ITI Promoting Innovation Worldwide

12 June 2020

ITI VIEWS ON THE EUROPEAN COMMISSION WHITE PAPER ON ARTIFICIAL INTELLIGENCE

EXECUTIVE SUMMARY

- Europe has an opportunity to take an international leadership role on Artificial Intelligence (AI) and other policy issues that are increasingly global, by building trust in the era of digital transformation while **preserving an enabling environment for innovation** to ensure its global competitiveness and security.
- Technology allows us to address the most pressing questions in today's society. Promoting these advances and **considering the potential economic and social harms of limiting the use of AI** is no less important than managing any potential harms arising from its application.
- **Context is key in identifying appropriate policies.** We support the EU's "human centric" approach, which promotes the ethical and responsible deployment of AI. However, not all AI applications are affecting fundamental rights. Also, many uses are already subject to sectoral regulation, sufficiently covering risks presented by AI applications.
- If regulatory shortcomings are identified following an assessment of applicable laws, adapting existing laws would be the appropriate approach. It additional legislation is absolutely necessary to fill an identified gap, policymakers should avoid regulating AI as such in observance of the technological neutrality principle, and **focus on governance in the use of technology, addressing potential issues arising in specific applications of AI in different sectors**.
- A clear, targeted scope focusing on those high-risk AI applications where issues are most likely to arise is key. The **categorisation of 'high-risk' and 'low-risk' applications** should consider use case, complexity of the AI system, probability of worst-case occurrence, irreversibility, scope of harm in worst case scenario and sector.
- Ideas for new ex-ante conformity assessments for '**high-risk' applications** should carefully consider the practicability, added value, and existing sectoral certification processes. A **combination of ex-ante risk self-assessment and ex-post enforcement for 'high-risk' applications** involving industry could be a viable solution.
- We support an **effective and balanced liability regime** that fosters trust in AI, provides a clear path for redress and adequately compensates victims for damages, while allowing for incremental improvements and innovations that come with placing AI systems on the market. Changes to the existing liability framework should be incremental and targeted at proven flaws.
- Requirements to **retrain AI on European data** raises several concerns, including preventing certain AI products from being made available in Europe. Rather than considering prescriptive requirements on the data, we suggest **focusing on the actual output of the AI systems**.
- As the AI ecosystem is global and the technology is not developed in regional siloes, **we welcome the Commission's commitment to international cooperation based on promotion of fundamental rights**. The most effective means of advancing Europe's AI agenda is to further the development and use of AI globally by cooperating with its international partners.

Global Headquarters
700 K Street NW, Suite 600
Washington, D.C. 20001, USA
+1 202-737-8888

Europe Office
Rue de la Loi 227
Brussels - 1040, Belgium
+32 (0)2-321-10-90

info@itic.org
www.itic.org
@iti_techtweets

1. INTRODUCTION

Europe has an opportunity to take an international leadership role on Artificial Intelligence (AI) and other policy issues that are increasingly global. We welcome the adoption on 19 February 2020 of the **White Paper on AI by the European Commission** and appreciate the opportunity to submit comments to the ongoing consultation.

The Information Technology Industry Council (ITI) is the premier advocate and thought leader for the global technology industry. ITI's membership comprises 70 of the leading technology and innovation companies from all corners of the ICT industry, including hardware, software, digital services, semiconductor, network equipment, cybersecurity, and Internet companies.

ITI and its members share the firm belief that building trust in the era of digital transformation is essential. We strongly believe it is important to **preserve an enabling environment for innovation** to ensure Europe's global competitiveness and security. Our industry acknowledges Europe's vision on creating a Trustworthy AI for Europe built around a human centric approach, and we want to be a constructive partner in realising this vision.

The benefits of AI are vast. AI-driven medical diagnostics can alert doctors to early warning signs to helping them treat patients more capably. Increasingly intelligent systems are capable of monitoring large volumes of financial transactions to identify fraud more efficiently. Small and medium-sized enterprises (SMEs) can gather new insights and improve their businesses by using AI and data analytics made available to them through online services. Therefore, it is crucial for Europe to **not only look at the potential harms of using AI, but also consider the potential economic and social harms of limiting the use of AI**, which may decrease its positive impact on our communities.

Technological innovations bring innumerable benefits to the European economy and society. We are already experiencing the benefits of AI in an array of fields. **Promoting these advances is no less important than managing any potential challenges.**

Moreover, the AI ecosystem is global and the technology is not developed in regional siloes. As such, the most effective means of advancing Europe's AI agenda is to maintain discussion beyond national borders. Many AI products and services used in Europe comprise both European and non-European elements developed in different locations made interoperable through the adoption of voluntary, international standards while complying with European rules and aligning with European values.

