ALGORITHM WATCH

/ Position by AlgorithmWatch

AlgorithmWatch's response to the European Commission's proposed regulation on Artificial Intelligence – a major step with major gaps

April 2021

Disclaimer: Our analysis below is a first take and a provisional analysis of selected issues that we find particularly relevant to our work. A profound and comprehensive assessment of the 100+-page document certainly requires more time and in-depth analysis.

Yesterday, the European Commission unveiled its eagerly awaited proposal for the regulation of Artificial Intelligence (AI) in Europe. It represents the Commission's attempt to protect fundamental rights while encouraging innovation in the field of AI and make Europe's economy "fit for the future". Likely, the new regulation will profoundly shape AI regulation in the next decades, not only in Europe but across the globe – through its direct extraterritorial effects as well as through its standard-setting potential. The proposal builds upon the White Paper on Artificial Intelligence, which was published by the European Commission in February 2020 and which was followed by a public consultation phase until June 2020. Read here AlgorithmWatch's contribution to the consultation process.

Overall, the new proposal reflects a shift in the Commission's narrative that we welcome. Whereas the White Paper's narrative suggested a worrisome reversal of EU priorities, putting global competitiveness ahead of the protection of fundamental rights, the new regulation sets out with the prohibition of AI practices which it declares to be in breach with the Union values and fundamental rights protected under Union Law, and devotes with Title III an entire section to the regulation of high-risk systems. Measures in support of innovation are not discussed until Title V. The new narrative, however, should not obscure the fact that many parts of the proposed legislation have severe loopholes, which very much contradict the idea of a regulation that puts fundamental rights and public interest first. It is now up to the Parliament and the Council to correct these shortcomings and to enshrine sufficient safeguards. Below you find our comments on selected issues in more detail.

Missed opportunity to clearly draw red lines and close loopholes

Over the past weeks the European Commission faced growing pressure to explicitly ban biometric mass surveillance technologies and to clearly draw legislative red lines for AI based systems which violate fundamental rights. Dozens of civil society organisations and digital rights activists, among them AlgorithmWatch, urged the European Commission to substantively enhance fundamental rights protections in the upcoming AI regulation and to close existing loopholes[1]. The call was furthermore supported by 116 Members of the European Parliament via an open letter to President von der Leyen. Meanwhile, more than 47.000 European citizens have signed a petition for a ban on biometric mass surveillance practices as part of the Reclaim Your Face campaign, and the number continues to grow.

Unfortunately, these calls have not sufficiently been taken into account: Although the proposed regulation comprises an entire section on prohibited AI practices and provides in Article 5 for the prohibition of 'real-time' remote biometric identification systems in publicly accessible spaces by law enforcement authorities, social scoring of natural persons by public authorities likely to result in detrimental or unfavourable treatment, or the distortion of human behaviour through the use of subliminal techniques or through exploiting human vulnerabilities, there are too many worrisome exceptions and catches.

First, the prohibition of 'real-time' remote biometric identification systems only applies to systems which are used for law enforcement purposes in publicly accessible spaces, thus neither to systems used by other public authorities nor to those used by private actors. Evidently, the major risks to fundamental rights such systems come with are not limited to law enforcement purposes – a fact which the proposal does not sufficiently reflect. Second, there is a range of exceptions to this prohibition, listed in Article 5 of the proposal, creating a number of loopholes which authorities could try to exploit. For example, the use of real-time biometric identification systems can be allowed for the "prevention of a specific, substantial and imminent threat to the life or physical safety of natural persons or of a terrorist attack", the interpretation of which leaves wide discretionary power to the authorities. While judicial authorization is generally necessary for allowing the use of these systems in such exceptional situations, it can be postponed in cases of urgency. In our view, the narrow applicatory scope of this prohibition of real-time biometric identification does not sufficiently consider that the wide-scale use of such systems may not only violate individuals' fundamental rights but also pave the way for indiscriminate mass surveillance and undermine fundamental principles of democratic societies. Similarly, the prohibition of AI systems used for social scoring purposes is also limited to those deployed by public authorities. Again, private actors are kept out of the line of fire. Third, AI systems that are used for predictive policing and which clearly violate fundamental rights and Union values are not included in the prohibition list but only declared 'high-risk'.

[1] A first letter was sent in January 2021; a second letter in April 2021.

Classification and assessment of high-risk AI practices is insufficient

These shortcomings are further compounded by the fact that the second regulative "line of defence" for the protection of fundamental rights in the current proposal - the classification of AI practices as high-risk - also shows serious loopholes. We are pleased to see that the high-risk approach has improved compared to the White Paper: The list of high-risk AI practices set out in Annex III includes AI systems used for recruiting, to evaluate creditworthiness, to determine access to social benefits, for predictive policing,to control migration and to assist judicial interpretation. Furthermore, the misleading criterion 'sector' to determine high-risk AI practices has been removed and replaced by criteria which include the extent of the use of the AI application and its intended purpose, the number of potentially affected persons and their vulnerabilities, the dependency on the outcome and the irreversibility of harms, as well as the extent to which existing Union legislation provides for effective measures to prevent or substantially minimise those risks. Moreover, we welcome the fact that providers of high-risk AI systems are subjected to a variety of requirements in terms of transparency and risk-assessment which they need to fulfil before putting these systems into service, which creates an incentive to promote compliance-by-design approaches.

