ITI Promoting Innovation Worldwide

06/08/2021

ITI's Views on the European Commission's Artificial Intelligence Act Proposal

Summary of Key Recommendations

- The Artificial Intelligence (AI) Act should provide **a targeted, flexible, and future-proof framework** to mitigate potential risks associated with some specific AI applications, while stimulating innovation in the field and encouraging the uptake of AI technologies.
- **The definition of Artificial Intelligence should be more targeted** to differentiate between traditional software that operates according to predictable and relatively static rules and AI systems capable of self-learning or self-adjusting, and exclude the former.
- **The definition of "provider" should be adjusted accordingly**, to avoid targeting actors in the AI value chain such as entities developing toolkits, or software libraries.
- There is a mismatch between the very targeted criteria for the **identification of high-risk AI applications** in article 7(2) and the variety of applications identified in annexes II and III. **The broad language of the annexes should be narrowed down to exclude non-problematic uses of AI** that may potentially be caught in scope.
- The Act should enshrine strict and meaningful **safeguards to allow for responsible deployment of real-time remote biometric identification** for national security or law enforcement purposes.
- **Providers of general-purpose software** (i.e., software that can be adapted for a variety of tasks) **should be explicitly excluded from the scope** when they are not the ones who directly develop or deploy the system as a high-risk AI application. Compliance with the requirements should be with the actors who are best placed to implement them.
- Requirements on data governance, recordkeeping, transparency and human oversight should be proportionate, realistic and reflect the diversity of applications in scope. **Requirements should be goal-oriented, rather than prescriptive,** so that companies can apply the most meaningful and appropriate processes to ensure compliance, without affecting the objectives of the regulation.
- The proposal should better **promote reliance on international industry-driven** consensus-based **standards** to avoid fragmentation with global regulatory environments, and avoid the development of region-specific technical specifications, that would harm the global competitiveness of companies developing AI in the EU.
- **Greater clarity and flexibility for the acceptance of international testing outcomes is needed in the proposal.** Considering that Conformity Assessment for AI is a nascent field, for which there is not a commonly understood practice or infrastructure available, this would help avoiding backlogs in testing procedures.
- Requirements for market surveillance investigations should be proportionate and should **not require companies to disclose sensitive information such as source code**. Authorities should be equipped with broad expertise to enforce the Regulation holistically
- The provisions on the work of the AI Board should ensure **structured and regular stakeholder involvement.**
- The AI Act should be leveraged to facilitate a **global conversation around AI governance with like-minded global partners**, to find alignment on key elements related to AI governance and promote open, non-discriminatory and principle-based cooperation and trade in the field of AI.

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

Introduction

The wide variety of possible uses of Artificial Intelligence (AI) represent a huge opportunity for the global economy and for societies around the world. Encouraging innovation and uptake of these technologies is rightly a fundamental public policy goal. At the same time, it is important to ensure a responsible rollout of the technology in a trustworthy manner.

The European Commission's AI Act is the world's first proposal for a horizontal regulatory framework on AI. As such, it is paramount that it provides a targeted, flexible, and future-proof framework. This framework should mitigate the potential risks associated with some specific AI applications, while at the same time stimulating innovation in the field and encouraging the uptake of AI technologies. These outcomes can be achieved by targeting regulation only to certain high-risk AI applications, and ensuring that requirements for providers are clear, proportionate and goal oriented.

In parallel, it is important that the AI Act takes into account the global dimension of the technology. AI is not developed in regional siloes, but rather in complex global supply chains. As such, global cooperation and innovative mechanisms to facilitate regulatory compatibility and open trade with like-minded global partners should be foundational elements of a successful regulatory framework. These elements are also essential to sustaining the availability of key technologies in the EU and ensuring Europe's global competitiveness. As a first mover and leader in the AI governance space, we strongly encourage the EU to adopt a global approach – including on issues such as standardisation and conformity assessment - that will enable convergence when other jurisdictions follow suit, in a manner that promotes innovation and avoids unnecessary and harmful discrepancies.

The Information Technology Industry Council (ITI) is the premier global advocate for technology, representing the world's most innovative companies from technology, hardware, software, services, and related industries. We welcome the Commission proposal of the AI Act and appreciate the opportunity to provide our detailed comments on the proposal in the following pages.

Coherence with Existing, Proposed and Expected Legislation

With the publication of the AI Act, the EU is a first mover in the field of AI regulation, introducing many new concepts and requirements for actors in the AI ecosystem. Given the ambitious legislative agenda of the European Commission on digital technologies, for instance with regard to the planned update of the Liability regime for Artificial Intelligence, or the ongoing review of the General Product Safety Directive (GPSD), it is fundamental to maintain an encompassing perspective on all these issues and ensure that concepts, definitions and requirements are coherent and non-contradictory across the various initiatives.

At the same time, policymakers should take into account other ongoing existing sectoral and legislative files which may have an impact on AI technologies, and ensure that they do not overlap and remain consistent. These may include the Data Governance Act (DGA), the Digital Services Act (DSA), the upcoming Data Act, the draft Regulation on General Product Safety

and the revision of Liability Rules and existing requirements by sectoral supervisory authorities. Establishing a structured dialogue to ensure cooperation between policymakers in the European Parliament and in the Council of the EU involved in these different files may be a good way to make sure that the EU's legislation for digital technologies develop into a coherent and easily applicable body of rules.