The significance of Europe's global partnerships, and the importance of shared values like trust, fairness, explainability, effectiveness, safety, and human oversight - the core principles that need to guide future policy action on AI – cannot be overstated. The EU therefore should work towards robust adoption of trustworthy AI technologies for the benefit of its citizens by ensuring its approach fosters the region's global competitiveness. This will in turn help Europe shape the global AI governance debate. The Commission's goal of achieving an 'ecosystem of excellence along the entire value chain' of AI can only be attained by working with governments and industry in the EU and beyond.

The following comments aim to address discrete aspects of the White Paper on AI, while also commenting on the economic and social implications of the technology. We additionally explore the role of the ICT industry, in a manner that supports innovation, while safeguarding the public and individual interests at stake.

2. A RISK-BASED APPROACH TO RULES AFFECTING AI

We appreciate the White Paper's approach suggesting that regulation should concentrate on how to minimise the various risks of potential harms that may emerge from high-risk AI applications (p. 10). We agree that risks need to be identified and mitigated and encourage policymakers to take a risk-based rather than overly precautionary approach to rules affecting AI. Given that the potential benefits of AI development are enormous, and that AI is a rapidly evolving technology, a legislative approach should be flexible enough to account for the rapidly changing and fast-paced technological advancement in this sector.

Technological innovations bring innumerable benefits to the European economy and society. Should the future European AI approach be too restrictive, there is also a risk of limiting the enablement of such technologies and miss opportunities for Europe and its citizens. We are already experiencing the benefits of AI in an array of fields – targeted services for small business customers to eliminate repetitive tasks and predict insights, fraud detection and prevention, or even making filing taxes easier. Start-ups, small and medium-sized enterprises (SMEs), and larger tech companies have all developed AI systems to help solve some of society's most pressing problems. Many others from across key European sectors are using AI to improve their business, provide better public services and advance ground-breaking research that saves lives. **Technology allows us to address the most pressing societal challenges in areas such as healthcare, public security, and disaster management**. Promoting these advances is no less important than managing the challenges.

- **Fundamental rights, including personal data and privacy protection and non-discrimination**

The White Paper raises concerns that AI development could lead to a variety of fundamental rights concerns. The tech industry is aware of and addressing the main challenges. For instance, the technology industry recognizes the need to mitigate bias, inequity, and other potential harms in automated decision-making systems. We share the goal of **responsible AI adoption and development**. As technology evolves, we take seriously our responsibility as enablers of a world with AI, including seeking solutions to address potential risks.

Context is key in identifying appropriate policies. We support the EU's "human centric" approach, which promotes the ethical and responsible deployment of AI. Our industry is committed to partnering with relevant stakeholders to develop a reasonable, effective, and balanced accountability framework that takes into account the different actors and phases of developing and deploying AI systems. As leaders in the AI field, ITI members recognize their important role in making sure technology is built and applied for the benefit of everyone. **Approaches must be context- and risk-specific** and should consider that not all AI applications are affecting individuals' fundamental rights. Such AI applications would not require an all-encompassing fundamental rights-based approach and often no additional regulatory intervention.

In fact, some basic AI uses have little or no impact on individuals' rights, such as in the context of industrial automation and the use of analytics to streamline automobile manufacturing or to improve baggage handling and tracking at busy European airports. AI development should not be disrupted with new stringent obligations that could significantly slow the adoption of AI and hamper innovation.

Where fundamental rights are affected, AI applications used in specific sectors (e.g. healthcare, financial services, transportation) are already subject to **sectoral regulation that is often already geared towards addressing risks to fundamental rights of individuals (e.g. Medical Device Regulation (EU) 2017/745, Payment Services Directive (EU) 2015/2366).** While it is important to assess if existing, domain-specific EU regulations are exhaustive, it is also important to further underline that they already cover many of the most common concerns, and therefore any future regulatory activities should be limited to address discrete and specific issues not covered by existing rules.

Further, AI applications can enhance access to products and services to currently underserved or disadvantaged groups. In the financial services context for example, AI can help increase access to capital for small businesses who would otherwise be unable to get a loan. This is because an algorithm can assess non-traditional factors to help candidates who might otherwise struggle to get a loan because of their background or their lack of a strong credit history. Also, when hiring or granting loans to individuals, AI-assisted decision-making with human oversight can help address unconscious bias.

