Access Now's submission to the Consultation on the "White Paper on Artificial Intelligence - a European approach to excellence and trust"

JUNE 2020

TABLE OF CONTENTS

I. INTRODUCTION 1

II. DO NOT PROMOTE THE INDISCRIMINATE UPTAKE OF ARTIFICIAL INTELLIGENCE 2

III. IMPLEMENT A RIGHTS-BASED APPROACH AND MANDATORY HUMAN RIGHTS IMPACT ASSESSMENTS 3

IV. ESTABLISH PUBLIC REGISTERS FOR AI/ADM SYSTEMS 7

V. BAN APPLICATIONS THAT ARE INCOMPATIBLE WITH FUNDAMENTAL RIGHTS SUCH AS BIOMETRIC TECHNOLOGIES THAT ENABLE MASS SURVEILLANCE 8

VI. ESTABLISH NATIONAL CENTRES OF AI EXPERTISE TO HELP EXISTING REGULATORS 11

VII. ENFORCE HIGH SCIENTIFIC STANDARDS 13

VIII. CONCLUSION 14

I. INTRODUCTION

Access Now defends and extends the digital rights of users at risk around the world. By combining direct technical support, comprehensive policy engagement, global advocacy, grassroots grantmaking, and convenings such as RightsCon, we fight for human rights in the digital age.

Access Now's European Policy Manager, Fanny Hidvegi, as a member of the European Commission's High Level Expert Group on Artificial Intelligence, advises the EU on its strategy for the development of AI. Daniel Leufer is a Mozilla Fellow hosted by Access Now.

With the increasing investment in and proliferation of automation-based technologies, the EU must enforce and develop the highest human rights compliance standards for emerging technologies and AI systems that are designed, developed, or deployed in the European Union.

Access Now has put forward concrete policy objectives for EU lawmakers on artificial intelligence in our European Human Rights Agenda in the Digital Age. As part of the consultation process on the European Commission's 'White Paper on Artificial Intelligence - A European Approach,' Access Now will submit this document outlining our **six key recommendations**:

1. Do not promote the indiscriminate uptake of artificial intelligence
2. Implement a rights-based approach and mandatory human rights impact assessments
3. Ban applications that are incompatible with fundamental rights such as biometric technologies that enable mass surveillance
4. Establish national centres of AI expertise to help existing regulators
5. Establish public registers for AI/ADM systems
6. Enforce high scientific standards

II. DO NOT PROMOTE THE INDISCRIMINATE UPTAKE OF ARTIFICIAL INTELLIGENCE

One of the three 'pillars' of the European Commission's Coordinated Plan on AI is to boost "AI uptake across the economy, both by the private and public sector," and throughout the White Paper on AI we find repetitions of this mantra. Indeed, it seems that the main reason the White Paper gives for promoting trustworthy AI is because "lack of trust is a main factor holding back a broader uptake of AI." What we see here is a complete reversal of priorities: the EU wants to have more AI uptake, so it is willing to make some effort to ensure it is trustworthy by mitigating risks. Instead, the EU should earn people's trust for its AI initiatives by putting the protection of fundamental rights ahead of concerns about global competitiveness in AI. The primary objective should be to avoid individual and societal harms, not to mitigate them.

What the White Paper fails to consider two key points: firstly, that not all AI uptake is compatible with fundamental rights; secondly, that AI-based systems are not always the best solution for a given problem. Indeed, while AI systems certainly have the potential to increase efficiency and improve aspects of our lives, their benefits must be analysed on a case by case basis and backed up by evidence. If and when AI is compatible with our rights and freedoms and has demonstrated benefits, then it should be promoted. When the opposite is true, and it threatens our rights and freedoms or presents no evidence of real benefit, the EU should actively oppose its uptake.

In the assessment of potential benefits we should build in checks and balances to ensure that we do not ignore or exacerbate existing systems of oppression and the division between different parts of society. Inclusion starts with having a seat at the table, but it is not enough merely to be consulted. Decisions about AI investment and promotion must be made by the people and communities impacted and with the meaningful involvement of experts and researchers in human rights, poverty, social welfare, labour rights, discrimination (particularly racial and gender) and other relevant fields. The White Paper fails to put forward policy suggestions to address the specific risks, harms and violations against disenfranchised communities and the role of automated systems - and technology more broadly - in systems of oppression.