Scope and Definitions

Definition of Artificial Intelligence

The proposed definition of AI, which refers to the work done in the OECD, can be very broad and could theoretically capture many different types of systems and processes. Carefully articulating the scope of Artificial Intelligence implicated by the Regulation is therefore essential to establishing a well-targeted regulatory framework. For instance, references to logic- and knowledge-based approaches, as well as the reference to statistical approaches, would include a wide variety of computer-based systems in the scope of the Regulation, that are not generally considered consistent with classic definitions or risks of AI. In addition, reference to 'content' ('software that...can...generate outputs such as content') is also extremely broad and could cover almost any output. Similarly, the term 'search and optimization' could apply to almost all data management, organisation or optimization methods that have nothing to do with AI as is intended in the Regulation.

There are thus many kinds of algorithmic systems that would fall under the AI Act's broad definition but do not present novel risks that are not already covered by an existing legislative framework. There is a difference between the latest wave of AI systems that learn and adjust their outputs over time based upon new and repeated data inputs, and traditional software and control systems that operate according to predictable and relatively static rules. The latter have long been embedded in a wide variety of high-risk systems from flight control to pacemakers to industrial settings and are already appropriately regulated. The current proposal could be interpreted to extend beyond 'AI', to cover any decision support system that is driven by algorithms. While continuously building on the OECD's proposed definition of AI, this regulation should have a more targeted scope by clarifying this definition and excluding traditional software and control systems. This would be beneficial for legal certainty and to reduce overlap with existing law. We recommend that the text differentiate between AI and machine learning/statistical modeling. One way to do so would be to view AI as a system capable of self-learning or self-adjusting, as opposed to systems whose outputs are determined through an initial process of data modeling but do not adjust over time.

Definitions

Recital 60 recognises the complexity of the AI value chain, made of "relevant third parties, notably the ones involved in the sale and the supply of software, software tools and components, pre-trained models and data, or providers of network services". The broad definition of "AI system" and subsequently of a "provider", makes it hard to effectively determine which AI systems and which entities (providers) would be in scope of the proposed Regulation. AI in particular relies on a vibrant community of open-source contributors who have helped build frameworks like Tensor Flow and PyTorch, which have been the foundation for many new AI technologies. It is important that the AI Act recognise the important role

open-source communities play in the development of AI systems and ensure that helping build these foundational tools (that can be used to make all kinds of AI systems) would not qualify developers by itself as "providers" for the purposes of this regulation. For this reason, the AI act should clarify the different roles across the AI value chain, so that entities developing toolkits, SW libraries, etc. are not considered "providers". For example, by stating *that these relevant third parties are not considered "providers of AI systems"* as defined by the proposed Regulation.

The draft Regulation utilises the concept of "safety component" in the determination of the level of risk of an AI system. The proposed definition of what constitutes a "safety component" would be open to interpretation and remain a source of uncertainty for the qualification of high-risk AI systems. To reduce this ambiguity, the assessment of a "safety component" should refer back to Union harmonised legislation to align with any relevant essential requirements. In other words, when assessing an AI system for the purposes of article 6(1), a safety component should be understood in the meaning of the relevant Union harmonisation legislation listed in Annex II.

Furthermore, the proposed AI Act extends the requirements of high-risk AI systems to certain systems that are already subject to conformity assessments under specified EU regulations. One of these regulations is for wireless radio equipment. This provision should be clarified to ensure innovation in devices which may use both wireless radio equipment (e.g., Wi-Fi and Bluetooth) and AI, but the AI is not related to the radio equipment. The AI Act should more clearly distinguish between AI systems that are directly related to, or are integral parts of, the radio equipment subject to regulation, as opposed to AI systems that are unrelated to the radio equipment but exist on the same product. Only AI that impacts safety of the radio product as defined in the Radio Equipment Directive (RED) should be in scope. Without that distinction, nearly all AI systems used on these devices would be considered high-risk AI systems. The same logic should apply to comparable other systems that fall under specific EU regulations.

Definition and Identification of High-Risk AI

AI encompasses a variety of technologies which can have several specific uses and applications, all of which present different levels of risk. The wide majority of its uses carry little or no risks to individuals or society at large. For example, an AI system that uses analytics to streamline automobile manufacturing or to improve baggage handling and tracking at busy European airports will not pose any risks beyond those covered by existing legislation. Several uses of AI are also already regulated by existing horizontal and sectoral legislations, such as the General Data Protection Regulation, the Machinery Directive, the Medical Devices Regulation or the Payment Services Directive, which manage risks that may arise in these contexts.

A narrow, clear scope focusing on specific high-risk AI applications that are not captured by other existing horizontal or sectoral legislation is therefore critical to ensure the effectiveness of the proposed regulatory framework. This will help to avoid overburdening AI developers and negatively impacting their ability to innovate.

Looking at Article 6, we support the fact that the draft AI Regulation builds on existing product safety laws, and welcome the targeted focus on the use of AI as a component of a product that is already required to undergo third party conformity assessment. This will be key to avoid duplication with existing product safety legislation in place. However, the reference to the AI system being itself a product could lead to confusion and should be clarified to ensure legal certainty, given that the current scope of product safety laws does not leave space for AI to be a product in itself.

We are aligned with the way in which the AI Act identifies high-risk applications in articles 6 and 7, and support the precise criteria included in article 7(2) outlining how the Commission might evaluate high-risk AI applications. We support this approach as it utilizes a narrow definition of high-risk systems that considers damages to health, safety and fundamental rights of persons but also severity, likelihood of their occurrence and plurality of potentially affected individuals.

When it comes to article 7, we see a mismatch between the very targeted criteria in article 7(2) and the variety of applications identified in annexes II and III. For instance, while it is clear that there are privacy and fundamental rights concerns associated with access to training, education, or creditworthiness evaluation, it is unclear how the criteria in article 7(2), such as those related to the extent and intensity of the potential harm or its potential impact on a plurality of individuals, would apply to these use cases.