Access to rich, robust data is important in developing accurate, robust AI systems, so it is essential to ensure availability of large and diverse high-quality datasets and increased computing power and flexible platforms for storing and managing such data (e.g., cloud, multi-cloud etc.). In this context, **the GDPR already provides important safeguards on the use of personal data**. Deploying techniques such as anonymization, pseudonymization, de-identification and other privacy enhancing techniques (PETs) are crucial to ensure data can be used to train algorithms and perform AI tasks without breaching privacy.

3. POSSIBLE ADJUSTMENTS TO EXISTING RULES

The White Paper provides a comprehensive overview of existing laws applicable to AI and identifies several areas where it sees a need to amend them, create new ones or use other legislative avenues. As mentioned above, context is key in identifying appropriate policies. Many uses – e.g. in medicine, financial services, or transport – are already subject to sectoral regulation. In many instances, the essential requirements already contained in harmonized legislation may be sufficient in covering risks presented by applications of AI.

A proper assessment of applicable laws should therefore precede new legislation, with a view toward evaluating whether new rules are actually needed, avoiding conflicts of law, and ensuring that both existing and forthcoming regulatory requirements prioritize international compatibility and reliance on global, industry-driven, voluntary, consensus-based standards. In cases where regulatory shortcomings are identified, adapting existing laws would be the appropriate way forward. It is determined that additional legislation is necessary to fill an identified gap, policymakers should avoid regulating AI as such, as this would run counter to the principle of technological neutrality and regulation would likely become obsolete and possibly disproportionate as the technology and use cases evolve. Instead, governments should work with industry and other AI stakeholders to focus on governance in the use of technology in order to address the potential issues arising in specific uses and applications of AI in different sectors.

Devising rules based on specific applications of the technology, effects, and governance approaches, rather than on regulating the underlying technology in abstract, would provide for more precise outcomes.

- **Scope of a possible EU regulatory framework for AI**

A clear, targeted scope focusing on those high-risk AI applications where issues are most likely to arise will be critical to the effectiveness of any future regulatory framework, and avoidance of over-regulation. Since there is no single agreed-upon definition of AI, it will be important for policymakers to provide greater clarity if they plan to seek specific rules for AI functions. An essential factor is to properly identify AI and its different categories, including the component parts of AI systems beyond algorithms, as well as to define related key terms such as machine learning. Some algorithms have been applied for decades but do not constitute "Artificial Intelligence" or "machine learning" systems. The first task is to determine what is AI and what is not. There is a difference between the latest wave of AI systems that learn from data and experience, and traditional software and control systems that operate according to predictable rules, which have long been embedded in a wide variety of high-risk systems from flight control to pacemakers to industrial settings. We noted that the risks associated with traditional software and control systems that make probability predictions are already adequately addressed by existing regulation.

- **A differentiated, risk-based approach: high-risk and low-risk applications**

We appreciate the White Paper's intention to lay out a differentiated approach based on risk, with a distinction between high- and low-risk AI applications, based on a number of criteria including the intended use. However, the use of "sectors" may lead to a too broad categorisation – it is important to use a sufficiently targeted and well-outlined classification to ensure this criterion does not become irrelevant. We encourage developing a categorization that takes into account sector, use case, complexity of the AI system, probability of worst case occurrence, irreversibility and scope of harm in worst case scenarios e.g. individual v. larger groups of people, and other criteria.

More specifically, we urge policymakers to consider the following specifications for high-risk AI applications in order to build on the White Paper's differentiation, and ensure for the development of principles-based rules:

- **Specify what constitutes high-risk AI applications based on probability and irreversibility:** To ensure proportionality, the definition should be augmented to better reflect well-established interpretations of risk as a function of severity and likelihood. For example, high-risk could be defined as AI systems that either (a) may cause catastrophic irreversible harm and there is a possibility that such harm may occasionally occur, and (b) may cause serious harm and such harm is probable. More clearly reflecting a nuanced understanding of high-risk within the framework would make clear that the objective of the framework is to mitigate harm for (a) and reduce the likelihood for (b).

- **Acknowledge and define AI's opportunity costs:** In several instances, using automated systems can greatly reduce risk. In air traffic control, using an automated tool paired with human oversight is an example of how AI can reduce risk as opposed to a situation in which air traffic controllers could make mistakes due to fatigue or distraction - factors that do not affect a machine. Analysis about the spread of pandemics is another example where limiting the use of AI is likely to lead to potentially bigger risks than possible negative consequences of the AI system being deployed for this purpose.