The uptake of any technology, particularly in the public sector, should not be a standalone goal or value in itself. In public procurement mechanisms (starting from the contract notice at the latest) AI solutions must be evaluated against non-AI approaches. In the Draft Guidelines for Public Procurement of AI published by the World Economic Forum and the UK's Office for AI, one of the fundamental guidelines provided is to focus on the challenge to be solved rather than on AI as a solution: in other words, not to prioritise AI uptake for its own sake. They underline that public authorities must evaluate whether they need an AI system to address a given issue, and further state that if "during the evaluation of the tender responses it becomes evident that another solution that does not incorporate AI is better able to address the problem, you should consider following this alternative delivery path." Public procurement rules must be improved and implemented to reflect human rights considerations and to meet transparency and open procurement requirements and standards.

The EU must stop being led astray by empty mantras and myths about promoting innovation at all costs or 'keeping up' in some dubious 'AI race' against the US and China. AI development is not a zero-sum game where we must embrace all AI applications or none; it is perfectly possible for the EU to remain competitive in certain branches and applications of AI while having the maturity and foresight to refuse to develop and deploy other branches and applications that threaten our rights. Applications of AI such as facial recognition pose such a threat to our rights that precaution must be put before innovation and competitiveness; in such cases, we need red lines rather than risk mitigation. If the EU wants to promote the idea of Trustworthy AI, it must demonstrate the resolve to not pursue the development and deployment of AI applications that undermine our rights.

III. IMPLEMENT A RIGHTS-BASED APPROACH AND MANDATORY HUMAN RIGHTS IMPACT ASSESSMENTS

Access Now believes that a human rights-based approach is essential to ensure the EU's attempt to build trustworthy AI actually deserves our trust and is not just an empty brand name. Where AI systems pose a threat to any of our fundamental rights, the EU must ensure that states uphold their obligation to protect and promote those rights and that companies conduct due diligence according to their responsibility.

Following a rights-based approach, the EU should consider the following regulatory levels:

1. Explicit bans

Adopt the **explicit policy objective to stop or ban applications** of automated decision-making (ADM) or AI systems in areas where mitigating any potential risk or violation is not enough and no remedy or other safeguarding mechanism could fix the problem. This approach is in line with the basic values of the European Union built on the Treaties and the EU Charter of Fundamental Rights. (See Section V below).

2. Mandatory human rights impact assessments for all applications

As opposed to a binary risk assessment approach, Access Now argues that for all applications in all domains, the burden of proof should be on the entity wanting to develop or deploy the AI system to demonstrate that it does not violate human rights via a mandatory human rights impact assessment (HRIA). This must apply both for the public sector and for the private sector, as part of a broader due diligence framework.

Regarding the private sector, we point to Mark Latonero's recommendations that "[t]echnology companies should find effective channels of communication with local civil society groups and researchers [and] should conduct Human Rights Impact Assessments (HRIAs) through the life cycle of their AI systems."[1] We also welcome the European Commission plans for a legislative proposal for mandatory human rights due diligence, and we hope that this consultation and process around AI systems will be taken into consideration and the Commission will have a consistent approach.

We furthermore support the Council of Europe's recommendations for private sector actors and AI systems, especially regarding human rights due diligence and human rights impact assessments:

> Human rights impact assessments should be conducted as openly as possible and with the active engagement of affected individuals and groups [...] the results of ongoing human rights impact assessments, identified techniques for risk mitigation, and relevant monitoring and review processes should be made publicly available, without prejudice to secrecy safeguarded by law.[2] When secrecy rules need to be enforced, any confidential information should be provided in a separate appendix to the assessment report. This appendix should be accessible to relevant supervisory authorities.[3]

[1] Mark Latonero, Governing Artificial Intelligence: Upholding Human Rights and Dignity, https://datasociety.net/wp-content/uploads/2018/10/DataSociety_Governing_Artificial_Intelligence_Upholding_Human_Rights.pdf

[2] The Council of Europe puts forward these specific recommendations for "deployment of high-risk algorithmic systems". We agree with the listed requirements but they should apply for all AI/ADM systems in case they are relevant. The result of a HRIA could be that there is no risk to mitigate.

