

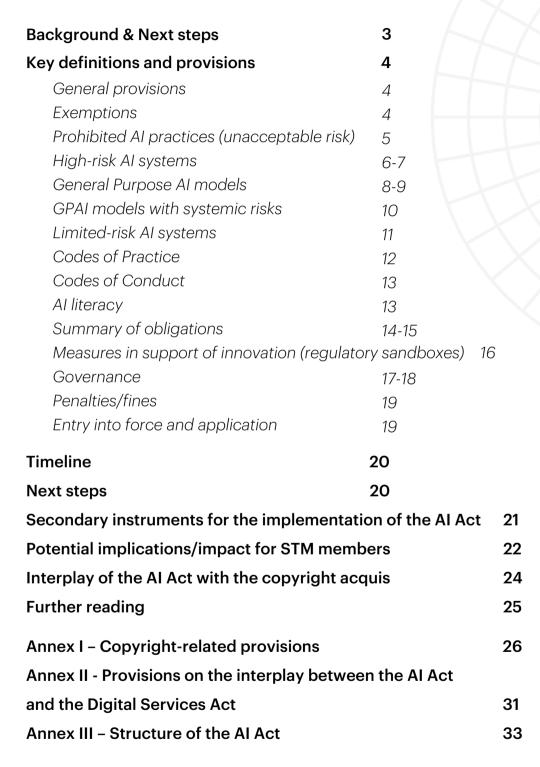


## STM MEMO ON THE EU AI ACT

August 2024 2nd edition

www.stm-assoc.org

## **CONTENTS**







The <u>proposal for an Artificial Intelligence Act</u> (Al Act) was first introduced by the European Commission (EU COM) on 21 April 2021. It is a Regulation, so it will be directly applicable in all Member States with no need for a national transposition.

The Council of the EU adopted their <u>General Approach</u> on 6 December 2022, just as ChatGPT was being released to the public. The exponential increase in awareness and interest in generative AI tools prompted the EU Parliament to prolong its internal discussions to include a proposed approach to regulating foundation models in their <u>negotiating position</u>, which was adopted on 14 June 2023.

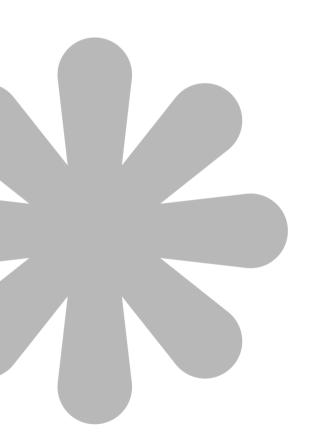
Trilogue negotiations to reconcile the EU COM, the Council and the EU Parliament's versions took place in the second half of 2023 under huge political pressure.

A provisional deal was announced on 8 December 2023 and was green lighted by EU ambassadors on 2 February 2024. The EU Parliament adopted it in its Committees on 13 February and in plenary session on 13 March, with a final corrigendum approved on 19 April.

The AI Act was published as **Regulation (EU)** 2024/1689 on the Official Journal on 13 June; the official text is here.

## **NEXT STEPS**

The AI Act enters into force 20 days after its publication on the Official Journal. The official countdown to the application of the new rules has started. The EU Commission and the AI Office will be working to produce the guidelines and Codes required by the Act.





## **KEY DEFINITIONS AND PROVISIONS**

## **General provisions**

The definition of AI uses the same language of the revised version of the OECD's definition of AI [1], only with a different ordering of some of the verbiage:

Art 3(1) An AI system is a machine-based system designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments.

The AI Act has extra-territorial effect and applies to all organisations that make AI systems available in the EU market, regardless of where they are established.

### **Exemptions**

The AI Act doesn't apply to:

- Al systems that have been developed exclusively for military and defence uses [Recital 24; Article 2(3)];
- Al systems used for the sole purpose of research and innovation (before the placing on the market/putting into service) [Recital 25, 97, 105, 109; Article 2() and (8)]; and
- Al models/systems provided under free and open-source licences [2]. These are always bound to respect copyright-relevant provisions; otherwise, the Al Act doesn't apply to them unless they are high-risk Al systems or if they entail systemic risks, in which case they are not exempt from the respective obligations [Recitals 89, 102, 103, 104; Art. 2(12), Art. 53(2), Art. 54(5)].

<sup>[1] &</sup>quot;An Al system is a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that [can] influence physical or virtual environments. Different Al systems vary in their levels of autonomy and adaptiveness after deployment.". Available at: <a href="https://oecd.ai/en/ai-principles">https://oecd.ai/en/ai-principles</a> (retrieved on 28/03/2024).

<sup>[2]</sup> When the parameters of these AI models, which include the weights, information on the model architecture, and details on model usage, are publicly available, and such models are not made available against a price or otherwise monetised (e.g. through data collection).



