BSA RECOMMENDATIONS ON THE AI ACT TRILOGUE NEGOTIATIONS

EXECUTIVE SUMMARY

BSA | The Software Alliance is the leading advocate for the global software industry before governments and in the international marketplace. Our members are at the forefront of software-enabled digital transformation that is fuelling global economic growth. BSA members include many of the world's leading suppliers of software, hardware, and online services to organizations of all sizes and across all industries and sectors. As leaders in AI development, BSA members have unique insights into both the tremendous potential that AI holds to address a variety of social challenges and the governmental policies that can best support the responsible use of AI and ensure continued innovation.

BSA has been a strong supporter of the risk-based approach of the original AI Act proposal, and we strongly recommend that the EU co-legislators ensure that it is reflected in the final version of the Act. At the same time, the European Parliament and the Council have brought forward significant changes that would ensure that the AI Act is in line with international work on Artificial Intelligence, with a commonly acknowledged definition of AI, and that its risk profiles definition provide the necessary legal certainty and flexibility.

BSA's recommendations for EU policymakers focus on:

- ***Adopting a technologically neutral and risk-based AI Act***

    1. Ensure that the definition of Artificial Intelligence is in line with other global partners.
    2. Establish a definition of high-risk that provides for the necessary flexibility and legal certainty.
    3. Include precise and workable definition of high-risk use cases.
    4. Maintain key scope clarifications for Open Source and R&D.
    5. Ensure that post-market enforcement is balanced and in line with EU law.

- ***Ensuring balanced obligations for the AI value chain***

    6. Ensure that bias management practices remain at the highest level in Europe.
    7. Ensure a balanced approach to the AI value chain, ensuring clear responsibilities for providers and deployers.
    8. Establish workable and balanced responsibilities for foundation models.
    9. Design obligations for providers and deployers that support AI uptake
    10. Establish clear obligations to ensure content authenticity.

ADOPTING A TECHNOLOGICALLY NEUTRAL AND INNOVATION-FRIENDLY AI ACT

The AI Act will be the first-of-its-kind regulatory framework applying to Artificial Intelligence. As such, the Act should be in line with international work on AI, especially in cooperation with the EU's global partners, to ensure interoperability and the possibility to set agreed standards on key aspects of AI policy, such as the definition of AI and the definition of high-risk. At the same time, as a horizontal regulatory framework, it is fundamental that the AI Act maintains a technologically neutral and innovation-friendly approach, ensuring that Open Source and Research and Development activities remain outside of the scope of the AI Act, and that the post-market enforcement of the Act is in line with EU law, and supports innovation and protection of trade secrets and IP.

1. Ensure that the definition of Artificial Intelligence is in line with other global partners

The definition of AI included in the original proposal of the AI Act was considered excessively broad by both the Council and the European Parliament. Similarly, many stakeholders – including BSA – had voiced concerns on the original definition of AI, which would have likely included many processes and software that are not traditionally considered Artificial Intelligence. The European Parliament's report put forward a definition of AI that is in line with the work of the OECD, and likely to become an internationally accepted definition of Artificial Intelligence. BSA recommends that the co-legislators ensure a definition of AI in line with other global partners and technological developments, including the OECD and the definition in the U.S. National Institute for Standards and Technology (NIST) AI Risk Management Framework, by confirming the European Parliament's report definition of AI in Art. 3(1).

Recommended Trilogue amendments – retain Parliament proposal

Article 3(1)

'artificial intelligence system' (AI system) means ***a machine-based system that is designed to operate with varying levels of autonomy and that can, for explicit or implicit objectives, generate outputs such as predictions, recommendations, or decisions that influence physical or virtual environments***.[1]

2. Establish a definition of high-risk that provides for the necessary flexibility and legal certainty

The risk-based approach of the AI Act hinges on the definition of high-risk provided by Art. 6(2) of the Act. Both the Council and the European Parliament saw fit to further define the definition of high-risk, to ensure that providers and deployers of AI have a clear understanding of what

[1] Hereinafter:
Bold Italics: language addition by Parliament or Council
Bold Italics Underlined: BSA suggested addition
~~Strikethrough~~: BSA suggested deletion

would constitute high-risk in the many diverse applications of AI that would fall within Annex III. By linking the concept of high-risk to the significance of harm, the European Parliament Report introduces a clear and established concept that is also flexible enough for supervisory authorities. While the Council had introduced similar language, its use of "purely accessory" as a concept, that would be further developed through Implementing Acts by the European Commission, would likely create more uncertainty than it would remove. BSA recommends that the co-legislators establish a definition of high-risk that provides for the necessary flexibility and legal certainty, as included in the European Parliament Report.