[3] https://search.coe.int/cm/pages/result_details.aspx?objectid=09000016809e1154

Although the obligation to perform HRIAs should be mandatory across both public and private sectors, additional obligations must apply to public sector usage, including public-private partnerships, and, following the Council of Europe *Recommendation CM/Rec(2020)1 of the Committee of Ministers to member States on the human rights impacts of algorithmic systems*, to cases where "private sector actors provide services that rely on algorithmic systems and that are considered essential in modern society for the effective enjoyment of human rights."

Contrary to the idea that AI uptake must be prioritized in the public sector, it is here that the utmost precaution must be taken. Moreover, in most cases of public sector use of AI systems, the technology is provided by a private entity. Such public-private partnerships give companies access to sensitive data and have the potential to erode public sector ownership of vital resources. If private entities are contracted to deploy AI systems for the public sector, they must be subjected to the highest levels of scrutiny and transparency. The protection of trade secrets must not undermine public sector transparency, and legislation must therefore enforce high transparency requirements for public-private partnerships including: mandating publicly viewable prior and post-hoc conformity, human rights impact assessments and, as we will outline in the next section (Section IV), making a publicly viewable register of AI systems. We must also build fundamental rights considerations, alternative and parallel modelling, and testing into all phases of public procurement processes ("Preparation and Planning, Publication, Selection Evaluation and Award, Contract Implementation") and into technical specifications.

Regardless of the public or private sector nature of the entity carrying out the HRIA, it should meet specific criteria. While this list needs to be developed further, Access Now puts forward the following criteria:

- The process by which an actor determines the risks and impacts of an AI system must be reliable, verifiable, trustworthy, contestable and should be reassessed throughout the system's life cycle. If other actors, in particular those affected by a given system, determine that a system does in fact carry risks or has had a negative impact on them despite the result of the initial assessment, mechanisms must be in place to facilitate a contestation of the initial assessment.
- Depending on the outcome of the HRIA, different safeguards or next steps should be applicable. We promote in this regard the Council of Europe recommendation, *Unboxing artificial intelligence: 10 steps to protect human rights*[4] by the Human Rights Commissioner, and the *Recommendation CM/Rec(2020)1 of the Committee of Ministers to member States on the human rights impacts of algorithmic systems* that have detailed suggestions for human rights impact assessments[5].

[4] https://www.coe.int/en/web/commissioner/-/unboxing-artificial-intelligence-10-steps-to-protect-human-rights

[5] https://www.coe.int/en/web/freedom-expression/-/algorithms-and-automation-council-of-europe-issues-guidelines-to-prevent-human-rights-breaches

- Human rights impact assessments must be made as public as possible in all cases. Risks identified and measures taken to avoid or mitigate those risks must be documented, and updated throughout the lifecycle of the system. In cases where secrecy is absolutely essential, confidential information should be contained in a separate appendix, which if necessary can be made accessible to relevant supervisory and regulatory authorities. Furthermore, for all public sector cases and for private sector cases with significant effect, the information on HRIAs can be incorporated into the public register, as discussed in the next section (Section IV). Meaningful involvement of human rights experts and civil society is necessary in both the development of criteria for HRIAs and their audits and oversight.
- To make human rights impact assessments meaningful and enforceable, further measures are necessary for the private and the public sector. Among other mechanisms, the European Commission - as mentioned above - should address due diligence processes systematically and include this consideration in its reform of liability regimes. Furthermore, there may be a need to complement EU rules on public procurement to implement mandatory HRIAs.
- As a standard for public sector HRIAs we recommend relying on the well-developed human rights standards of necessity and proportionality and related concepts.
- In defining the potential outcomes of a HRIA, special consideration must be given to mandatory notification for affected individuals and groups. In this regard, we agree with the Council of Europe's Commissioner for Human Rights that an individual who has been subject to a decision by a public authority that is solely or significantly informed by the output of an AI system should be notified without delay.
- As we detail in Section VI, National Centres of Expertise should have an aiding role in the development of HRIAs alongside regulators and human rights enforcement bodies (including data protection authorities and equality bodies).
- The proposed AI register (see Section IV) should be used to make public specific aspects of the Human Rights Impact Assessments (HRIA) in addition to any supplementary assessments, such as Algorithmic Impact Assessments (AIA).