## **KEY DEFINITIONS AND PROVISIONS**

con't

## Prohibited AI practices (unacceptable risk)

These are listed under Chapter II, Article 5.
The following uses have been identified as yielding unacceptable risk and will thus be banned under the AI Act:

- Facial recognition in public spaces (with exemptions for law enforcement; if pre-authorisation has been granted by a court, and only for specific crimes e.g., terrorism, human trafficking);
- Bulk scraping of facial images from the internet or CCTV footage to create facial recognition databases (e.g., such as Clearview AI):
- Biometric categorisation based on sensitive characteristics;
- Emotion recognition in the workplace and educational institutions;
- Social scoring;
- Al systems that manipulate humans into behaviours that may cause physical or psychological harm;
- Al that is used to exploit vulnerabilities (e.g. age, disability);
   and
- Predictive policing, unless it is used with clear human assessment and objective facts.



## **KEY DEFINITIONS AND PROVISIONS**

con't

### **High-risk AI systems**

High risk systems and related obligations are listed under Chapter III, Articles 6 to 39.

Some areas where AI applications could have significant adverse societal impact have been identified as high-risk:

- Medical devices;
- Vehicles:
- Employment;
- Education and vocational training;
  - Al systems intended to be used to determine access or admission or to assign natural persons to educational and vocational training institutions at all levels;
  - Al systems intended to be used to evaluate learning outcomes, including when those outcomes are used to steer the learning process of natural persons in educational and vocational training institutions at all levels;
  - Al systems intended to be used for the purpose of assessing the appropriate level of education that an individual will receive or will be able to access, in the context of or within educational and vocational training institutions;
  - Al systems intended to be used for monitoring and detecting prohibited behaviour of students during tests in the context of or within educational and vocational training institutions.
- Influencing elections and voters;
- Access to services (e.g., insurance, banking, credit, benefits, etc.);
- Critical infrastructure management (e.g., water, gas, electricity);
- Emotion recognition systems;
- · Biometric identification;
- Law enforcement, border control, migration and asylum;
- · Administration of justice;
- Specific products and/or safety component of specific products.





Al systems falling in these categories will be authorised but subject to requirements including data governance, transparency, human oversight, fundamental rights impact assessment and conformity assessment, implementation of risk management and quality management systems, accuracy and robustness through testing, registration in a public EU database, and post-market monitoring by surveillance authorities.

To avoid a blanket categorisation, a **filtering system** was added so that even if a system falls in one of the abovelisted areas, it loses its classification if one of the following conditions applies:

- The AI system is intended to perform a "narrow procedural task";
- The Al system is meant to review or improve the result of a previously completed human activity;
- The AI system is purely intended to detect decision making patterns or deviations from prior decision making patterns;
- The AI model is used to perform only preparatory tasks to an assessment relevant to the critical use cases.

Because of this additional filter, the applications falling into the high-risk category will be in the lower range of the number initially forecast by the EU Commission, which was in between 5 to 15% [3].

The right was introduced for natural or legal persons to lodge complaints to the relevant market surveillance authority [Article 85] concerning non-compliance with the AI Act and to receive explanations about decisions based on high-risk AI-systems that affect their rights [Article 86]. It will be interesting to watch how any potential overlap between this provision, the General Data Protection Regulation (GDPR), the Data Act and the upcoming rules on AI liability will play out.

<sup>[3]</sup> European Commission, SWD(2021) 85 final, Executive Summary of the Impact Assessment accompanying the Al Act, available at <a href="https://digital-strategy.ec.europa.eu/en/library/impact-assessment-regulation-artificial-intelligence">https://digital-strategy.ec.europa.eu/en/library/impact-assessment-regulation-artificial-intelligence</a>.





## Other definitions relevant in the Alcontext are as follows:

Art 3(29) 'training data' means data used for training an Al system through fitting its learnable parameters;

Art 3 (33) 'input data' means data provided to or directly acquired by an Al system on the basis of which the system produces an output; Art 3(60) 'deep fake' means Algenerated or manipulated image, audio or video content that resembles existing persons, objects, places or other entities or events and would falsely appear to a person to be authentic or truthful; Recital 99: Large generative Al models are a typical example for a general-purpose Al model, given that they allow for flexible generation of content, such as in the form of text, audio, images or video, that can readily accommodate a wide range of distinctive tasks.

### **General Purpose AI models**

In the final rounds of the negotiations, lawmakers agreed to add a stand-alone chapter [Chapter V in the final version] on top of the risk categories to regulate general purpose AI (GPAI) models and systems, defined as follows:

Art 3 (63) 'general purpose Al model' means an Al model, including when trained with a large amount of data using self-supervision at scale, that displays significant generality and is capable to competently perform a wide range of distinct tasks regardless of the way the model is placed on the market and that can be integrated into a variety of downstream systems or applications. This does not cover Al models that are used before release on the market for research, development and prototyping activities.