We recommend that lawmakers further clarify and narrow the language in annex III, taking into account the diversity of applications that may fall under some of the definitions. In the field of access to employment, for instance, for jobs that require specific skills like coding, many employers may give candidates automated evaluation tests that use AI to help identify errors, etc. In such case there is likely minimal risk, yet the definitions of AI and the scope of employment captured in annex III will likely trigger the high-risk threshold for such an application. A second example could involve AI applications that are used in critical infrastructure management to perform minor back-office functions. These sorts of applications would also likely be caught in the scope of the Regulation despite their limited risk factor. Similarly, some video security systems often employ AI to detect motion or unknown objects, or to help users cope with large amounts of video and human deficiencies. It is however unclear whether such a video security system when used to help monitor a critical infrastructure facility would be automatically considered 'high risk'. Finally, the inclusion of "AI systems intended to be used as safety components in the management and operation of road traffic" in annex III should be clarified by specifying that it refers only to systems which are not integrated in vehicles; otherwise, these would conflict with Regulation (EU) 2018/858 and Article 2.2.

Annex III also includes biometric identification as a high-risk use. For this case, the proposed Act says that a system is a biometric identification system if it operates "without prior knowledge of the user of the AI system whether the person will be present and can be identified." Based on this language, it is not clear the extent to which users can consent to the use of biometric identification systems, what kind of consent would be sufficient, and whose consent and when would be necessary. The answers to these questions should be calibrated

carefully to not foreclose beneficial uses of AI to which people would be willing to consent, if given the appropriate opportunity.

Article 7 details a procedure to update the list of high-risk applications in Annex III via delegated acts. Dialogue between industry, civil society and the Commission will be fundamental to ensure that the regulation remains up to date in the years to come and captures the relevant technological developments in a fast-changing market. While delegated acts provide for engagement among the European institutions in principle, in practice and as a general matter, these instruments provide little opportunity for public scrutiny and stakeholder input. Only in rare instances Commission proposals of delegated acts are subject to meaningful scrutiny and review. Given the importance of these annexes, any procedure leading to updating them should as minimum be subject to public consultation, to strengthen transparency and increase dialogue between policymakers and the interested parties.

Extraterritorial Application of the Regulation

Article 2(1)c of the proposal includes in the scope of the proposal providers and users of an AI system located outside the EU, but "where the output produced by the system is used in the Union." While this language intends to ensure that all AI systems entering the EU market are subject to the same requirements, this article reads broadly, and risks creating legal uncertainty. For instance, it is not clear what constitutes an "output" in this context, and this may contribute to uncertainty for many providers based outside of the EU as well as users in the EU. The reference to the "use" of the output also reads broad and should more precisely refer to the extent in which an AI system would affect the rights of European citizens. At the same time, it is fundamental that these measures aim to preserve interoperability to the extent possible, to avoid cases in which providers end up being impacted by several and conflicting legal requirements.

It is also unclear how this provision would work in practice, given that it appears to extend the regulatory framework beyond EU borders. Given the global nature of supply chains, including those that feed into the development of AI systems, it is key that the regulation defines roles and responsibilities of relevant actors in a manner that is clear, proportionate, and unambiguous to avoid situations of legal uncertainty.

In order to simplify the requirements, policymakers should clarify for instance how the regulation would apply to scenarios in which AI systems used outside the EU would affect the rights of European citizens, as well as the opposite scenario where an AI system used in the EU would not affect individuals in Europe.

Prohibited AI Uses

Article 5 proposes a ban of AI applications that are found to pose an unacceptable level of risk. In order to ensure legal certainty, scope of these applications need to be as clear as possible. These applications include, among others, AI systems that deploy "subliminal techniques beyond a person's consciousness" in order to materially distort their behaviour in article 5(1)(a). The definition of "subliminal technique" for this case is unclear, and it should be made explicit which exact uses of AI are meant to be targeted with this article. A precise

definition of "subliminal technique to influence a person's behaviour" is fundamental to avoid cases where unproblematic uses become covered by the scope of these provisions. Here, lawmakers should also provide guidance on what legal standard or process would be used to determine and enforce this provision, particularly in relation to demonstration of 'psychological harm'.

With regard to real-time remote biometric identification in publicly accessible spaces, we acknowledge the real risks to fundamental rights that can be posed by government use of AI for surveillance purposes. At the same time, it is also important to recognise the important public safety and national security benefits allowing responsible deployment with strict, meaningful safeguards. Managing risks in these operations is possible through clearly defined processes and controls such as human review, sufficient confidence scoring (for instance by assigning a percentage of accuracy to any output), judiciary supervision, clear use policies, reasonable boundaries around data retention, and transparency measures. Additional transparency requirements on the user of the AI system (for instance related to when, where and how the AI system is used, how the data is processed and stored and for how long) may be a solution to enhance safeguards for the safe and responsible deployments of such systems. In order to increase legal clarity, it would be important to further clarify and specify the notion of "publicly accessible spaces" under Article 5.1 d).

Requirements for High-Risk AI Applications and Obligations

Chapter 2 introduces a variety of requirements for high-risk AI applications with the goal of increasing transparency, explainability, fairness and cybersecurity. Our industry greatly values these goals as they are key enablers of trust in AI technologies and therefore play an important role in facilitating their adoption.

While we recognise the value of the goals of this section of the proposal, some requirements, for instance on data governance, human oversight or record keeping, seem to be based on a one-size-fits all approach. Excessively prescriptive process requirements may be difficult to implement in the same way by all the actors in scope. Given the diversity of the products and software impacted by this regulation, ranging from industrial machinery to financial services software, it would be more practical that these requirements are more goal-oriented so that companies are able to adapt the most meaningful and appropriate processes to ensure compliance, without weakening the goals of the regulation. A goal-oriented approach would entail establishing obligations to reach certain outcomes (e.g., mitigate bias to the extent possible) without prescribing how a certain goal should be achieved. Reliance on industry-driven standards will also help providers achieve compliance with the regulation. Such approach would certainly be beneficial for innovation and reflect the complexity of the sector.

