
- Healthcare
- Social networks/media, internet intermediaries
- Environment and climate
- Election monitoring
- Public administration
- Employment
- No opinion
- Other

Limited characters

Section 3: Potential Gaps in Existing Binding Legal Instruments Applicable to AI

In the following section, please indicate to what extent you agree or disagree with the following statements or if you have no opinion on a given issue.

19. Self-regulation by companies is more efficient than government regulation to prevent and mitigate the risk of violations of human rights, democracy and the rule of law.

1 I completely disagree	2 I rather disagree	3 Indifferent	4 I rather agree	5 I fully agree	No opinion

20. Self-regulation by companies is sufficient to prevent and mitigate the risk of violations of human rights, democracy and the rule of law

1 I completely disagree	2 I rather disagree	3 Indifferent	4 I rather agree	5 I fully agree	No opinion

21. Which of the following instruments of self-regulation do you consider to be the most efficient?

○ Ethics guidelines

○ Voluntary certification

○ Other

Limited characters

○ No opinion

22. Existing international, regional and/or national binding and/or non-binding legal instruments are sufficient to regulate AI systems in order to ensure the protection of human rights, democracy and the rule of law.

1 I completely disagree	2 I rather disagree	3 Indifferent	4 I rather agree	5 I fully agree	No opinion

23. Please provide examples of existing international, regional and/or national (binding and/or non-binding) instruments that in your view are effective in guiding and regulating the design, development and use of AI systems to ensure compatibility with the standards for human rights, democracy and the rule of law:

Limited characters

24. If you responded disagree/completely disagree to question 22, please indicate why existing international, regional and/or national (binding and/or non-binding) legal instruments are not sufficient to regulate AI systems (select all you agree with):

- There are too many and they are difficult to interpret and apply in the context of AI.
- They provide a basis but fail to provide an effective substantive protection of human rights, democracy and the rule of law against the risks posed by AI systems.
- They lack specific principles for the design, development and application of AI systems.
- They do not provide enough guidance to the designers, developers and deployers of AI systems.
- They do not provide for specific rights (e.g. transparency requirements, redress mechanisms) for persons affected by AI.

- They create barriers to the design, development and application of AI systems.

25. Please indicate other *specific* legal gaps that in your view need to be addressed at the level of the **Council of Europe**

Limited characters

Section 4: Elements of a Legal Framework on AI Systems

In relation to some AI systems, we can reasonably foresee a significant risk to human rights, democracy and the rule of law. Bearing this in mind, in the following section, please indicate to what extent you agree or disagree with the following statements or if you have no opinion on a given issue.

26. Individuals should always be informed when they interact with an AI system in any circumstances.

1 I completely disagree	2 I rather disagree	3 Indifferent	4 I rather agree	5 I fully agree	No opinion

27. Individuals should always be informed when a decision which affects them personally is made by an AI system.

1 I completely disagree	2 I rather disagree	3 Indifferent	4 I rather agree	5 I fully agree	No opinion

28. Individuals should always be informed when an AI system is used in a decision-making process which affects them personally.

1 I completely disagree	2 I rather disagree	3 Indifferent	4 I rather agree	5 I fully agree	No opinion

29. Individuals should have a right to a meaningful explanation of algorithmic based decisions, in particular how the algorithm reached its output.

1 I completely disagree	2 I rather disagree	3 Indifferent	4 I rather agree	5 I fully agree	No opinion

30. Individuals should always have the right that any decision taken by an AI system in the framework of judicial proceedings are reviewed by a “human” judge.

1 I completely disagree	2 I rather disagree	3 Indifferent	4 I rather agree	5 I fully agree	No opinion

31. Individuals should have a right to demand the review of an algorithmic based decision by a human being.

1 I completely disagree	2 I rather disagree	3 Indifferent	4 I rather agree	5 I fully agree	No opinion

32. There should always be a person responsible for reviewing algorithmic based decisions in the public sector and private companies.

1 I completely disagree	2 I rather disagree	3 Indifferent	4 I rather agree	5 I fully agree	No opinion

33. Public institutions should not use AI systems to promote or discredit a particular way of life or opinion (e.g. “social scoring”).

1 I completely disagree	2 I rather disagree	3 Indifferent	4 I rather agree	5 I fully agree	No opinion

34. States should be obliged to design, develop and apply sustainable AI systems that respect applicable environmental protection standards.