3. Safeguards and a risk-based approach

The outcome of human rights impact assessments should lead to different responses, safeguards assigned to the specific risks and impacts established in the process. In this context, we can envision a limited application for a risk-based approach as a secondary mechanism with assigned levels of safeguards and actions.

We caution against any regulatory model, however, that is based on a rigid distinction between high- and low-risk applications. As we have seen with the GDPR, a significant loophole is left open by allowing the data controller to determine whether a system poses a high risk and whether a data protection impact assessment is needed. We must therefore avoid a situation in which those responsible for deploying an AI system can shirk their responsibilities by ignoring risks.

In addition to the human rights considerations reflected in human rights impact assessments, the Commission must work towards clear and coherent criteria as to when an AI/ ADM system has significant effect on an individual, a specific group or society at large. In line with our colleagues at AlgorithmWatch, we suggest considering the following aspects: (a) the potential impact an ADM system has on people's life chances and social participation (e.g. e-recruiting systems as opposed to traffic lights); (b) the number of individuals concerned by a decision taken by an ADM; and (c) whether or not decisions are based on correlation or causality; correlation-based decisions obviously raise more concerns.

4. Personal and non-personal data in the context of AI/ADMs

The European Commission should dedicate resources to strengthen the enforcement of human rights and the GDPR. While AI presents issues for meaningful consent, objection, data minimisation, purpose limitation, and the right to explanation, a potential new legislation must not replace or undermine the GDPR.

Instead, it should complement the broad interpretation of the GDPR by including affinity profiling, sensitive inferences and the collective impact of data processing.

Finally, the European Commission should pay particular attention to harms, risk and violations that arise from the processing of non-personal data, inferred data or different forms of anonymisation, pseudo-anonymisation and aggregation. The protection of personal data is a necessary but not sufficient tool to address individual and collective harms.[6] A data protection law in itself cannot adequately address the use of non-personal data, nor can it address the collective impact of AI, such as furthering over-policing, surveillance, and inequality. The current legislative framework does not, for example, prohibit AI-driven discrimination on non-protected grounds such as someone's financial status.

Estelle Masse, Access Now's global lead on data protection serves on the European Commission Multistakeholder expert group on the Application of the GDPR. She will bring further comments on this specific issue to the attention of the group.

IV. ESTABLISH PUBLIC REGISTERS FOR AI/ADM SYSTEMS

Without the ability to know whether AI/ ADM systems are being deployed, all other efforts for the reconciliation of fundamental rights and AI/ ADM systems are doomed to fail. Access Now and AlgorithmWatch therefore jointly call for a mandatory disclosure scheme for AI/ADM systems deployed in the public sector. We **ask for legislation to be enacted at the EU level** to mandate that member states **establish public registers of AI/ADM systems**. They should come with the legal obligation for those responsible for the AI/ADM system to disclose and document the purpose of the system, an explanation of the model (logic involved) and the information on who developed the system. This

[6] https://luminategroup.com/posts/blog/data-isnt-the-new-oil-its-the-new-co2

information has to be made available in an easily-readable and accessible manner, including structured digital data following on a standardised protocol. Whereas disclosure schemes on AI/ADM systems should be mandatory for the public sector in all cases, these **transparency requirements should also apply to the use of AI/ADM systems by private entities, when an AI/ADM system has a significant effect on an individual, a specific group or society at large** (see Section III), and should include a mandatory notification requirement to the relevant authorities. As mentioned above, such registers can also be used to make public the results of Algorithmic Impact Assessments (AIA) / Human Rights Impact Assessments (HRIA) undertaken by public authorities. Consideration should also be given to whether the register should include specific details of HRIAs by private entities in case they have a significant effect.

The system should include a notification and coordination mechanism for relevant authorities and the potential new centre of expertise (see section VI below). The register system should complement the minimum standards provided by national freedom of information laws and transparency requirements of public procurement processes.