This is fundamental because many of the requirements as they are now may involve significant costs to achieve compliance, reflect misunderstandings of how AI systems or data sets are used in practice, do not provide the required legal certainty, or simply do not match with technical characteristics of some AI applications. The impact on some sensitive technologies may in practice result in a de facto ban because compliance with such requirements would be too expensive or burdensome. More broadly, increased compliance

costs would probably result in higher prices for users, or make it too hard for smaller players to innovate. This could lead to a vicious circle of reduced consumption, and further cost increases, and result in more rudimentary, less AI-enabled technologies than in other parts of the world.

The Recitals make numerous references to protecting "health, safety and fundamental rights," but the conceptual structure of the proposal is built on existing market surveillance schemes derived product safety legislation. The proposal seems to combine two concepts that are fundamentally distinct: AI systems that are high-risk in the context of safety and health under the product safety legislative framework and stand-alone high-risk AI-systems that may otherwise impact people's lives or pose risks to fundamental rights. We question whether the same requirements designed for product safety will indeed result in the protection of those rights (such as non-discrimination, fairness, etc.). For instance, although the proposed market surveillance approach might make sense with respect to products that pose risks to health or safety, it works less well regarding risks to fundamental rights (e.g., would a discriminatory decision to deny a loan provide a basis for authorities to order the withdrawal of the system from the market?). Risks to fundamental rights would be better addressed through risk management systems by the provider, combined with appropriate transparency and accountability mechanisms by the user.

Another key element that requires clarification is how all these requirements, as well as the obligations in Chapter 3, would interact with general purpose software. Many companies in the B2B space produce one general purpose software that will then be used by customers to develop, train and deploy a variety of AI applications, some of which may fall under the definition of high-risk AI. It is important that the regulation clarify how the chain of responsibilities for the application of the requirements would work in these cases between users and providers. Certain general purpose software providers may in fact have little control over, or even knowledge of, the intended uses that are made of their software. Unlike standalone AI systems, it would not be feasible or desirable for providers to prescribe specific intended uses of these systems, as this would unduly restrict the customer's ability to innovate. Compliance with the regulation would therefore be impossible. In order to reflect the complexity of AI supply chains as referred to in recital 60 of the Commission's proposal, it is important that the final regulation explicitly excludes providers of general-purpose (AI) software from the definition of "providers" when they are not the ones who directly develop or deploy the system as a high-risk AI application. A practical solution could be that when general purpose software is developed or used as high-risk AI, compliance with the relevant provisions should fall on those actors who are best placed to implement the requirements.

It is also not clear in the Regulation how providers of component parts should be treated (e.g., the provider of speech recognition technology in a broader AI system). The Regulation should address component providers explicitly, making it clear that suppliers of AI components and their customers would have the freedom to allocate responsibilities between them by contract.

Data Governance Requirements

Article 10 lists a variety of data governance requirements intended to ensure quality of the training, validation, and testing data sets. Some requirements appear potentially disproportionate or could present significant implementation concerns for companies, particularly smaller ones. This is particularly problematic because violations of Article 10 are subject to the Act's most stringent penalties.

For instance, the requirement in article 10(3) for the data to be "relevant, representative and free of error and complete" is unrealistic in practice, as it is impossible to identify and eliminate all errors in a data set. In some cases, it may be more useful for models to learn with errors in the data, so they become more robust and better able to handle data encountered in the real world, which is unlikely to be sanitized and perfectly accurate. Similarly, the requirement to identify "any possible data gaps or shortcomings" in article 10(2) is too generic and very difficult to obtain in practice. This type of "absolute" requirement is extremely burdensome and impractical for AI developers, and as phrased now, may disincentivise innovation. All software and computer systems, including AI, will always contain bugs. Even the most complete coding process with associated QA controls cannot possibly identify all bugs prior to deployment. The standard needs to take full account of this reality. In addition, the definition of *relevance* is unclear. In fact, some AI providers will not know the relevance of the data until the algorithm is built. For this reason, the requirements of article 10(3) may make exploratory and development analyses more difficult to justify. Data governance requirements should therefore take into account the difference between training data sources and operational data sources, as excessive requirements for the former would impact companies' capacity to innovate.

More clarity is also needed on the provisions in article 10(4) that would require testing, training, and validation data to take into account for the purpose of bias monitoring the specific "geographical, behavioural and functional settings" in which the AI system is meant to be used. Providers should be required to make reasonable efforts to address the concerns of this article, without being faced with overly burdensome requirements. 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. Also, many AI systems are developed for a global audience and therefore would not necessarily need to reflect specific geographical settings.

Finally, we appreciate that article 10(5) introduces the concept of processing special categories of data mentioned in the GDPR and the Data Protection Law Enforcement Directive in order to monitor for bias. Some providers of AI need to collect sensitive data such as demographic data to help effectively detect and mitigate bias. However, further clarification on the processing of special categories of personal data under Article 10 (5) would be helpful, particularly with regard to the criteria which would make processing of these data 'strictly necessary' as per article 10(5). The introduction of an explicit lawful basis in the body of the Regulation would also help increase legal certainty. Policymakers should also consider how this article would interact with upcoming legislation such as the e-Privacy Regulation and include similar caveats for that piece of legislation.

Technical Documentation

Legislators should take into consideration the feasibility of compliance with prescriptive documentation requirements in article 11 and annex IV for the variety of systems in scope. The requirements in point 2 of Annex IV to provide a detailed description of the general logic of the algorithm, or extensive information on datasets could lead to the revealing of potentially sensitive information. At the same time, such burdensome reporting requirements may be difficult for smaller providers to comply with and thus may risk disincentivizing innovation.

