Structure for the White Paper on artificial intelligence – a European approach

1.

Artificial Intelligence is developing fast and will change our lives by improving diagnosis and healthcare, increasing the efficiency of farming, fighting climate change, increasing the security of Europeans and in many other ways that we can only begin to imagine. At the same time, there are fears that Artificial Intelligence entails a number of potential risks, such as racial discrimination, gender bias or intrusion in our private lives.

The European Commission is committed to enable scientific breakthrough, preserve the EU's technological leadership and ensure that new technologies are at the service of Europeans – improving their lives while respecting their rights.

To maximise the benefits and address the challenges of Artificial Intelligence, Europe has to act as one and define its own way to promote the development and deployment of AI based on European values.

Commission President Ursula von der Leyen announced in her Political Guidelines a coordinated European approach on the human and ethical implications of artificial intelligence as well as a reflection on the better use of big data for innovations.

The Commission supports a regulatory and investment approach that promotes the uptake of AI while addressing the risks associated with this new technology. The purpose of this White Paper is to set out policy options on how to achieve these objectives.

Stakeholders, including social partners, civil society organisations, Member States, researchers etc are invited to react to the presented options in order to inform future Commission proposals.

2. INTRODUCTION

As digital becomes part of every aspect of people's lives, trust in Artificial Intelligence becomes a prerequisite for its uptake. This is an opportunity for Europe, given its strong attachment to values and the rule of law and its proven capacity to build safe and reliable complex products and services from aeronautics to energy, automotive and medical equipment.

In the past, our economies have grown on the basis of raw materials, Europe's current and future economic growth and societal well-being are increasingly being built on creating value from data. Artificial intelligence is one of the first applications of the data economy. While today most data is related to consumers and is stored on central cloud-based infrastructures, a large part of tomorrow's far more abundant data will come from industry and businesses, and will be stored on a variety of systems, notably computing devices working at the edge of the network. This opens up new opportunities for Europe, which has a strong position in digitised industry and business-to-business applications, but is relatively weak in consumer platforms.

Simply put, AI is the combination of data, algorithms and machine learning, as well as computing power. Advances in computing and the availability of data are therefore key drivers of the current upsurge of artificial intelligence. Europe can combine its technology and industrial strengths with a world-class

digital infrastructure, and a regulatory framework based on its fundamental values to **become a global leader in innovation in the data economy and its applications** as set out in the European data strategy[1]. On that basis, it can develop an artificial intelligence ecosystem that brings the benefits of the technology to the whole of the society and economy. Thus, Europe can achieve a triple win:

- for **citizens** benefitting for example from improved health care (more precise diagnosis and prevention of diseases, new generation of medicines), safer and cleaner transport systems, better public services and improved safety;
- for **business** developing for example a new generation of products and services in the areas where Europe is particularly strong (machinery, transportation and construction industry, health care and high value added sectors like fashion and tourism); and
- for the **public interest**, for example by reducing the costs of providing services (transport, education, energy and waste management) and by improving the sustainability and energy efficiency of products.

Given the major impact that AI can have on our society, it is vital that European AI is grounded in our values of freedom, human dignity and privacy protection. It is paramount to define a common European approach to AI. Indeed, failing to do so, Member States will take (and are already starting to take) national measures which will result in a patchwork of national rules and a fragmented European framework, which will ultimately hamper legal certainty for industry and citizens' trust alike and prevent the emergence of a dynamic European AI industry.

This White Paper presents policy options to enable a rapid and ethical development of artificial intelligence in Europe, in full respect of values and rights of European citizens. The main building blocks of this White Paper are:

- The policy framework setting out measures to federate efforts at European and national level. In partnership with the private sector, the framework aims 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 adoption of solutions based on artificial intelligence, including by small and medium-sized undertakings (SMEs).
- Key elements of a future regulatory framework for artificial intelligence in Europe that will create a unique **'ecosystem of trust'.** To do so, the respect of EU rules must be ensured in particular for AI systems operated in the EU that constitute a high risk. Thus, the ecosystem of trust should give citizens the confidence to welcome artificial intelligence and give companies the legal certainty to innovate with artificial intelligence. The Commission strongly supports a human-centric approach that will be based on the Communication on Building Trust in Human-Centric Artificial Intelligence[2] and the input obtained during the piloting phase of the Ethics Guidelines prepared by the High-Level Expert Group on Artificial Intelligence.

Finally, the European strategy for data, which accompanies this White Paper, aims to enable Europe to become the most attractive, the most secure and the most dynamic data hub in the world. The strategy sets out a number of policy measures, including mobilising private and public investments, needed to achieve this goal.

---

1 COM(2020) XXX final.
2 COM(2019) 168 final.

3. CAPITALISING ON STRENGTHS IN INDUSTRIAL AND PROFESSIONAL MARKETS

Europe is well placed to benefit from the potential of artificial intelligence, not only as a user but also as a creator and producer of this technology. It has excellent research centres which publish more scientific articles related to artificial intelligence than any other region in the world. Europe has a world-leading position in robotics and competitive manufacturing and services sectors, from automotive to healthcare, from energy to financial services to agriculture. Europe has also developed strong computing power which is essential to the functioning of AI. Europe also holds large amounts of public and industrial data, the potential of which is yet under-used, and has well-recognised industrial strengths in safe and secure digital systems with low-power consumption, that are essential for the further development of artificial intelligence.

Europe should leverage its strengths to expand its existing market position along the value chain, from hardware manufacturing through software all the way to services. This is already happening to an extent. **Europe produces more than a quarter of industrial and professional service robots** (e.g. for precision farming, security, health, logistics), and plays an important role in the development and exploitation of **software applications for companies and organisations (business-to-business)**, as well as applications to progress towards e-government and the "intelligent enterprise".

Europe is also leading the way in deploying **artificial intelligence in manufacturing**: with more than half of its top manufacturers implementing at least one case of artificial intelligence in manufacturing operations[3].