**Art 3 (66)** 'general-purpose AI system' means an AI system which is based on a general-purpose AI model, that has the capability to serve a variety of purposes, both for direct use as well as for integration in other AI systems;

All GPAI models and systems built on top of them will have to meet some basic requirements when they're put on the market [Article 53]; this will also allow those who licence or run a GPAI and build systems/applications on top of it (downstream providers) to be able to meet obligations themselves. Those requirements include:

- a) drawing up technical documentation,
- b) making available information and documentation to downstream providers,
- c) putting in place a policy to respect Union copyright law in particular to identify and respect, including through state-of-the-art technologies where applicable, the reservations of rights for text and data mining expressed pursuant to Article 4(3) of the DSM Directive [4], d) drawing up and making publicly available a sufficiently detailed summary about the content used for training of the general-purpose AI model, according to a template provided by the AI Office.

[4] <u>Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market</u> (DSM Directive)



Provisions c) and d) are a potentially great step forward for rightsholders including academic and scholarly publishers. Whilst no new rights are introduced and publishers can still rely on existing copyright rules, these provisions enable the enforcement of such rights, in particular the right to reserve rights on content to potentially prevent unauthorised extraction and use for text and data mining [5].

The shape that the "template provided by the Al Office" will take will be fundamental in determining how the implementation of the new rules will move forward, so that will be an engagement priority for STM.

[5] As defined under Art 2(2) of the Directive EU 2019/790 for Copyright in the Digital Single Market: "text and data mining' means any automated analytical technique aimed at analysing text and data in digital form in order to generate information which includes but is not limited to patterns, trends and correlations;"

The language of the Recitals provides more information about the characteristics of the template. In particular, Recital 107 reads that the "summary should be generally comprehensive in its scope instead of technically detailed to facilitate parties with legitimate interests, including copyright holders, to exercise and enforce their rights under Union law [...] and by providing a narrative explanation about other data sources used".

Mention is made of "trade secrets and confidential business information", which can be expected to be used as pretext by AI companies not to release information about training data to the level of granularity that rightsholders would like. Here, another portion of the Recital helps as it clearly links the summary to the possibility for copyright holders to "exercise and enforce their rights under Union law". It is however going to be difficult to require an itemised list through the template as Recital 108 specifies that the "AI Office should monitor whether the provider has fulfilled those obligations without verifying or proceeding to a work-by-work assessment of the training data in terms of copyright compliance".

For the integral text of the provisions described in this section, please consult Annex I.



### **GPAI** models with systemic risks

A stricter regime will apply to GPAI models posing "systemic risk" [Recital 110, 111, 112, 113, 114, 115; Articles 3(65), 51, 52, 55], i.e. those satisfying either or both these conditions:

- having high impact capabilities evaluated on the basis of appropriate technical tools and methodologies, including indicators and benchmarks;
- based on a decision of the Al Office, ex officio or following a qualified alert by the scientific panel.

In any case, models that were trained with computing power above 10^25 FLOPs (Floating-Point Operations, roughly equivalent to calculation steps) [which currently would only cover OpenAl's GPT-4 and potentially Alphabet's Gemini] are going to be considered as having systemic risk.

The provider should notify the AI Office at the latest two weeks after the requirements are met or it becomes known that a GPAI model will meet the requirements. Thus, similarly to the list of Very Large Online Platforms and Search Engines under the Digital Services Act (DSA), the AI Office will be a closed list of GPAI models posing systemic risk.





Recital 110 clarifies that systemic risk could also include "the dissemination of illegal, false, or discriminatory content".

The EU Commission (or the AI Office, depending on the exact division of tasks, tbc) will be empowered to adapt the thresholds listed in the paragraphs above, as well as to supplement benchmarks and indicators in light of evolving technological developments.

Additional requirements for the models with systemic risks include:

- Performing model evaluation;
- Systemic risk assessment and mitigation;
- Monitoring and reporting serious incidents to the EU COM and national authorities:
- Conducting and documenting adversarial testing of the model (so-called "red-teaming");
- Ensuring an adequate level of cybersecurity protection;
- Monitoring and reporting energy consumption.



The Al Office will encourage and facilitate the drawing up of codes of practice to facilitate the implementation of the obligations regarding the detection and labelling of artificially generated or manipulated content.