Recommended Trilogue amendments – retain Parliament proposal

Article 6(2)

In addition to the high-risk AI systems referred to in paragraph 1, AI systems ***falling under one or more of the critical areas and use cases*** referred to in Annex III shall be considered high-risk ***if they pose a significant risk of harm to the health, safety or fundamental rights of natural persons***.

3. Include precise and workable definition of high-risk use cases

The list of high-risk use cases provided for by Annex III is a key element in the structure of the AI Act. Both the Council and the European Parliament have amended the language of Annex III, to reflect a more specific description of the use cases that would fall within the scope of the Act. BSA is supportive of these changes, and would recommend retaining the amendments made by the European Parliament in Paragraph 2 of Annex III, which provides the necessary clarity for AI deployed in Critical Infrastructure.

Similarly, BSA recommends retaining the language proposed by the European Parliament Report for Paragraph 4 of Annex III (Employment) , in particular the changes made to letter b) with the inclusion of "materially influence", which should also be reflected in letter a) of the same paragraph. This would ensure consistency for deployment of AI in the employment sector, and the necessary legal certainty for AI providers and deployers in qualifying when an AI would fall within the scope of the AI Act.

Recommended Trilogue amendments - retain Parliament proposal with additional amendments

Annex III

2. Management and operation of critical infrastructure:

(a) AI systems intended to be used as safety components in the management and operation of road, rail and air traffic ***unless these are regulated in harmonisation or sectoral legislation***.

(aa) AI systems intended to be used as safety components in the management and operation of the supply of water, gas, heating, electricity ***and critical digital infrastructure***.

4. Employment, workers management and access to self-employment:

(a) AI systems intended to be used ***to make or materially influence decisions affecting*** the recruitment or selection of natural persons, notably ***for placing targeted job advertisements***, screening or filtering applications, evaluating candidates in the course of interviews or tests;

(b) AI ***systems*** intended to be used ***to make or materially influence*** decisions ***affecting the initiation***, promotion and termination of work-related contractual relationships, ***task allocation based on individual behaviour or personal traits or characteristics***, or for monitoring and evaluating performance and behavior of persons in such relationships.

4. Maintain key scope clarifications for Open Source and R&D

The original Commission proposal included language in Art. 2 that would partially limit the scope of the AI Act to R&D activities. Both the Council and the European Parliament have further improved on that language, clarifying that research activities in the AI field would not be included in the scope of the Act, until they were placed on the market as a high-risk AI, or subject to other specific obligations provided for by the Act. BSA is strongly supportive of this approach. Furthermore, the European Parliament included helpful language that would exempt Open Source AI from the scope of the Act, unless they are placed on the market or put into service directly as a high-risk AI system. These changes would ensure that the EU remains a leader in the development of AI and supports Open Source research and activities. At the same time, as it will be further explained below, it would be equally important to acknowledge the vastly different functioning of Open Source in the context of foundation models, and the ability of Open Source developers to control downstream use of the AI systems they have contributed to developing.

Recommended Trilogue amendments = retain Parliament proposal with amendments

Article 2 (Scope)

5d. This Regulation shall not apply to research, testing and development activities regarding an AI system prior to this system being placed on the market or put into service, provided that these activities are conducted respecting fundamental rights and the applicable Union law.