One reason for Europe's strong position in terms of research is the EU funding programme that has proven instrumental in federating efforts, avoiding duplications, and leveraging public and private investments in the Member States. Over the past two years, EU funding for research and innovation for artificial intelligence has gone up by €1.5 billion, i.e. **an increase of 70% compared to the previous period.**

However, investment in research and innovation in Europe is still a fraction of the public and private investments in other regions of the world. Some €3.2 billion were invested in artificial intelligence in Europe in 2016, compared to around €12.1 billion in North America and €6.5 billion in Asia. To respond to the challenge, Europe needs to increase investment levels significantly. The **Coordinated Plan on artificial intelligence**[4] developed with Member States is proving invaluable in building stronger cooperation on artificial intelligence in Europe and in creating synergies for maximising investments into the artificial intelligence value chain.

4. SEIZING THE OPPORTUNITIES AHEAD: THE NEXT DATA WAVE

While Europe currently has a **weaker position in consumer applications and online platforms,** which results in a competitive disadvantage in data access, **important changes are underway.** Whereas around 80% of the current 40 zettabytes of data is stored in data centres, many of which are controlled by non-European operators, the advent of the Industrial Internet of Things and edge computing will result in a radical change in the distribution of data and in the nature of data collected. As a result, 80% of the 175 zettabytes of data that is expected to be available in 2025 will not be stored on platform, but rather should be stored locally at the edge of networks in factories, hospitals, etc. Platforms will no longer be dominant in this area[5].

---

[3] Followed by Japan (30%) and the US (28%). Source: CapGemini (2019)

[4] COM (2018) xxx

[5] IDC Data age 2025 study

Europe has world leadership in low power consumption electronics. Similarly, Europe should expand this strength into the area of specialised processors for artificial intelligence. Currently, third countries dominate this market. This could change with the help of initiatives such as the European Processor Initiative, which addresses the development of **low-power computing systems** for both edge and next generation high performance computing, and the efforts of the Key Digital Technology Joint Undertaking, to start in 2021, which will make edge computing its main objective for the next phase of low-power microelectronics.

Moreover, Europe is leading in **neuromorphic solutions**[6] that are ideally suited to automating industrial processes (industry 4.0) and transport modes. They can improve energy efficiency by several orders of magnitude.

Recent advances in **quantum computing** will allow for exponential increases in processing capacity[7]. Europe can be at the forefront of this technology thanks to its academic strengths in quantum computing, as well as European industry's strong position in quantum simulators and programming environments for quantum computing. European initiatives aimed at increasing the availability of quantum testing and experimentation facilities will help applying these new quantum solutions into various industrial and academic sectors.

In parallel, Europe will continue to lead progress in the **algorithmic foundations of artificial intelligence**, building on its own scientific excellence. There is a need to build bridges between existing silos, such as machine learning and deep learning (characterised by limited interoperability, the need of large amount of data to run the models and learning through correlations) and symbolic approaches (where rules are created through human intervention). Combining symbolic reasoning with deep neural networks may help us improve explainability of AI outcomes.

5. AN ECOSYSTEM OF EXCELLENCE

To build an ecosystem of excellence that can support the development and uptake of artificial intelligence across the EU economy and public administration, there is a need to step up action at multiple levels.

A. Working with Member States

Delivering on its strategy on artificial intelligence (adopted in April 2018), in December 2018 the Commission presented a Coordinated Plan - prepared with Member States - to foster the development and use of AI in Europe[8].

This plan proposes some 70 joint actions for closer and more efficient cooperation between Member States, and the Commission in key areas, such as: Research, Investment, Market uptake, Skills and talent, Data and International cooperation. The plan is expected to run until 2027 and to be monitored and reviewed regularly.

We want to further **maximise the impact of investments** in research, innovation and deployment, assessing national AI strategies and building on and extending the Coordinated Plan on artificial intelligence with Member States:

---

[6] To be added a definition of "neuromorphic solutions"

[77] Quantum computers will have the capacity to process in less than seconds many fold larger data sets than today's highest performance computers allowing for the development of new AI applications across sectors.

[8] https://ec.europa.eu/digital-single-market/en/news/coordinated-plan-artificial-intelligence

- *Action 1: The Commission, taking into account the results of the public consultation on the White Paper, will propose to the Member States a revision of the Coordinated Plan to be adopted by end 2020*

EU-level funding in artificial intelligence should attract and federate investments in areas where the efforts required go beyond what any single Member State can achieve. The objective is to attract **more than €20 billion** of total investment per year in artificial intelligence in the next decade in the EU.

B. Federating and focusing the efforts of the research and innovation community

Europe cannot afford to maintain the current dispersed landscape of centres of competence with none of them reaching the scale to be at the level of the world leading institutes. It is imperative to create more synergies between the various European research centres on artificial intelligence and to federate their efforts to improve their excellence, retain and attract the best researchers and develop the best technology. Europe needs a lighthouse centre of research, innovation and expertise that should be a world reference of excellence in AI and that can attract investments and best talents in the field.

The centres and the network should concentrate in sectors where Europe has the potential to become a global champion such as **industry, health, transport, agrifood chains, energy/environment, earth observation and space**. In all these domains, the race to become global leaders is ongoing, and Europe has significant potential, knowledge and expertise. Equally important is the creation of testing and benchmarking sites that allow validation and certification of products and services relying on artificial intelligence.

- *Action 2: the Commission, as part of the Industrial Strategy, will propose in the second quarter of 2020 a **legal instrument** to facilitate the creation of excellence and testing centres that can enjoy the seal of "EU excellence" and can federate European, national and private investments. The Commission has proposed an ambitious and dedicated amount to support world reference testing centres in Europe under the Digital Europe Programme as part of the Multiannual Financial Framework for 2021 to 2027.*

C. Focus on SMEs

It will also be important to ensure that SMEs can access and use artificial intelligence. To do this, the Digital Innovation Hubs and the artificial intelligence-on-demand platform should be strengthened further. The Digital Europe Programme will be instrumental. While all innovation hubs should provide support to SMEs to understand and adopt artificial intelligence, it will be important that at least one innovation hub per Member State has a high degree of specialisation in artificial intelligence.

SMEs and start-ups will need access to finance in order to adapt their processes or to innovate using artificial intelligence. Building on the forthcoming pilot investment fund of €100 million in artificial intelligence and blockchain, the Commission will further scale up the access to finance in artificial intelligence under InvestEU.

The InvestEU programme aims to give an additional boost to investment, innovation and job creation in Europe by triggering at least 650 billion euros in additional investment (on the basis of the Commission proposal). It will support sustainable infrastructures; research innovation and digitalisation; SMEs; and social investment and skills.

