



INT/894
White Paper on Artificial Intelligence

OPINION

European Economic and Social Committee

White paper on Artificial Intelligence – A European approach to excellence and trust
[COM(2020) 65 final]

Rapporteur: **Catelijne MULLER**

Referral	Commission, 09/03/2020
Legal basis	Article 304 of the Treaty on the Functioning of the European Union
Section responsible	Section for the Single Market, Production and Consumption
Adopted in section	25/06/2020
Adopted at plenary	15/07/2020
Plenary session No	553
Outcome of vote (for/against/abstentions)	207/0/6

1. Conclusions and recommendations

- 1.1 The EESC congratulates the Commission for its strategy, laid out in the White Paper on Artificial Intelligence (AI), to encourage the uptake of AI technologies while also ensuring their compliance with European ethical norms, legal requirements and social values.
- 1.2 The EESC also welcomes the aim to capitalise on European strengths in industrial and professional markets and stresses the importance of enhancing **investment, infrastructure, innovation and skills** so that businesses, including SMEs, and society as a whole can seize the opportunities of AI. AI innovation should be fostered to maximise the benefits of AI systems, while at the same time preventing and minimising their risks.
- 1.3 However, it considers the focus on mere data-driven AI too narrow to make the EU a true leader in cutting-edge, trustworthy and competitive AI. **The EESC urges the Commission to also promote a new generation of AI systems that are knowledge-driven and reasoning-based, and that uphold human values and principles.**
- 1.4 The EESC calls on the Commission to: (i) **foster multidisciplinary research**, by involving other disciplines such as law, ethics, philosophy, psychology, labour sciences, humanities, economics, etc.; (ii) **involve relevant stakeholders** (trade unions, professional organisations, business organisations, consumer organisations, NGOs) in the debate around AI and as equal partners in EU-funded research and other projects such as the Public Private Partnership on AI, sector dialogues, and the *Adopt AI* programme in the public sector and the lighthouse centre; and (iii) keep **educating and informing the broader public** on the opportunities and challenges of AI.
- 1.5 The EESC urges the Commission to consider in more depth the **impact of AI on the full spectrum of fundamental rights and freedoms**, including – but not limited to – the right to a fair trial, to fair and open elections, and to assembly and demonstration, as well as the right not to be discriminated against.
- 1.6 The EESC continues to **oppose the introduction of any form of legal personality for AI**. This would hollow out the preventive remedial effect of liability law and poses a serious risk of moral hazard in both the development and use of AI, where it creates opportunities for abuse.
- 1.7 The EESC asks for **a continuous, systematic socio-technical approach**, looking at the technology from all perspectives and through various lenses, rather than a one-off (or even regularly repeated) prior conformity assessment of high-risk AI.
- 1.8 The EESC warns that the "high-risk" sector requirement could exclude many AI applications and uses that are intrinsically high-risk, in addition to biometric recognition and AI used in recruitment. The EESC recommends that the Commission draw up a list of **common characteristics of AI applications or uses that are considered intrinsically high risk**, irrespective of the sector.

- 1.9 The EESC strongly suggests that any use of biometric recognition only be allowed: (i) if there is a scientifically proven effect, (ii) in controlled environments, and (iii) under strict conditions. **The widespread use of AI-driven biometric recognition for surveillance or to track, assess or categorise humans or human behaviour or emotions, should be prohibited.**
- 1.10 The EESC advocates **early and close involvement of the social partners** when introducing AI systems at workplaces, in line with the applicable national rules and practices, in order to ensure that systems are usable and comply with worker rights and working conditions.
- 1.11 The EESC also advocates early and close involvement of those employees that will ultimately be working with the AI system, as well as employees with legal, ethical and humanities expertise, when introducing AI systems, in order to ensure that systems align with the law and ethical requirements, but also with workers' needs, so that workers retain autonomy over their work and AI systems that enhance workers' skills and work satisfaction.
- 1.12 **AI techniques and approaches used to fight the coronavirus pandemic should be robust, effective, transparent and explainable. They should also uphold human rights, ethical principles and existing legislation, and be fair, inclusive and voluntary.**
- 1.13 The EESC calls on the Commission to assume a leadership role so as to ensure better coordination within Europe of applied AI solutions and approaches used to fight the coronavirus pandemic.

2. **EU White Paper on AI**

- 2.1 The EESC is pleased to note that the European Commission takes up many of the recommendations from earlier EESC opinions and the High-Level Expert Group on AI, encouraging the uptake of AI technologies while also ensuring their compliance with European ethical norms, legal requirements and social values, underpinned by what it calls an "ecosystem of excellence and of trust".
- 2.2 The EESC welcomes the proposals aimed at businesses, including SMEs, and society as a whole seizing the opportunities of the development and use of AI. The EESC stresses the importance of enhancing investment, infrastructure, innovation and skills to improve the EU's competitive success at global level.