## **Limited-risk AI systems**

All systems escaping the categories described above – effectively the majority of Al uses, such as recommender systems and spam filters – won't be imposed any obligation except for some **minimal transparency obligations** [Recital 133, 134; Article 50] in specific cases:

- 1. Inform natural persons that they're interacting with an Al system, when those are intended to interact directly with natural persons.
  - This obligation is with providers
  - Unless this is obvious from the point of view of a natural person who is reasonably well-informed, observant and circumspect, taking into account the circumstances and the context of use.
- 2. Watermark in machine readable format Al-generated or manipulated output (audio, image, video or text)
  - This obligation is with <u>providers</u> of AI systems including GPAI to design systems in a way that their output is marked accordingly
  - Unless AI systems perform an assistive function for standard editing or do not substantially alter the input data provided by the deployer or the semantics thereof
- 3. Clearly disclose Al-manipulation (e.g., through label or disclaimer)
  - This obligation is with <u>deployers</u> who use an Al system to generate or manipulate:
  - a. <u>Image</u>, <u>audio</u> or <u>video</u> content constituting a deep fake
    - Where the content is part of an evidently creative, satirical, artistic or fictional work or programme, the disclosure can happen in an appropriate manner that does not hamper the display or enjoyment of the work.
  - b. <u>Text</u> which is published with the purpose of informing the public on matters of public interest
    - Where the Al-generated content has undergone a process of human review or editorial control and where a natural or legal person holds editorial responsibility for the publication of the content, the obligation does not apply.



#### **Codes of Practice**

The AI Office (part of the EU COM, more about it below) will be responsible for facilitating the drawing up Codes of Practice, to be drawn up by providers of GPAI models, with national competent authorities contributing to the process.

GPAI providers will effectively be the ones drafting, and this will become a sort of "industry code". Civil society organisations and other relevant stakeholders may have an opportunity to support the process, and the AI Office will have to guarantee that the Codes reflect the state of the art and take into account international approaches. The Codes should include key performance indicators and regular reporting on the implementation of the commitments.

Providers of GPAI models with systemic risk may rely on Codes of Practices to comply with requirements until harmonised standards are established; at that point, the compliance with a European harmonised standard will grant providers the presumption of conformity instead.

To enable providers to demonstrate compliance in time, Codes of Practice should be ready at the latest 3 months before the entry into application of the relevant provisions [Recital 179], so that means within 9 months after the entry into force for Codes of Practice bridging compliance with GPAI provisions (12 minus 3 months), within 21 months (24 minus 3 months) for the other provisions, and within 33 months for some high-risk system requirements (36 minus 3 months).

The AI Office may also encourage and facilitate the drawing up of Codes to facilitate the effective implementation of the obligations regarding the detection and labelling of artificially generated or manipulated content [Recital 135; Article 50(7)]. The EU COM will then have the power to approve these Codes through implementing acts.





#### **Codes of Conduct**

Codes of Conduct [Recital 165; Article 95] will be used as voluntary instruments to drive the take up of requirements and principles identified by the AI Act by those actors who are not legally bound to respect them. For example, Codes of conduct may be drawn up to advance AI literacy among persons dealing with the development, operation and use of AI, or to foster the voluntary application of some/all the high-risk AI requirements by non-high-risk systems, e.g., governance mechanisms.

They will be developed with the involvement of relevant stakeholders, and the EU COM may start initiatives, including for specific sectors, to facilitate the lowering of technical barriers to exchange of data for AI development across borders.

The EU COM should evaluate the impact and effectiveness of voluntary Codes of Conduct to foster the application of the requirements within two years after the entry into application and every three years thereafter.

### **Al literacy**

Per Article 4, all providers and deployers of AI systems will have to take measures to ensure, to the best of their abilities, a sufficient level of AI literacy of their staff and other persons dealing with the operation and use of AI systems on their behalf.

This also applies to publishing houses; when making available Al models and systems to their employees, they will have a duty to instruct them on how to use them.

## **Summary of obligations**



In summary, here is what obligations will cover what type of AI model/system. Please note that this is not legal advice and is meant to provide a high-level overview. Some use cases may not be captured by the categories laid out below:



Article	Obligation	High- Risk Al	General Purpose AI with systemic risk	General Purpose Al	Low- Risk Al
Art. 9 + Art. 17	Risk + quality management	×			
Art. 10	Data governance (including data collection, preparation, de-biasing, etc.)	×			
Art. 11 + Art. 18	Draw up & keep up-to- date technical documentation + documentation keeping	×	×	(Art. 53)	
Art. 12 + Art. 19	Record-keeping + keeping the automatically generated logs	×			
Art. 13	Transparency (understandability)	×			
Art. 14	Human oversight	×			
Art. 15	Accuracy, robustness & cybersecurity (through testing)	×	(Art. 55d)		
Art. 20	Corrective actions & duty of information in case of non-conformity	×			
Art. 21	Cooperation with competent authorities	×	×	×	
Art. 22	Authorised representatives	×	×	(Art. 54)	
Art. 25	Make available the necessary information downstream	×	×	Art. 53(1)b	