- **Remove "exceptional instances" clause:** We support the notion of clear and predictable cumulative criteria, as well as clarity over what constitutes a high-risk use of AI, and ideally the negative impact or concrete consequences that can reasonably be expected on affected parties. However, the "exceptional instances" clause is too open-ended and should be removed to avoid legal uncertainty. For example, the notion that applications affecting consumer rights could potentially fall in the high-risk category seem overbroad, unjustified and against the objective of focusing only on well-defined areas of risk. In addition, the instances to which the White Paper appears to refer seems to be appropriately covered by existing legislation (non-discrimination provisions in labour law and consumer laws).

- **Align references to damages with the PLD:** The 1985 Product Liability Directive (PLD) empowers European consumers to receive compensation for damage caused by defective products. The PLD applies to any product sold in the European Economic Area with a 3-year limit for the recovery of damages. The PLD defines damage as death, personal injury, or damage to the product in questions or other products of a consumer. This definition could be a good basis to support the definition of what constitutes high-risk AI. We would advise avoiding references to immaterial damages that could potentially lead to waves of compensation claims for producers on illegitimate grounds and lead to a backlog in assessing cases all while mostly being covered already by existing legislation in the fields of data protection, non-discrimination and freedom of expression.

- **Ensuring an effective and balanced liability regime**

AI presents great opportunities for society in different fields yet raises valid concerns around responsible and safe deployment. The clarification of rules around liability, currently designed for physical products, is an appropriate area of focus. There are also important considerations about finding the appropriate balance of ex-ante, preventive rules, and ex-post remedies. We support an effective and balanced liability regime that fosters trust in the use of AI, provides a clear path for redress and adequately compensates victims for damages, while allowing for incremental improvements and innovations that come with placing AI systems on the market. In many cases the existing liability framework will be easily applied in an AI context and we suggest that the EU maintain a strong presumption against altering it except in response to significant and demonstrable shortcomings. Should a need for future action be identified in areas that involve increased risks for end-users of AI applications, it should be addressed in a sector-specific manner, with new regulation or suggested legislation filling clearly identified gaps and designed to avoid overreach. Sector-specific safe harbour frameworks or liability caps are also worth considering in domains where there is a concern that liability laws may otherwise discourage socially beneficial innovation. Updating such sector-specific regulation, rather than adopting sweeping changes to general product liability frameworks, would allow for more precise targeting of remedies for identified gaps in liability coverage.

If the existing liability regime falls short of addressing new challenges arising from specific applications, legislative intervention should be limited to filling in the gaps while avoiding an overhaul of the existing framework, which has proven to provide an adequate balance in protecting consumers while encouraging the launch of innovative products in the market.

It is also important to acknowledge that the liability rules for digital products might be challenging to apply to AI. Digital products are developed through a trial and error process aimed at constantly

improving products and services, including their safety and security, even after they are made available to the public. However, the nature of the process by which security updates are delivered may introduce a new dynamic into the liability framework. For instance, if a vulnerability or a harmful exploit is detected in a product or service in the market, even though developers send out patches to mitigate such risks, in some instances users can choose to delay installing or in some instances not install patches at all, raising questions around responsibilities between producer and user.

While the Commission is consulting on the opportunity to expand the PLD's strict liability framework to damages stemming from high-risk AI, it seems reasonable that if a product containing AI technologies causes harm to a business or individual, and the business/individual can fulfil the requirements of the PLD, they should be able to recover damages. In this sense, amendments to the PLD to cover embedded AI are unnecessary, since the directive is technology-neutral and strikes the right balance between the obligations of consumers and producers, thereby creating legal certainty. Still, it is crucial to recognise that there is a fundamental difference between on the one hand a hypothetical, undefined risk, and on the other the danger based on the product's fault or its use in a specific context. Strict liability frameworks like the one set up by the Product Liability Directive (PLD) remove any consideration of intent or negligence. Manufacturers should instead be equipped with a right to cure or correct identified violations with consumers directly. This would also be in the interest of fostering consumer trust in AI applications.

An expansion of the scope of the PLD to introduce strict liability for all AI-based technologies beyond those embedded in a product would disproportionately spread liability throughout the AI development and supply chain, exposing to liability also actors that could not reasonably be expected to bear responsibility for situations beyond their control. This becomes especially critical if AI technology is purposefully misused, e.g. by bad actors for illegitimate surveillance, or for consciously discriminating in hiring processes etc. While AI systems can be well developed, bad actors might abuse these tools for their own agendas, and if liability rules are not carefully designed, they could unjustly expose developers, even when the cause of harm was not the AI system, but its misuse.

