BRIEFING
Requested by the ITRE Committee

European Parliament

The White Paper on Artificial Intelligence

KEY FINDINGS

On 19 February 2020, the EU Commission published the White Paper on *Artificial Intelligence - A European approach to excellence and trust*, together with the report *The safety and liability aspects of IA* and the communication *A European strategy for data*, as part of a wide package on Artificial Intelligence. This briefing summarizes the main aspects of the White Paper on AI, which is currently undergoing a public consultation process open until 31 May 2020.

The Commission's White Paper on Artificial Intelligence

The White Paper on Artificial Intelligence (AI), published by the European Commission on 19 February 2020, contains proposals for EU action in the area of AI, with the purpose to launch a debate with the public, stakeholders, the European Parliament and the Council in order to arrive at a political consensus. The White Paper on AI is structured around two main dimensions: **promoting excellence** and **building trust in AI**; it proposes:

- **Excellence**: measures for AI capacity-building to streamline research, foster collaboration between Member States and increase investment into AI development and deployment; the aim of the framework is to mobilise resources to achieve an **'ecosystem of excellence' along the entire value chain**, starting in research and innovation, and to create the right incentives to accelerate the adoption of solutions based on AI, including by small and medium-sized enterprises (SMEs).

- **Trust**: policy options to enable a **trustworthy and secure development of AI** in Europe to determine the future regulatory framework and the legal requirements that would apply to relevant actors, with a particular focus on **high-risk** applications and the key elements of a future regulatory framework for AI in Europe that will create a unique **'ecosystem of trust'.** The build of an ecosystem of trust is a policy objective in itself, and should give citizens the confidence to take up AI applications and give companies and public organisations the legal certainty to innovate using AI.

Policy Department for Economic, Scientific and Quality of Life Policies
Directorate-General for Internal Policies

Author: Matteo Ciucci, Frédéric Gouardères

PE 648.773 - April 2020

EN

An ecosystem of excellence: key actions on AI

The following **AI actions** to build an **ecosystem of excellence** in AI, complementary to other data actions presented in the European data strategy, have been proposed.

The Commission will:

- **Action 1**: propose to the Member States a revision of the Coordinated Plan on Artificial Intelligence to be adopted by end 2020 with the objective of attracting over **€20 billion** per year of **total investment** in the EU in AI over the next decade.
- **Action 2**: facilitate the creation of **excellence and testing centres** that can combine European, national and private investments, possibly including a new legal instrument.
- **Action 3**: establish **networks of leading education institutes** to attract the best professors and scientists and offer world-leading masters programmes in AI through the advanced skills pillar of the Digital Europe Programme complemented by the research and innovation actions of Horizon Europe.
- **Action 4**: work with Member States to ensure that at least one digital innovation hub per Member State has a high degree of specialisation on AI under the Digital Europe Programme; with the European Investment Fund, launch a pilot scheme of €100 million in Q1 2020 to provide equity financing for innovative developments in AI and with InvestEU scale it up significantly from 2021.
- **Action 5**: set up in the context of Horizon Europe a new **public private partnership** in **AI, data and robotics**; ensure coordination of research and innovation in AI; collaborate with other public-private partnerships in Horizon Europe; work together with the testing facilities and the digital innovation hubs mentioned above.
- **Action 6**: initiate open and transparent **sector dialogues** giving priority to **healthcare, rural administrations and public service** operators in order to present an action plan to facilitate development, experimentation and adoption. The sector dialogues will be used to prepare a specific '*Adopt AI programme*' that will support public procurement of AI systems, and help to transform public procurement processes themselves.

On international aspects, the EU will continue to cooperate on AI with **like-minded countries and global players** with an approach based on EU rules and values, will closely monitor the policies of third countries that limit data flows and will address undue restrictions in bilateral trade negotiations and through action in the World Trade Organization.

An ecosystem of trust: AI regulatory framework

To address the EU citizens' concerns on AI and ensure **trust**, the Commission set out an AI strategy on 25 April 2018, agreed on 7 December 2018 a Coordinated Plan on AI strategies with the Member States and established a High-Level Expert Group on AI that published non-binding guidelines on **trustworthy AI** on 8 April 2019.