It will be important to accelerate the development and uptake of artificial intelligence applications by European SMEs and to foster a stronger EU market for applications of artificial intelligence in these value chains. These actions should also include other intermediaries such as industrial clusters, regional competence centres and intelligent cities' infrastructures.

- *Action 3: the Commission will work with Member States to ensure that at least one digital innovation hub per Member State has a high degree of specialisation on artificial intelligence. Subject to an agreement on the next Multiannual Financial Framework, €900 Million are foreseen for Digital Innovation Hubs under the Digital Europe Programme.*
- *The Commission and the European Investment Fund will launch a pilot scheme of €100 million in Q1 2020 to provide equity financing for innovative developments in artificial intelligence. This will be scaled up at least 10 times from 2021 through InvestEU.*

D. Partnership with the private sector

It is also essential to make sure that the private sector is fully involved in setting the research and innovation agenda and provides the necessary level of co-investments. This requires both: to set up a large public private partnership, and the commitment of the top management of companies.

- *Action 4: the Commission will establish a new public private partnership in artificial, intelligence and robotics, which can develop into a Joint Undertaking in artificial intelligence that federates efforts, ensures coordination of research and innovation in artificial intelligence, co-investments in deployment of relevant infrastructures and testing facilities as well as roll out in public sector.*

E. Promoting the adoption of artificial intelligence by the public sector

It is essential that public administrations, hospitals, utility and transport services and other areas of public interest rapidly begin to deploy products and services that rely on artificial intelligence into their activities. A specific focus will be in the areas of health care and transport where technology is mature for large scale deployment.

- *Action 5: The Commission will initiate sector dialogues giving priority to health care and public service operators in order to present an action plan by mid-2020 to facilitate adoption. The sector dialogues will be used to prepare a specific 'Adopt artificial intelligence programme' that will support public procurement of artificial intelligence systems and help to transform public procurement processes themselves.*

F. Securing access to data and computing infrastructures

The actions set out in this White Paper are complementary to the plan presented in parallel in the European data strategy. Improving access to and management of data is a fundamental issue. Without data, the development of artificial intelligence and other digital applications is not possible. More than 33 zettabytes of data were generated worldwide in 2018, a number that is set to grow to 175 by 2025. This enormous amount of new data yet to be generated constitutes an opportunity for Europe to reassert itself at the forefront of the data and artificial intelligence revolution. Equally important is investment in key computing technologies and infrastructures.

The Commission has proposed more than €4 billion to support high performance and quantum computing including edge computing as well as the artificial intelligence, data and cloud infrastructure in the Digital Europe programme. The European data strategy develops these issues further.

G. Skills

The European approach to artificial intelligence will need to have a strong skills dimension. While upskilling and education policies are mainly Member States' competences, the Commission can be

instrumental in mobilising efforts and supporting best practices to be replicated at national level. The Commission will soon present an update of the Skills Agenda, which aims to ensure that everyone in Europe can benefit from the green and digital transformations of the EU economy.

Reinforcing the skills that are relevant for artificial intelligence as well as adapting education systems and upskilling the workforce to make them fit for the artificial intelligence-led transformation will be a priority of the revised Coordinated Plan on artificial intelligence and of the future Digital Education Action Plan. This will include transforming the assessment list of the ethical guidelines into a proper curriculum for developers of artificial intelligence.

In addition, by creating a centre of excellence for AI in Europe, talents from all over the world will be attracted by such a centre and the possibilities it offers. Such a centre also develops and reproduces an excellence of skills which can spread and nourish across Europe (see action 2)

- *Action 6: Establish and support through the advanced skills pillar of the Digital Europe Programme networks of leading universities and higher education institutes to attract the best professors and scientists and offer world-leading masters programmes in artificial intelligence..*

H. International aspects

The EU's work on artificial intelligence has already influenced international discussions. When developing its ethical guidelines, the High-Level Expert Group involved a number of non-EU organisations (from US and Canada) and several governmental observers (from Japan and Singapore). In parallel the EU was closely involved in developing the Organisation for Economic Co-operation and Development's ethical principles for artificial intelligence.[9] The G20 subsequently endorsed these principles.

The EU will continue to cooperate with global players on artificial intelligence based on an approach that promotes the EU's interests (e.g. mainstreaming European standards, accessing key resources including data, creating a level playing field). The Commission is convinced that international cooperation must be based on an approach that promotes the respect of fundamental rights including human dignity, pluralism, gender equality, non-discrimination and protection of privacy.[10]

6. AN ECOSYSTEM OF TRUST: REGULATORY FRAMEWORK FOR ARTIFICIAL INTELLIGENCE

As any new technology, the use of artificial intelligence brings both opportunities and risks. In addition to a lack of investment, the main factor holding back a broader uptake of artificial intelligence is a lack of trust[11]. Citizens fear having to deal with unaccountable machines taking incomprehensible decisions while companies are concerned by legal uncertainty.

That is why the Commission set out an AI strategy[12] on 25 April 2018 addressing the socio-economic dimension in parallel with the reinforcement of investment in research, innovation and AI-capacities across the EU. A Coordinated Plan[13] was set up with the Member States to align strategies. The

---

9 Source to be added

10 Under the Partnership Instrument, the Commission will finance a €2.5 million project that will facilitate cooperation with like-minded partners, in order to promote the EU artificial intelligence ethical guidelines and to adopt common principles and operational conclusions.

11 Eurobarometer…
IBM conducted a research which found that while 82% of enterprises are considering deployments of artificial intelligence, 60% fear liability issues and 63% lack the in-house talent to confidently manage the technology. (link?)

12 COM(2018) 237

13 COM(2018) 795

Commission established also a high-level expert group that drafted Guidelines for the relevant legal and ethical challenges. The Group published Guidelines on trustworthy AI in April 2019.[14] The Commission welcomed its seven key requirements in a dedicated Communication[15]:

- Human agency and oversight,
- Technical robustness and safety,
- Privacy and data governance,
- Transparency,
- Diversity, non-discrimination and fairness,
- Societal and environmental well-being, and
- Accountability.