Human-in-command approach

- 2.3 The White Paper is, however, also slightly "fatalistic" in tone, suggesting that AI "overcomes us" which, leaves us no other option than to regulate its use. The EESC truly believes in the EU's commitment to ensure that Europe only accepts AI that is trustworthy and therefore should dare to take a much stronger stance here. The EESC thus urges the Commission to always keep the option open at all times of not accepting a certain given type of AI (-use) at all. This is what the EESC has been calling the **"human-in-command"** approach to AI that we need to cultivate.

Capitalising on AI in Europe – a forward-looking definition

- 2.4 The working definition of AI in the White Paper is "a collection of technologies that combine data, algorithms and computing power". Later in the text, **data** and **algorithms** are defined as the main elements of AI. However, that definition would cover any piece of software ever written, not just AI. There is still no universally excepted definition of AI, which is a generic term for a range of computer applications.
- 2.5 The **mere focus of the White Paper on data-driven AI is too narrow to make the EU a true leader in cutting-edge, trustworthy and competitive AI**. The White Paper excludes many promising AI systems from consideration, and thus from being governed and regulated. The EESC urges the Commission to also promote a new generation of AI systems that integrate data-driven approaches with **knowledge-driven, reasoning-based approaches**, so-called hybrid systems. The White Paper does acknowledge the need for hybrid systems for purposes of **explainability**, but the advantages of hybrid systems go beyond explainability: they can speed up and/or restrain learning, and validate and verify the machine-learning model.
- 2.6 The White Paper focuses only on bias in relation to data, but not all biases are the result of low-quality or limited data. **The design of any artefact is in itself an accumulation of biased choices**, ranging from the inputs considered to the goals set to optimise for. All these choices are in one way or another driven by the inherent biases of the person(s) making them.
- 2.7 Most importantly, however, AI systems are more than just the sum of their software components. **AI systems also comprise the socio-technical system around them**. When considering AI governance and regulation, the focus should thus also be on the ambient social structures around it: the organisations and enterprises, the various professions, and the people and institutions that create, develop, deploy, use, and control AI, and the people that are affected by it, such as citizens in their relations with governments, businesses, consumers, workers, and even society as a whole.
- 2.8 It should also be noted that **legal definitions (for the purpose of governance and regulation) differ from pure scientific definitions**, whereas a number of different requirements must be met, such as inclusiveness, preciseness, permanence, comprehensiveness, and practicability. Some of these are legally binding requirements and some are considered good regulatory practice.

Bringing all forces together

- 2.9 The EESC welcomes the effort to address the fragmented AI landscape in Europe by bringing together AI researchers, focusing on SMEs and partnering with the private and public sectors. In addition, the EESC would recommend: (i) fostering multidisciplinary in research, by involving other disciplines such as law, ethics, philosophy, psychology, labour sciences, humanities, the economy, etc.; (ii) involving relevant stakeholders (trade unions, business organisations, consumer organisations, NGOs) in the debate on AI, but also as equal partners in EU-funded research and other projects such as the Public Private Partnership on AI, the sector dialogues,

the *Adopt AI* programme in the public sector and the lighthouse centre; and (iii) continuing to educate and inform the broader public on the opportunities and challenges of AI.

AI and the law

- 2.10 The White Paper acknowledges the fact that **AI does not operate in a lawless world..** The EESC particularly welcomes the emphasis on the implications of AI for fundamental rights and recommends that the Commission considers more in-depth the AI impacts on a broad set of fundamental rights and freedoms such as freedom of speech and expression, and the right to respect for private life (which goes far beyond protecting people's data), to a fair trial, to fair and open elections, to assembly and demonstration, and to not be discriminated against.
- 2.11 The EESC welcomes the clear stance taken in the White Paper on the applicability of existing liability regimes to AI and the effort to build on those regimes so as to address the new risks AI can create, tackling enforcement lacunae where it is difficult to determine the actual economic operator responsible, and making regimes adaptable to the changing functionality of AI systems.
- 2.12 The Commission should also recognise that AI knows no borders and that the efforts cannot and should not be confined to Europe. A general worldwide consensus should be reached, drawing on discussions and research by legal experts, in an effort to establish a common international legal framework.
- 2.13 In any case the EESC **continues to firmly oppose the introduction of any form of legal personality for AI.** This would hollow out the preventive remedial effect of liability law and poses a serious risk of moral hazard in both the development and use of AI, where it creates opportunities for abuse.