The proposal should focus on feasible documentation appropriate to the use case, rather than rigorous proofs of quality. Since AI systems are software processes and not products, a check-box exercise of requirements is unlikely to adequately serve the purpose of a future-proof framework. Data quality and data provenance vary greatly, and thorough documentation is not always possible or necessary in light of risks. We recommend allowing for more flexibility in demonstrating compliance with the AI Act, using relevant documentation that could come from a company's internal practices as well as from documentation required by international standards.

Record Keeping

The record keeping requirements in article 12 are also based on a prescriptive approach, and present difficulties for uniform implementation by the variety of providers in scope of the AI Act. Article 13(1) for instance mandates an automatic recording of logs on high-risk AI systems. As mentioned, there should be flexibility on *how* to achieve some of the goals, to ensure that the regulation is as future-proof as possible and allows the necessary flexibility for providers to find the most appropriate and efficient solutions for their specific AI product. In this sense, a prescriptive mandatory requirement to have automated log-recording in place would place a significant burden on providers. We thus suggest avoiding mandatory record-keeping requirements on an ongoing basis as provided by article 16(d) and article 20, as some providers may have capacity problems in storing and maintaining these large amounts of data.

Looking at previous experiences, requirements imposing record keeping, for instance for medical devices for a period of time equivalent to the design and expected life of the device can generate high administrative burden and investments for an outcome that is uncertain. Specific, measured requirements for enumerated high-risk AI applications should be considered instead. Often, a better solution is to support the development of standardised testing and require specific types of testing for high-risk AI applications. The EU should establish (or fund/support establishment of) these standards and benchmarks, including tests, for high-risk scenarios, and ensure AI systems that will be used in those scenarios meet these standards. This would be a far more effective method of addressing these concerns than requiring years of recordkeeping and making proprietary data/programming/algorithms available.

Transparency

Transparency is an important aspect of and helps facilitating trust in AI systems. Still, it does not automatically equate to better control of automated decisions by the user. 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 detailed information of the workings of the technology to use it responsibly.

Transparency requirements should thus vary according to the diversity of applications in scope of the AI Act. In considering AI explanations, value to the consumer is key – one of the benefits of transparency should be to help individuals understand how the use of AI will benefit them. It is also important to strike a balance so that users and consumers do not experience "decision fatigue" and can understand the use of the AI technology without being bogged down in technical details. In addition, there should be a differentiation of transparency requirements between consumer facing AI and B2B products. For instance, in B2B scenarios excessive sharing obligations might impact IP rights and contractual arrangements between business partners.

Transparency is best achieved by 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 prioritised to build consumer trust. Policymakers should avoid governance that creates an environment where outliers are viewed as a flaw in an overall AI system. If an outlier is indeed an outlier, then the algorithm will learn and dismiss it in later iterations so no "explanation" is necessary. As such, when and how an "explanation" may be required is highly contingent on the stage of an AI system's developmental lifecycle, the context in which a later-stage model is deployed, the purposes for which it is deployed, and numerous other factors. Any guidelines related to transparency or explainability should capture a statistically meaningful number of results to ensure uncertain results are actual concerns and not just isolated anomalies. 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.

The approach to transparency in article 13 seems prescriptive and potentially burdensome, and the value of some of the provisions to further understandability is unclear. For instance, we do not see how the provisions in article 13(3)(b)(v) requiring transparency on the specifications for the input data or "any other relevant information in terms of the training, validation and testing of data sets" would contribute to enhancing the user's understanding of the AI system's output. In addition, mandating strict transparency requirements on datasets can pose risks in terms of revealing sensitive data and/or revealing a business' IP. This risk is heightened by the requirements in article 60 and annex VIII to make this information publicly available, in the case of standalone high-risk AI systems, in an EU database. It is thus important that the provision of information is balanced with protection of sensitive information, such as personal data and commercially sensitive information, in combination with making such data available only to supervisory authorities who should confirm for which purposes each data set would be required.

The requirement to disclose any "known or foreseeable circumstance" which may lead to health and safety risks related to the use of AI in accordance with its intended purpose or "under conditions of reasonably foreseeable misuse" in 13(3)(b)(iii) is very broad and may be difficult for providers to identify. In addition, article 13(3)(e) requires providers to disclose "the expected lifetime of the high-risk AI system." This may not be feasible in practice, in particular for AI systems offered as services, because these systems are often updated and improved on an ongoing basis so that they remain active and functional for as long as possible; they seldom have an expected "expiry date" from the outset.

Rather, providers should be allowed to implement the practices and processes that make the most sense for their specific product and business model, while at the same time meeting the goal of transparency. The overall approach of combining ex *ante* self-assessment with significant ex post penalties is premised on the belief that providers best understand their products and systems and are best positioned to implement to the practices and process that best suited for their systems. The same should hold true for these requirements.

Human oversight

Human oversight is crucial to reap the full benefits of AI while controlling for potential risks. The value of human involvement is different for each specific use case. For example, it is proportionate to have a human monitoring 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. For other, less critical situations, detailed human involvement may not be necessary or proportionate.

The appropriate degree of human involvement in reviewing machine-generated decisions should therefore be determined based on the specifics of the individual use cases. In some cases, human oversight can lead to delays, in others, accuracy of outputs could even be undermined by human interventions (for example for mathematical calculations).

For these reasons, the human oversight requirements in article 14 should allow for the necessary flexibility to implement the most appropriate solutions for the diverse uses in scope of the regulation. Mandating specific solutions, such as the case of the "stop button" in article 14(4) is overly prescriptive and will not only be difficult for all providers to implement in the same way, but also may not necessarily lead to the best outcomes in terms of reaching the desired goal. Finally, the instructions to users and the requirement that human oversight enable the user to "fully understand the capacities and limitations of the AI system" in Art.14.4.(a) seem difficult to achieve in practice, since a provider cannot guarantee what a user will understand.