5e. This Regulation shall not apply to AI systems or components provided under free and open source licences except to the extent they are placed on the market or put into service by a provider as part of a high-risk AI system or of an AI system that falls under Title II or IV. ~~This exemption shall not apply to foundation models as defined in Art 3.~~

5. Ensure that post-market enforcement is balanced and in line with EU law

The original Commission proposal provided for very broad powers for post-market monitoring and enforcement, including the possibility to request access to source code in Art. 64(2). The Council has partially confirmed this possibility, with language that would slightly limit the possibility to access source code. At the same time, the European Parliament has instead amended the language in Art. 64(2) removing references to source code, introducing the possibility to request access to trained models and model parameters. From a technical standpoint, the European Parliament approach would lead to better enforcement and eliminate

excessively pervasive requirements. Access to source code would likely not lead to a better understanding of possible concerns, since it would necessitate a very high level of technical understanding of the specific AI system, and at the same time jeopardize key trade secrets and IP. From a legal standpoint, requesting access to source code would also likely run afoul of EU trade commitments, and in many instances counter to EU priorities in trade negotiations. BSA strongly recommends confirming the European Parliament language in Art. 64(2). At the same time, BSA would also recommend deleting the sentence referring to "simpler software systems" in Recital 79, whereby source code access could still be granted. In first instance, it is not clear what would constitute a "simpler software system", which in many cases could be precursors to AI systems and essentially constitute a request for access to part of a source code of an AI, and similarly, such requirement would not technically improve enforcement and likely be equally counter to EU trade commitments and law.

Recommended Trilogue amendments - retain Parliament proposal with additional amendments

> (79) ~~***In cases of simpler software systems falling under this Regulation that are not based on trained models, and where all other ways to verify conformity have been exhausted, the national supervisory authority may exceptionally have access to the source code, upon a reasoned request.***~~
>
> *Article 64 (Access to data and documentation)*
>
> 2. Where necessary to assess the conformity of the high-risk AI system with the requirements set out in Title III, Chapter 2***, after all other reasonable ways to verify conformity including paragraph 1 have been exhausted and have proven to be insufficient,*** and upon a reasoned request, the ***national supervisory authority*** shall be granted access to the training and ***trained models*** of the AI system***, including its relevant model parameters. All information in line with Article 70 obtained shall be treated as confidential information and shall be subject to existing Union law on the protection of intellectual property and trade secrets and shall be deleted upon the completion of the investigation for which the information was requested.***

ENSURING BALANCED OBLIGATIONS FOR THE AI VALUE CHAIN

The original AI Act proposal sought to establish a complex system of obligations and responsibilities for providers and deployers of AI systems, which was often tailored towards a very significant role for providers, often in the design and development phase, with less of a focus on the deployment phase, and the role of deployers. Both the Council and the Parliament have equally sought to rebalance this approach, in particular with regards to the very diverse AI value chain, which is rarely made of a binary relationship between one provider and one developer. BSA recommends ensuring that the AI value chain obligations established by the AI Act are balanced, and seek to allocate responsibilities and obligations on the entities best placed to comply with them, mitigate risks and understand the specific context and use-case.

6. Ensure that bias management practices can remain at the highest level in Europe

The original Commission proposal included in Art. 10(5) the possibility to process special categories of personal data when strictly necessary for the detection of bias monitoring, detection and correction. This possibly, confirmed by the Council General Approach, would be limited to very strict circumstances, and goes in the right direction in ensuring that AI providers and deployers have the necessary tools to detect unfair bias. The European Parliament significantly modified Art. 10(5), with the addition of more stringent conditions for processing, including limitation on transmission – which may impede the highest level of bias management practices in Europe. The Parliament proposal adds significant requirements for processing personal data for the purposes of bias management, which are likely to overlap with similar requirements already included in the forthcoming NIS 2 Directive and GDPR. BSA strongly recommends that the original version of Art. 10(5) is retained, to ensure that the AI Act is designed to protect against unfair bias, and providers and deployers can use the necessary tools to monitor, detect, correct and minimize unfair bias.