In addition, the Guidelines contain a concrete assessment list for practical use by companies. During the second half of 2019, this assessment list has been tested by around 350 organisations. The High-Level Group is in the process of revising its guidelines in light of this feedback. The key result of the feedback process is that businesses find transparency and human oversight important but that in many economic sectors these are not yet specifically required by existing legislation.

Beyond this set of non-binding Guidelines, a clear European regulatory framework could help enhance the trust of consumers and businesses in artificial intelligence, and therefore speed up the uptake of the technology. The regulatory framework inspired by the key requirements of the High Level Expert Group has to be consistent with efforts to promote Europe's innovation capacity and competitiveness in this field. It has to ensure socially and economically optimal outcomes and the respect of EU-legislation, principles and values.

Already today, developers and deployers of artificial intelligence are subject to European legislation on fundamental rights as spelled by e.g. data protection, privacy, non-discrimination, gender equality, consumer protection, and product safety and liability rules. At the same time, consumers expect the same rights and safety level no matter whether a product relies on artificial intelligence or not. However, the specific features of artificial intelligence can make the application and enforcement of this legislation more difficult. For this reason, there is a need to examine whether existing legislation deals sufficiently with the risks of artificial intelligence or whether it needs to be changed.

Given how fast Artificial Intelligence evolves, the regulatory framework must leave room for further developments. Any changes should be limited to clearly identified problems for which feasible solutions exist.

In the absence of a common European framework, it is to be expected that Member States will introduce different regulatory frameworks at national level. For example, German authorities have called for a European regulatory framework; the German Data Ethics Commission has called for a five-level risk-based system of regulation that would go from no regulation for the most innocuous artificial intelligence systems to a complete ban for the most dangerous ones. Denmark has just launched the prototype of a Data Ethics Seal. Malta has introduced a certification system for AI.

A solid European regulatory framework for trustworthy AI will help to create a frictionless internal market for the further development and uptake of Artificial Intelligence as well as the strengthening of Europe's industrial basis in AI.

---

[14] file:///C:/Users/siolylu/AppData/Local/Microsoft/Windows/INetCache/IE/Q7POTYT3/AIHLEG_EthicsGuidelinesforTrustworthyAI-ENpdf.pdf

[15] COM(2019) 168

A. PROBLEM DEFINITION

While artificial intelligence can do much good, including by making products and processes safer, it can also do harm. This harm might be both material (safety and health of individuals, including loss of life, damage to property, etc.) and immaterial (loss of privacy, limitations to the right of freedom of expression, human dignity, discrimination, etc.), and can relate to a wide variety of risks. A regulatory framework should concentrate on how to minimise the various risks of potential harm.

Particular concerns arise because AI systems may be perceived as something of a "black box": the understanding of the logic applied by the system is very limited. Even the original developers may not know why a certain decision is reached. More generally, the reasoning behind a given output may be very difficult and resource intensive to trace. This particular complexity of AI systems make the appropriate application and enforcement particularly challenging.

At the same time, artificial intelligence systems based on machine learning may make applying and enforcing existing EU law more difficult. In addition, enforcement authorities may lack the means to verify how a given automated decision was taken, or whether existing rules were respected. A regulatory framework should contribute to the seamless enforcement of existing EU law.

The main risks concern data protection, fundamental rights and safety.

i. [16]Risks for fundamental rights, including, privacy data protection and discrimination

The use of artificial intelligence may lead to breaches of fundamental rights[17], including the freedom of expression, the freedom of assembly, human dignity, non-discrimination based on sex, race, ethnic origin and other grounds data protection, private life or the right to an effective judicial remedy and a fair trial, and consumer protection. These risks might be the result of flawed design of artificial intelligence systems (e.g. errors in speech or image recognition) or the use of data that is biased (e.g. the system is trained using only data from men).

Artificial intelligence can perform many functions that previously could only be done by humans. As a result, citizens will increasingly be subject to decisions by machines and may find it sometimes difficult to have these decisions reviewed by human operators. Moreover, AI gives organisations the possibility to track and analyse the daily habits of people. It allows state authorities or other entities to use artificial intelligence for mass surveillance. By analysing large amounts of non-personal data and identifying links among them, artificial intelligence can also be used to retrace and de-anonymise personal data about certain people, creating new personal data protection risks even in respect to datasets that per se do not include personal data.

> An employer advertises for a job opening in a male-dominated industry via a social media platform. The platform's ad algorithm brings the job opening only to the attention of men in a mistaken attempt to maximise returns on the number and quality of applicants. *Source: Noam Scheiber, "Facebook Accused of Allowing Bias Against Women in Job Ads," The New York Times, September 18, 2018.*

[16] (Explanation from HLG tob e inserted.)

[17] Council of Europe research shows that a large number of fundamental rights could be impacted from the use of artificial intelligence, https://rm.coe.int/algorithms-and-human-rights-en-rev/16807956b5.

Bias and discrimination are inherent risks of any societal or economic activity. Human decision-making is prone to mistakes and biases. However, the same level of bias when present in artificial intelligence could affect and discriminate many people without the social control mechanisms that govern human behaviour[18]. This can also happen when the system 'learns' while in operation. In such cases, and provided that such outcome could have not been prevented or anticipated at the design phase, the risks will not stem from a flaw in the original design of the system but from the practical impacts of the correlations or patterns that the system identifies in a large dataset.

> Certain artificial intelligence programmes for facial analysis display gender and racial bias, demonstrating low errors for determining the gender of lighter-skinned men but high errors in determining gender for darker-skinned women. *Source: Larry Hardesty, "Study Finds Gender and Skin-Type Bias in Commercial Artificial-Intelligence Systems," MIT News, February 11, 2018.*

Artificial intelligence might also give rise to risks for the freedom of association, the freedom of expression, the privacy and the protection of personal data.[19] For example, employers can use artificial intelligence to observe how their employees behave in the workplace.

ii. Risks for safety and the effective functioning of the liability system

Artificial intelligence technologies may present new safety risks for users when they are embedded in products and services. For example, as result of a flaw in the object recognition technology, an autonomous car can wrongly identify an object on the road and cause an accident involving injuries and material damage. As with the risks to fundamental rights, these risks can be caused by flaws in the design of the artificial intelligence technology, related to problems with the availability and quality of data, or other problems stemming from machine learning. While some of these risks are not limited to products and services that rely on artificial intelligence, the use of artificial intelligence may increase or aggravate the risks.