V. BAN APPLICATIONS THAT ARE INCOMPATIBLE WITH FUNDAMENTAL RIGHTS SUCH AS BIOMETRIC TECHNOLOGIES THAT ENABLE MASS SURVEILLANCE

Perhaps the greatest threat posed by technologies that fall under the umbrella of artificial intelligence comes from their ability to facilitate the mass surveillance of populations. Advances in machine learning in particular have significantly reduced the cost of processing huge amounts of data, including video footage, and have thereby reduced the cost of mass surveillance. However, mass surveillance of populations should never be allowed, no matter how efficient, cost effective or accurate it may become due to technological advances. Mass surveillance constitutes one of the most egregious violations of our fundamental rights and freedoms and technologies that enable it must be banned outright.

If the EU truly wants to show leadership in promoting rights-respecting, trustworthy AI, then it must ban the development and deployment of such applications of AI. The EU cannot 'remain in the race' with China and the US when it comes to developing mass surveillance technologies and still claim to promote trustworthy AI. Rather, the EU must establish red lines to ban applications of AI which are incompatible with fundamental rights.

For the establishment of red lines, we underline that a ban is necessary in areas where mitigating any potential risk or violation is not enough and no remedy or other safeguarding mechanism could fix the problem. Second, we point to the human rights standard of the necessity and proportionality test to develop a criteria for what falls in that category. This approach is in line with the basic values of the European Union built on the Treaties and the EU Charter of Fundamental Rights. While the cornerstone of our legal system is the enforcement by courts, in these areas both the European

Commission (as the guardian of the Treaties) and EU Member States have obligations to respect and promote human rights. The individual and societal impact of certain applications calls for conscious policy decisions at this very moment. Finally, we highlight the following applications that should be banned:

- indiscriminate biometric surveillance and biometric capture and processing in public spaces or installed on wearable devices;
- use of AI to solely determine access to or delivery of essential public services (such as social security, policing, migration control);
- uses of AI which purport to identify, analyse and assess emotion, mood, behaviour, and sensitive identity traits (such as race, gender, disability) in the delivery of essential services;
- uses of AI to make behavioural predictions with significant effect on people based on past behaviour, group membership, or other characteristics such as in predictive policing systems;
- use of AI systems at national borders or in testing on marginalised groups, such as undocumented migrants;[7]
- use for autonomous lethal weapons and other uses which identify targets for lethal force (such as law and immigration enforcement);
- use for general-purpose scoring of citizens or residents, otherwise referred to as unitary scoring or mass-scale citizen scoring; and
- applications of automation that are based on flawed scientific premises, such as inferring emotion from analysis of facial configuration or determining sexual orientation from analysis of face images.

To this end, Access Now has joined a network of 44 civil society organisations to call on EU bodies - including the European Commission, the European Parliament, plus all EU Member States - to ensure that such technologies are comprehensively and indefinitely banned in both law and practice. Given that the current regulatory and enforcement framework has not been successful in preventing Member States from deploying unlawful biometric mass surveillance systems, we urge the Commission to act now.

The use of biometric technologies for the untargeted mass processing of special categories of personal data, in particular biometric data in public places, creates serious risks of mass surveillance. This unjustifiably infringes on fundamental rights including privacy, data protection, equality, freedom of expression and information, freedom of assembly and association, due process and more. Such uses of biometric processing constitutes an objectification of people's intimate and personal qualities in a way that is so intrusive as to infringe on their human right to dignity.

The use of untargeted mass biometric processing systems - whether by law enforcement, public authorities (such as schools or local councils), or private actors - does not meet the required justifications or thresholds of necessity or proportionality to be considered lawful for the level of

[7] University of Toronto (2019). 'Bots at the Gate: A Human Rights Analysis of Automated Decision Making in Canada's Immigration and Refugee System'.https://citizenlab.ca/wp-content/uploads/2018/09/IHRP-Automated-Systems-Report-Web-V2.pdf

violation and intrusion they create. This threshold is demanded by the Charter of Fundamental Rights, the General Data Protection Regulation (GDPR) and the Law Enforcement Directive (LED). The legal frameworks within which such activities take place often do not meet the requirements of "prescribed for by law" established under the European Convention on Human Rights (ECHR) and the EU Charter of Fundamental Rights, and fail to provide adequate, effective remedies against untargeted, unnecessary, disproportionate surveillance.