Recommended Trilogue amendments (unchanged from Commission proposal):

> *Article 10 (Data and data governance)*
>
> To the extent that it is strictly necessary for the purposes of ensuring bias monitoring, detection and correction in relation to the high-risk AI systems, the providers of such systems may process special categories of personal data referred to in Article 9(1) of Regulation (EU) 2016/679, Article 10 of Directive (EU) 2016/680 and Article 10(1) of Regulation (EU) 2018/1725, subject to appropriate safeguards for the fundamental rights and freedoms of natural persons, including technical limitations on the re-use and use of state-of-the-art security and privacy-preserving measures, such as pseudonymisation, or encryption where anonymisation may significantly affect the purpose pursued.

7. Maintain a balanced approach to the AI value chain, ensuring clear responsibilities for providers and deployers

The regulatory requirements for the various entities responsible for designing, developing and deploying AI should account for the unique roles and capabilities of the entities that may be involved in an AI system's supply chain. Any obligation (and associated liabilities) should fall on the entity that is best positioned to both identify and efficiently mitigate the risk of harm associated with a given AI system or use case. We continue to support an AI legislation that promotes accountability for both AI developers and deployers, as each entity in the design, development and deployment of AI should have clear responsibilities and obligations. At the same time, BSA would caution against establishing a clear-cut dichotomy between AI providers and AI deployers, which would be not representative of the much more diverse reality of AI development and deployment – where multiple entities are involved in the design, development, training and deployment of AI, and as such their responsibilities and obligations should be better reflected in the Regulation.

To this end, the original AI Act proposal included language in Art. 28 that would require compliance with the AI Act for the entity that decides the destination into a high-risk use of the AI system. While the original language of Art. 28 was not clear on certain aspects – for example

it did not clarify the obligations for a non high-risk AI that is deployed in a high-risk scenario – the fundamental structure ensured a balanced approach and clearer distinctions for when the obligations of the AI Act would be triggered. The Council General Approach upended this system, with the introduction of Art. 4a, 4b and 4c, which would essentially equate general purpose AI systems – regardless of their design and development and use – to high-risk AI systems. The European Parliament introduced a separation between so-called multi-purpose AI systems, and foundation models. For the former, the obligations would remain similar to the original structure of the AI Act, with the addition of communication obligations to support downstream compliance in case of high-risk placement, and the possibility to settle these obligations contractually. BSA is strongly supportive of this approach, and recommends that the European Parliament amendments of Art. 28 are retained in the final version of the AI Act.

Additionally, BSA recommends deleting the suggested Art. 4a, 4b and 4c of the Council General Approach, which would directly counter the Parliament proposal and establish legal obligations akin to those of high-risk AI for AI systems without an intended purpose. This important point is included in the European Parliament suggested Recital 60g, which clarifies that "multi-purpose" AI systems should not be included in the scope of the AI Act, unless they are integrated into or placed on the marked as a high-risk AI system.

Recommended Trilogue amendments – retain Parliament proposal with amendments in Recital 60g

Article 28 (Responsibilities along the AI value chain of providers, distributors, importers, deployers or other third party)

1. Any distributor, importer, ***deployer*** or other third-party shall be considered a provider ***of a high-risk AI system*** for the purposes of this Regulation and shall be subject to the obligations of the provider under Article 16, in any of the following circumstances:

(a) they ***put their name or trademark on a high-risk AI system already placed*** on the market or put into service;

(b) they ***make a substantial modification to*** a high-risk AI system ***that has*** already ***been*** placed on the market or ***has already been*** put into service ***and in a way that it remains a high-risk AI system in accordance with Article 6***;

(ba) they make a substantial modification to an AI system, including a general purpose AI system, which has not been classified as high-risk and has already been placed on the market or put into service in such manner that the AI system becomes a high risk AI system in accordance with Article 6.

2. Where the circumstances referred to in paragraph 1, point ***(a) to (ba)*** occur, the provider that initially placed the AI system on the market or put it into service shall no longer be considered a provider ***of that specific AI system*** for the purposes of this Regulation. ***This former provider shall provide the new provider with the technical documentation and all other relevant and reasonably expected information capabilities of the AI system, technical access or other assistance based on the generally acknowledged state of the art that are required for the fulfilment of the obligations set out in this Regulation***.

Paragraph 2 shall also apply to providers of foundation models as defined in Article 3 when the foundation model is directly integrated in an high-risk AI system.