A lack of clear safety provisions may create legal uncertainty for businesses that are marketing their artificial intelligence products in the EU. Market surveillance and enforcement authorities may find themselves in a situation where they are unable to intervene because they are not empowered to act. Legal uncertainty may therefore reduce overall levels of safety, and may complicate liability claims, as the expected level of safety of such technologies is unclear.

---

[18] The European Commission's Advisory Committee on Equal Opportunities for Women and Men is currently preparing an "Opinion on Artificial Intelligence" analysing inter alia the impacts of artificial intelligence on gender equality which is expected to be adopted by the Committee in early 2020.

[19] The General Data Protection Regulation and the ePrivacy Directive (new ePrivacy Regulation under negotiation) address these risks but there might be a need to examine whether artificial intelligence systems pose additional risks. The Commission will be monitoring and assessing on a continuous basis if the General Data Protection Regulation needs to be complemented to tackle these challenges.

If the safety risks materialise, the lack of clear requirements and the characteristics of artificial intelligence technologies mentioned above, including complexity, autonomy[20] and opacity ('black box-effect'), make it difficult to trace back problematic decisions that artificial intelligence systems make. This in turn may make it difficult for victims of damages to obtain compensation under the current EU and national liability legislation.[21].

> Under the Product Liability Directive, a consumer can claim compensation for damage caused by a defective product. However, in the case of an AI based system such as autonomous cars, it may be difficult for a consumer 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 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.

Victims of damage may not have effective access to the evidence that is necessary to build a case in court and be less protected compared to situations when the damage is caused by traditional technologies. These risks of harm occurring will increase as the use of artificial intelligence becomes more widespread.

By the same token, the challenges with the difficulty of tracing back problematic decisions made by artificial intelligence systems apply equally to the risks for fundamental rights referred to in the previous section, including discrimination, privacy and data protection.

B. EXISTING EU LEGISLATIVE FRAMEWORK FOR ARTIFICIAL INTELLIGENCE

An extensive body of existing EU product safety and liability legislation, further complemented by national legislation and relevant standards is relevant and potentially applicable to a number of emerging artificial intelligence applications. There is also EU legislation (including the General Product Safety Directive, the Machinery Directive, the Radio Equipment Directive and the Product Liability Directive) to address the risks of safety and liability for defective products. This safety and liability framework normally applies when artificial intelligence systems fall under the category of regulated products (for instance, a car, a drone, an industrial robot or a medical device).[22]

The Report on the broader implications of artificial intelligence, Internet of Things and other digital technologies for the safety and liability framework, which accompanies this White Paper, provides an overview of relevant EU and national legislation. It identifies their shortcomings with respect to the specific risks posed by artificial intelligence and other digital technologies.

The Commission is in the process of considering adaptations of the relevant safety and liability legislation to address *inter alia* the risks of artificial intelligence and other new technologies. The legislator has already tasked the Commission to do this for the approval of autonomous and connected cars[23].

---

[20] When presented with a given input, very complex artificial intelligence systems can be difficult to predict in terms of output that they will generate. This means that artificial intelligence can sometimes display surprising "behaviour". The possible "behaviour" of an artificial intelligence system in its environment is nevertheless largely defined and constrained by its developers. Humans determine and programme the goals which a system should optimise for.

[21] The implications of artificial intelligence, Internet of Things and other digital technologies for safety and liability legislation are analysed in the Commission Report accompanying this White Paper.

[22] The EU legal framework for product safety consists of the General Product Safety Directive and a number of sector-specific rules covering different categories of products ranging from machines, planes and cars to toys and medical devices. Product liability law is complemented by different systems of civil liability for damages caused by products or services.

[23] New Vehicle General Safety Regulation (EU) 2144/2019. The Commission will now develop the relevant implementing legislation by Q1/2021.

As regards the protection of fundamental rights and consumer rights, the EU legislative framework includes legislation such as the Race Equality Directive[24], the Employment Equality Directive[25], the Access to goods and services Directive[26], a number of consumer protection rules[27] as well as rules on personal data protection and privacy, notably the General Data Protection Regulation. In addition, fundamental rights need to be respected when implementing other EU legislation, including in the field of financial services, migration or responsibility of online intermediaries.

While the EU legislation remains in principle fully applicable, it is important to assess whether it can be enforced adequately to address the risks that artificial intelligence systems create, or whether adjustments are needed.

The Commission is of the opinion that the legislative framework could be improved in the following areas:

- *Effective application and enforcement of existing EU and national legislation*: the key characteristics of artificial intelligence create challenges for ensuring the proper application and enforcement of EU and national legislation. The lack of transparency (opaqueness of AI) makes it difficult to prove possible breaches of rules, attribute liability and meet the conditions necessary 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.

- *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 there are cases where software per se is regulated by EU product safety legislation[28], generally speaking software is not regarded as a product in EU product safety legislation, even if it is relevant for product safety. Moreover, the existing legislation does not apply to services, and thus not to services based on artificial intelligence (e.g. health services, financial services, transport services).

- *Changing nature of AI systems*: the integration of software, including artificial intelligence, into products can modify the functioning of AI 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.

- *Uncertainty as regards the allocation of responsibilities between different economic operators in the supply chain*: EU legislation on product safety does not always cover all economic actors who develop and integrate artificial intelligence into systems. The rules can for example become unclear if artificial intelligence is added after the system is placed on the market by a party that is not 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 artificial intelligence 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 safety risks (linked for example to new applications of AI such as to the home appliances) risks that result from loss of connectivity, etc. These risks may be

---

[24] Directive 2000/43.
[25] Directive 2006/54.
[26] Directive 2004/113.
[27] Such as the Unfair Commercial Practices Directive and the Consumer Rights Directive.
[28] For instance software intended by the manufacturer to be used for medical purposes is considered a medical device under the Medical Device Regulation (Regulation (EU) 2017/745).

present at the time of placing products on the market or arise as a result of software updates or machine learning when the product is being used.

As indicated earlier, several Member States are already exploring options for national legislation to address the challenges created by artificial intelligence. This raises the risk that the single market may be fragmented. Divergent national rules are likely to create obstacles for companies that want to sell and operate artificial intelligence systems in the single market. Ensuring a common European approach would enable European companies to benefit from smooth access to the single market and support their competitiveness on global markets.