We repeat here our call to permanently stop all biometric processing in public and publicly-accessible spaces, wherever it has the effect or potential effect to establish mass surveillance. This requires that:

1. EU Member States immediately halt all biometric processing that could amount to mass surveillance in public spaces, ensuring that both current and future deployments are included. This should be supported by a political debate by the European Council on the fundamental rights impacts of biometric mass processing in Member States;

2. EU Member States, under the auspices of the European Data Protection Board (EDPB) and national Data Protection Authorities (DPAs), publicly disclose all existing and planned activities and deployments that fall within this remit;

3. EU Member States cease all planned legislation which establishes biometric processing that could lead to mass surveillance in public spaces. Instead, clear and foreseeable laws should only allow for targeted identification checks that are proportionate to the issues and context, and provide for effective remedies against abuse. DPAs can play a role by advising national regulators and requesting action from their national governments;

4. The European Commission, in particular Directorate-General (DG) HOME and with reference to DG RTD for the Horizon2020 Programme, ensure that funding given to Member States for biometric research or deployment is for activities which are fully compliant with the Charter, including immediately ceasing all funding for biometric processing programmes which could contribute to mass surveillance in public spaces. All EU bodies who give operational support or advice to EU institutions, including but not limited to Europol, Frontex and the Fundamental Rights Agency (FRA), ensure that Member States cannot use these technologies in a manner that gives way to fundamental rights abuses;

5. The European Commission, under the auspices of the EDPS's advisory role, review and ex post facto evaluate on fundamental rights and data protection grounds all laws covering EU biometrics that contribute to or amount to mass surveillance and, as appropriate, recast, repeal or provide appropriate guidance to Member States about safeguards;

6. The European Commission (in particular DGs GROW, CNECT and JUST as the Directorate-Generals leading the Commission's work on the White Paper on Artificial Intelligence (AI) and DG HOME in its capacity on borders) implement, through legislative and non-legislative means and, if necessary, infringement proceedings and Court action, an immediate and indefinite ban on biometric processing

that leads to mass surveillance in public spaces. This process must be done under the supervision and/or support of the European Data Protection Supervisor (EDPS), the European Data Protection Board (EDPB), the FRA and DPAs. It is the role and responsibility of the European Union, in particular the European Commission, the Council of the EU and the European Parliament, with the support of the European Data Protection Board which also includes the European Data Protection Supervisor, the EU Fundamental Rights Agency (FRA), national Member States, the national Data Protection Authorities (DPAs) of every EU Member State and any other oversight bodies, to determine the appropriate methods to ensure that biometric mass surveillance is comprehensively stopped and banned in law, and in practice, across the EU.

VI. ESTABLISH NATIONAL CENTRES OF AI EXPERTISE TO HELP EXISTING REGULATORS

Access Now calls for the establishment of **independent centres of expertise on AI** on a national level to monitor, assess, conduct research, report on, and provide advice to government and industry in coordination with regulators, civil society, and academia about the societal and human rights implications of the use of AI/ADM/ systems. The overall role of these centres is to create a meaningful accountability system that links the objectives for the "ecosystem of excellence" and the "ecosystem of trust".

The centre should be an independent statutory body that would have a central role in coordination, policy development and national strategy relating to AI, and in helping build the capacity of existing regulators, government and industry bodies to respond to the increased use of AI systems.

These national centres of expertise can provide oversight for investment and research funding according to public interest criteria, as well as issuing guidelines for the development, procurement and deployment of AI systems in different sectors. The roles and responsibilities of these centres should include, but go beyond, what the White Paper suggests for the ecosystem of excellence. These institutions could have the role of testing centres (described in Action 2), contribute to the objectives of developing skills (Action 3) and carry out a quality control and advisory role for objectives related to investment (Action 4 and 5).

As we described in Section II, the promotion of AI uptake, particularly in the public sector, must be in line with the highest standards, including for public procurement. These centres would ensure that across domains (see Action 6). The HLEG's Policy and Investment Recommendations and its upcoming publication for sector-specific recommendations should follow similar principles.

