III

(Preparatory acts)

EUROPEAN ECONOMIC AND SOCIAL COMMITTEE

547TH EESC PLENARY SESSION, 30.10.2019-31.10.2019

Opinion of the European Economic and Social Committee on 'Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions — Building trust in human-centric artificial intelligence'

(COM(2019) 168 final)

(2020/C 47/09)

Rapporteur: **Franca SALIS-MADINIER**

Referral	European Commission, 3.6.2019
Legal basis	Article 304 of the Treaty on the Functioning of the European Union
Section responsible	Single Market, Production and Consumption
Adopted in section	18.10.2019
Adopted at plenary session	30.10.2019
Plenary session No	547
Outcome of vote (for/against/abstentions)	198/1/4

1. **Conclusions and recommendations**

1.1.   Artificial intelligence (AI) is not an end in itself, but a tool that can deliver far-reaching positive change and involve risk, which is why its use must be regulated.

1.2.   The Commission should take measures with regard to forecasting, preventing and prohibiting the malicious use of AI and machine learning and better regulate the placing of products with malicious intent on the market.

1.3.   The EU should, in particular, promote the development of AI systems that focus on specific applications to speed up the ecological and climate transition.

1.4.   It is important to identify which challenges can be met by means of codes of ethics, self-regulation and voluntary commitments and which need to be tackled by regulation and legislation supported by oversight and, in the event of non-compliance, penalties. AI systems must always comply with existing legislation.

1.5.   AI requires an approach which covers technical as well as societal and ethical aspects. The EESC is pleased that the EU intends to build a human-centric AI approach which is in line with its fundamental values: respect for human dignity, freedom, democracy, equality and non-discrimination, the rule of law and respect for human rights.

1.6.   The EESC reiterates (¹) the need to consult and inform workers and their representatives when AI systems are introduced that are likely to alter the way work is organised, supervised and overseen, as well as worker evaluation and recruitment systems. The Commission must promote social dialogue with a view to involving workers in the uses of AI systems.

(¹) OJ C 440, 6.12.2018, p. 1.

1.7. The EESC stresses (²) that trustworthy AI presupposes that humans have control of machines and that citizens are informed about its uses. AI systems must be explainable or, where this is not possible, citizens and consumers must be informed about their limitations and risks.

1.8. The EU needs to address the emerging risks (³) in the area of health and safety in the workplace. Standards must be established to avoid autonomous systems causing harm or damage to people. Workers must be trained to work with machines and to stop them in an emergency.

1.9. The EESC calls for the development of a robust certification system based on test procedures that enable companies to state that their AI systems are reliable and safe. The transparency, traceability and explainability of algorithmic decision-making processes are a technical challenge which needs to be supported by EU instruments such as Horizon Europe.

1.10. Privacy and data protection will determine how far citizens and consumers trust AI. Data ownership and the control and use of data by companies and organisations have yet to be resolved (particularly in relation to the Internet of Things). The EESC urges the Commission to review the General Data Protection Regulation (GDPR) (⁴) and related legislation on a frequent basis in the light of developments in technology.

1.11. The EESC believes that consideration must be given to the contribution that AI systems can make to reducing greenhouse gas emissions, particularly in industry, transport, energy, construction and agriculture. It calls for the climate and digital transitions to be interlinked.

1.12. The EESC believes that oversight of AI systems may not be sufficient to define who is responsible and build trust. The EESC recommends that, as a priority, clear rules be drawn up assigning responsibility to natural persons or legal entities in the event of non-compliance. The EESC also calls on the Commission, as a priority, to examine the fundamental question of the insurability of AI systems.

1.13. The EESC proposes developing, for companies which comply with the rules, a European trusted-AI Business Certificate based partly on the assessment list put forward by the high-level experts' group on AI (high-level group).

1.14. By promoting work in this area in the G7 and G20 and in bilateral dialogues, the EU must endeavour to ensure that AI regulation goes beyond the EU's borders. We need an international agreement on trustworthy AI, which will develop international standards and carry out frequent checks on the relevancy of those standards.