In conclusion, **expanding this type of legal exposure to AI system developers and others playing an intermediate role in the value chain would be disproportionate to the goals the Commission is seeking to achieve. It would also have a significant chilling effect on innovation and competition**, and most likely negatively affect European SMEs. If Europe were to become the first global player to apply strict liability to services and software, the roll-out and uptake of AI-based technologies would also be hindered, hitting hard businesses and start-ups operating in Europe and opposing the ambitions to create an ecosystem of excellence as expressed in the White Paper.

- **Software and AI (-services) as products**

Any clarifications in the definition of a product under the PLD would require great care and precision, limiting the types of software that would be included. Software standing alone typically does not pose the same type of heightened risks associated with traditional physical products. When software is integrated into a product rather than a service, there may be a higher risk potential for physical damage to persons or property, thus providing a potential rationale for extension of the PLD. There are already special instances where software is treated as a quasi-product under EU law - for example under the Medical Devices Regulation. This precedent could provide a sensible middle ground between special regulation and the PLD, clarifying where special regulation should treat software as a product, and only including those applications under a framework of strict liability.

- **Cyber vulnerabilities as a defect**

Cyber vulnerabilities should not per se be classified as defects in AI systems, as these are dynamic risks that can in most instances be mitigated through responsible system configuration to enable remote updates and responsible cyber hygiene practices by consumers. In particular, discovered vulnerabilities in software products can be patched after they have been placed on the market via patches developed in a timely manner by the manufacturer. However, software producers do not fully control in all instances whether updates are installed – oftentimes, it falls to the user to install or accept these updates and in such cases vulnerabilities can either go unnoticed or are not fixed, with users maintaining some level of responsibility for mitigation. The imperative of user responsibility also underscores a particular challenge in the use of existing product testing and certification regimes, which are largely geared toward the assessment of static product safety risks, to fully assess dynamic risks such as cyber vulnerabilities. We recommend that AI systems designers prioritise configuring AI systems so as to enable automated remote updates to mitigate discovered vulnerabilities, and also that we prioritise educating consumers of AI systems regarding responsible cyber hygiene practices so they are aware of the importance of updating AI systems in those instances where automated remote updates are unavailable.

4. TYPES OF REQUIREMENTS

The White Paper outlines suggested mandatory legal requirements, several of which we believe would hinder the development of beneficial AI applications.

Standards

Standardisation can help form a bridge between AI regulations and practical implementation. The EU should **support global, voluntary, industry-led standardisation,** and safeguard the work and processes of international standards development bodies. Global AI standards can help establish consensus around technical aspects, management, and governance of the technology, as well as frame concepts and recommended practices to establish trustworthiness of AI inclusive of privacy, cybersecurity, safety, reliability, and interoperability. Standards must not establish market access barriers or preferential treatment; rather, they should work for the benefit of the international community and be applicable without prejudice to cultural norms and without imposing the culture of any one nation in evaluating the outcomes/use of AI.

We agree with the White Paper's desire to limit and appropriately scope the creation of mandatory standards. We anticipate that standards will support the regulatory framework. However, we would emphasize that standards are another key area for industry leadership to ensure that appropriate technical specifications are generated on the basis of market demands.

Standardisation is an opportunity to reach consensus on those aspects of a technology that need to work seamlessly with other products and services and function across markets. Standards enable consumer use of technology around the globe. This occurs through robust engagement by both public and private sectors contributing and competing to determine the most appropriate standards for the current technology and markets.

Policymakers may reference standards as the basis for technical regulations, where appropriate. However, policies and regulations that require or preference implementation of specific standards may preclude regulators, companies, and customers alike from relying on the most fit-to-purpose means of demonstrating compliance with a specific requirement or set of requirements. For instance, a government-prescribed standard intended to increase security may unintentionally prohibit use of the most appropriate standard and in doing so, negatively affect the security of the product or service. Considering the rapid pace of innovation and change, it may not be feasible for government to keep pace in choosing what it perceives to be the most appropriate standard(s) for compliance. This work is best left to the technical experts, who can choose from a variety of standards and determine which ones to implement to achieve compliance. The Commission should use its approach to AI as a way to incorporate greater flexibility into its approach to standardisation, thus recognising the value of standardisation for new technology and enabling companies to determine the most appropriate international standards, or other implementations, to use in complying with mandatory regulations. A greater degree of flexibility with respect to the standards that may be used to demonstrate compliance with relevant AI-specific requirements would yield positive outcomes for both domestic and global innovation, regulatory protections, and market openness.