Regulating high-risk AI

- 2.14 The EESC welcomes the risk-based approach to controlling the impacts of AI. The Commission announces a regulatory framework for "high-risk AI", which would need to comply with requirements regarding robustness, accuracy, reproducibility, transparency, human oversight, and data governance. According to the White Paper, two cumulative elements constitute high-risk AI: (i) a high-risk sector and (ii) high-risk use of an AI application. The White Paper adds two examples of AI applications or uses that could be considered intrinsically high-risk, i.e. irrespective of the sector. It also qualifies biometric recognition as an intrinsically high-risk application. The exhaustive list of high-risk sectors (whilst periodically reviewed) now includes the following sectors as potentially high-risk: healthcare, transport, energy, and parts of the public sector.
- 2.15 The second criterion, that the AI application is used in a risky manner, is looser, suggesting that different risk levels could be considered. The EESC suggests adding society and the environment as impact areas here.
- 2.16 Following the White Paper's logic, a **high-risk** AI application used in a **low-risk** sector will in principle not be subject to the regulatory framework. The EESC stresses that undesirable

adverse effects of high-risk AI in a low-risk sector could exclude AI applications or uses from regulation, providing a "window" for circumventing rules: think of targeted advertising (a low-risk sector), which has been shown to have potentially segregating, discriminatory and dividing effects, for example during elections or with personalised pricing (a high-risk use or effect). **The EESC recommends drawing up common characteristics of AI applications or uses that are to be considered high risk "as is", irrespective of the sector in which it is being used.**

- 2.17 While the EESC acknowledges the need for conformity testing of AI, it fears that a one-off (or even a regularly repeated) **prior conformity assessment** will not suffice to guarantee the trustworthy and human-centric development, deployment and use of AI in a sustainable manner. **Trustworthy AI needs a continuous, systematic socio-technical approach**, looking at the technology from all perspectives and through various lenses. For policy-making, this requires a multidisciplinary approach where policy-makers, academics from a variety of fields, the social partners, professional organisations, professionals, businesses, and NGOs work together continuously. Especially when it comes to public interest services related to the health, safety and well-being of people and based on trust, it has to be guaranteed that AI systems are adapted to practical requirements and cannot overrule human responsibility.

Biometric recognition

- 2.18 The EESC welcomes the Commission's invitation to open a public debate on the use of AI-driven biometric recognition. Biometric recognition of micro-expressions, gait, (tone of) voice, heart rate, temperature, etc. is already being used to assess or even predict our behaviour, mental state, and emotions, including in recruitment processes. To be very clear, **no sound scientific evidence exists to suggest that a person's inner emotions or mental state can be accurately "read" from their facial expression, gait, heart rate, tone of voice or temperature, let alone that future behaviour could be predicted by it.**
- 2.19 **It should also be noted that the GDPR only restricts the processing of biometric data to some extent.** The GDPR defines biometric data as "personal data resulting from specific technical processing relating to the physical, physiological or behavioural characteristics of a natural person, which allow or confirm the unique identification of that natural person". Many biometric recognition technologies, however, are not designed to uniquely identify a person, but only to assess a person's behaviour or emotions. These uses might not fall under the definition of biometric data (processing) under the GDPR.
- 2.20 AI-driven biometric recognition also affects our broader right to respect for private life, identity, autonomy and psychological integrity by creating a situation in which we are (constantly) being watched, followed and identified. **This could have a psychological "chilling effect", where people might feel inclined to adapt their behaviour to a certain norm.** This constitutes an invasion of our fundamental right to privacy (moral and psychological integrity). Furthermore, AI-driven biometric recognition could affect other fundamental rights and freedoms, such as freedom of assembly and the right not to be discriminated against.
- 2.21 The EESC recommends that any use of biometric recognition **only be allowed if there is a scientifically proven effect, in controlled environments, and under strict conditions.**

Widespread use of AI-driven biometric recognition to conduct surveillance, track, assess or categorise humans or human behaviour or emotions should not be allowed.

Impact of AI on work and skills

- 2.22 The EESC notes that the White Paper lacks a strategy on how to address the impact of AI on work, whereas this was an explicit element of the 2018 European Strategy on Artificial Intelligence.
- 2.23 The EESC advocates **early and close involvement of workers and service providers of all types, including freelancers, the self-employed and gig workers** – not just people who design or develop AI, but also those who purchase, implement, work with or are affected by AI systems. **Social dialogue must take place before** the introduction of AI technologies in the workplace, in line with the applicable national rules and practices. In the workplace, access to and governance of worker data should be guided by principles and regulations negotiated by the social partners.
- 2.24 The EESC would like to draw special attention to **AI used in hiring, firing and worker assessment and evaluation processes**. The White Paper mentions AI used in recruitment as an example of a high-risk application that would be subject to regulation irrespective of the sector. The EESC recommends extending this area of use to include AI used in firing and in worker assessment and evaluation processes, but also to explore the common characteristics of AI applications that would entail a high-risk use in the workplace, irrespective of the sector. AI applications that have no scientific basis, such as emotion detection through biometric recognition, should not be allowed in workplace environments.
- 2.25 The maintenance or acquisition of AI skills is necessary in order to allow people to adapt to the rapid developments in the field of AI. But policy and financial resources will also need to be **directed at education and skills** development in areas that will not be threatened by AI systems (i.e. tasks in which human interaction is vital, such as public interest services related to the health, safety and well-being of people and based on trust, where humans and machines cooperate, or tasks we would like human beings to continue doing).