Article	Obligation	High- Risk Al	General Purpose Al with systemic risk	General Purpose Al	Low- Risk Al
Art. 27	Fundamental rights impact assessment (only for public sector)	×			
Art. 40	Harmonised standards & standardisation deliverables	×			
Art. 43	Conformity assessment	×			
Art. 47 + Art. 48	EU declaration + marking of conformity				
Art. 49 + Art.71	Registration in public EU database	×			
Art. 50(1)	Informing users they're interacting with AI	×	×	×	×
Art. 50(2)	Watermarking Al- generated content	×	×	×	×
Art. 53 (1)(c)	Policy to respect copyright including Art4 of the DSM Directive		×	×	
Art. 53 (1)(d)	Sufficiently detailed summary of the content used for training		×	×	
Art. 55 (1)(a)	Model evaluation (standardised protocols, including adversarial testing)		×		
Art. 55 (1)(b)	Systemic risk assessment & mitigation		×		
Art. 73	Reporting serious incidents to EU COM/competent authorities	×	X Art. 55(1)c		
Art. 56	Codes of Practice to prove compliance	×	×	×	
Art. 72	Post-market monitoring	×	Art. 75	Art. 75	

## Measures in support of innovation (regulatory sandboxes)

Regulatory sandboxes [Articles 3(55); 57, 58 and 59] will facilitate responsible innovation and the development of compliant AI systems, by enabling the establishment by national authorities of a controlled environment for the development, testing and validation of innovative AI systems. This should include testing in real-world conditions to train, learn and check a model's behaviour before placing it on the market or putting it into service.

Al operators will be able to apply to participate in the relevant sandbox and the national authorities will select them against certain criteria; operators will have to sign up to and respect terms and conditions. In the course of their participation in the sandbox, no penalties are foreseen for breach of the Al Act requirements, but participants will maintain liability for any damage caused.





#### Governance

The AI Act is a first in terms of provisions and obligations imposed on Almodels and systems, but also in terms of setting up an architecture fit for continuous oversight and adaptation of such rules. The main components of the governance structure are the following:

Al Office [Article 64]: with an EU COM <u>Decision</u> adopted on 24
January and entering into force on 21 February 2024, the Al
Office was established as part of the administrative structure
of DG CONNECT; 20 staffers will come from DG CONNECT
itself, and other 80 will be employed in between temporary
staff members and seconded national experts.

The Office is tasked with oversight of general-purpose Al models and their advancements, and with developing tools to assess their capabilities; it will be coordinating compliance, implementation, and enforcement.

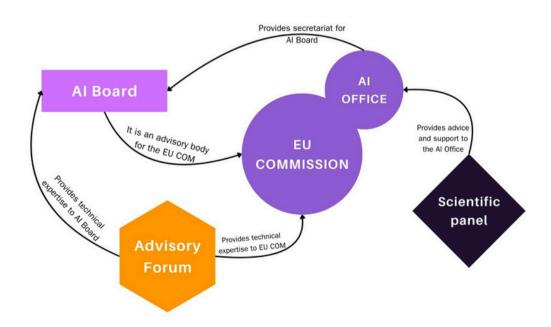
The Office will be empowered to conduct ex-post evaluations on GPAI models to assess their compliance, investigate systemic risks at EU level, and request GPAI model providers to take appropriate measures to comply with their obligations or to implement mitigation measures where there are substantiated concerns of systemic risk at EU level, including restricting availability of the model or withdrawing it from the market.

They will be contributing to developing and encouraging the adoption of Codes of Practice and standards. They are explicitly tasked to cooperate with stakeholders (Article 4 of the Decision) by establishing fora for cooperation with providers of Al models/systems, consulting stakeholders and establishing a forum for cooperation with the open-source community. The Al Office will also have a "global vocation" and contribute to international initiatives.

They will have a duty to cooperate with and conduct regular consultation of stakeholders, so this will be an important engagement target for STM to build connections with.

## Governance con't

- Al Board [Article 65]: it will be made up of member states' representatives (one per EU country appointed for three years); it will serve as coordination for Member States, advisory body to the EU COM and will play a key role, together with and facilitating the Al Office, in the implementation and oversight of GPAI rules. It will have the possibility to issue recommendations and written opinions on the implementation of the legislation. It will have two standing subgroups with market surveillance and notifying authorities, and its secretariat will be provided by the Al Office.
- Advisory Forum for stakeholders [Article 67]: made up of industry representatives, SMEs, start-ups, civil society and academia to provide technical expertise to the Board and Al Office. It will have two co-chairs with a two-year term, renewable once, and it will meet twice per year. It may be interesting for STM and/or members to be represented on the Forum.
- Scientific panel of independent experts [Article 68]: the EU COM will better specify provisions on the establishment of the scientific panel through an implementing act. The panel will advise the AI Office, especially around GPAI models, will contribute to developing methodologies and identifying relevant benchmarks for evaluating the models' capabilities and advise on classifying models with systemic risks.