C. *SCOPE OF THE REGULATORY FRAMEWORK*

From the discussion above, the Commission concludes that one option would be to complement existing legislation applicable to artificial intelligence with additional rules with a view to make it fit for the current technological level of development.

A key issue for the future regulatory framework is to determine the scope of its application. The working assumption is that the regulatory framework would apply to products and services relying on artificial intelligence. For the purpose of ensuring legal certainty, it is important to define the term 'artificial intelligence' [29].

Products or services that rely on Artificial Intelligence use algorithms implemented in software and/or hardware, that identify the actions the product or service should take in order to achieve a given goal, by interpreting, reasoning, and learning from data (often large amounts of data).

The two main elements that compose Artificial Intelligence are therefore "data" and "algorithms". Algorithms are "trained" using the data in order to determine the actions needed to achieve a given goal, and may continue to learn when in use.

> In autonomous driving for example, the algorithm uses, in real time, the data from the car (speed, engine consumption, shock-absorbers, etc..) and from the sensors scanning the whole environment of the car (road, signs, other vehicles, pedestrians etc..) to derive which direction, acceleration and speed the car should take to reach a certain destination. By learning from the data observed, the algorithm adapts to the situation of the road and to the outside conditions including other drivers' behaviour to derive the most comfortable and safest drive.

The definition of artificial intelligence must be sufficiently flexible to accommodate technical progress while providing the necessary legal certainty.

In order to respect subsidiarity and proportionality, it may be necessary to consider some exceptions to the scope of the regulatory framework including in particular the development and use of artificial intelligence for military purposes.

---

[29] A functional definition of artificial intelligence should look at the characteristics that differentiate artificial intelligence from more general terms, such as software. While the term 'software' is used in several acts of EU law, such as the Directive on the legal protection of computer programs, there is no established definition. The Directive specifies that the term 'computer program' refers to programs in any form, including those which are incorporated into hardware[29].
Other, more technical definitions are possible as well: Such alternative technical approaches to the definition would for instance focus on systems that are trained with the machine learning technique, covering inter alia: deep learning and back-propagation, supervised learning, unsupervised learning, reinforcement learning, generative adversarial networks and symbolic reasoning.

D. POSSIBLE SITUATIONS WHERE THE REQUIREMENTS WOULD APPLY

As a matter of principle, the horizontal framework for AI **should not be excessively prescriptive** so that it could create a disproportionate administrative burden, especially for SMEs.

The Commission is of the view that **it should focus on "high-risk" applications.**

A differentiated risk-based approach would indeed help to ensure that the regulatory intervention is proportionate, but it also requires clear criteria to differentiate between the different AI applications along a scale from very 'low-risk' to very 'high-risk' systems. The determination of what is a high-risk application should be easy for companies to understand and to make.[30] This is necessary to ensure a smooth implementation by all relevant economic actors as well as national competent authorities.

The criteria to determine high levels of risk could include the following:

- Defining high-risk sectors (e.g. healthcare, transport), possibly combined with an indicative or exhaustive list with the possibility for the Commission to amend the list; or
- Defining high-risk uses (e.g. predictive policing), possibly combined with an indicative or exhaustive list with the possibility for the Commission to amend the list; or
- (Self-)Identifying the level of risks through a risk assessment carried out by the developer and/or user of artificial intelligence; or
- Other types of criteria taking into account the context:
    - whether the individual or legal entity cannot avoid being affected by the output of an artificial intelligence system, or risk suffering negative consequences as a result of the decision to 'opt out' (e.g. healthcare applications); or
    - how important the output of the artificial intelligence system is for an individual or legal entity (e.g. social security benefits); or
    - whether the output of the artificial intelligence system with a significant effect for an individual or legal entity is irreversible (e.g. collision avoidance in self-driving vehicles[31]); or
    - whether the individuals or legal entities affected by the output of the artificial intelligence system are in a specific area with a high risk, for example of prohibited discrimination (e.g. recruitment proceedings discriminating women or workers based on ethnicity).

Having considered the different options, **the Commission is of the opinion that the definition of 'high-risk' applications should rely on the cumulative application of two criteria**:

- an exhaustive list of sectors (e.g. healthcare, transport, police, judiciary) that would be specified in an annex of the legislative instrument and subject to possible amendments by the Commission if necessary, and

[30] A well-tailored legal provision for that purpose could combine two criteria: (i) a general definition anchored around any decision, made by an AI or by a human following a recommendation by an AI, that produces legal effect for an individual or a legal entity or otherwise poses risks of injury, death or other significant material damage; and (ii) a list of sectors where risks are particularly acute (e.g. healthcare, education, finance, transport, police, judiciary) with in-built flexibility mechanisms for adaptations on a need basis.]

- in addition, a more abstract definition of 'high-risk' use of artificial intelligence used or intended for use in the identified sectors. For instance, it could cover applications of artificial intelligence which can: produce legal or similarly significant effects to the rights of an individual or a company; or which pose risk of injury, death or significant material or immaterial damage; or applications where their effects cannot be avoided by individuals or legal entities; or where there is a risk of serious negative consequences as a result of the decision to 'opt out'.

A combined application of the two criteria would ensure that the scope of application is focused and would provide a reasonable level of legal certainty for relevant economic operators. Mandatory requirements would apply only to those applications that fall under both criteria. It would be indeed difficult to rely on a closed list of applications, as the level of risk will often depend on the actual use of artificial intelligence. This is reflected in the choice of sectors as one of the criteria for determining the level of risk. At the same time, not all applications of artificial intelligence in a given sector will pose a high-level of risk to protected legal interests. For example, a flaw in an automated parking management system in a hospital will probably not pose significant risks. Therefore, the second criterion would help to ensure that only those applications are covered which pose a significant risk to rights or of creating harm, or that have legal or similarly significant effects for an individual or a legal entity.

For applications that do not fulfil the combination of the two criteria, the existing provisions of EU legislation would continue to apply. That legislation includes for example the provisions of the General Data Protection Regulation on the information the data subject must receive about the use of automated processing, including profiling. However, should any AI applications supplier want to demonstrate it respects European rules, there would be a voluntary scheme of labelling as described in section H.

