© shutterstock
European Economic and Social Committee
European Parliament
Conference on the Future of Europe
AI STAKEHOLDER SUMMIT 2021
#AISummitEU

AI made in Europe: empowering and protecting the public and businesses

Brussels & online | 08.11.2021 | 9.30 – 16.00

CONCLUSIONS AND RECOMMENDATIONS

The main purpose of this event was to present the EESC's positions on Artificial Intelligence and to contribute to the current interinstitutional regulatory process on the recently proposed AI Act (AIA). The event was divided into three separate panels and each panel was dedicated to a specific topic with an overall focus on the AIA. The event was organised together with the European Parliament's Special Committee on Artificial Intelligence in a Digital Age (AIDA) and brought together around 400 participants (15 speakers, a delegation of 24 EESC members, MEPs and participants from the European Parliament, and 321 remote participants). The morning was organised by the EESC on its premises, and in the afternoon, an Interparliamentary Committee meeting took place at the EP. The full programme is available here and the recording of the conference here.
You will find below the summary of the main conclusions drawn and recommendations made around different topics covered at the morning part of the summit.

❖ AI ACT

It is important to understand that the AIA doesn't regulate AI technology itself. The main focus is on "**use cases**" - the use of AI systems in a given context.

Any discussion about the AIA should start with the **definition of AI for legal purposes**. This discussion has to involve both technicians and legal experts. From a legal perspective, it is important to look at the meaning of the text to interpret what is covered. This seems to be quite difficult with the current definition proposed in the AIA. If a company declares its products as software (rather than an AI system), it can easily argue that is not covered by the AIA.

The global nature of the AIA is a very strong point. AI systems that are developed outside of the EU have to meet the same legal standards if they are deployed or having an impact within the EU. This rules out debates that if we regulate too strongly in the EU, AI will be developed elsewhere in the world.

The **risk-based approach** and **risk pyramid** on which the AIA is built raise many questions. Some uses are categorised as a minimum risk but they could have a huge impact. For instance, recently shown President Macron's deep fake could have enormous impact on the next French elections. However, this concrete use was put on the low risk list. We need to have a conversation with people who know how to make risk assessments properly and how to categorise AI risk and perhaps reframe the whole pyramid or set very strict boundaries. The risk pyramid also overlooks non-risk, unintentional risk and misuse and is not complete. It needs to be rephrased and thoroughly thought out.

High-risk uses of AI are listed in the Annex to the regulation. But the underlying question is whether the fact that a use on this list meets some requirement means that we can normalise the use? This opens the door to dangerous interpretations.

Involvement of stakeholders was, is and will remain key. Negotiating parties have to keep up exchanges and continue to listen to all interested parties. The Commission plans to set up an expert group which will focus, amongst other things, on identifying new use cases to be added to the list of high-risk AI applications.

Work has also to be done on **empowering citizens, workers and consumers on their rights** and the ownership and control of their data. The social partners have a crucial role to play in this respect.

Both the European Commission and the European Parliament agree on the fact that the key issue in the AI Act negotiations will be **finding the right balance** between the freedom to innovate and safeguards we need to provide to respect fundamental rights and to implement the necessary checks before an AI product is put on the market. However, their views diverge when it comes to applying **self- or third-party conformity assessment procedures**. The European Parliament is in favour of the latter, while the Commission advocates a less burdensome approach with self-assessments. According to the Commission, there are already many self-assessment procedures in internal market legislation, and combined with standards companies can comply with, they can simplify procedures for putting products on the market.

The Commission advocates a functional equivalence approach in the AI Act. This means existing internal market approaches for safety follow the same logic and can be applied to fundamental rights.

The question of **liability and unintended harm** will be dealt with in the ongoing revision of the Product Liability Directive. However, this issue is not new and does not apply exclusively to AI.

To build excellence in the EU, the **role of SMEs and MSMEs** is crucial. Some of the most pressing issues for SMEs are the need to train and educate workers on AI, facilitating access to finance, and access to high-quality and safe data.