2. **Summary of the Commission proposal**

2.1. This communication builds on the work of the high-level group which the Commission appointed in June 2018. In this communication, the Commission identifies seven key requirements for achieving trustworthy AI, which are listed in point 4,

2.2. The Commission has launched a pilot phase involving stakeholders on a broad scale. This exercise focuses in particular on the assessment list drawn up by the high-level group for each of the key requirements. At the beginning of 2020, this group will review and update the assessment list and if appropriate the Commission will propose further measures.

2.3. The Commission wants to take its AI approach international and will continue to play an active role, including in the G7 and G20.

(²) OJ C 288, 31.8.2017, p. 1, OJ C 440, 6.12.2018, p. 1.
(³) https://osha.europa.eu/en/emerging-risks
(⁴) Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (OJ L 119, 4.5.2016, p. 1).

3.   **General comments**

3.1.   Human-centric AI needs an approach covering technical, societal and ethical issues. The EESC is pleased that the European institutions intend to build an AI approach which is in line with the values underpinning the EU: respect for human dignity, freedom, democracy, equality and non-discrimination, the rule of law and respect for human rights. As the Commission points out (⁵), AI is not an end in itself, but a tool that can deliver far-reaching positive change. Like any tool, it creates both opportunities and risks, which is why the EU has to regulate its use and clearly establish just who is responsible.

3.2.   Trust in human-centric AI will be forged by affirming values and principles and providing a well-established regulatory framework and ethical guidelines setting out key requirements.

3.3.   It is important to work with all stakeholders to identify which of the many challenges posed by AI need to be tackled by regulation and legislation supported by regulatory oversight mechanisms and, in the event of non-compliance, penalties, and which can be tackled by means of codes of ethics, self-regulation and voluntary commitments. The EESC is pleased that the Commission has taken on board some of the principles originally raised by the EESC, but considers it unfortunate that it has not yet proposed any specific measures to address legitimate concerns (as regards consumer rights, system security and liability).

3.4.   AI systems must comply with the existing regulatory framework, particularly as regards protection of personal data, product liability, consumer protection, non-discrimination, professional qualifications and information and consultation of workers in the workplace. It is important to make sure that this legislation is adapted to the new challenges of digitalisation and AI.

3.5.   As the Commission notes, 'processes to clarify and assess potential risks associated with the use of AI systems, across various application areas, should be put in place' (⁶). The EESC attaches the utmost importance to the future arrangements for this assessment and to the establishment of indicators that could be used to perform it. The assessment list proposed by the high-level group is a starting point for implementing these processes.

3.6.   This also concerns the question of fair distribution of the expected added value of AI systems. The EESC believes that the beneficial transformation which AI has the potential to bring in terms of economic development, sustainability of (particularly energy) production and consumption processes and better use of resources must benefit all countries and all citizens.

4.   **Specific comments**

4.1.   *Human agency and oversight*

4.1.1.   The Commission wants to be sure that the use of AI systems will never undermine human autonomy or give rise to adverse effects. The EESC supports this approach of human oversight of machines, as it has already stated in previous opinions.

4.1.2.   Under this approach, citizens also have to be properly informed about the uses of these systems. The systems have to be explainable or, where this is not possible (in the case of deep learning, for instance), the user has to be informed about the system's limitations and risks. In any event, people have to retain the freedom to decide differently from the AI system.

4.1.3.   In businesses and public administrations, workers and their representatives must be properly informed and consulted when AI systems are introduced that are likely to alter the way work is organised and to affect them (in terms of supervision, oversight, evaluation and recruitment). The Commission must promote social dialogue with a view to involving workers in the uses of AI systems.

4.1.4.   With regard to human resources, particular attention must be paid to the risks of misuse of AI systems, such as unlimited surveillance, collection of personal and health data, and sharing of these data with third parties, and to the emerging risks in terms of health and safety in the workplace (⁷). Clear standards must be established to ensure that human-machine collaboration does not cause damage to humans. The International Organization for Standardization (ISO) standard on collaborative robots (⁸), which is aimed at manufacturers, integrators and users, provides guidelines for the design and organisation of a collaborative workspace and the reduction of the risks to which people can be exposed. Workers must be trained to use AI and robotics, to work with them and, in particular, to stop them in an emergency ('emergency brake principle').