Yet, the devil is in the detail: Article 7 on the definition of adverse impact and harm remains vague and nebulous. As a result, it is neither clear on what basis the current list of high-risk applications has been compiled, nor what will be considered as sufficient ground with respect to future risk classification. Moreover, the clause "AI systems *intended* to be used" applied throughout Annex III grants wide scope for interpretation. Third, and most worrisome, Article 43 stipulates that only AI systems intended to be used for the 'real-time' and 'post' remote biometric identification of natural persons will be subject to third party conformity assessment, while all other AI systems classified as high-risk will be subject to self-assessment by the provider[2], including those systems used for predictive policing, migration control and recruitment. In our view, it is unacceptable to leave such an important assessment solely to corporate actors who have a great self-interest in the deployment of these systems.

Moreover, it cannot be ruled out that systems not classified ex ante as high-risk under the current proposal will turn out to have severe and detrimental impacts on individuals and societies. While there are some transparency obligations for specific systems not considered high-risk, namely those interacting with natural persons (such as chatbots), emotional recognition and biometric categorisation systems, as well as systems used to generate or manipulate content (such as 'deepfakes'), in our view, these obligations do not go far enough. Emotional recognition and biometric categorisation systems as well as those involving deepfakes are all applications which are likely to come with a high potential of severe harm on individuals and democratic societies. What is more, the scientific basis of especially emotional recognition systems is highly disputed. The already minimal

[2] According to the definition in Article 3(2) the 'provider of an AI system' is a natural or legal person, public authority, agency or body who develops or has developed an AI system and makes it available under its own name or trademark. In most cases this is a corporate actor.

transparency obligations applying to this set of AI systems are further weakened by a range of exemptions (such as for the use of chatbots or biometric categorisation systems for the prevention, investigation or prosecution of crimes).

EU database is a first step towards greater transparency

While we are concerned about the lack of clarity and transparency with respect to the high-risk classification, we would like to applaud the European Commission for introducing the idea of an EU database for high-risk AI practices. According to Article 60, the database is supposed to contain data on all stand-alone AI systems that are considered high-risk, and all information processed in the database shall be accessible to the public. This is a promising first step towards greater transparency and corresponds to our calls for public registers and improved data access for public interest research. In order to make full use of the transparency potential of publicly accessible registers, the EU database, however, should not only include data made available by providers of high-risk systems but be complemented by a list of all AI-systems that are in use by public authorities regardless of their assigned risk level[3]. In our view, the information provided must include the purpose of the system, an explanation of the model (logic involved) and details on who developed the system, as well as the results of any algorithmic impact assessment / human rights impact assessment undertaken by public authorities. These requirements should apply to all systems regardless of their purpose. We believe that exceptions for certain areas of application, such as those provided for in Annex 8(11), are not the right way to go. Furthermore, it is important that the information be available in an easily-readable and accessible manner.

EU AI Board: Strengthening oversight across the Union?

With regard to the enforcement of the regulation, it is in our view critical that appropriate and reliable accountability frameworks be set up. According to the proposal, the enforcement of the regulation lies to a great extent with Member States – similar to what is the case for the GDPR. However, the proposal also suggests the creation of a European Artificial Intelligence Board (EAIB), which comprises one national supervisory authority per EU country, the EU's Data Protection Supervisor, and a European Commission representative who chairs the Board. The Board’s mandate is to supervise and facilitate the consistent application of the legislation and to share best practices. The new structure is very reminiscent of the European Board for Digital Services, which the Commission suggested as part of the DSA. Both boards strengthen the Commission’s oversight and supervision power. In the event that national competent authorities either lack sufficient expertise or resources, or are unwilling to do so, the Commission itself can intervene to secure a consistent application of the law.

[3] For further information on the idea of public registers of AI systems used by the public sector see our joint call with Access Now.

While some may read this as an attempt to undermine the EU’s Data Protection Supervisor, it can also be interpreted as a signal that the Commission is serious about harmonizing enforcement and making the deployment of AI systems more accountable, and that the Commission is not afraid of confronting Member States and corporate actors alike. Whether or not this is a sufficient and effective way remains to be seen.

Put people first

The Commission set out with the claim to set up a human-centric AI Framework that puts people first. The more surprising it is that the current proposal completely ignores the perspective of those affected by the output of AI systems. If such systems have consequential effects on people's lives, they must not only be granted transparency with regard to the deployment of these systems but also have the possibility to challenge outcomes. Thus, there must be easily accessible and legally guaranteed options for affected individuals and groups to contest such decisions and, where appropriate, to demand reversal, reconsideration through a different procedure, or compensation. Mere technological solutions do not suffice in order to ensure that AI systems are used to the benefit of the many, not the few. Accountability frameworks, empowering those directly affected by such systems, are an essential aspect in this regard. We sincerely hope that the Parliament and the Council will work towards ensuring that those with less bargaining power have a voice, too.