3. AI and coronavirus

- 3.1 AI can contribute to gaining a better understanding of coronavirus and COVID-19, as well as protect people from exposure, help find a vaccine, and explore treatment options. But it is still important to be open and clear about what AI can and cannot do though.
- 3.2 **Robustness and effectiveness:** data-driven AI to forecast the spread of coronavirus is potentially problematic, because there is too little data about coronavirus for AI to have reliable outcomes. Moreover, the little data that has become available is incomplete and biased. Using this data for machine-learning approaches could lead to many false negatives and false positives.

- 3.3 **Transparency** on the data and the models used, as well as **explainability** of outcomes, are paramount. At this moment in particular the world cannot afford to take decisions based on "black boxes".
- 3.4 In using AI to combat this pandemic, **respect for human rights, ethical principles and existing legislation** are more important than ever. In particular when AI tools potentially infringe on human rights, there must be a legitimate interest for their use, which must be strictly necessary, proportionate and, above all, time-limited.
- 3.5 Finally, we need to ensure **fairness and inclusion**. The AI systems being developed to fight the pandemic should be bias-free and not discriminate. Moreover, they should be available to all and take account of the societal and cultural differences of the different countries affected.

Track-and-trace and health-monitoring apps

- 3.6 According to virologists and epidemiologists, opening up society and the economy from lockdown requires efficient tracking, tracing, monitoring and protecting of people's health. Currently, many **apps** are being developed for tracking, tracing and performing health checks, activities that have usually (and historically) been carried out by professionals. Worldwide, many governments have placed a large amount of trust in tracking and tracing apps as a means of opening up societies again.
- 3.7 The deployment of these kinds of apps is a very radical step. It is therefore important to critically examine the **usefulness, necessity and effectiveness** of the apps, as well as their societal and legal impact, before a decision is made to use them. There must still be the option of not using the apps, and less invasive solutions should be prioritised.
- 3.8 The **effectiveness and reliability** of tracking and tracing apps is extremely important, because ineffectiveness and unreliability can lead to many false positives and false negatives, a false sense of security, and thus a greater risk of contamination. Initial scientific simulations raise serious doubts as to whether a tracking app will have any positive effect on the spread of the virus at all, even with 80% or 90% use. Also, an app cannot register specific circumstances, such as the presence of plexiglass and windows or wearing of personal protective equipment.
- 3.9 Moreover, **these apps lead to the (partial) setting aside of various human rights and freedoms**, as they touch on our freedom of association, right to safety, to non-discrimination, and to privacy.
- 3.10 While very important, privacy is about much more than our personal data and anonymity. Privacy is also about the right not to be followed, tracked, and put under surveillance. It has been scientifically proven that when people know they are being followed, they start to behave differently. According to the European Court of Human Rights, this "chilling effect" is an invasion of our privacy. The same broad concept of privacy should be included in the AI debate.
- 3.11 There is a risk that data collected (now or in the future) will not only be used to fight the current pandemic, but also to profile, categorise and score people for different purposes. In the more

distant future it is even possible to imagine that "**function creep**" could lead to unwanted types of profiling in supervision and surveillance, acceptance for insurance or social benefits, hiring or dismissal, etc. The data collected using such apps may therefore under no circumstances be used for profiling, risk scoring, classification, or prediction.

- 3.12 Moreover, **any AI solution deployed under these extraordinary circumstances and even with the best of intentions, will set a precedent**, whether we like it or not. Previous crises have shown that, despite every good intention, such measures will in practice never go away.
- 3.13 The use of AI during this pandemic should thus always be measured and weighed against several considerations, such as: (i) is it effective and reliable? (ii) do less invasive solutions exist? (iii) do its benefits outweigh societal, ethical and fundamental rights concerns? and (iv) can a responsible trade-off be achieved between conflicting fundamental rights and freedoms? Moreover, these kinds of systems **may not be deployed under any form of obligation or coercion**.
- 3.14 The EESC urges policy-makers **not to succumb to techno-solutionism too readily**. Given the gravity of the situation, we recommend that applications linked to projects designed to help control the pandemic be grounded in sound research in epidemiology, sociology, psychology, law, ethics and systems sciences. Before deciding on the use of these systems, efficacy, necessity and sensitivity analysis and simulations need to be conducted.

Brussels, 15 July 2020

Luca JAHIER

The president of the European Economic and Social Committee