As a not exhaustive illustration, of high-risk applications, possible options could be high-risk uses in the following sectors:

- in system and products that endanger the human life, public health or present severe risks for the public order (for example medical equipment)
- in national security (predictive policing)
- in the provision of services by the public sector, including in public procurement, for example traffic management systems
- in the provision of services of public interest (such as airports, private hospitals and other services provided by private operators with a public impact)
- or for the following high-risk uses:in the use of biometric identification systems
- in recruitment processes

E. TYPES OF REQUIREMENTS

When designing the future regulatory framework for artificial intelligence, it will be necessary to decide on the types of **mandatory legal requirements** to be imposed on the developers and professional users of artificial intelligence and at which level these requirements will apply. It will have to differentiate also between high risk AI systems and others, and adapt the level and intrusiveness of the regulatory intervention. In particular, **with regard to high-risk AI systems (as defined under point F) operated in the EU, it is paramount that EU rules are respected by all actors, regardless whether they are EU based or not.**

The behaviour of any AI systems will very much depend on the data set on which the algorithm has been trained. An algorithm trained on non-European data will give a different result than the same algorithm trained on European data. Therefore, for a European approach to AI to emerge that respects European values, ethics and rules, it is necessary to set appropriate requirements at the level of the data training set and the data which have fed the algorithm.

Taking into account the guidelines of the High Level Expert Group and the findings of the pilot phase of these guidelines and following the above reasoning, the requirements for high-risk applications could address data, traceability, transparency and human oversight through a conformity process controlled by public authorities in the EU through the following **options (alternative or cumulative):**

a) Data quality and traceability obligations

- The EU framework on data quality and traceability will apply to all AI, in a non-discriminatory way. All AI applications and algorithms are welcome to service the European market as long as they comply with EU rules and values.
- To ensure that AI is developed in full respect of European values from the very outset, high-risk AI systems must be trained through quality data that have been gathered and used according to European rules and requirements[32].
- This means that the data must:
    - respect fundamental rights as set out in the Charter of Fundamental Rights of the European Union and in existing implementing legislation;
    - respect privacy, confidentiality and integrity of the data sets in conformity with GDPR and any other applicable rules;
    - respect specific requirements in terms of accuracy, replicability and reproducibility, as defined by the High Level Expert Group in the Ethics Guidelines for Trustworthy AI (April 2019). The regulatory framework will prescribe that accuracy, replicability and reproducibility shall be properly documented and that relevant data, documentation and software must be available for test and inspection by competent Authorities.
- If compliance with the above requirements in accordance with the conformity assessment laid out under chapter G for systems developed outside the EU cannot be determined, the systems must be retrained in Europe and with European data sets in order to ensure traceability and compliance.
- This ex ante check of compliance with European rules and values will be done, at least for high-risk applications, according to the conformity assessment laid out under chapter G on Enforcement.
- Accurate records regarding the data used for training AI, including a description of how the data set was selected and of the methodologies used for de-biasing and mitigating the risk of discriminatory or unfair outcomes, should be maintained.

b) Transparency and human oversight obligations

- High-risk AI systems should maintain transparency about the respect of European rules through documentation of the quality of the datasets, accuracy, programming methodologies and techniques

[32] Footnote: The approach is similar to the scheme on the ban on animal testing foreseen in the cosmetics Regulation 1223/2009.

used to build, test and validate the AI system, including – where relevant - with considerations regarding bias.

- AI system's capabilities and limitations, including the purpose for which the AI can be used, the conditions under which it can function as expected and the level of accuracy have to be communicated to authorities, operators or end-users.
- Citizens should be made aware when they are interacting with an AI system and not a human.
- An appropriate human oversight must be ensured: a human should review the automated decision or recommendation produced by an AI in cases where core individuals' rights are at stake or where the protection offered by GDPR is considered not sufficient.

c) ***Specific requirements for Remote Biometric Identification Systems***

The way biometric data are obtained and used – potentially without consent or opportunities to opt out – [33] carries specific risks for fundamental rights[34]. The fundamental rights implications of using biometric identification systems vary considerably depending on the purpose, context and scope of the use. Some of the fundamental rights implications stem from the technology's lack of accuracy. Accuracy has strongly increased, but the technologies still always come with a certain rate of error, which can negatively impact fundamental rights. Moreover, importantly, several fundamental rights concerns would remain even if there were a complete absence of errors. There is not much information available about the actual use of Biometric Identification Systems in the EU, this lack of more comprehensive information limits the opportunities to analyse its implications.

- In addition to the requirements in a) and b), guidance should be provided under which conditions the processing of biometric data is proportionate. This assessment will depend on the purpose for which the technology is used and on the safeguards in place to protect individuals. In case biometric data are used for mass surveillance, there must be clear criteria about which individuals should be identified. Finally, depending on the use of the system and the level of risk, there should be certain accuracy requirements.

Another option, either alternative of cumulative, would be a **voluntary scheme** whereby interested entities can commit to comply with the legal requirements in return for a quality label for their AI applications. This could help them to signal that their AI products and services are trustworthy. If it operates in addition to a complemented regulatory framework, this option would consist of a legal instrument that sets out a voluntary labelling framework for developers and deployers of artificial intelligence systems that would not be considered as high-risk.

The voluntary scheme either could consist in the above mentioned requirements or it could revolve around the assessment list of the ethics guidelines of the High Level Expert Group. If the developer or deployer complied with these voluntary requirements, they would be allowed to use the label of 'trustworthy artificial intelligence'. While participation in the labelling scheme would be voluntary, once the developer or professional user opted to use the label, the requirements would be binding. This

[33] Biometric data is defined 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, such as facial images or dactyloscopic [fingerprint] data." (Law Enforcement Directive, Art. 3 (13); GDPR, Art. 4 (14); Regulation (EU) 2018/1725, Art. 3 (18).

[34] For example on people's dignity. Relatedly, the rights to respect for private life and protection of personal data are at the core of fundamental rights concerns when using facial recognition technology. Moreover, freedom of expression, association and assembly must not be undermined by the use of the technology. There is a potential impact on non-discrimination and rights of special groups, such as children, older persons and persons with disabilities. See: Facial recognition technology: fundamental rights considerations in the context of law enforcement, https://fra.europa.eu/en/publication/2019/facial-recognition

scheme, which would complement the legal framework, would have to include measures to ensure enforcement.

