cepAdhoc

10 February 2020

European Strategy on artificial intelligence

An Assessment of the European Commission's leaked White Paper on AI

Alessandro Gasparotti and Lukas Harta, LL.M.

In mid-January a draft of the Commission's White Paper on Artificial Intelligence (AI) was leaked. It outlines weaknesses in the current legislation and introduces three main regulatory options for future action on AI. The cep assesses these options as follows:

- **Option 1: Voluntary labelling scheme.** This option would respect the freedom to conduct a business, because companies are free to decide whether they want to use the label or not. Yet, it might not be enough to address e.g. safety and liability issues.
- **Option 2: Sectorial requirements for public authorities.** This option would require public authorities to publish information on the effectiveness of the AI applications they use. This is appropriate as public authorities have greater power to interfere with people's fundamental rights than private actors have.
- **Option 3: Mandatory risk-based requirements for high-risk AI applications.** This option would regulate high-risk AI applications in both the private and the public sector. Without a precise definition of high-risk, companies will have an incentive to downplay the possible risks of their AI application so that their products do not have to comply with the standards for high-risk AI applications.

Alessandro Gasparotti and Lukas Harta are analysts in the internal market department at cep| gasparotti@cep.eu

cep | Kaiser-Joseph-Straße 266 | 79098 Freiburg | Germany | Telephone +49 761 38693-107 | www.cep.eu

1 Political context

On 17 January 2020, a paper of the European Commission entitled "Structure for the White Paper on artificial intelligence – a European approach" was leaked.[1] The White Paper's official publication is to be on 19 February 2020. It forms part of the Commission's broader strategy for artificial intelligence (hereinafter: "AI"), which includes the Communication on AI for Europe[2], the Coordinated Plan on AI[3] and the Communication on Building Trust in Human-Centric AI[4] (cf. cep**PolicyBriefs** on investment,[5] on education and social systems,[6] on legal and ethical rules,[7] and on the ethics guidelines[8]). In this regard, President of the European Commission Ursula von der Leyen has pledged to "forward legislation for a coordinated European approach on the human and ethical implications of Artificial Intelligence"[9] within her first 100 days in office. Furthermore, on 23 January 2020, the European Parliament's Internal Market and Consumer Protection Committee approved a resolution "on automated decision-making processes: Ensuring consumer protection, and free movement of goods and services".[10] The resolution stresses, inter alia, the need for a risk-based approach to AI regulation and calls the Commission to develop a risk assessment scheme for AI to ensure a consistent approach to the enforcement of product safety legislation in the internal market. In addition, in an exchange of views with the European Parliament's Committee on Legal Affairs on 27 January 2020, Margrethe Vestager, Executive Vice-President of the European Commission for a Europe Fit for the Digital Age, underlined the need to establish high standards for AI in the EU, particularly high transparency and accountability standards for AI technologies used in the public sector.[11]

The leaked White Paper states that "the objective of the European approach is to promote the development and uptake of artificial intelligence across Europe, while ensuring that the technology is developed and used in a way that respects European values and principles."[12] According to the Commission, three fields of activity are key in order to achieve these objectives: investment, access to data, and AI regulation. To foster investment in AI, the Commission intends to use EU-level funding in order

- to establish "a world-leading artificial intelligence computing and data infrastructure in Europe",[13]
- to strengthen digital innovation hubs that will improve the uptake of AI, and
- to ensure access to finance for AI innovators.