The recently published White Paper on AI proposes that '*it **may be** needed, in addition to the possible adjustments to existing legislation, **a new legislation specifically on AI***' to address **two kind of risks** which exist in AI-based applications:

- **Risks to fundamental rights**: the use of AI can lead to breaches of fundamental rights, including the rights to freedom of expression, freedom of assembly, human dignity, non-discrimination based on sex, racial or ethnic origin, religion or belief, disability, age or sexual orientation, as applicable in certain domains, protection of personal data and private life, or the right to an effective judicial remedy and a fair trial, as well as consumer protection. These risks might result from **flaws in the overall design of AI systems**, including as regards human oversight, or from the **use of biased data**.

- **Risks to safety and liability**: AI technologies may present new safety risks for users when they are embedded in products and services. As with the risks to fundamental rights, these risks can be caused by **flaws in the design** of the AI technology, be related to **problems with the availability and quality of data** or to other problems stemming from machine learning. While some of these risks are not limited to products and services that rely on AI, the use of AI may increase or aggravate the risks.

  The Commission explicitly mentions the Product Liability Directive (PLD) as an act requiring revision and possibly extensions due to certain AI specificities: *'Under the PLD, a manufacturer is liable for damage caused by a defective product. However, in the case of an AI based system such as autonomous cars, it may be difficult to prove that there is a defect in the product, the damage that has occurred and the causal link between the two. In addition, there is some uncertainty about how and to what extent the Product Liability Directive applies in the case of certain types of defects, for example if these result from weaknesses in the cybersecurity of the product'*.

Addressing five areas of **risk** and situations which derive from AI:

- **Effective application and enforcement of existing EU and national legislation**: The lack of transparency or AI opaqueness makes it difficult to identify and prove possible breaches of laws, including legal provisions that protect fundamental rights[1], attribute liability and meet the conditions to claim compensation. Therefore, in order to ensure an effective application and enforcement, **it may be necessary to adjust or clarify existing legislation in** certain areas, for example on **liability** as further detailed in the report on liability.

- **Limitations of scope of existing EU legislation**: an essential focus of EU product safety legislation is on the placing of products on the market. While in the EU product safety legislation, software, when is part of the final product, must comply with the relevant product safety rules, it is an **open question** whether **stand-alone software** is covered by EU product safety legislation, outside some sectors with explicit rules (i.e. medical). General EU **safety legislation currently in force applies to products and not to services**, and therefore in

[1] As regards the protection of fundamental rights and consumer rights, the EU legislative framework includes legislation such as the Race Equality Directive (Directive 2000/43/EC), the Directive on equal treatment in employment and occupation (Directive 2000/78/EC), the Directives on equal treatment between men and women in relation to employment (Directive 2004/113/EC) and access to goods and services (Directive 2006/54/EC), the Unfair Commercial Practices Directive (Directive 2005/29/EC) and the Consumer Rights Directive (Directive 2011/83/EC), the General Data Protection Regulation Directive and other sectorial legislation covering personal data protection, such as the Data Protection Law Enforcement Directive (Directive (EU) 2016/680). In addition, as from 2025, the rules on accessibility requirements for goods and services, set out in the European Accessibility Act will apply (Directive (EU) 2019/882).

principle not to services based on AI technology either (e.g. health services, financial services, transport services).

- **Changing functionality of AI systems**: the integration of software, including AI, into products can modify the functioning of such products and systems during their lifecycle. This is particularly true for systems that require **frequent software updates** or which rely on machine learning. These features can give rise to **new risks** that were not present when the system was placed on the market. These risks are **not adequately addressed** in the existing legislation, which predominantly focuses on **safety risks** present at the time of placing on the market.

- **Uncertainty as regards the allocation of responsibilities between different economic operators in the supply chain**: in general, EU legislation on **product safety** allocates the responsibility to the producer of the product placed on the market, including all components e.g. AI systems. However, the rules can for example become unclear if **AI is added after the product is placed on the market by a party that is not the producer**. In addition, EU product liability legislation provides for liability of producers and leaves national liability rules to govern liability of others in the supply chain.

