EUROPEAN DIGITAL RIGHTS

Public consultation on the "White Paper on Artificial Intelligence - A European Approach"

EDRi Answering Guide

The European Commission has launched a **public consultation** to ask your views about the regulation of artificial intelligence (AI). This guides proposes answers to the consultation based on upholding human rights first and foremost.

AI is a very broad term including a range of processes and technologies which enable computers to complement or replace specific tasks otherwise performed by humans, such as making decisions and solving problems, or to do them at a scale that humans cannot. As it functions today, AI involves the computerised analysis of large data sets to analyse, model, and predict an issue or scenario – although experts can disagree on what exactly would be considered "AI".

In February 2020, the Commission set out its own ideas in the White Paper on Artificial Intelligence, highlighting intentions for more investment in AI and promote AI in the public sector, whilst also ensuring that AI development and deployment is 'trustworthy'.

Unfortunately, the White Paper failed to strike the right balance. Issues of fundamental rights, trust, privacy and equality cannot be an afterthought to 'innovation'. The increased use of AI in all areas of life – such as in policing, healthcare, welfare systems, inside the workplace, deciding which ads you see online, in migration control and more – raises serious concerns for people and societies. It will mean increased monitoring and surveillance, the potential for greater discrimination and inequality, less control over our data and privacy, and greater influence (and profits) for technology and security firms. The European Commission needs to properly address these risks, or risks 'stumbling zombie-like into a digital dystopia'.

We have prepared this guide to make it easier for you to respond. This guide is designed for people who choose the option to respond as an individual. For all of the written questions in the consultation there is a 500 *character* limit – this is really short, so do not plan long answers. It should take 10-30 minutes to complete, depending on how passionate you are about AI regulation. If you need some guidance about specific wording, please see EDRi's response to the consultation here. You can also read EDRi's explainer on the human rights risks of AI here.

GETTING STARTED AND ABOUT YOU

Go to the consultation page – first you will be asked to create an account for the EU survey portal if you don't already have one. Once registered, you will be asked to identify yourself, the capacity in which you are answering the consultation - either 'EU citizen' or 'Non-EU citizen' (sigh) - country of origin and email address, all of which are mandatory. The Commission is unlikely to contact you using this data, except to confirm your response.

It will also ask you about whether it can publish your answers – you can choose to stay anonymous or not.

SECTION 1 – *An Ecosystem of Excellence*

These questions are heavily based on the the European Commission's general aim for Europe to become "a leader in AI", therefore the questions are biased toward market, economic and industrial policy. Questions of 'trust' come up later and are considered separate to 'excellence'. For the EDRi network, however, trust and human rights are the real markers of excellence.

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

The European Commission has been vocal about 'partnerships with the private sector' and 'promoting AI in the public sector'. We recommend selecting **1** for both of these categories to signal a need for caution here.

As the increased resort to AI systems by public and state authorities is likely to pose a number of risks to our rights, the justification and the benefits of these systems need to be really clear before they are adopted or promoted, particularly when used in public functions such as assigning benefits or predicting risks of committing a crime. What's more, there are countless examples of harms caused by increased private involvement in our public services by providing "data-driven services" - a lack of transparency and accountability, risks for privacy and data protection, and in general profit, rather than people, driven practices.

Are there other actions that should be considered? 500 characters max

The Commission assumes generally in the White Paper, without evidence, that AI should be promoted, especially in the public sector. In response, it is important to highlight AI should not be promoted just for the sake of it. Furthermore, if governments use AI for public functions (such as the delivery of essential services like healthcare, housing, education, social welfare, transport), there must be:

1. clear, published reasons to justify the use of AI
2. scientific evidence that the technology works, and,
3. particularly where the technology will play an important role in determining people's access to vital services or to enjoy their fundamental rights and freedoms, **people should have a say in whether or not AI can be acceptably used in a democratic society**.

In our opinion, the European Commission should include actions such as democratic oversight of AI in the public sector, consultations with civil society, the general public and affected communities, as well as stringent human rights safeguards.

Revising the Coordinated Plan on AI (Action 1)

The coordinated plan on AI is a document which sets out the EU's overall strategy with respect to AI, including its investment, research, and EU member state engagement plan.

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

Here we suggest selecting 1 (not important at all) for 'promoting the uptake of AI by business and public sector' because, again, AI – like any other technology – is not inherently good. It depends on how it is developed, how it is deployed, by who, for what purpose, and so much more. Promoting AI regardless of the risks it poses could also be irresponsible, especially in the public sector and in the delivery of essential services. Just because we can, does not mean we should.

Are there other areas that that should be considered?

We propose that the coordinated plan on AI is updated to include criteria (scientific and policy) about how the EU will allocate its resources of AI. The plan should include a section on human rights, societal impacts of AI and automation, and how to ensure democratic oversight for the application of AI systems.