### Penalties/fines

The fines for non-compliance [Article 99, 101] are quite significant and tiered as follows, depending on the severity of the infringement and the size of the company (with lower fine caps for SMEs and start-ups):

- Failure to provide accurate information: Up to 1.5% of the company's annual worldwide turnover or EUR 7.5 million.
- Failure of system/model providers and GPAI providers to comply with obligations: Up to 3% of the company's annual worldwide turnover or EUR 15 million.
- Prohibited Al practices & non-compliance with data requirements: Up to 7% of the company's annual worldwide turnover or EUR 35 million.

Non-complying systems will (after a grace period) be barred from the EU market.

## **Entry into force and application**

The AI Act will enter into force 20 days after its publication on the Official Journal of the EU and will become applicable as follows:

- Prohibitions will apply after 6 months;
- GPAI rules will apply after 12 months;
- Most provisions will apply after 24 months;
- Some requirements for high-risk systems will not apply until 36 months after official publication.

To bridge the transitional period before the Regulation becomes generally applicable, the EU COM aas launched Pactching an Al Pact. It will convene Al developers from Europe and around the world who commit on a voluntary basis to implement key obligations of the Al Act ahead of the legal deadlines.

The EU Commission published a call for interest for organisations that want to get actively involved in the AI Pact. As a next step, they will bring together interested parties in the first half of 2024 to discuss the ambitions of the Pact and collect preliminary ideas and best practices which could inspire future pledges. Following the formal adoption of the AI Act, the AI Pact will be officially launched and "frontrunner" organisations will be invited to make their first pledges public. Organisations, including STM members, will be able to find more information and express their interest at this page.

To promote rules on trustworthy AI at international level, the EU will continue to work in fora such as the G7 (e.g., under the Hiroshima AI process on International Guiding Principles and voluntary Code of Conduct for Advanced AI systems), the OECD, the Council of Europe, the G2O and the United Nations (UN). Just recently, we supported the agreement by G7 leaders.



### TIMELINE

With 1 August 2024 as date of entry into force of the AI Act, here's a timeline for the staggered application of the AI Act and the release of Codes of Practice.

It is important to note that a special regime [Article 111(3)] applies for those models that are already on the market at the time of entry into force of the AI Act. Providers of GPAI models that have been placed on the market 12 months from the date of entry into force of the AI Act are expected to comply with the new obligations laid down in this Regulation by 36 months from the date of entry into force.

The EU COM will have to evaluate and review this Regulation by three years after the date of entry into application and every four years thereafter and report to the European Parliament and the Council. The scope of the transparency obligations will have to be reviewed within two years after entry into application and every four years thereafter.

#### 1 August 2024:

entry into force

## (or earlier) 2 May 2025:

release Code of Practice bridging compliance with GPAI provisions (within 9 months)

#### (or earlier) 2 May 2026:

release Code of Practice for most AI Act provisions (within 21 months)

#### 2 June 2027:

release Code of Practice for some high-risk system requirements (within 33 months)

#### 2 August 2027:

application of obligations to GPAI models already on the market by 2 August 2025.

#### 2 February 2025:

application of prohibitions (after 6 months)

#### 2 August 2025

application of GPAI provisions (after 12 months)

#### 2 August 2026

application of most AI provisions (after 24 months)

#### 2 August 2027:

application of some high-risk system requirements (after 36 months)

## SECONDARY INSTRUMENTS FOR THE IMPLEMENTATION OF THE ALACT

**Delegated acts** are used, typically, when legislative acts have to be (regularly) adapted to take account of technical and scientific progress. The EU COM is empowered to adopt delegated acts [Recital 173] to update e.g., the following:

- · Definition of Al system;
- Criteria and use cases for high-risk Al;
- · Threshold for GPAI with systemic risk;
- Technical documentation requirements for GPAI;
- Transparency information for providers of GPAI;
- · Conformity assessments;
- EU declaration of conformity.

**Implementing acts** are adopted where uniform conditions for implementation are needed. The EU COM can adopt implementing acts for e.g., the following:

- Establish modalities and criteria to set up regulatory sandboxes;
- Make provisions for the establishment of the scientific panel of independent experts;
- To adopt Codes of Practice;
- To establish modalities and conditions of the evaluations of GPAI models by the AI Office;
- To establish modalities and practical arrangements for the fine proceedings for GPAI.