- **Changes to the concept of safety**: the use of AI in products and services can give rise to risks that EU legislation currently does not explicitly address. These risks may be linked to **cyber threats**, personal security risks (linked for example to new applications of AI such as to home appliances), risks that result from loss of connectivity, etc. These risks may be present at the time of placing products on the market or arise as a result of software updates or self-learning when the product is being used. The EU should make full use of the tools at its disposal to enhance its evidence base on potential risks linked to AI applications, including using the experience of the EU Cybersecurity Agency (ENISA) **for assessing the AI threat landscape**.

Scope of the future AI regulation: high-risk AI applications

The Commission is of the view that AI regulation should follow a **risk-based approach**, with clear criteria to differentiate between **low-risk**[2] or **high-risk** AI applications. High-risk AI applications should meet the following **two cumulative criteria**:

- **Sector**: the AI application appears on a **high-risk sector list**, which should be periodically reviewed and amended where necessary in function of relevant developments (including for instance, healthcare, transport, energy, and parts of the public sector).

- **Impact**: the AI application in the high-risk sector is, in addition, used in such a manner that **significant risks** are likely to arise[3]. The **assessment of the level of risk** of a given use **could be based on the impact on the affected parties.** For instance, uses of AI applications that produce legal or similarly significant effects for the rights of an individual or a company; that pose risk of injury, death or significant material or immaterial damage; that produce effects that cannot reasonably be avoided by individuals or legal entities.

[2] The Commission's White Paper never uses the term low-risk, but instead refers to 'application not qualified as high-risks'.

[3] This second criterion reflects the acknowledgment that not every use of AI in the selected sectors necessarily involves significant risks. For example, whilst healthcare generally may well be a relevant sector, a flaw in the appointment scheduling system in a hospital will normally not pose risks of such significance as to justify legislative intervention.

Low-risk AI applications, which do not qualify as high-risk, remain **entirely** subject to already existing EU-rules.

Exceptions: exceptional high-risk instances might also apply irrespective of the sector concerned[4]; they could deal, for instance, with employment equality, i.e. the use of AI applications for recruitment processes, as well as in situations impacting workers' or consumers' rights.

Remote biometric identification exception: the Commission proposes that AI for the purpose of remote biometric identification and other intrusive surveillance technologies **could always be considered high-risk,** and therefore the below requirements would at all times apply.

Requirements for high-risk AI applications

The requirements for high-risk AI applications, which will be further specified in **clear benchmarks** for all involved actors, could consist of the following **key features**:

Training data:

- **Safety requirements** aimed at providing reasonable assurances that the subsequent use of the products or services that the AI system enables is safe;
- **Non-discrimination requirements** to take reasonable measures aimed at ensuring that such subsequent use of AI systems does not lead to outcomes entailing prohibited discrimination (gender, ethnicity, and other grounds), including obligations to use data sets that are sufficiently representative;
- **Privacy requirements** aimed at ensuring privacy and personal data protection during the use of AI-enabled products and services.

Data and record-keeping: requirements aimed at keeping, during a limited and reasonable time period and to be made available upon request for testing or inspection by competent authorities:

- Accurate **records** regarding the data set used to train and test the AI systems;
- In certain justified cases, the **data sets** themselves;
- **Documentation** on the programming and training methodologies, processes and techniques used to build, test and validate the AI systems. Special arrangements should be made to protect confidential information, such as trade secrets.

Information: requirements aimed at ensuring:

- **Clear information** to be provided as to the AI system's capabilities and limitations, in particular the purpose for which the systems are intended, the conditions under which they can be expected to function as intended and the expected level of accuracy in achieving the specified purpose;
- Additional information to citizens about any not immediately obvious interaction with an AI system.

Robustness and accuracy: requirements aimed at ensuring:

[4] Other pieces of EU legislation may also apply: for example, when incorporated into a consumer product, the General Product Safety Directive may apply to the safety of AI applications.

- Robustness and accuracy, or at least correctly reflecting their level of accuracy, during all life cycle phases;
- Reproducibility of outcomes;
- Adequate error or inconsistency management during all life cycle phases;
- Resiliency against attacks and data or algorithm manipulation attempts, and mitigating measures.