To this end the Commission is expected to release a review of the Coordinated Plan on AI.[14]

In addition, the Commission aims at developing common European data spaces and intends to adopt, by early 2021, an implementing act on high-value public sector datasets. These datasets should be available for

1 www.euractiv.com/wp-content/uploads/sites/2/2020/01/AI-white-paper-EURACTIV.pdf.
2 COM(2018) 237.
3 COM(2018) 795.
4 COM(2019) 168.
5 Centre for European Policy: cep**PolicyBrief** No. 2019-10.
6 Centre for European Policy: cep**PolicyBrief** No. 2019-12.
7 Centre for European Policy: cep**PolicyBrief** No. 2019-13.
8 Centre for European Policy: cep**PolicyBrief** No. 2019-16.
9 Von der Leyen: Political Guidelines for the next European Commission 2019-2024, available at https://ec.europa.eu/commission/sites/beta-political/files/political-guidelines-next-commission_en.pdf p. 13.
10 2019/2915(RSP).
11 www.euractiv.com/section/digital/news/eus-vestager-calls-on-public-sector-to-establish-particularly-high-ai-standards/.
12 European Commission: Leaked White Paper on AI, available at www.euractiv.com/wp-content/uploads/sites/2/2020/01/AI-white-paper-EURACTIV.pdf p. 1.
13 Ibid. p. 5.
14 COM(2018) 795.

free and in machine-readable format. The main part of the leaked White Paper deals with the third field of activity: AI regulation. Therefore, AI regulation is the focus of this cep**Adhoc**.

2 The draft White Paper on AI regulation

2.1 Weaknesses in the current legislation on AI

AI brings many opportunities, e.g. performing complex tasks in a fraction of the time required by a human, but poses challenges in relation to safety and liability concerning products equipped with such technology. The challenges result, i.a., from the autonomy of AI-enabled products – i.e. when AI-enabled products perform their tasks without human supervision – and from the opacity of AI decision-making – i.e. when understanding the process that led to a specific outcome is difficult or even impossible.

The Commission recognizes in the White Paper that there is already a robust body of legislation at EU and national levels that applies to AI.[15] The two main pieces of EU legislation regulating the safety requirements and liability regime for the use of AI are the General Product Safety Directive[16] and the Product Liability Directive.[17] Nevertheless, the Commission also points out that – due to the fast development of AI – the existing legislation might not cover all the specific risks that are bound to arise with the widespread use of AI. After a first round of consultation with Member States, businesses and other stakeholders, the Commission identified, i.a., the following three weaknesses in the current legislation:

(1) Aggravation of risks due to autonomous decision-making by AI-enabled products

The Commission mentions, i.a., personal safety risks, cyber threats and risks associated with loss of connectivity, especially if an AI-enabled product relies on cloud computing to operate. If, e.g., a car driver uses a car navigation system, a loss of connectivity does not cause severe safety risks; it is still the car driver who steers the car and not the navigation system. If, however, an autonomously driving car loses connectivity the car receives no more input on its current position, the course of the road ahead, the road condition or the traffic conditions. This can lead to an inappropriate speed. Similarly, while a hacked car navigation system guiding the driver incorrectly can be very inconvenient for the user, a hacked autonomous car could be used to bring about serious accidents or even terrorist attacks. Thus, cyber threats are as well aggravated by autonomous decision-making.

(2) Changing nature of AI-enabled products during their lifecycle

AI-enabled products are likely to change during their lifecycle, notably due to machine-learning – i.e. when an AI application carries out a given task without being explicitly programmed on it, using patterns and inference instead – or substantial updates of the database that AI-enabled products use to train themselves.

An AI-enabled product could meet safety standards when first placed on the market, which is the relevant point in time under both the Product Liability Directive[18] and the General Product Safety Directive.[19] However it might not respect these standards at a later stage due to an evolution in its behavior, e.g. if it comes to different results because of new data. For example, if an AI-enabled fitness watch that has been trained on

---

15 European Commission: Leaked White Paper on AI, available at www.euractiv.com/wp-content/uploads/sites/2/2020/01/AI-white-paper-EURACTIV.pdf p. 10.

16 Directive 2001/95/EC of the European Parliament and of the Council of 3 December 2001 on general product safety.

17 Directive 85/374/EEC on the approximation of the laws, regulations and administrative provisions of the Member States concerning liability for defective products.