These requirements should be consistent with the existing rules of Art. 22 GDPR. The degree of oversight should be adapted to the specific risks, the level of automation, and context of the AI system in order to avoid hindering automated processes. It will be important to note that bias could also be introduced by human developers. Therefore, the added value of human oversight measures should be re-considered for some instances instead of promoted in the proposal. Furthermore, we would strongly recommend specifying which human

oversight approach (human-in-the-loop, human-on-the-loop, or human-in-command) and at which step of the high-risk AI system's lifecycle should be adopted.

Conformity Assessment, Standards and Marking

With the Conformity Assessment and Standards requirements in chapters 4 and 5, the proposal lays out a template for application of the New Legislative Framework – the EU's three legislative acts governing standardisation, conformity assessment, and accreditation across industrial goods sectors – to high-risk applications of AI. These provisions should take into account the complex nature of global AI supply chains and fundamental differences between AI systems, which by default are continuously "evolving," and the more "static" industrial goods. Relying on voluntary industry-driven consensus-based international standards, ensuring alignment on base standards, and ensuring that there are innovative mechanisms in place to easily and flexibly accept international testing outcomes are fundamental elements to avoid regulatory fragmentation and ensure that the EU does not become an "island" in the global AI market.

Standards

The proposal must rely on voluntary industry-driven consensus-based international standards as key to establishing consensus around technical aspects, management, and governance of the technology, as well as framing concepts and recommended practices to underpin trustworthiness of AI inclusive of privacy, cybersecurity, safety, reliability, and interoperability. The quality management and risk management systems in the AI Act as well as the data governance requirements should thus be based on relevant international standards. Standards particularly when used as technical regulations, must not create market access barriers or preferential treatment; rather, they should work for the benefit of society, consumers and the wider ecosystem of the international community and European regulatory authorities, and be applicable without prejudice to cultural norms to evaluate the outcomes/use of AI. The proposal should safeguard against the risks for fragmentation between the EU and global regulatory environments, which may impact the availability of AI products in the EU market and therefore decrease innovation in the field, uptake of the technology and consumer choice.

The provisions concerning exclusive reliance on harmonised European standards as a means of demonstrating compliance with corresponding requirements derive from the New Legislative Framework, and we appreciate the Commission's efforts to build upon a standardisation framework that has provided predictability in a number of industrial goods sectors. However, there are many standardisation activities taking place outside of European Standardisation Organizations (ESOs) that may be relevant for the purposes of the AI Act.[1] Sole reliance on European standards not only creates the possibility of future divergence in different regions' approaches to AI governance, but also limits the tools EU policymakers have at their disposal for ensuring that the most fit-to-purpose solutions may be used in a timely

[1] Relevant examples can be found in the list of the InterNational Committee for Information Technology Standards (INCITS) on common reference standards for AI and Biometrics: https://www.incits.org/contentAsset/raw-data/688802a6-4aeb-4333-9782-0c8b855ba040/reportFile/bd433a99-972d-4af0-8c2f-cc712a82516d.pdf.

manner to demonstrate compliance with regulatory requirements, particularly where digital services and new technologies are concerned. As a first mover and leader in the AI governance space, EU should adopt a global approach to standardisation that will enable other jurisdictions to follow suit in a manner that does not detract from innovation or lead to unnecessary divergences. In addition, global standards help enable interoperability, establish a common understanding and set a level playing field for AI based products and services, which are key to enable not only the Single European market but also to reduce barriers for international trade.

Article 41 grants the Commission powers to adopt common specifications via implementing acts in cases where relevant harmonised European standards do not exist or are found to be insufficient for the protection of fundamental rights. Following this path despite the existence of suitable standards from standards bodies other than the ESOs or ISO/IEC carries a risk for potential divergence from international standards and can adversely impact the ability of European companies to compete in global markets and also reduce consumer benefit. We urge lawmakers to avoid in all instances the development of any bespoke (and therefore region-specific) technical specifications, and instead where necessary rely exclusively on international standards. Adopting technical specifications outside of an open and consensus-based model can result in frameworks that are not future and technology-proof. Rather, the EU should rely on global standards developed in organisations such as ISO/IEC JTC1 SC 42 which champion an inclusive, open and diverse approach to standards creation and are built on a consensus basis by technical subject matter experts, thus providing above outlined benefits to all stakeholders in the ecosystem.

Third-Party Conformity Assessment

Conformity assessment for Artificial Intelligence technologies is a nascent field for which there is neither a commonly understood practice nor the established conformity assessment infrastructure necessary to carry out the requisite assessments contemplated by the proposed Act. For this reason, there are significant practical and logistical concerns regarding precisely how Notified Bodies, once identified, accredited, and designated, would carry out the task of assessing the conformity of certain high-risk AI systems with the broad requirements outlined in the Act. Given that tools and processes for assessing compliance in this field are still emerging, it is unclear how existing facilities would have to be transformed to perform these tasks in a timely way and with the needed skill and expertise, and what type of guidance would be needed to ensure appropriate capacity of the testing bodies. Logistical problems, including the lack of sufficient designated Notified Bodies, may also lead to backlog for testing bodies, which could significantly slow down the adoption of certain AI technologies in the EU market.

Decision 768/2008 and the corresponding provisions contained in the AI Act require that any third-party conformity assessment be carried out by a Notified Body established under EU law (i.e., located in the territory of an EU member state). While this requirement derives from the New Legislative Framework, in an area in which reliance on conformity assessment yields a number of technical and practical questions, we would strongly encourage lawmakers to in-build greater flexibility as concerns the acceptance of international test results, including as a

means of demonstrating to third countries the importance of avoiding localised testing requirements.