2a. The provider of a high risk AI system and the third party that supplies tools, services, components or processes that are used or integrated in the high risk AI system shall, by written agreement specify the information, capabilities, technical access, and or other assistance, based on the generally acknowledged state of the art, that the third party must provide in order to enable the provider of the high risk AI system to fully comply with the obligations under this Regulation.

2b. For the purposes of this Article, trade secrets shall be preserved and shall only be disclosed provided that all specific necessary measures pursuant to Directive (EU) 2016/943 are taken in advance to preserve their confidentiality, in particular with respect to third parties. Where necessary, appropriate technical and organizational arrangements can be agreed to protect intellectual property rights or trade secrets.

Recital 60g

Pretrained or retrained ("fine tuned") models developed for a narrower, less general, more limited set of applications that cannot be reasonably adapted for a wide range of tasks such as simple multipurpose AI systems should not be considered foundation models for the purposes of this Regulation, because of their greater interpretability which makes their behaviour less unpredictable.

8. Establish workable and balanced responsibilities for foundation models

The European Parliament introduced new regulatory requirements for Foundation Models, as a parallel compliance framework that would run partially parallel to that of high-risk AI systems. BSA is supportive of establishing balanced and workable responsibilities for foundation models, that would ensure the highest level of trust in Artificial Intelligence in Europe. At the same time, it is of paramount importance to ensure that these requirements are not overly burdensome, take into account a very diverse AI ecosystem and, chiefly, are workable for developers and deployers of foundation models.

The European Parliament proposal moves on two separate tracks in addressing foundation models. On one hand, requiring the provision of necessary information, should these models be integrated in a high-risk AI, and therefore form part of compliance requirements for a provider under the AI Act. Under this lens, the provision of information on the model's performance, technical documentation and identification of known risks would be workable requirements for developers. The Parliament proposal goes much further in some instances, requiring developers to describe reasonably foreseeable risks – a task close to impossible for many developers who would not be best placed to describe the risk profile of AI systems that are general in design – or requiring lifecycle-long level of performance – a similarly extremely complex task. On the other hand, the Parliament proposal also seeks to include very comprehensive fundamental rights and independent audit requirements that would often go much further for foundation models, than they do for high-risk AI in the AI Act. These would include requirements to assess for risks to rule

of law, democracy and the environment, independent experts auditing and quality management systems.

A key element in the development and deployment of foundation models is the degree of control that a developer may retain on the model. The AI ecosystem for foundation models is very diverse, and many business models are employed in this context. By far and large, developers of foundation models retain little to no control over further uses of the model, and therefore have no legal recourse or control over the model after it has been deployed by a third party. This is additionally problematic for Free or Open Source models, which by definition do not entail any control over the model after it has been made available. **BSA strongly recommends that language to reflect this diversity is included in the AI Act, to ensure that the compliance obligations for foundation models are allocated to the entities best placed to comply with them**. Moreover, developers of foundation models should be required to assess for known risks, and not for reasonably foreseeable risks. The Parliament proposal would establish very excessive obligations on foundation models developers, which would be required to assess for every possible risk, even those they may not be aware of. In many cases, the deployer would be better placed to carry out a more detailed risk assessment, depending on the specific context of use.

In a similar vein, obligations that would require a pervasive degree of control on downstream developer or deployers would be often unworkable for foundation models. Depending on the specific context of use, the original developer will often not retain contractual or legal control over the model, and would therefore have no access to the model to establish lifecycle-long performance monitoring, and especially quality management systems intended as defined by the AI Act. **BSA strongly recommends removing requirements that would oblige the developer to retain control over the model, often in contrast with contractual agreements or other legislation.** For example, forcing a developer to maintain a quality management may contrast GDPR – as it would mandate a degree of control over personal data that the developer may not have – cybersecurity and contractual agreements. Moreover, Free and Open Source models would be directly countering the terms of their functioning, as these business models entail no control – and often no contractual relationship – over the model after it has been made available.