18 Art. 7 lit. b of Directive 85/374/EEC.

19 Art. 3 para. 1 of Directive 2001/95/EC.

balanced data is placed on the market, it might, with time, gather data of mainly active and healthy people. If the watch uses these data in order to suggest a workout to the users, the watch might encourage old people or unfit ones to train too much, with risks for their health. Furthermore, if the watch autonomously controls e.g. an electrical muscle stimulation device, users could be directly harmed by the device without having control over the decision made by the AI.

(3) Difficulties linked to enforcement

AI-enabled products are often no longer based on an easy to read code. While the output in most cases is more precise, understanding causality or the process of decision-making of the AI is not always possible. For example, if an AI-enabled software interviews applicants to a job offer, the parameters that lead to its decision can be opaque. Therefore, if an applicant thinks he was discriminated against, the discovery and possible redress of the discrimination could be unattainable.

2.2 Regulatory options for future legislation

In order to address the weaknesses of the current legislation, the White Paper presents three possible regulatory options:

Option 1: Voluntary labelling scheme

AI developers that comply with certain conditions would be allowed to use the label of "ethical/trustworthy AI". Compliance with the scheme should be enforced. Furthermore, the Commission believes that a labelling scheme would help Europe play an important role in international discussions on ethical and trustworthy AI.[20] However, the Commission states that voluntary schemes might not be enough to solve e.g. questions of safety and liability.

cep**Assessment**: This approach is the least burdensome for developers and users of AI, because the use of the label is voluntary. Therefore, option 1 respects the freedom to conduct a business. Developers of AI will only incur extra costs if they decide to comply with the scheme. However, if the label is highly valued by consumers, companies will be pushed to adhere to the voluntary scheme, so to signal consumers their "trustworthiness". Also in this case, however, the labelling scheme is appropriate because it leads companies to meet the wishes of consumers and increases transparency for consumers.

The issue which authority should be responsible for granting and enforcing such labels is not discussed in the White Paper. The certification scheme introduced by the General Data Protection Regulation (hereinafter: "GDPR")* which proofs that companies comply with its rules shows however that this aspect is very important. The GDPR certification scheme is rarely used, because it is not applied consistently across the EU** as national authorities interpret the GDPR differently and set different standards for obtaining the certification. This distorts competition within the internal market and reduces legal certainty for companies certified in one country when they operate in another one. It is therefore key to create harmonised and clear requirements to comply with in order to obtain the EU label for trustworthy AI.

* Regulation (EU) 2016/679 of the European Parliament and 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.

** Centre for European Policy: cep** PolicyBrief** No. 2020-01.

[20] Centre for European Policy: cep**Input** No. 2019-07.

Option 2: Sectorial requirements for public authorities and facial recognition

This approach sets out requirements for public authorities when they use AI-enabled products. Like the Canadian directive on automated decision-making, these requirements entail rules for

- impact assessments of the algorithms used (i.e. requirements for assessing the impacts of algorithms on administrative decisions, whereby more important decisions would receive closer scrutiny),
- quality assurance (i.e. requirements ensuring that the data used by the AI application is tested for unintended data biases and other factors that may distort the outcomes),
- redress mechanisms, and
- reporting (e.g. requirements for public authorities to publish information on the effectiveness and efficiency of the AI applications in meeting public authorities' objectives on a website).

The aim of option 2 would be to ensure that public authorities use AI in a way that reduces risks to public institutions and leads to more efficient, accurate, consistent, and interpretable decisions.

cep**Assessment**: The advantage of this option is the sectorial and problem-specific regulation of AI. Given that public authorities have significantly greater power to interfere with people's fundamental rights than private actors have, regulating the use of AI by public authorities is more urgent than between private actors. Also, the clear definition of the scope of application of AI legislation would help AI developers and users to determine whether such regulation (e.g. the reporting obligation) applies to them or not. Public authorities would know for sure that it does, private users would know for sure that it does not. While only regulating the use of AI by public authorities, the regulation under option 2 could nevertheless have a signalling effect on the private sector, encouraging companies to comply with the essential standards of the public sector in order to display "trustworthiness" to consumers. Some consumers might prefer products that fulfil the requirement for public authorities over cheaper products that do not. Such an approach on regulation might obtain, in the private sector, the outcome envisaged by a voluntary labelling scheme.