In this spirit, the proposal should provide greater clarity with regard to the acceptance of testing results produced by competent testing facilities located outside of the EU. Article 39 of the proposal allows for recognition of third-country test results in cases in which there is an "agreement" in place between the EU and the third country. However, the specific nature of such an agreement is unclear, and given the ubiquity of AI applications that could potentially require third party conformity assessment, EU lawmakers should consider reliance on existing international accreditation schemes and proven international mutual recognition arrangements as a means of facilitating the acceptance of test results developed by competent facilities based outside of the EU. Beyond more limited government-to-government arrangements (including both bilateral mutual recognition agreements and conformity assessment protocols), reliance on more innovative pathways for acceptance of test results developed by testing bodies based outside of the EU would facilitate innovation and regulatory compatibility without detracting from the regulatory oversight of European authorities.

Finally, while at the moment only biometric identification would have to undergo third-party conformity assessment among the standalone high-risk applications, according to article 43(6) the Commission is empowered to adopt delegated acts to extend third-party conformity assessment to other applications in annex III. Such decisions should be taken in close consultation with industry and other stakeholders to foster transparency and better take into account legitimate concerns around feasibility or other logistical issues.

Conformity Assessment of AI Systems Already in the Market
Article 83 lays out requirements for the application of the Regulation to AI systems already in the market. Specifically, article 83(3) establishes that the regulation applies to systems that were put in the market before the entry into force of the AI Act only if these systems are subject to "significant changes in their intended purpose" from that date. This provision should be clarified as it may lead to substantial legal uncertainty. At a minimum, we recommend that the AI Act defines "significant change" in the context of this article, as it is not clear what constitutes a significant change. For example, significant change can refer to a change that would have a material impact on the intended use such as triggering a new risk or harm or substantially modify the intended use.

Conformity Marking
Any trust or conformity marking requirements should be meaningful, easily understandable and meet user needs first and foremost, and align with already established norms and requirements of sectoral supervisory authorities, rather than prioritising a "one size fits all" solution across all industries with CE marking.

Post-market Monitoring and Market Surveillance

Post-market monitoring

According to article 61, providers would have to actively and systematically collect documents and data provided by users or collected via "other sources" to evaluate the performance of the AI system and their compliance with the requirements of this regulation. As previously noted, this type of approach is prescriptive and does not seem to account for the burden on providers, especially smaller ones, which may not have the capacity to carry out such extensive monitoring. These requirements should always account for flexibility reflecting the difference between the types of AI system in scope.

The estimated compliance costs in the Explanatory Memorandum (Section 3.3 - Impact Assessment) could be understated, particularly for post-market obligations on providers to monitor the performance of high-risk AI systems for the lifecycle of the system. For example, requiring providers to collect data on an AI system's operation, to ensure 'continuous compliance' and take corrective actions as needed, would represent a significant shift of compliance responsibilities and cost from the user to the provider of AI systems. In some cases, this may be impossible for the provider to do, since some AI systems are installed in sensitive environments and, by design for security reasons, a provider cannot access them or exfiltrate data. Data-sharing stipulations mean that many AI systems in government use do not permit human access under normal circumstances. Any monitoring and specifically 'human oversight' would need a new mode of operation, with significant resource costs to examine and analyse data on a per customer basis. Any kind of automation or metrics would have to be tailored to the specific workflow task or use case. This would need to be set out in detail to make it actionable for companies.

Incident reporting

Article 62 details reporting obligations in cases of "serious incidents or any malfunctioning" which constitute a breach of EU law on fundamental rights. However, it is not clear what the definition of "serious incident" is. In order to comply with such a provision, industry needs additional certainty as to what constitutes such an incident. This should be accompanied by a discussion on Art. 3(44) defining 'serious incident'. Here, the concept of 'serious incident' is defined too broadly as any "incident that […] indirectly leads, might have led or might lead" to certain undesirable events. Liability requires a reasonable standard of causation where the product's defect must be the (reasonably likely) legal cause of the harmful result. However, in this case, the causality threshold is weakened, thus potentially putting excessive responsibilities on the providers and contradicting the chain of causation.

It is equally important to clarify what is meant by the notion of malfunctioning, which lacks qualifiers and could be given broad meaning to apply to any circumstance in which the AI system does not perform as intended. This threshold seems unreasonably low and would create large administrative costs for AI providers and market surveillance authorities. The first would have to incur large compliance costs to closely monitor and report all cases of malfunctioning and the second would have to assess and decide on appropriate measures in every such cases. It seems reasonable that only serious malfunctioning that breaches EU law should be covered under Art. 62.

In addition, policymakers should clarify how incident notification under the AI Act relates and potentially overlaps with other legislative frameworks such as the GDPR, the DORA Regulation

and the proposed NIS2. Each of these instruments specify different reporting times and require that notification be made to different regulators, which will be tasked with assessing the incident and deciding on appropriate measures, including investigations. It is thus fundamental that lawmakers clarify which reporting system applies under which circumstances where overlapping obligations exist.

Market surveillance

The section of the proposal on market surveillance, and especially article 64, gives market surveillance authorities broad powers to access data and documentation to carry out their activities. Article 64(1) for example would require providers to grant "full access to training, validation and testing datasets", remotely or through APIs. Given the sensitivities of granting access to such data as well as the risk to reveal confidential or protected information stemming from exposure, such requests should always be sufficiently motivated and proportionate.