Human oversight: human oversight could have the following, non-exhaustive, manifestations:

- Human validation of the output of AI systems before becoming effective;
- Human intervention afterwards the output of the AI system becomes immediately;
- AI monitoring while in operation and ability to intervene in real time and deactivate;
- Imposition of operational constraints on the AI system in the design phase.

Specific requirements for particular AI applications: in accordance with the current EU data protection rules and the Charter of Fundamental Rights, AI can only be used for **remote biometric identification** purposes **where such use is duly justified, proportionate and subject to adequate safeguards**. The Commission will launch a broad European debate on the specific circumstances, if any, which might justify such use, and on common safeguards.

A future regulatory framework for AI

Addresses: in a future regulatory framework, each obligation should be addressed to the actors who are **best placed to address any potential risks**; it is paramount that the requirements are applicable to all relevant economic operators providing AI-enabled products or services in the EU, regardless of whether they are established in the EU or not.

Prior conformity assessment for high-risk AI applications: to verify the compliance with the former requirements, the Commission proposes to perform a prior conformity assessment, including procedures for testing, inspection or certification, with checks of the algorithms and of the data sets used in the development phase. When designing and implementing a system relying on prior conformity assessments, particular account should be taken of the following:

- Not all requirements outlined above may be suitable to be verified through a prior conformity assessment. For instance, the requirement about information to be provided generally does not lend itself well for verification through such an assessment;
- Particular account should be taken of the possibility that certain AI systems evolve and learn from experience, which may require repeated assessments over the life-time of the AI systems in question;
- The need to verify the data used for training and the relevant programming and training methodologies, processes and techniques used to build, test and validate AI systems;
- In case the conformity assessment shows that an AI system does not meet the requirements for example relating to the data used to train it, the identified shortcomings will need to be remedied, for instance by re-training the system in the EU in such a way as to ensure that all applicable requirements are met.

The conformity assessments would be mandatory for all economic operators addressed by the requirements, **regardless of their place of establishment.**

Voluntary labelling for no-high risk AI applications

For **low-risk** AI applications that do not qualify as 'high-risk' an option would be, in addition to applicable legislation, to establish a **voluntary labelling scheme**. Under the scheme, interested economic operators that are not covered by the mandatory requirements could decide to make themselves subject, **on a voluntary basis**, either to those requirements or to a specific set of similar requirements especially established for the purposes of the voluntary scheme. The economic operators concerned would then be awarded a quality label for their AI applications. While participation in the labelling scheme would be voluntary, once the developer or the deployer opted to use the label, the requirements would be **binding**. The combination of ex ante and ex post enforcement would need to ensure that all requirements are complied with.

Governance

The Commission lets the option of a European governance structure on AI in the form of a framework for cooperation of national competent authorities **open to evaluation**.

The potential European governance structure could have as main activity an advisory task on standardisation and certification and the facilitation of the implementation of the legal framework, such as through issuing guidance, opinions and expertise. It should rely on a network of national authorities, as well as sectorial networks and regulatory authorities, at national and EU level; a committee of experts could provide assistance to the Commission. The carrying out of **conformity assessments** could be entrusted to **notified bodies designated by Member States**. Testing centres should enable the independent audit and assessment of AI-systems in accordance with the requirements outlined above. The governance structure relating to AI and the possible conformity assessments at issue here would leave the powers and responsibilities under existing EU law of the relevant competent authorities in specific sectors or on specific issues (finance, pharmaceuticals, aviation, medical devices, consumer protection, data protection, etc.) unaffected.

IP/A/ITRE/2020-05; Manuscript completed: April 2020; Date of publication: April 2020
Administrators responsible: Frédéric GOUARDÈRES, Matteo CIUCCI; Editorial assistant: Catherine NAAS
Contact: Poldep-Economy-Science@ep.europa.eu
This document is available on the internet at: www.europarl.europa.eu/supporting-analyses

Print ISBN 978-92-846-6470-2 | doi: 10.2861/312540 | QA-04-20-162-EN-C
PDF ISBN 978-92-846-6469-6 | doi: 10.2861/614816 | QA-04-20-162-EN-N