Guidelines [Article 96] consist of explanatory notes and other documents produced by the Commission services in order to provide practical and informal guidance about how particular provisions should be applied. The EU COM can issue guidelines on e.g., the following:

- High-risk Al requirements and obligations;
- · Prohibited practices;
- Responsibilities along the value chain (Article 25);
- Practical implementation of the provisions related to substantial modification;
- Practical implementation of transparency obligations (Article 50);
- Relationship of the AI Act and its enforcement with other EU law;
- Application of the definition of an Al system.

The Al Office is expected to play a key role in e.g.,

- Codes of Practice to cover obligations for providers of GPAI;
- Templates, e.g., for sufficiently detailed summary to provide information about data used to train GPAI models; for a questionnaire to facilitate carrying out a fundamental rights impact assessment.



## POTENTIAL IMPLICATIONS/IMPACT FOR STM MEMBERS

Please note that the following is not legal advice; this aims to inform members of STM's analysis of the AI Act as of the date of this memorandum. STM does not, by way of this document, cover the myriad use cases and circumstances that STM members will encounter with respect to compliance with the EU AI Act and encourage members to review the text of the AI Act and/or to seek the advice of counsel for specific questions. It is recommended that individual members review and assess the AI models they develop, licence, distribute and/or put in use.

We anticipate that the STM membership's use of Al wouldn't fall under prohibited uses nor, for the vast majority, under high-risk uses (especially for applications for educational and vocational purposes); however, please ensure to conduct a thorough review to confirm that is the case.

The most relevant rules will most typically be the ones concerning general purpose AI (GPAI), both a) vis-à-vis providers insofar as taking steps to ensure that the providers' GPAI training and refinement activities are copyright-compliant and ensuring that publishers' content is used lawfully, properly licenced and cited; and b) internally when building a house's own model or when building an application on top of a licenced/open source one.

With respect to a), where **STM members play the role of content providers**, it is recommended that they consider ways to adopt content protection technologies and reserve text and data mining rights for commercial purposes, as foreseen by Article 4 of the DSM Directive. STM has formulated <u>separate guidelines</u> on this matter.

With respect to b), where STM members play the role of deployers/providers of Al models and systems, it is recommended that they consider duties and opportunities regarding the following:

- Be transparent about the data used in training and/or finetuning a GPAI model/system;
- Design/deploy models to ensure that Al-generated or manipulated content is marked in a machine-readable format (unless Al systems perform an assistive function for standard editing or do not substantially alter the input data provided by the deployer or the semantics thereof);
- Disclose Al-manipulation (e.g., through label or disclaimer) on text, unless the Al-generated content has undergone a process of human review or editorial control and where a natural or legal person holds editorial responsibility for the publication of the content.
- Make sure to inform customers/readers when they're interacting with an AI system (e.g., an AI-powered chatbot);
- Put measures in place to guarantee the Al literacy of staff;
- Consider signing up to the Al Pact to fast-track compliance;
- Take into account the following ethical principles in the design and use of AI models, and in the establishment of internal AI principles/guidelines: human agency and oversight; technical robustness and safety; privacy and data governance; transparency; diversity, non-discrimination and fairness; societal and environmental well-being and accountability.

## INTERPLAY OF THE AI ACT WITH THE COPYRIGHT ACQUIS

Please be reminded that the AI Act is a horizontal piece of legislation that comes on top of and neither substitutes nor makes void any of the existing copyright acquis. Remedies for copyright infringement will still be available.

Please also note that no mention is made in the Al Act of licencing models.



## **FURTHER READING**

https://artificialintelligenceact.eu/ai-act-explorer/

https://ai-act-law.eu/

https://artificialintelligenceact.eu/assessment/eu-ai-act-compliance-checker/

https://crfm.stanford.edu/2023/11/18/tiers.html https://futureoflife.org/ai-policy/miles-apart/

https://iapp.org/media/pdf/resource\_center/eu\_ai\_act\_cheat\_sheet.pdf

https://iapp.org/media/pdf/resource\_center/eu\_ai\_act\_timeline.pdf

https://thefuturesociety.org/heavy-is-the-head-that-wears-the-crown/

https://thefuturesociety.org/a-blueprint-for-the-european-ai-office/

https://thefuturesociety.org/wp-content/uploads/2023/12/EU-AI-Act-Compliance-Analysis.pdf

https://papers.ssrn.com/sol3/papers.cfm?abstract\_id=4539452

https://papers.ssrn.com/sol3/papers.cfm?abstract\_id=4316944

https://barrysookman.com/2024/04/07/34308/

https://privacymatters.dlapiper.com/2024/04/europe-the-eu-ai-acts-relationship-with-data-protection-law-key-takeaways/

## ANNEX I - COPYRIGHT-RELATED PROVISIONS

## **Chapter V, General Purpose AI Models**

Section II – Obligations for providers of general-purpose AI models

#### Article 53(1)(c) and (d)

Obligations for providers of general-purpose AI models 1. Providers of general-purpose AI models shall:

[...]