A united and strengthened research and innovation community striving for excellence

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

This question is quite narrowly framed - there are no preferred answers here from a digital rights perspective. However there are a lot of things missing - see next question.

EDRi suggests: Support to establishment of a lighthouse research (3); Network of existing AI research (4); Set up a public private partnership (1).

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

The Commission highlights research and innovation as a priority of its strategy, however there is little mention of research into human rights and the societal impact of AI. We have major concerns about potential discrimination, disinformation, a lack of transparency – all of these issues require further research, and to consider which uses of AI are impermissible? Where should we draw the lines?

EDRi suggests the following::

- Funding for EU projects on AI should be conditional on meeting the EU's own ethical standards for AI and fundamental rights laws.
- EU funds, such as the Horizon2020 fund, should comply and immediately stop funding for projects which pose a risk to fundamental rights, such as iBorderCtrl - which aims to use facial and emotion recognition technology to supposedly detect lies in the course of visa applications, but is not substantiated by scientific evidence and significantly infringes upon human dignity.
- Other projects which could facilitate mass surveillance should be similarly ceased, as highlighted in EDRi's paper Ban Biometric Mass Surveillance.

Focusing on Small and Medium Enterprises (SMEs)

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

There are no preferred answers here from a digital rights perspective. However, watch out for the implicit assumption that AI should be 'promoted' to small and medium sized businesses – see next question.

EDRi suggests **1** for 'help raise SME awareness about potential benefits'. We want to challenge the idea that automation or technology in general is **in itself** a solution to deep-seated societal problems. This idea, often referred to as 'technosolutionism', can be damaging, which is why we argue that AI should not be promoted in the absence of scientific evidence, democratic engagement, safeguards.

For more information, the AI White Paper proposes support and finance to small and medium business to enable them to take up AI. They want to to this via Digital Innovation Hubs, which are centres for business advice, skills training and other business support.

Are there any other tasks that you consider important for specialised Digital Innovations Hubs?

Here we make the point that small businesses should not enjoy any exemptions to protecting human rights. Some of the biggest AI-related scandals have involved small companies, and so the EU should ensure that when small businesses to take up AI, whatever the circumstances, they should respect data protection, privacy and other fundamental rights. There should be no blanket exemptions to fundamental rights standards.

One of those examples is Clearview AI – a relatively small company or 'startup' which famously mined data – mainly photos online – to build a vast facial recognition database to sell to law enforcement agencies without peoples' consent.

SECTION 2 – *An Ecosystem of Trust*

For those that care about tech, digital and human rights – this is the juicy section. This is where we highlight risks but also suggestions for a system which protects rights.

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

All of these options here present major concerns. EDRi have scored **5** for all except 'AI is not always accurate', which we score **4**.

Accuracy is an important concern, however we want to avoid the common argument put forward by the tech and security industries, that the biggest problems of AI are to do with data and tool quality, all of which can be improved by further uptake, use and minor improvements in the technology. This can be very problematic in the case of AI, for example when suggestions of improvement rely on collecting even more data for training, which can have privacy implications, or which would allow some uses of AI to be even more accurate at surveillance, targeting and profiling people.

Do you have any other concerns about AI that are not mentioned above? Please specify:

So many concerns, so little space! Here we recommend focusing on any issues in particular that matter to you, whether it be concerns for certain groups (e.g. impact on migrants/ people of colour, people with disabilities, women, workers), in different sectors (e.g. AI use in policing, recruitment, healthcare), or simply general risks.

There are countless concerning examples. From the development of hiring tools proposing to identify disabilities, to worker surveillance software, tools claiming to predict high crime areas and the use of emotion recognition technology supposedly to detect lies in visa applications, there is an urgent need for the public to engage with conversations about the red lines for AI. What uses are impermissible? Which should we ban?

To know more about the main human rights risks linked to AI, read our short explainer here.

Here are the general concerns EDRi highlights (we use less words). For this question, it may be wise to focus on just a few concerns here or in the explainer:

- The deployment of AI in sensitive areas (public services) without democratic oversight, transparency or sufficient evidence to justify the need/ purpose.
- Increasing use of opaque, privately-developed technology in the public sphere , which do not meet transparency requirements (when they exist).
- The conscious avoidance of liability for harms produced by AI technology..
- AI posing collective or societal-level harms which don't have remedies in anti-discrimination of data protection frameworks (often because these frameworks focus on the individual).
- Companies and governments using excuse of 'innovation' to justify trials without safeguards.
- The characteristics of machine learning can lead to unauthorised use/ purpose and function creep.

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

EDRi recommends selecting '**other**' or '**no**'. Only if you select '**other**' do you get to explain using 500 characters.

Other, please specify:

EDRi selected '**other**' because, we believe that the GDPR must be reinforced, not undermined by any AI regulation.