❖ AI IN THE WORLD OF WORK - HIGH RISK AI

Debate with workers on AI is needed to understand in a simple way how AI systems work, how they interact with humans and what could be their impact.

AIA provides a ceiling rather than a threshold. AIA is presented as the regulation ensuring ethical and trustworthy AI. However, in reality, it follows an internal market approach using a product safety model and is not based on **fundamental rights**. This is reflected in the text. The big problem with this approach is that, even with high requirements for high-risk AI, its use may not prevent violation of fundamental rights. It is not because we diminish the risk, we can use these tools. An obligatory fundamental rights impact assessment should be part of the AIA, both *ex-ante* and *ex-post*.

The AIA allows uses of AI systems in **human resources management** which might be very problematic. Systems such as monitoring tone of voice, facial scans during interviews, keyboard activity monitoring and automatic scanning of CVs should simply **not be allowed**, even if they were transparent and reliable. This deserves more attention than AIA currently provides and real dialogue with social partners. However, the AIA includes no role for social partners and this needs to be changed.

A list of completely prohibited uses should be proposed, including **algorithmic worker surveillance**. There has always been some monitoring of workers, but today we are harvesting workers most intimate data (behaviour data, emotions, sentiment analysis). The AI systems extract only one component of the whole person and deduce something from it without a context. (The whole person is a context).

The underlying, fundamental question is why we want these tools, which monitor, for instance, an employee's heartbeat at work? The fact that these programmes are marketed points to a complete **ethical breakdown in the tech sector**.

The question of **algorithmic training** of workers should be asked the other way around. Teachers, lawyers and doctors should be trained to understand systems they use, but most importantly, it is crucial to train engineers and developers of these tools in human rights. They don't know how an AI system can affect fundamental rights, because they don't consider this perspective and are not trained to do so. This is one of the biggest gaps when we talk about skills.

The right of **explainability**, as proposed in Article 22 of the GDPR, is of crucial importance here. The AIA should include the same article, along with more details of how it should be applied in the workplace. Management in companies often don't understand the software. They need to know what kind of software they are acquiring and for what purpose.

The harmonising effect of AIA is, of course, beneficial for the internal market but can have also negative consequences if replacing more protective legislation of workers at national level. Currently, national legislation in many countries allows workers' representative to have their say over whether certain technology should be used in the workplace or not, especially if it allows worker surveillance. The AIA is much less protective and we should ensure that it does not replace these more protective national regulations.

Standardisation and technical documents will drive compliance with the AIA. However, the big problem is that standardisation bodies are private companies and technical documents are drafted by members of these committees. They can't be considered to be a democratic institution. Moreover, workers are not included in these organisations, which means in practice that standards will not cover or represent the views of workers.

Besides the AIA proposal, the current **regulatory framework** for protecting workers' rights when AI is used at work is completely **fragmented**. There needs to be a single EU framework for protecting workers' rights and safety when technology is deployed, especially AI.

❖ BIOMETRIC IDENTIFICATION SYSTEMS

Biometric identification systems represent an enormous threat for our **democratic future and the rule of law**. Deploying remote biometric identification systems in publicly accessible spaces is a direct attack on our privacy and means in practice the end of anonymity in these spaces. It can also create a huge threat to fundamental rights, such as freedom of assembly or freedom of expression.

The pandemic has greatly exacerbated the impact of these problems (such as disinformation, polarisation, lack of privacy). The main cause is within the system, as the technology is not designed with a human-centric approach. The question we have to ask ourselves is whether we can sustain the use of these technologies in a limited setting or whether this will lead to a **function creep** where more and more areas of our lives will be impacted. If, for example, a piece of behavioural analytics software works well for serious crimes, do we allow it also to be used subsequently for other crimes? Once these technologies are deployed, the temptation to use them for other reasons is huge and it is very difficult to get rid of it. We should ban this technology upfront, discuss the issue and, if we consider some uses to be democratically sound, then allow them.