1 I completely disagree	2 I rather disagree	3 Indifferent	4 I rather agree	5 I fully agree	No opinion

35. The code behind AI systems used in the public and private sectors should always be accessible to the competent public authorities for the purposes of external audit.

1 I completely disagree	2 I rather disagree	3 Indifferent	4 I rather agree	5 I fully agree	No opinion

36. There should be higher transparency standards for public entities using AI than for private entities.

1 I completely disagree	2 I rather disagree	3 Indifferent	4 I rather agree	5 I fully agree	No opinion

37. There should be higher standards for access to an effective remedy for individuals in relation to decisions informed and made by an AI system in the field of justice than in the field of consumer protection.

1 I completely disagree	2 I rather disagree	3 Indifferent	4 I rather agree	5 I fully agree	No opinion

38. Member States should establish public oversight mechanisms for AI systems that may breach legally binding norms in the sphere of human rights, democracy and the rule of law.

1 I completely disagree	2 I rather disagree	3 Indifferent	4 I rather agree	5 I fully agree	No opinion

39. Errors and flaws discovered in AI systems which have led or could lead to the violation of human rights, democracy and the rule of law must be reported to the competent authorities.

1 I completely disagree	2 I rather disagree	3 Indifferent	4 I rather agree	5 I fully agree	No opinion

40. The use of facial recognition in public spaces should be prohibited.

1 I completely disagree	2 I rather disagree	3 Indifferent	4 I rather agree	5 I fully agree	No opinion

41. The information obtained through the use of facial recognition systems should always be reviewed by a human being before being used for purposes that have an impact on individual freedom, such as in relation to a person boarding an airplane, upon police arrest or in the framework of judicial proceedings.

1 I completely disagree	2 I rather disagree	3 Indifferent	4 I rather agree	5 I fully agree	No opinion

42. The use of AI systems in democratic processes (e.g. elections) should be strictly regulated.

1 I completely disagree	2 I rather disagree	3 Indifferent	4 I rather agree	5 I fully agree	No opinion

43. Should a future legal framework at Council of Europe level include a specific liability regime in relation to AI applications?

- ○ Yes
- ○ No
- ○ No opinion

44. If yes, what aspects should be covered?

Limited characters

Section 5: Policies and Measures for Development

45. In your opinion, how useful would the following compliance mechanisms be in preventing and mitigating the risks to human rights, democracy and the rule of law arising from the design, development and application of AI?

	1 Not useful	2 Rather not useful	3 Indifferent	4 Rather useful	5 Highly useful	No opinion
Human rights, democracy and rule of law impact assessments						
Certification and quality labelling						

Audits and intersectional audits[15]						
Regulatory sandboxes						
Continuous automated monitoring						

46. Please indicate what combination of mechanisms should be preferred to efficiently protect human rights, democracy and the rule of law (select 3 maximum).

- Human rights, democracy and rule of law impact assessments
- Certification and quality labelling
- Audits and intersectional audits
- Regulatory sandboxes
- Continuous automated monitoring
- Other

Limited characters

47. Please select which mechanism(s) should be part of either a binding instrument or a non-binding instrument to best protect human rights, democracy and the rule of law.

	Binding instrument	Non-binding instrument	No opinion
Human rights, democracy and rule of law impact assessments			
Certification and quality labelling			
Audits and intersectional audits			
Regulatory sandboxes			
Continuous automated monitoring			
Other [limited characters]			

48. In your opinion, how useful would the following follow-up activities be if implemented by the Council of Europe?

[15] [definition]

	1 Not useful	2 Rather not useful	3 Indifferent	4 Rather useful	5 Highly useful	No opinion
Monitoring of AI legislation and policies in member States						
Capacity building on Council of Europe instruments, including assistance to facilitate ratification and implementation of relevant Council of Europe instruments						
AI Observatory for sharing good practices and exchanging information on legal, policy and technological developments related to AI systems						
Establishing a centre of expertise on AI and human rights						

49. What other mechanisms, if any, should be considered?

Limited characters

50. Are there any other issues with respect to the design, development and application of AI systems in the context of human rights, democracy and the rule of law that you wish to bring to the attention of the CAHAI?

Limited characters