(⁵)   COM(2019) 168 final.
(⁶)   COM(2019) 168 final, p. 5.
(⁷)   See in particular *OSH and the future of work: benefits and risks of artificial intelligence tools in workplaces*.
(⁸)   ISO/TS 15066, 2016.

4.2.   *Technical robustness and safety*

4.2.1.   The EESC calls for the introduction of European security standards and the development of a robust certification procedure based on test procedures that would enable companies to state that their AI systems are reliable. The EESC would also like to stress the importance of the insurability of AI systems.

4.2.2.   The Commission pays scant attention to the issue of forecasting, preventing and prohibiting the malicious use of AI and machine learning, against which many researchers have issued warnings [9]. Their recommendations should be taken into account, particularly those concerning the dual use of these technologies which can potentially touch on digital security (increase in cyber attacks, exploitation of human and AI vulnerabilities, data poisoning), physical security (hacking of autonomous systems, including autonomous vehicles, drones and automatic weapons) and political security (mass collection of personal data, targeted propaganda, video manipulation, etc.). Researchers, engineers and public authorities must work closely to prevent these risks; for their part, experts and other stakeholders such as users and consumers must be involved in discussions on these issues.

4.3.   *Privacy and data governance*

4.3.1.   The Commission calls for access to data to be 'adequately governed and controlled' [10]. The EESC believes that we need to go further than general statements. The degree of trust that people have in AI systems will also determine their development. The issues of data ownership, and the control and use of data by companies and organisations have yet to be resolved. The amount of data transmitted for example by cars to car manufacturers and the type of data transmitted are startling [11]. Despite the concept of privacy by design, with which connected objects have to comply under the GDPR, we can see that consumers have very little or no information on this subject and no means of controlling these data. The EESC therefore urges the Commission to review the GDPR and related legislation in the light of developments in technology [12].

4.4.   *Transparency*

4.4.1.   The EESC believes that the explainability of algorithmic decision-making processes is key to understanding not the mechanisms but the underlying logic of the decision-making processes and how they are influenced by AI systems. Developing standard test procedures for machine learning systems continues to be a technical challenge which needs to be supported by EU instruments such as Horizon Europe.

4.4.2.   The EESC agrees with the Commission that AI systems must be identifiable as such, 'ensuring that users know they are interacting with an AI system' [13], including in the context of relations between patients and health professionals and professional services linked to citizens' health and well-being. The EESC also stresses that users and consumers must also be able to be informed about the services performed by human beings. Many AI systems actually involve large amounts of human work, which is often hidden from end-users [14]. There is the underlying issue here of the lack of transparency towards users and consumers of services, and a form of usage of concealed and unrecognised work.

4.4.3.   In addition, the EESC believes that consumers must always be informed when AI systems are integrated into the products they buy, and must always be able to access and control their data.

4.5.   *Diversity, non-discrimination and fairness*

4.5.1.   Risks in the form of discrimination are present in some AI applications which profile citizens, users and consumers (for example for recruitment, letting property and certain personal services). The EU has adopted a body of legislation on equal treatment and non-discrimination [15] and AI systems must comply with it. However, this legislation must also be adapted and, if appropriate, bolstered (including in terms of enforcement) in order to cope with new practices. There is a real danger that algorithmic profiling could become a new and powerful tool of discrimination. The EU must prevent this danger.

---

[9]   See report on *The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation*, February 2018.

[10]   COM(2019) 168 final, p. 6.

[11]   *Your car knows when you gain weight*, The New York Times (International Edition), 22.5.2019.

[12]   OJ C 190, 5.6.2019, p. 17.

[13]   COM(2019) 168 final, p. 6.

[14]   See for instance *A white-collar sweatshop: Google Assistant contractors allege wage theft*, The Guardian, 29.5.2019 and *Bot technology impressive, except when it's not the bot*, The New York Times (International Edition), 24.5.2019.

[15]   OJ L 180, 19.7.2000, p. 22; OJ L 303, 2.12.2000, p.16; OJ L 373, 21.12.2004, p. 37; OJ L 204, 26.7.2006, p. 23.