Although these centres of expertise should not have regulatory powers, they can provide essential expertise to aid and coordinate among regulatory bodies, human rights bodies, data protection authorities, etc, in their work. They should also monitor and publish quarterly reports on AI enforcement across sectors. Individual complaints, collective redress mechanisms and ex officio

investigations related to violations or risks in the jurisdiction of existing human rights bodies or other regulators should remain in their powers.

The primary nature of these cases - even if they are about rule-based algorithms or other forms of automated decision-making systems as opposed to state of the art machine learning systems - is not about a specific technology but rather about how that technology is being used by individuals or institutions. These cases should be understood in the complexity of such mechanisms and how they impact individual and collective rights. Existing enforcement bodies are best placed to make these determinations. They are, however, already under-resourced for their existing tasks and responsibilities both in terms of capacity and specific needs in expertise. A potential centre of expertise would be suitable to address these needs across the public sector.

The proposed centres would also enable a systematic approach to the promotion and protection of human rights and would be able to complement the limitations of existing bodies focusing on individual human rights enforcement and could ensure the prevention of community and societal harms. The national centres could further be responsible for aiding regulatory and governance bodies in carrying out their functions in conjunction with new technologies. Furthermore, establishing an EU-level coordinator between these centres would allow for coordination and the publication of an annual report on the 'state of play' of AI across the EU.

The national centres of expertise **should further support small and medium-sized enterprises (SMEs) in fulfilling their obligations under human rights due diligence** (see Section III), including the aforementioned step of conducting a human rights impact assessment (HRIA,) and in registering ADM/AI systems in the public register discussed in Section IV. This support of SMEs can be bolstered by involving civil society organisations, stakeholder groups and existing enforcement bodies such as DPAs and National Human Rights Bodies, in the centres of expertise. Ultimately, such collaboration between diverse stakeholders will benefit all aspects of the ecosystem and build trust, transparency and cooperation between all actors.

Access Now welcomes the Australian Human Rights Commission's approach on this subject reflected in their discussion paper published in December 2019, entitled Human Rights and Technology.[8] We further welcome the initiative taken by the UK's Office for Artificial Intelligence to issue guidance on the use of AI in the public sector in their guidelines on Assessing if artificial intelligence is the right solution. In contrast to the White Paper's misguided promotion of indiscriminate AI Uptake, the UK guidelines urge public authorities to simultaneously consider non-AI solutions to problems as well as to consider the broad infrastructure and personnel requirements for deploying an AI system. The UK has established multiple institutions to address issues related to AI systems and with its departure from the EU has weakened the level of expertise, capacities and resources embedded and shared across EU and Member State institutions, and the EU must address this loss. One example is the

[8] https://tech.humanrights.gov.au/sites/default/files/2019-12/TechRights_2019_DiscussionPaper.pdf

Information Commissioner's Office (ICO) which was one of the most well-resourced data protection authorities that provided much-needed technical expertise for complex investigations.

VII. ENFORCE HIGH SCIENTIFIC STANDARDS

According to the High Level Expert Group's Ethics Guidelines for Trustworthy AI, it is essential that Trustworthy AI is robust, meaning that an AI system should "perform in a safe, secure and reliable manner, and safeguards should be foreseen to prevent any unintended adverse impacts."[9] Fundamental to this aim is that AI developed and deployed in the EU conforms to high scientific standards, as no system can be safe, secure and reliable if it is based on flawed scientific premises. Unfortunately, this has not been the case to date, not only with AI applications developed by private companies, but also with AI applications funded by the EU.

One example of this failure to conform to high scientific standards is the proliferation of companies and researchers developing so-called 'emotion detection' applications. Among those who have received EU funding for such developments are the SEWA (Sentiment Analysis in the Wild) project, funded under Horizon 2020, which aims to develop an advertisement recommendation engine based on automated sentiment analysis. Such projects aim to 'detect emotion' in a variety of ways, but typically claim to be able to detect emotion from an analysis of images or video of people's faces. This is known as the 'basic emotions' approach or 'common view.'

A recent review of scientific evidence for this view (Barrett et al. 2019) discusses real-world applications of the 'common view', among which they mention that "[t]echnology companies are investing tremendous resources to figure out how to objectively "read" emotions in people by detecting their presumed facial expressions, such as scowling faces, frowning faces, and smiling faces, in an automated fashion," further noting that "[s]everal companies claim to have already done it."