Training data

We fully acknowledge the importance of training data sets in the development and deployment of AI. More important is though how the data is used and the outcomes that it leads to. Therefore, requiring that data sets be unbiased, "sufficiently representative" or "sufficiently broad and cover all relevant scenarios needed to avoid dangerous situations" (p.19) is an unfit approach because it may not be possible to achieve unbiased datasets in reality.

Rather than focusing on the data sets themselves, which often will reflect biases that exist in the real world, we suggest focusing on testing outcomes of the AI systems before deployment or applying safeguards against biased outcomes after deployment. Stereotypes can get perpetuated either in recommendations, searches, or quality of tool so considered quality-control and review processes should be in place and outputs tested to protect against this. This requires testing and human involvement in the development of AI with diverse teams that are continually evaluating in the development and innovation of AI.

Also, many AI systems are developed for a global audience, and retraining these systems with only European data would make them uneconomical and would delay – or in a worst case scenario prevent – certain AI products from being made available to European consumers.

–At times, strictly controlling training data sets could be at odds with compliance with other laws including the GDPR. Some AI models may require less strict requirements to data sets if they are designed with the appropriate caveats and caution – requirements should be set in connection to the purpose and what is required to ensure non-discrimination in relation to that purpose. GDPR stipulates for example that developers should not have access to attribute to ethnicity, unless explicit consent is given by data subjects. We see an urgent need for the Commission to specify how training data requirements will interact with GDPR, including the right to be forgotten and data minimisation. We also urge the concept of 'sufficiently representative' to be defined more clearly.

Keeping of records and data

We caution against the introduction of mandatory record-keeping requirements for datasets used to test, train, or operate AI systems on an ongoing basis. If keeping of data could lead to revealing details of AI systems or underlying code, this could risk undermining privacy, copyright, and trade secrets, infringe on IP rights, and heighten cybersecurity risks, privacy, and data manipulation risks. Instead, we urge the Commission to assume an approach that looks at **outcomes** rather than process.

Keeping vast amounts of data would be unworkable for many companies given how AI is developed in a constantly iterative way. For example, there may be numerous data sets used to train AI systems which could not be recreated, and AI systems may be ingesting continuous flows of historic or real-time data over time. Other specific AI system learning techniques are built to protect privacy (federated learning) and disclosure obligations could undermine this crucial goal.

Lastly, it is important to note that there are no common data naming conventions, no formatting standards or concurrent versioning systems used for data; these factors would further complicate mandatory sharing requirements in this field.

Information provision, transparency, understandability & interpretability

We would encourage a harmonised and clear definition of and a more in-depth discussion around **transparency** throughout the White Paper. Transparency does not automatically equate to better control or decisions of AI systems. For example, the driver of a car does not need to fully understand the systems in a vehicle to be able to drive the vehicle safely. Similarly, users of AI would in most cases not need to have a detailed understanding of the workings of the technology to use it responsibly.

Transparency, in our view, is best achieved through ensuring understandability and interpretability. **Understandability** should allow users of AI to understand broadly how an AI application works and how their data is being used to create a better user experience for them individually. Rather than introducing obligations to disclose technical features, we recommend an approach in which understandability is prioritized to build consumer trust. **Interpretability** on the other hand is geared towards allowing technical experts to understand the rationale behind an AI's decision/outcome. Both aspects are important, and we encourage policymakers to think of transparency in these terms to make explicit the objective of any potential transparency requirements.

We further urge that there should be a differentiation for transparency requirements for AI in high-risk applications being used in consumer-facing v. B2B products and services. For **B2B scenarios**, we do not see reason to share such information unless the information in question is deemed to be critical for public interests including safety. This is because excessive sharing obligations might risk IP rights and contractual arrangements between business partners. Further, an organization that develops AI cannot proactively monitor the way its customers are using AI.

As a general principle, if AI is playing a substantive role within a high-risk AI application, that fact should be easily discoverable along with some insight into the nature of the role AI is playing by those who have a legitimate interest.

Public disclosure will typically be appropriate for applications designed for or affecting consumers (e.g. government services or healthcare). However, public information about B2B use of AI should not be required except in case of clear public interest.

Lastly, the development stage of a machine learning system in its lifecycle, the context in which it is or will be deployed, the deployment purpose, as well as other factors should be factored into **evaluation of decisions** made by a system.