Other requirements, such as the one in article 64(2) to grant access to source code of the AI systems in some specific cases are extremely far-reaching and disproportionate. The AI Act currently provides for broad powers for market surveillance authorities to request access, and does not clarify how companies may seek remedies, or how such requests would be issued and justified by the issuing body, and possibly reviewed by the competent judiciary authorities. Disclosure of source code could seriously put at risk important trade secrets and IP rights and contravenes widely accepted best international practices for digital trade and should therefore never be requested by surveillance authorities.

It is also unclear how the provision in article 64(5) that would allow market surveillance authorities to organise "testing through technical means" of AI systems in cases where the relevant documentation provided is insufficient would work in practice. It is important that such measures remain proportionate to avoid becoming a disincentive for companies to innovate.

Governance

The AI Board is a good initial step toward achieving the key goal of ensuring that governance and enforcement of the AI Act is overseen by policymakers with broad expertise that are committed to enabling consistency across Member States and fostering the necessary exchanges between industry, stakeholders and regulators. For this reason, we would we recommend the introduction of a more structured and regular dialogue between Board members and third parties. Such a dialogue within the governance mechanisms can help ensure that public authorities and stakeholders can more effectively follow technological developments and ensure the regulatory framework remains up to date. Given the Board's competence to issue opinions on the use of harmonised standards and the definition of technical specifications, it is fundamental that it engages with global standardisation organisations to ensure that global approaches to standardisation remain aligned with the goal of avoiding regulatory divergence.

Leaving the Board aside, the governance structure proposed by the AI Act is an intricate system with each Member State designating national competent authorities and a national supervisory authority. This complex structure risks leading to differing interpretations of the provisions of the AI Act, creating unnecessary legal complexity that will frustrate the development of a robust AI ecosystem in Europe and could run counter the goal of further harmonising the Single Market. This architecture would pose the most significant burdens on SMEs across Europe. While large entities may have the legal compliance capacity to navigate the conflicting interpretations and enforcement standards across Member States, this fragmentation may severely limit SMEs abilities to develop innovative new AI systems in Europe.

In order to approach AI governance more holistically, it is important that national supervisory authorities are equipped with a diverse array of expertise and perspectives. To avoid fragmentation in the Single Market, there should be a lead national supervisory authority in Member States which is equipped with the necessary expertise and designated through convergent criteria via additional mechanisms that lawmakers should develop. This can help ensure appropriate technical expertise and further legal predictability. Having different market surveillance authorities for different high-risk applications may be problematic, as it could give rise to *case-by-case* decisions made by authorities with different expertise.

Other measures

Codes of Conduct

We recognise that provisions governing Codes of Conduct as introduced in title IX replicate comparable elements of other recently introduced EU legislation. However, rather than developing region-unique codes of conduct or technical specifications through processes that lack the open participation and due process of international standards development activities, we strongly encourage the EU to rely exclusively on industry-driven international standards and best practices where appropriate as a means of establishing a pathway to demonstrating conformance with a given requirement(s). This will also help to avoid fragmentation that might occur.

Avoiding reliance on region-specific codes of conduct or technical specifications would be crucial to keep with the risk-based approach of this act, which we agree should underpin the EU's approach to regulation of AI, as other approaches would damage the dynamism of EU developers' scene and risk damaging the potential of EU providers to innovate.

Support to Innovation and Small Businesses

Title V of the proposal introduces innovative and flexible regulatory approaches such as regulatory sandboxes to encourage and stimulate AI innovation. These measures are extremely valuable for smaller businesses as they reduce the potential burden of compliance for certain requirements and therefore enable innovation, new products, and opportunities for growth. For this reason, the priority access for smaller businesses to regulatory sandboxes as per article 55 is a welcomed initiative.

Collaborative, multi stakeholder policy prototyping can provide a safe space to explore and develop innovative regulatory and co-regulatory tool. It is important that sandboxes provide a mechanism to share expertise and technical knowledge with policymakers in a collaborative effort. Investing in stronger sandboxes is a way to form policy recommendations that is more suited to the fast-developing technology industry.

The regulation therefore needs to more clearly lay out what the incentives are for companies to join sandboxes, and what outcomes they can expect. It would also be important to propose concrete solutions to advance uptake of AI technologies by SMEs in the EU and increase trust in the technology. For instance, providing small businesses with easy access to information on the necessary steps to comply with the AI regulatory framework, as well as on the benefits of adopting AI solutions.[2] We encourage the EU policymakers to work with civil society and industry to think about the specific functioning of sandboxes and setting them up in a way which truly helps companies to drive innovation in a protected environment to unearth learnings for all stakeholders involved.

Global Considerations

The AI Act should be leveraged to facilitate a broader global conversation around AI governance. In close cooperation with like-minded global partners, the EU should strive to find alignment on key elements related to AI governance, in an effort to promote open, non-discriminatory and principle-based cooperation and trade in the field of AI.

As noted above, the AI ecosystem is global, and the technology is not developed in regional siloes. Thus, the most effective means of advancing Europe's AI agenda is to expand the discussion beyond national borders. The EU should engage beyond the borders of the single market to further the development and use of AI globally by cooperating with its international partners. International cooperation on AI should be based on promotion of respect of fundamental rights, non-discrimination and protection of privacy.

This also means recognising the significance of Europe's mutual interdependence with like-minded global partners, and the importance of shared common values like trust, fairness, explainability, effectiveness, safety, and human oversight. There is a valuable opportunity in working together to shape balanced solutions and ensure that policy options on AI that are being considered globally remain aligned, coherent and interoperable at a global level.

***

[2] As proposed by the Progressive Policy Institute's Report "Encouraging AI Adoption by EU SMEs", (2021). https://eadn-wc05-3904069.nxedge.io/cdn/wp-content/uploads/2021/03/PPI_Encouraging-AI-adoption-by-EU-SMEs-3.24.21-2.pdf.