We encourage you to emphasise this – AI regulation should not provide loop-holes to data protection legislation, or other frameworks, like discrimination law.

Saying this, we add:

- Current law does not address use of non-personal data for AI, types of data which do not fall under the GDPR
- AI can have huge collective impacts, such as furthering overpolicing, surveillance, inequalities – all not addressed in existing legal frameworks but are still major issues
- AI can lead to discrimination on financial status and other grounds which are usually not protected in discrimination law

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

Your answer here creates options:

1. If you select '**Yes**' an additional question will appear: '*do you agree with the approach to determine "high-risk AI applications proposed in section 5.B of the white paper?"*'
2. If you select' **no**', no additional question appears
3. If you select 'other' do you get to explain using 500 characters

We recommend selecting '**no**' or '**other**'. EDRi answers **Other**.

The European Commission's approach is to categorise all AI into two binary categories:High risk and low risk. The White Paper says that systems to be considered 'high risk' are decided according to (a) sector (without much information as to how 'high risk sectors are determined') and (b) if the application poses 'significant risks'.

We argue that this approach is over-simplified and narrow – it risks enabling some applications to avoid scrutiny. It also skips an important step – what are the red lines for AI? What are the legal boundaries within which this technology should operate?

Instead, EDRi writes:

1. new rules should clearly outline criteria to determine which AI systems are legal and which are not. Such criteria should be based on proving that they work and are needed, conducting mandatory fundamental rights impact assessment for all applications, and ensuring democratic oversight.
2. Uses of AI which breach fundamental rights - like biometrics/ facial recognition for mass surveillance - should be banned outright.

We include more about which systems we think should be banned two questions on.

conditional question **Do you agree with the approach to determine "high-risk" AI applications proposed in Section 5.B of the White Paper?**

You will only see this if you answered '**yes**' to the previous question and believe that the requirements for AI should be limited to high risk applications. It could be interesting to raise concerns about the sectoral criteria for 'high risk'. How will these criteria be decided? Will areas like policing, migration control, and other public sector areas be high-risk? What about social media platforms?

If the Commission maintains the high-risk/ low-risk distinction, fundamental rights and the human impacts should be the determining factors, rather than sector. More categories of risk other than just high and low might help to add nuance and ensure a more effective regulation system than this blunt binary choice

Therefore your answer could:

1. outline concerns with sectoral definitions of risk, particiualrly questioning by impact on people, human rights, safety are not the primary issue
2. challenge the Commission by showing that areas, *particularly public services*, such as healthcare, policing, migation, social welfare should all be considered very high risks
3. note that it is important that the public and civil society are consulted on what may constitute a high-risk use. And
4. highlight the overly simplistic nature of the high-low risk binary.

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

Here is the moment to highlight the most problematic uses of AI.

EDRi makes clear in its answer that **these uses are unacceptable because they are incompatible with fundamental rights**. They are beyond 'risk'- they should be banned. We call on the Commission to urgently rethink its risk-based approach and consider which uses of AI will never be acceptable, legal or complaint with human rights.

Here we include applications such as those involving predictive policing, autonomous lethal weapons, , the use of AI to determine access to or delivery of essential public services, and biometric processing of any kind for mass surveillance, as they are all incompatible with EU fundamental rights and values and should be banned by default.

EDRi's argues that these uses of AI are incompatible with human rights and should be banned:

- The use of AI to determine delivery of essential public services,
- predictive policing

- autonomous lethal weapons
- identification/ analysis of emotion and identity traits,
- and indiscriminate biometric surveillance.

Determining 'risk' should be rights and outcomes focused, not according to the sector. We also note a major concern with systems which impact fair trial, in migration control and policing, and systems which may perpetuate inequalities in hiring.

In your opinion, how important are the following mandatory requirements of a possible future regulatory framework for AI (as section 5.D of the White Paper) (1-6: 1 is not important at all, 6 is very important)?

We rated all of the requirements as *very* important (**6**).

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

We strongly recommend the answer 'biometric identification systems should never be allowed in publicly accessible spaces.'

This is a very important and increasingly relevant question as states and private actors across Europe increasingly use facial recognition, speaker recognition, and other forms of biometric processing in publicly-accessible spaces, posing a myriad of risks to our fundamental rights and freedoms, and potentially leading to mass surveillance. You can read more about why we need a ban on biometric mass surveillance here.