However, despite this enthusiasm from technology companies, the conclusion of Barrett et al. is that "the science of emotion is ill-equipped to support any of these initiatives."[10] The very premise of such emotion detection technologies is flawed, and given that there is no scientific evidence to support this type of automated emotion detection technology, the EU must stop funding initiatives which attempt to implement it.[11]

Beyond just spurious emotion detection, however, EU funding has supported other dubious or even pseudo-scientific applications of AI. Foremost among these is iBorderCTRL, a Horizon 2020-funded

[9] High Level Expert Group on AI, 'Ethics Guidelines for Trustworthy AI,' p.7

[10] Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional Expressions Reconsidered: Challenges to Inferring Emotion From Human Facial Movements. Psychological Science in the Public Interest, 20(1), 1–68. https://doi.org/10.1177/1529100619832930, p. 48

[11] In their 2019 annual report, AI Now have also called for such systems to be banned in high stakes use cases: "Regulators should ban the use of affect recognition in important decisions that impact people's lives and access to opportunities," https://ainowinstitute.org/AI_Now_2019_Report.pdf

project that aimed to create an automated border security system to detect deception based on facial recognition technology and the measurement of micro-expressions. In essence, the EU spent €4.5 million on a project that 'detects' whether a visitor to the continent is lying or not by asking them 13 questions in front of a webcam.

However, the historical practice of lie detection is lacking in substantial scientific evidence and the AI technologies being used here to analyse micro expressions are just as questionable. To make matters worse, the Commission has ignored the transparency criteria outlined in the HLEG's Ethics Guidelines by refusing to publish certain documents, including an ethics assessment, "on the grounds that the ethics report and PR strategy are "commercial information" of the companies involved and of "commercial value"."

Beyond these two examples of scientifically dubious AI applications, we have seen researchers and companies propose a range of pseudoscientific and dangerous applications: predicting sexual orientation from facial analysis; predicting criminality from personal data; and even predicting whether someone is a 'terrorist' or paedophile from facial analysis. European citizens and consumers will never be able to trust AI developed and deployed in the EU unless rigorous scientific standards are enforced to stop such applications.

VIII. CONCLUSION

Access Now welcomes the opportunity to submit a response to the public consultation on the "White Paper on Artificial Intelligence - a European approach to excellence and trust".

The regulatory and investment space around artificial intelligence and other technologies should be built on existing values enshrined in the Treaties and the EU Charter of Fundamental Rights. The ecosystem of excellence and trust should be created and maintained for and by people and **should not have the objective of indiscriminate uptake of any technology**. The two pillars must go hand in hand with accountability mechanisms for **high scientific standards** for investment, research, innovation and policymaking alike.

The European Union must create an ecosystem that is centered on individual and collective rights and that takes into account overall societal benefits as well as impacts on disenfranchised people and communities. The design, development and deployment of any technologies must respect human rights. In the context of AI/ADM/systems, the EU should promote a rights-based instead of a risk-based approach. **Mandatory human rights impact assessments** both in the public and private sector are among the key policy tools to achieve this objective. We also ask for legislation to be enacted at the EU level to mandate that member states **establish public registers of ADM systems used by the public sector (and in some cases by the private sector),** as without the ability to know whether AI/ ADM systems are being deployed, all other efforts for the reconciliation of fundamental rights and AI/ ADM systems are doomed to fail. Finally, the European Union should explicitly recognise that some applications, use cases or even risks of potential violations or abuses are in conflict with our values

and rights enshrined in the EU Charter of Fundamental Rights. Applications such as biometric recognition that enable or contribute to mass surveillance should therefore be banned, stopped and not pursued under any circumstances.

Access Now defends and extends the digital rights of users at risk around the world. By combining direct technical support, comprehensive policy engagement, global advocacy, grassroots grantmaking, and convenings such as RightsCon, we fight for human rights in the digital age.

For more information, please contact:
Fanny Hidvegi, Europe Policy Manager, Access Now, fanny@accessnow.org
Daniel Leufer, Mozilla Fellow at Access Now, daniel.leufer@accessnow.org