Under option 2 the Commission also discusses whether specific rules for the use of facial recognition systems in public spaces should be introduced, or the use of such technology in public spaces be prohibited for a period of, e.g., three to five years. During this period, a sound methodology for assessing the impacts of this technology and possible risk management measures were to be identified and developed. The ban on facial recognition systems would apply equally to public and private actors. Exceptions should be considered for, e.g., Research & Development and security purposes.

cep**Assessment**: While it is true that facial recognition technology poses more challenges on fundamental rights than many other AI applications, a far-reaching measure such as a complete ban on facial recognition technology might hamper its development within the EU, as recognised by the Commission itself. This is particularly true since this market is already dominated by non-EU companies. Furthermore, apart from banning facial recognition technology, AI legislation would not cover private actors at all.

Option 3: Mandatory risk-based requirements for high-risk AI applications

New AI legislation would only apply to high-risk AI applications while the existing legislation – e.g. the GDPR – would continue to apply to all applications. One way to define high-risk AI applications would be to assess whether an application cumulatively

- falls within one of the particularly sensitive sectors that would be clearly specified (e.g. healthcare, transport, police, and judiciary), and
- fulfils a more abstract definition of "high-risk" application; this definition could read as follows: "High-risk applications means applications of artificial intelligence which can produce legal effects for the individual or the legal entity or pose risk of injury, death or significant material damage for the individual or the legal entity."[21]

cep**Assessment**: The main advantage of this option is that it could cover all high-risk AI applications – also those in the private sector – in a dynamic and flexible way. The interpretation of "high risk" is however discretional, and if not precisely defined will give companies an incentive to downplay the possible risks of their product so that their AI applications do not have to comply with the standards for high-risk applications.* Also, the leaked White Paper leaves open what the content of the risk-based requirements could be, merely stating that such an instrument might set out transparency and accountability requirements.

* Centre for European Policy: cep**PolicyBrief** No. 2019-16.

2.3 Additional regulatory options

In addition, the Commission considers two further regulatory options that can be combined with any of the aforementioned three options:

(1) Safety and liability legislation

The Commission considers amending the existing safety and liability legislation – including the General Product Safety Directive, the Machinery Directive,[22] the Radio Equipment Directive[23] and the Product Liability Directive – in order to address the specific risks of AI-enabled products.

cep**Assessment**: The advantage of this approach is that it addresses those weaknesses in the current legislation that the Commission has identified, without introducing AI-specific legislative acts. This avoids creating specific obligations for different sectors, actors, or categories of AI applications.

(2) Governance

Member States should mandate existing authorities or establish new ones with the task of monitoring the application and enforcement of the future regulatory framework for AI.

cep**Assessment**: As the application of the GDPR has shown, it is key for national authorities to be sufficiently funded and have the instruments to cooperate with other authorities in the EU. Moreover, certain methods of the national authorities should be harmonised, otherwise legal fragmentation could distort competition within the internal market, e.g. regarding imposing sanctions.*

* Centre for European Policy: cep**PolicyBrief** No. 2020-01.

[21] European Commission: Leaked White Paper on AI, available at www.euractiv.com/wp-content/uploads/sites/2/2020/01/AI-white-paper-EURACTIV.pdf p. 16.
[22] Directive 2006/42/EC of the European Parliament and the Council of 17 May 2006 on machinery.
[23] Directive 2014/53/EU of the European Parliament and the Council of 16 April 2014 on the harmonisation of the laws of the Member States relating to the making available on the market of radio equipment.

2.4 Final remarks

The Commission seems to be in favour of a risk-based approach (option 3). Such approach could be enforced by national authorities and coupled with updated safety and liability legislation. Interestingly, the Commission discourages a ban on facial recognition technologies, favouring instead its regulation through the full implementation of the relevant GDPR provisions.