The Parliament proposal also seeks to introduce very pervasive fundamental rights requirements, which are not required for high-risk AI, namely assessment for rule of law, democracy and environment. **BSA strongly recommends removing these requirements, as they are excessive for AI that is not considered high-risk.** In particular, requiring private actors to assess for rule of law and democracy, a complex task that is traditional left to the EU Institutions at the highest level, would be undesirable and excessively burdensome. Similarly, assessments for environmental impact are not best placed with the developer of a system, who is not aware of every possible placement after the model has been made available. Lastly, the requirement for independent experts, which was deemed excessive for high-risk AI, is equally excessive for foundation models. As the field of AI auditing is nascient, BSA strongly recommends against including such requirements, which would create overly burdensome obligations in a field with developing expertise.

A key element of the proposal should also be the safeguard of trade secrets and important development information. The Parliament proposal requires to share training resources, size and power, and extensive technical documentation. All this information may constitute a trade secret, and may be used to obtain proprietary information on the model. BSA

strongly recommends that these requirements are tempered with references to trade secrets and proprietary information.

The Parliament proposal also seeks to specifically address generative AI, in particular with reference to the data used to train the model. In first instance, BSA recommends including a clear distinction between generative AI systems provided to consumers and those provided in the context of business-to-business relations. Paragraph 4 adds specific requirements that are seem to seek to address challenges possibly posed by consumer-facing AI systems, without providing any distinction for very different uses in B2B settings. Given the significant differences in circumstances between a generative AI system used in an industrial setting and one used in consumer-facing cases, BSA recommends adding language that would clarify that the requirements set in Paragraph 4 would only apply to the latter. Furthermore, the requirement to provide an sufficiently detailed summary of any possible copyrighted material used to train the model is extremely burdensome, as it would require developers to prepare immensely detailed documents, sometimes on extremely large datasets. Given the recently approved Copyright Directive, which clearly allows for text and data mining activities, as long as materials are legally accessed, it remains unclear how this addition would measurably add to the already existing legislation, except for creating an unreasonable burden on developers and deployers, as the model is further customized and trained. **BSA strongly suggests removing this language, to ensure full harmonization of the AI Act with existing EU legislation.**

Recommended Trilogue amendments - retain Parliament proposal with additional amendments

Article 28b (Obligations of the provider of a foundation model)

1. A provider of a foundation model shall, prior to making it available on the market or putting it into service, ensure that it is compliant with the requirements set out in this Article, ***to the best of their ability, and taking into account the degree of control over other downstream providers and deployers and the specificities of the AI system.*** ~~regardless of whether it is provided as a standalone model or embedded in an AI system or a product, or provided under free and open source licences, as a service, as well as other distribution channels~~.

2. For the purpose of paragraph 1, the provider of a foundation model shall:

(a) demonstrate through appropriate design, testing and analysis that the identification, the reduction and mitigation of ***identified risks*** ~~reasonably foreseeable risks~~ to health, safety, fundamental rights, ~~the environment and democracy and the rule of law~~ prior and throughout development with appropriate methods ~~such as with the involvement of independent experts~~, as well as the documentation of remaining ~~non-mitigable~~ ***known*** risks after development;

(b) ~~process and incorporate only datasets that are subject to appropriate data governance measures for foundation models, in particular measures to examine the suitability of the data sources and possible biases and appropriate mitigation~~ ***employ processes to address unsuitability of data sources or possible biases in the training data when appropriate***;

c) design and develop the foundation model ~~in order~~ ***to the best of their ability*** to ~~achieve~~ ***aim for*** ~~throughout its lifecycle~~ appropriate levels of performance, predictability, interpretability, ~~corrigibility~~, safety and cybersecurity assessed through appropriate methods such as ~~model evaluation with the involvement of independent experts,~~ documented analysis, and extensive testing during conceptualisation, design, and development;

(d) design and develop the foundation model, making use of applicable standards to reduce energy use, resource use and waste, as well as to increase energy efficiency, and the overall efficiency of the system. This shall be without prejudice to relevant existing Union and national law and this obligation shall not apply before the standards referred to in Article 40 are published. ~~They shall be designed with capabilities enabling the measurement and logging of the consumption of energy and resources, and, where technically feasible, other environmental impact the deployment and use of the systems may have over their entire lifecycle~~;