Transparency is key for any use of AI systems in publicly and privately accessible spaces. We have to know when and for what purposes AI is being used. Citizens have to know that this technology is being used and be given the possibility of redress.

Even if the use of this technology is **government driven**, people are not willing to share their personal data. According to a 2019 Fundamental Rights Agency survey of 35 000 respondents, fewer than 2 out of 10 people are willing to share their images with public authorities for identification purposes. Processing of facial images by public authorities may affect human dignity. People may feel uncomfortable about going to public places under surveillance and they may change their behaviour under public surveillance.

The discussions around **facial recognition** are always linked to the accuracy of the software. Facial recognition software does not deliver results, only probabilities, meaning there is always a percentage of a likelihood of a match. It is also important to consider the absolute number of people being subjected to the technology. If there were 100 000 people in the scheme and the rate of false positive was only 1%, we would still have 1 000 wrongly flagged people. **Accuracy rates** comes from training and test databases and our knowledge about the accuracy in real life situations is quite limited. However, even with 100% accuracy, this technology would still pose fundamental rights issues.

It is important to distinguish emotion recognition from behavioural recognition. There is a fundamental problem with **emotion recognition** which presumes one can infer emotions from observing people's behaviour. It is a false assumption coming from already discredited behaviourism theory from the 1950's. When observing people's behaviour, we are dependent on the context, cultural specificities and other elements. The reliability of this technology is highly questionable.

Behavioural recognition can be carried out a little more easily than emotion recognition because we do not deal with inner mental states. It poses less risks to our privacy as some behaviours are easily observable. However, only the most superficial behaviours can be determined with some confidence - we can determine that someone is running but not fleeing, because fleeing makes reference to inner mental state (that requires understanding of intent or emotions).

With behavioural analytics and facial recognition, it is also important to consider **private uses of these technologies** and what should be allowed. Private surveillance for marketing purposes is very dangerous. Highly detailed profiles of people are built up with the goal of manipulating them to spend more money and buy products and this can be also used for political manipulation of voters.

Our society also has to be sustainable from the human point of view. Principles of dignity of individuals have to be built inside the technologies and with the AIA, we need **human dignity by design**.

❖ FORESIGHT PERSPECTIVE ON AI

When we talk about AI today, we should define it as **divorce between agency and intelligence**. We have the ability today to solve complex tasks with technology that has no intelligence.

There are **5 major trends**:

1. shift from logic to statistics, from deduction to association - before we had a recipe to get the meal and today we have machine learning, meaning you show the final meal and ingredients and the machine learns the recipe (the machine can cook the meal independently on the basis of statistics);
2. enveloping world around the machines - all available data has been generated by the current generation;
3. from difficult to complex - as we move with the development of AI, AI tends to transform difficulty into complexity (when thinking about replacing human tasks with robots, we do not build the robot in the environment but the environment round the robot);
4. constitutive vs regulative rules (constitutive rules apply for chess as a chess activity cannot exist without rules, and regulative rules apply for football as there is an activity and you shape the activity thanks to the rules but the activity precedes the rules);
5. historical data - we move now from historical to synthetic data (the future of AI will be increasingly searching for synthetic data).

When speaking about AI and ethics, we have to mention opportunities, risks but also **opportunity cost**, which means underuse of AI because people are scared to use it.

The **main challenges** linked to AI are:

- making AI work against wrong-doing ;
- making AI enhance human decision/control;
- making AI support human responsibility (AI helps us to care about the world, AI for SDGs);
- making AI make us more human (we are dealing with agency without intelligence that can erode our autonomy);
- making AI work for humanity.

The new challenge is not digital innovation but the **governance of the digital**. This governance of the digital requires a new **human project - the green and the blue** (the Green of our habitats - natural, synthetic, and artificial, from the biosphere to the infosphere, from urban environments to economic, social, and political circumstances — and the Blue of our digital technologies, from mobile phones to social platforms, from the Internet of Things to Big Data, from AI to future quantum computing). The marriage between the Green and the Blue, with its advantages, counterbalances the divorce between Agency and Intelligence, with its risks.