4.5.2.   The Anti-Racism Directive ([16]) and the Directive on equal treatment for men and women beyond the workplace ([17]) provide for the creation of special bodies responsible for promoting gender equality. The EESC calls for these bodies to play an active role in monitoring and overseeing AI systems with regard to the risks of direct or indirect discrimination.

4.6. *Societal and environmental well-being*

4.6.1.   The Commission does not propose any specific ways to link up the climate transition and the digital transformation, particularly as regards the use of AI systems. Consideration must be given to the contribution that AI systems can make to reducing greenhouse gas emissions, particularly in industry, transport, energy, construction and agriculture.

4.6.2.   The Commission points out that AI systems can be used to enhance social skills but they could also lead to a deterioration in this area. The EESC feels that the EU must be more proactive in gauging certain societal challenges. For example, studies have shown that some applications incorporating AI systems are designed to keep users of online services (social networks, games, videos, etc.) connected for as long as possible. The aim is to be able to collect as much data as possible on their behaviour; the strategies used range from endless transmitting of algorithmic recommendations to reminders and notifications, games, etc. The effects on children of the excesses of connection and solicitation have been studied ([18]) and the findings have shown an increase in anxiety, aggression, sleeplessness and an impact on education, social interaction, health and well-being. In order to build trustworthy AI, the EU must take these effects into account and prevent them.

4.6.3.   Lastly, one of the elements of societal well-being is related to a sense of security at work. The effects of digitalisation can undermine security and cause stress ([19]), and so strategies are needed to anticipate change before any restructuring occurs and provide ongoing training for all workers. This requires a high standard of social dialogue in companies between employers and workers' representatives, involving in particular inclusive deployment of new technologies, especially AI and robotics. To consolidate trust between management and workers, IA systems in the area of management, evaluation and oversight of workers must be explainable, their parameters must be known and the way they work must be transparent.

4.7. *Accountability*

4.7.1.   The decisions taken by machine learning systems cannot be explained in simple terms; moreover, they are updated regularly. The EESC believes that oversight of AI systems may not be sufficient to define who is responsible and build trust. It therefore recommends that rules be drawn up assigning responsibility to natural persons or legal entities in the event of non-compliance. The EESC recommends relying more on trustworthy companies or professionals than on algorithms, and proposes developing, for companies which comply with all the rules, a European trusted-AI Business Certificate based partly on the assessment list suggested by the high-level group.

4.7.2.   The Product Liability Directive ([20]) establishes the principle of strict liability for European producers: where a defective product causes harm to a consumer, the producer can be held liable even when there is no fault or negligence on their part. The increasingly widespread design, deployment and use of AI systems mean that the EU needs to adopt adapted liability rules for situations where products with digital content and consumer services can be dangerous and harmful. Consumers must be able to take legal action in the event of harm caused by an AI system.

5. **The need for regulation beyond Europe**

5.1.   In a global context, AI regulation must go beyond Europe's borders. Europe should promote a broad consensus on AI in the G7 and G20 and keep up bilateral dialogue so that a majority of countries can participate in AI standardisation processes and verify their relevance on a regular basis.

Brussels, 30 October 2019.

The President
of the European Economic and Social Committee
Luca JAHIER

([16]) Council Directive 2000/43/EC of 29 June 2000 implementing the principle of equal treatment between persons irrespective of racial or ethnic origin (OJ L 180, 19.7.2000, p. 22).
([17]) Council Directive 2004/113/EC of 13 December 2004 implementing the principle of equal treatment between men and women in the access to and supply of goods and services (OJ L 373, 21.12.2004, p. 37).
([18]) See Kidron, Evans, Afia (2018), *Disrupted Childhood — The Cost of Persuasive Design*, 5Rights Foundation.
([19]) Report by the high-level group on the impact of the digital transformation on EU labour markets, 2019.
([20]) Council Directive 85/374/EEC of 25 July 1985 on the approximation of the laws, regulations and administrative provisions of the Member States concerning liability for defective products (OJ L 210, 7.8.1985, p. 29).