Robustness and accuracy

A safety-by-design approach should be implemented for all high-risk AI applications. Internal documentation and monitoring will be key for companies developing high-risk AI applications to assure their customers of the product's quality and security. Mandating specific reporting, documentation or additional techniques would be premature and could hinder industry from finding best-suited solutions to challenging, complex processes. A rigid system could in fact risk longer term safety of products and accuracy of AI systems if innovation is not incentivized.

Proposing requirements *"ensuring that outcomes are reproducible*" (p. 20) is problematic because often it is not possible to achieve this. AI systems change over time and outcomes are not reliably reproducible, therefore compliance with requirements around reproducibility would be virtually impossible for many AI applications. Reproducibility of outcomes may require exactly reproducing the entire dynamic environment and the entirety of the data flows used to train the model. In practice this could lead to AI systems being only able to be built on very basic techniques, such as simple decision trees, as reproducibility of more complex systems would not be possible in practice.

Human oversight

We fully agree with the Commission and the HLEG that human oversight is crucial to reap the full benefits of AI while controlling for potential risks. Humans are in fact at the core of developing AI, beginning with the concept phase, the development phase, the training stage, the product roll-out and monitoring and quality control.

As mentioned earlier in this paper and as also noted by the Commission itself, we need to be mindful of different AI application areas and to what extent humans need to be involved throughout the lifecycle of an AI application. For example, it is useful to have a human monitor an automated decision in an air traffic control tower and override decisions made by the AI if necessary (for example in an emergency). In such a case, the AI de facto replaces the human and therefore, human oversight is needed continuously. However, for other, less critical situations, we may not require detailed human involvement e.g. for handling baggage at an airport.

The overarching point is that individual use cases should determine the degree of involvement of humans in reviewing machine-generated decisions. In some cases, human oversight can not only lead to delays, in others, accuracy of outputs could even be undermined by human interventions (for example for mathematical calculations).

Specific requirements for remote biometric identification

Our industry takes this issue seriously and recognises our important role in making sure AI technologies, like facial recognition technology, are built and applied in a way that benefits everyone. It is critical that society, governments, and the technology sector work together to begin to solve some of the most complex issues, including this one. New regulations and policies should be

compatible with existing rules like GDPR to protect users without causing harm or unintended consequences.

5. Compliance and enforcement

Ideas for new ex-ante conformity assessments that include independent audit and testing by public authorities to ensure that high-risk AI applications adhere to EU rules should carefully consider the practicability and added value of such an approach, taking into account existing sectoral certification processes.

While we appreciate the need for strong assurances, it is not clear that the existing conformity assessment infrastructure could effectively and efficiently carry out prescribed testing on what are often among the most socially valuable applications of AI. This is particularly the case if such evaluations could only be undertaken by Notified Bodies, given the lack of expertise needed to evaluate datasets or algorithms in sufficient depth as well as the volume of requests would create significant practical and capacity challenges, – especially if "repeated assessments over the lifetime of AI systems" as contemplated in the White Paper are required. Moreover, the legal requirement that Notified Bodies be established under EU Member State law suggests that only local testing and certification bodies could carry out necessary assessments. The extension of this requirement to the assessment of AI applications would exacerbate what are already likely to be significant capacity constraints, generate backlogs and market access barriers for non-EU firms, and run counter to the EU's long-standing international position against localisation requirements for testing bodies. Finally, the conformity assessment scheme would require developers to provide an independent assessment body access to the underlying data used to train a model, including algorithms, source code, or other proprietary information, this could create an untenable situation that may run afoul of intellectual property laws in a variety of contexts, leading to conflicts of laws, potentially contravene existing EU international trade commitments, and damage companies commercial interests.

A combination of ex-ante risk self-assessment and ex-post enforcement for high risk AI applications would likely achieve intended results within much faster timeframes and without hampering innovation or creating unnecessary burdens. For instance, requiring organisations to carry out and document risk assessments would be analogous to the data protection impact assessment under GDPR. Such an approach would also build on existing industry practices, including the ethical, legal, and due diligence practices that guide the responsible and trustworthy development of AI. To clarify compliance and facilitate accountability, regulators should assess what actors are best to act at what stage in the AI lifecycle. For example, the developer of AI is responsible for conceptualizing and training the AI, whereas deployers have the best visibility of the use case for the AI.