Please specify your answer:

Here we outline why we think this. There are a range of reasons to contest biometric identification and other biometric processing systems. Our reasons are:

1. Their use in public spaces will lead to mass surveillance;
2. This will irreversibly limit our fundamental rights to privacy, freedom of assembly, expression, non-discrimination, data protection, dignity and the right to a fair trial, creating societies of suspicion; and,
3. **Even uses which do not contribute directly or indirectly to mass surveillance in public spaces still pose significant threats to privacy, data protection, non-discrimination, and dignity.**

EDRi's Ban Biometric Mass Surveillance report outlines that biometric identification is unlawful and incompatible with human rights law. Societally, biometric identification is likely to un-

dermine peoples' freedom and ability to engage in public life, and that there is a risk that such systems will disproportionately be used to target already over-policed and surveilled groups, including racialised **groups**, migrants, working class and poorer communities.

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

The European Commission is proposing that 'high-risk' systems are subject to mandatory requirements. For everything else, developers or users of AI systems could choose to make themselves subject to requirements on a voluntary basis. They would then receive a quality label of some kind after passing the test.

EDRi selected '**rather not**' for this question for the reasons stated next.

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

EDRi strongly argues against voluntary labelling systems, because:

1. Self-labelling systems can be confusing for people and may give a false sense of security since it is the same company that develops a product the one saying that it is safe.
2. We believe that the high/low risk distinction is overly simplistic and could very well allow for loop-holes for systems with potentially very significant impacts on peoples' safety and rights. This is especially so if 'low risk' systems are only voluntarily controlled, this is essentially asking us to let big tech companies regulate themselves.
3. In addition, there is a concern about how those harmed by these low-risk systems might seek redress in such a scenario.

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

Here you can select multiple answers. The various options refer to different methods of regulation, including: regulations or requirements prior to putting AI systems on the market (ex ante), which is a more monitored method of regulation; ex post measures (after being put on the market); and introducing different measures depending on risk.

EDRi selected '**a combination of ex ante compliance and ex-post enforcement mechanisms**' and '**other**' - there is some value in external assessments but again, we are sceptical of the narrow 'high-risk' approach.

Selecting '**other**' also triggers a text box for further explanation.

Please specify any other enforcement system:

Here EDRi specifies that we think all systems should undergo a **mandatory ex**

ante human rights impact assessment from an external body.

This recommendation specifically addresses the need to make human rights a priority in the AI regulation, and ensure that there are no loop-holes just because a system falls into a low-risk category.

Do you have any further suggestion on the assessment of compliance?

Here EDRi emphasises that compliance should be external. We need this to guarantee fundamental rights are protected, we cannot rely on self-regulation for this.

SECTION 3 – *Safety and Liability Implications of AI, IoT and Robotics*

This section is about the potential impact of AI on the safety of products with digital components, and how to make sure that people can seek redress or compensation in the event of damage. The section is slightly less related to digital rights and privacy, so EDRi has less to say here!

If you are interested in more information about AI and consumer rights, take a look at the AI page of BEUC, the European consumer organisation.

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

We suggest that you select all options (**cyber risks, personal security risks, risk related to the loss of personal connectivity, mental health risks**) to encourage deeper reflection by legislators about the real life implications of AI, ensuring there is really a 'human centred' approach.

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

Just to make it really clear, we highlight here again the potential risks of discrimination posed by AI systems. In particular, the use of AI in online products and services requires collection and use of data lending toward discrimination in many fields related to targeted advertising. This poses risks of differentiated pricing, discrimination and financial detriments, the risk of creating filter bubbles, interfences in the political process, all based on sensitive inferences or associations.
EDRi also highlights that in certain cases AI may impact on accessibility and other rights of persons with disabilities.

Do you think that the safety legislative framework should consider new risk assessment

procedures for products subject to important changes during their lifetime?

We suggest that you select **yes**, as AI brings complexity to the development of products, and serves as products themselves, there is likely to be added complexity and opacity, and a strengthening of procedures of redress in the event of harm.

Do you have any further considerations regarding risk assessment procedures?

EDRi argues that internal supervisors, such as Data Protection Officers under GDPR should included and and asked for advice.

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

We suggest that you select **yes**, for the same reasons outlined in the previous queston.

Do you have any further considerations regarding the question above?

We think that AI developers and deployers should be accountable for harm generated by their products, and that products developed using AI should not enjoy exceptions to any EU laws, whether it be discrimination, data protection, or product liability.

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

We suggest that you select **yes for all applications** because at times, it can be difficult to demonstrate that harm or damage was linked to particular AI systems, because their code is protected by copyright protections. This prevents oversight of these applications in terms of their quality and their potential harm, and needs to be addressed.

Do you have any further considerations regarding the question above?

EDRi suggests that you highlight that the EU address copyright and database protections which prevent proper oversight of AI applications. Liability rules should provide incentives for openness.

Thank you for your contribution to this questionnaire. In case you want to share further ideas on these topics, you can upload a document below.

You're done! The questionnaire gives you an opportunity to upload other documents alongisde your answers. We recommend you do this if you work for an organisation or collective with a position on AI regulation.

READ MORE

If you would like to view EDRi's additional document with core recommendations for AI regulation, you can find them here.