(e) draw up ~~extensive~~ ***appropriate*** technical documentation and intelligible instructions for use in order to enable the downstream providers to comply with their obligations pursuant to Articles 16 and 28.1., ***this requirement does not in any way constitute an obligation to share information or documentation that would constitute a trade secret***;

~~(f) establish a quality management system to ensure and document compliance with this Article, with the possibility to experiment in fulfilling this requirement,~~

(g) register that foundation model in the EU database referred to in Article 60, in accordance with the instructions outlined in Annex VIII paragraph C.

When fulfilling those requirements, the generally acknowledged state of the art shall be taken into account, including as reflected in relevant harmonised standards or common specifications, as well as the latest assessment and measurement methods, ~~reflected notably in benchmarking guidance and capabilities referred to in Article 58a (new)~~.

3. Providers of foundation models shall, for a period ending ~~10~~ ***7*** years after their foundation models have been placed on the market or put into service, keep the technical documentation referred to in paragraph 1(c) at the disposal of the national competent authorities;

4. Providers of foundation models used in ***high-risk*** AI systems specifically intended to ***be provided to consumers and*** generate, with varying levels of autonomy, content such as complex text, images, audio, or video ("generative AI") and providers who specialise a foundation model into a generative AI system ***used in a high-risk case as per Art. 6 of this Regulation and provide it to consumers***, shall in addition

a) comply with the transparency obligations outlined in Article 52 (1),

b) train, and where applicable, design and develop the foundation model ***to guard*** ~~in such a way as to ensure adequate safeguards~~ against the generation of content in breach of Union law in line with the generally

acknowledged state of the art, ~~and~~ without prejudice to fundamental rights, including the freedom of expression, ***and to the extent that their ability to address these risks is in their control***,

~~c) without prejudice to national or Union legislation on copyright, document and make publicly available a sufficiently detailed summary of the use of training data protected under copyright law.~~

Annex VIII – Information to be submitted upon the registration of High Risk Systems in accordance with Article 51

Section C - The following information shall be provided and thereafter kept up to date with regard to foundation models to be registered in accordance with Article 28b (e):

1. Name, address and contact details of the provider;

2. Where submission of information is carried out by another person on behalf of the provider, the name, address and contact details of that person;

3. Name, address and contact details of the authorised representative, where applicable;

4. Trade name and any additional unambiguous reference allowing the identification of the foundation model

5. ~~Description~~ ***Characterisation*** of the data sources used in the development of the foundational model

6. Description of the ***intended*** capabilities and ***known*** limitations of the foundation model, including the ~~reasonably foreseeable~~ ***known*** risks and the measures that have been taken to mitigate them as well as remaining ~~non-mitigated~~ ***known*** risks with an explanation on the reason why they cannot be mitigated

~~7. Description of the training resources used by the foundation model including computing power required, training time, and other relevant information related to the size and power of the model~~

8. Description of the model's performance~~, including on public benchmarks or state of the art industry benchmarks~~

9. Description of the results of relevant internal and external testing ~~and optimisation~~ of the model

10. Member States in which the foundation model is or has been placed on the market, put into service or made available in the Union;

11. URL for additional information (optional).

9. Design obligations for providers and deployers that support AI uptake and protect fundamental rights

The European Parliament Report introduced significant changes to the functioning of the AI Act, including additional responsibilities for deployers of AI. In particular, the Parliament

report introduced a requirement for consultation of workers every time a high-risk system is deployed in the workplace. **BSA would strongly caution against retaining this requirement.** In first instance, the added language does not clarify whether workers would need to be consulted every time a high-risk AI is deployed, or exclusively when that AI may affect them directly. For example, an AI that supports critical infrastructure management would necessitate workers consultation under this language, and every time there is a substantial modification, such obligation would be triggered again. If the obligation would instead apply only when the AI would affect directly workers, the Parliament Report already provides for transparency requirements and communication obligations – as does other EU legislation such as GDPR – elsewhere. This requirement may prove to be either excessively burdensome, or a repetition of existing obligations.