Should the Commission pursue an ex-ante regulatory approach involving third-party conformity assessment, in addition to broad concerns noted above, we would note the following issues with the framework as currently proposed:

- **Products already on the market:** If it were deemed necessary for existing products in market to retroactively undergo conformity assessments, it would exacerbate what is already likely to be a significant backlog for testing bodies. A grandfathering clause, as used in other sectors for products already in the market, would solve this at the outset.

- **R&D and early stage products:** In the early stages of development there will often not be a clear view as to the ultimate shape of a product. It is therefore important that confidential piloting of an AI application be allowed prior to any conformity assessment, within the bounds set by existing sectoral regulation. If such testing were not permitted, it may result in organisations being deploying extra caution in the form of expensive extra measures, which could lead to delays in developing products and hinder innovation.

- **Products being altered/updated significantly:** For certain particularly high-risk AI applications going through significant changes during their lifetime via software updates after they have been put onto the market, there is a reasonable need for repeat assessment procedures. In these cases, we encourage the Commission to develop clear guidance regarding such a process. The role of users in these reassessments would also need to be considered as they are often required to individually perform software updates on their devices.

- **Requirement to retrain on European datasets:** The White Paper raises the possibility of requiring AI systems to be retrained using European data or in Europe, if developers are unable to prove that the original dataset used met European requirements. This raises significant concerns:

    - Requiring European iterations of AI systems to be retrained based on European data sets would in many instances not be possible and lead to certain products not being made available in the European Single Market. This is because training datasets can include a mix of own, third party and open source data. Requiring retraining could inadvertently lead to smaller datasets being used and AI products hence not being as advanced as if they were trained on larger datasets.

    - We therefore encourage a global focus to ensure fair and diverse user experience, without regional training dataset restrictions in order to make good on the global transformative benefits that AI promises consumers and industries.

    - In cases where an AI application is found to be faulty regarding certain European standards or existing laws, the developer should be entitled to fix the mistake in the way best fit to address the detected fault. Retraining AI applications on new datasets is often the last step in several other steps that can mitigate problems beforehand.

- **Voluntary labelling for non-high-risk applications**

While we generally agree that voluntary labelling can promote consumer trust, given that AI systems themselves do not constitute a physical product or service, we have questions about how such labelling could be applied in a manner that achieves the objective of facilitating information to consumers. Broadly speaking, any voluntary labelling approach should not become a de-facto market entry requirement for AI products and services in Europe. Further, technical challenges such as a lack of allowance for electronic labelling (e-labelling), would need to be broadly addressed for products and services that do not have a physical shape on which to affix a label; therefore, flexibility should be considered in those scenarios. As a general matter, we strongly encourage the Commission to adopt international best practices for e-labelling in allowing the display of regulatory and other required information via electronic means. Additionally, voluntary labels should be based on international standards and recognized by all EU Member States, as described in the next section.

6. Governance and conclusion

Assessing the need for upgrading the regulatory framework to enable AI to fulfil its potential in Europe is crucial to identify what legislative gaps exist and the extent to and manner in which any such gaps should be filled. We value the evaluation of sector-specific legislation that is being carried out by the European Commission. Several ITI members have also engaged in the European Commission's High-Level Expert Group (HLEG) on AI and helped create the ensuing ethics guidelines and policy recommendations; several our members have also partaken in the AI piloting phase. We agree with the Commission's view that, in a future regulatory framework, each obligation should be addressed to the actors who are best placed to address potential risks. We support the White Paper's suggestion to further involve stakeholders from industry in an open and inclusive way in the crafting of the European AI approach, including any regulation.

As the AI ecosystem is global and the technology is not developed in regional siloes, the most effective means of advancing Europe's AI agenda is to expand the discussion beyond national borders. We recommend the EU engages beyond the borders of the single market, to **further the development and use of AI *globally* by cooperating with its international partners**, and welcome the commitment to international cooperation on AI based on promotion of respect of fundamental rights, non-discrimination and protection of privacy (p.9).

The EU should work towards trustworthy AI for its citizens by ensuring its approach fosters the region's global competitiveness, in turn helping Europe shape global AI governance. We reiterate the need for global, voluntary, industry-driven standards development to support the deployment and uptake of AI in Europe and beyond.

This also means recognising the significance of **Europe's mutual interdependence** with like-minded democratic countries, and the importance of shared common values like trust, fairness, explainability, effectiveness, safety, and human oversight - the core principles that need to guide future policy action on AI. There is a valuable opportunity in working together to shape balanced solutions in situations where the application of some of these values conflicts in practice – for example, when explainability (through simpler algorithms) can conflict with accuracy, or human intervention reduces quality results (e.g. in misreading medical scans).

* * *