The objective of a voluntary labelling scheme would not be to address comprehensively the risks for safety, liability or fundamental rights. Instead, its purpose would be to incentivise the development of 'trustworthy artificial intelligence' in Europe and globally. It would allow users to recognise easily artificial intelligence systems that have been developed in compliance with the relevant standards for trustworthy artificial intelligence. This would help enhance the trust of users in artificial intelligence systems and promote the overall uptake of the technology. It would complement the existing mandatory conformity assessment mechanisms that apply under the EU product safety legislation.

A voluntary framework would also strengthen EU leadership in the discussions on 'trustworthy' artificial intelligence at the international level, while allowing the European and foreign developers and professional users of artificial intelligence to apply these additional rules in a flexible manner.

F. ADDRESSEES

Many economic actors are involved in the lifecycle of an artificial intelligence system. These include the developer of the algorithm, the deployer (the operator who brings the product or service to the market) and potentially others (producer, distributor or importer of a product based on artificial intelligence, the provider of services based on artificial intelligence, the professional or private user of a product or the recipient of a service relying on artificial intelligence).

In the future regulatory framework the responsibility should lie with the actor(s), who is/are best placed to address any potential risks. While the developers of artificial intelligence may be best placed to address risks arising from the development phase, their ability to control risks during the use phase may be more limited. In that case the deployer should be responsible.

G. ENFORCEMENT

The requirements above should be checked through a combination of ex ante and ex post assessments, possibly adapted to the level of risk of the application.

- In the case of high-risk AI applications (such as applications linked to public security) on the European market, an ex-ante conformity assessment will be carried out in the EU to control the respect of European rules. Such an assessment would build on the conformity mechanisms that already exist for a number of products and services entering the European market and should take place independently from the origin of the developer. The testing procedure would include an independent audit and assessment by testing centres of the algorithm and the data. In order to limit the administrative burden on SMEs, some support structure might be envisaged.
    - When such a conformity assessment process is established, it should take into account the further development of the system through learning.
    - In case a conformity assessment is not feasible, e.g. no possibility to determine the way the data has been gathered or the algorithms tested, an obligation of training the algorithms in Europe on data in Europe will be envisaged.
    - AI systems originating from third countries must comply with the ex ante conformity assessment process or benefit from an adequacy decision confirming that the certification procedures in these countries ensure a comparable level of quality.

This ex ante assessment is without prejudice to ex post controls by national authorities where a presumed violation has been reported.

- For lower risk applications, requirements are checked ex-post after a presumed violation has taken place (or an abuse in case of a voluntary labelling system); deterrent fines should be introduced to ensure compliance.

H. Governance

A European governance structure is necessary to avoid fragmentation of responsibilities, increase competence in Member States on AI, and make sure that Europe equips itself progressively with the capacity needed for testing and certification of AI products and services.

A European governance structure should be also relying on a network of national authorities. The governance structure should have a variety of tasks including being a forum for a regular exchange of information, identifying emerging trends, advising on standardisation activity as well as on certification and also playing a key role in facilitating the implementation of the legal framework, such as through issuing guidance, opinions and expertise, following the model of the national data protection authorities and the European data Protection Board.

The EU level coordination should be non-bureaucratic and agile. A permanent committee of experts that would provide assistance to the Commission could be an option. The committee could draw expertise in different sectors and be organised around different work streams.

Given the emerging intersection between AI and existing regulatory frameworks such as in finance, pharmaceuticals, medical devices, consumer products, etc., the governance structure should not duplicate functions but establish close links with other EU and national competent bodies in the various sectors (pharmaceuticals, aviation, finance, consumer products, IP, food safety) to complement existing expertise and help existing bodies in monitoring and oversight of AI enabled products and services.

Member States should entrust the conformity assessment to dedicated bodies. Testing centres should enable the independent audit and assessment of algorithms, data and design processes. Independent assessment will increase trust and allow oversight authorities to fulfil their tasks. The EU enjoys the world's best testing and assessment centres and should maintain its primacy also in area of AI. In any case, the requirements for independent assessments must be proportionate to the risk that the AI systems pose.

The governance structure should guarantee maximum Stakeholders Participation. It should consult stakeholders who may directly or indirectly be affected by the system throughout its life cycle. It should regularly seek the feedback of Consumers and Workers organisations, NGOs and civil rights organisations

7. CONCLUSION

Artificial intelligence is a strategic technology that offers many benefits for citizens and the economy, provided it is human-centric, ethical and respects fundamental values. Artificial intelligence offers important efficiency and productivity gains that can strengthen the competitiveness of European industry and improve the wellbeing of citizens. It can also contribute to finding solutions to some of the most pressing societal challenges, including the fight against climate change, the challenges linked to sustainability and demographic changes, and the protection of our democracies.

For Europe to seize fully the opportunities that artificial intelligence offers, it must develop and reinforce the necessary industrial and technological capacities. As set out in the accompanying European strategy for data, this also requires measures that will enable the EU to become a global hub for data.

The European approach for artificial intelligence aims to promote Europe's innovation capacity in the area of artificial intelligence while supporting the development and uptake of ethical and trustworthy artificial intelligence across the EU economy. Artificial intelligence should work for people and be a force for good in society.

With this White Paper, the Commission launches a broad consultation of all relevant stakeholders, including Member States, civil society, industry and acamdemics of concrete proposals for a European approach to artificial intelligence. These include both policy means to boost investments in research and innovation, enhance the development of skills and support the uptake of artificial intelligence by SMEs as well as proposals for key elements of a future regulatory framework. This consultation will allow a comprehensive dialogue with all concerned parties that will inform the next steps of the Commission.

The Commission invites for comments on the proposals set out in the White Paper. They may be sent by XXX 2020, either by e-mail to: YYY or by post to: ZZZ.

It is standard practice for the Commission to publish submissions received in response to a public consultation. However, it is possible to request that submissions, or parts thereof, remain confidential. Should this be the case, please indicate clearly on the front page of your submission that it should not be made public and also send a non-confidential version of your submission to the Commission for publication.