The European Parliament has also introduced a Fundamental Rights Impact Assessment, that would require deployers to carry out an assessment of the impact of high-risk AI. This requirement is also largely overlapping with existing legislation, such as GDPR and other sectoral legislation, and is likely to duplicate and complicate compliance obligations. **BSA recommends a more thorough analysis of the duplication of obligations brought about by a Fundamental Rights Impact Assessment, and consider deletion.**

Art. 29 (Obligations of deployers)

~~5a. Prior to putting into service or use a high-risk AI system at the workplace, deployers shall consult workers representatives with a view to reaching an agreement and inform the affected employees that they will be subject to the system.~~

10. Establish clear obligations to ensure content authenticity

The original Commission proposal included specific language in Art. 52 that would establish a series of obligations to ensure content authenticity. While the Commission language was a step in the right direction, and similarly the Council General Approach improved on the original proposal, the European Parliament further improves the Commission proposal adding necessary requirements to clarify what would constitute inauthentic content, and how AI developers and deployers can disclose when content has been manipulated or generated through an AI system. BSA recommends ensuring that the definition of deep fake is in line with the obligations set forth by Art. 52(3), including the labeling of content. The current definition would seek to include “text” in the scope of this Article, which would prove a significant burden for developers and deployers, as it is not clear when text would be construed as a deep fake, and whether all text that is generated with, or supported by, an AI would then have to be labeled. For these reasons, BSA recommends removing the reference to text in the definition of a deep fake. Additionally, BSA would also recommends partially amending the language of Paragraph 3, as it refers first to “artificially generated or manipulated content”, and in the subsequent explanation of disclosure to the public to “inauthentic content”. BSA would recommend, for the sake of consistency and for ensuring that this proposal applies to all artificially generated and manipulated content, to refer exclusively to the “artificially generated or manipulated content” and not introducing a new concept such as “inauthentic content”.

BSA and its Members remain at the forefront of content authenticity, and strongly recommend that the co-legislators adopt the language of the European Parliament Report, with the suggested changes.

Recommended Trilogue amendments - retain Parliament proposal with additional amendments

Article 52 (Transparency obligations)

3. Users of an AI system that generates or manipulates ~~text~~, audio or visual content that would falsely appear to be authentic or truthful and which features depictions of people appearing to say or do things they did not say or do, without their consent ('deep fake'), shall disclose in an appropriate, timely, clear and visible manner, including to the recipient of such content, that the content has been artificially generated or manipulated, as well as, whenever possible, the name of the natural or legal person that generated or manipulated it. ~~Disclosure shall mean labelling the content in a way that informs that the content is inauthentic and that is clearly visible for the recipient of that content~~. To label the content, users shall take into account the generally acknowledged state of the art and relevant harmonised standards and specifications.

3a. Paragraph 3 shall not apply where the use of an AI system that generates or manipulates text, audio or visual content is authorized by law or if it is necessary for the exercise of the right to freedom of expression and the right to freedom of the arts and sciences guaranteed in the Charter of Fundamental Rights of the EU, and subject to appropriate safeguards for the rights and freedoms of third parties. Where the content forms part of an evidently creative, satirical, artistic or fictional cinematographic, video games visuals and analogous work or programme, transparency obligations set out in paragraph 3 are limited to disclosing of the existence of such generated or manipulated content in an appropriate clear and visible manner that does not hamper the display of the work and disclosing the applicable copyrights, where relevant. It shall also not prevent law enforcement authorities from using AI systems intended to detect deep fakes and prevent, investigate and prosecute criminal offences linked with their use.

3b. The information referred to in paragraphs 1 to 3 shall be provided to the natural persons at the latest at the time of the first interaction or exposure. It shall be accessible to vulnerable persons, such as persons with disabilities or children, complete, where relevant and appropriate, with intervention or flagging procedures for the exposed natural person taking into account the generally acknowledged state of the art and relevant harmonized standards and common specifications.

For further information, please contact:

Matteo Quattrocchi,
Director, Policy – EMEA
matteoq@bsa.org