- (c) put in place a policy to comply with Union copyright law, and in particular to identify and comply with, including through state of the art technologies, a reservation of rights expressed pursuant to Article 4(3) of Directive (EU) 2019/790;
- (d) draw up and make publicly available a sufficiently detailed summary about the content used for training of the general-purpose AI model, according to a template provided by the AI Office.

### **Accompanying Recitals**

(105) General-purpose models, in particular large generative models, capable of generating text, images, and other content, present unique innovation opportunities but also challenges to artists, authors, and other creators and the way their creative content is created, distributed, used and consumed. The development and training of such models require access to vast amounts of text, images, videos, and other data.



## Accompanying Recitals con't

Text and data mining techniques may be used extensively in this context for the retrieval and analysis of such content, which may be protected by copyright and related rights. Any use of copyright protected content requires the authorisation of the rightsholder concerned unless relevant copyright exceptions and limitations apply. Directive (EU) 2019/790 introduced exceptions and limitations allowing reproductions and extractions of works or other subject matter, for the purpose of text and data mining, under certain conditions.

Under these rules, rightsholders may choose to reserve their rights over their works or other subject matter to prevent text and data mining, unless this is done for the purposes of scientific research. Where the rights to opt out has been expressly reserved in an appropriate manner, providers of general-purpose AI models need to obtain an authorisation from rightsholders if they want to carry out text and data mining over such works.

(106) Providers that place general purpose AI models on the EU market should ensure compliance with the relevant obligations in this Regulation. For this purpose, providers of general purpose AI models should put in place a policy to respect Union law on copyright and related rights, in particular to identify and respect the reservations of rights expressed by rightholders pursuant to Article 4(3) of Directive (EU) 2019/790.

Any provider placing a general purpose AI model on the EU market should comply with this obligation, regardless of the jurisdiction in which the copyright-relevant acts underpinning the training of these general purpose AI models take place. This is necessary to ensure a level playing field among providers of general purpose AI models where no provider should be able to gain a competitive advantage in the EU market by applying lower copyright standards than those provided in the Union.

(107) In order to increase transparency on the data that is used in the pre-training and training of general-purpose AI models, including text and data protected by copyright law, it is adequate that providers of such models draw up and make publicly available a sufficiently detailed summary of the content used for training the general-purpose model.

While taking into due account the need to protect trade secrets and confidential business information, this summary should be generally comprehensive in its scope instead of technically detailed to facilitate parties with legitimate interests, including copyright holders, to exercise and enforce their rights under Union law, for example by listing the main data collections or sets that went into training the model, such as large private or public databases or data archives, and by providing a narrative explanation about other data sources used.

It is appropriate for the AI Office to provide a template for the summary, which should be simple, effective, and allow the provider to provide the required summary in narrative form.

(108) With regard to the obligations imposed on providers of general-purpose Al models to put in place a policy to comply with Union copyright law and make publicly available a summary of the content used for the training, the Al Office should monitor whether the provider has fulfilled those obligations without verifying or proceeding to a work-by-work assessment of the training data in terms of copyright compliance. This Regulation does not affect the enforcement of copyright rules as provided for under Union law.

(109) Compliance with the obligations applicable to the providers of general-purpose AI models should be commensurate and proportionate to the type of model provider, excluding the need for compliance for persons who develop or use models for non-professional or scientific research purposes, who should nevertheless be encouraged to voluntarily comply with these requirements.

Without prejudice to Union copyright law, compliance with these obligations should take due account of the size of the provider and allow simplified ways of compliance for SMEs, including start-ups, that should not represent an excessive cost and not discourage the use of such models. In the case of a modification or fine-tuning of a model, the obligations for providers should be limited to that modification or fine-tuning, for example by complementing the already existing technical documentation with information on the modifications, including new training data sources, as a means to comply with the value chain obligations provided in this Regulation.





### Recitals - Free and open source licences

(102) Software and data, including models, released under a free and open-source licence that allows them to be openly shared and where users can freely access, use, modify and redistribute them or modified versions thereof, can contribute to research and innovation in the market and can provide significant growth opportunities for the Union economy.

General-purpose AI models released under free and open-source licences should be considered to ensure high levels of transparency and openness if their parameters, including the weights, the information on the model architecture, and the information on model usage are made publicly available.

The licence should be considered to be free and open-source also when it allows users to run, copy, distribute, study, change and improve software and data, including models under the condition that the original provider of the model is credited, the identical or comparable terms of distribution are respected.

(103) Free and open-source AI components covers the software and data, including models and general-purpose AI models, tools, services or processes of an AI system.

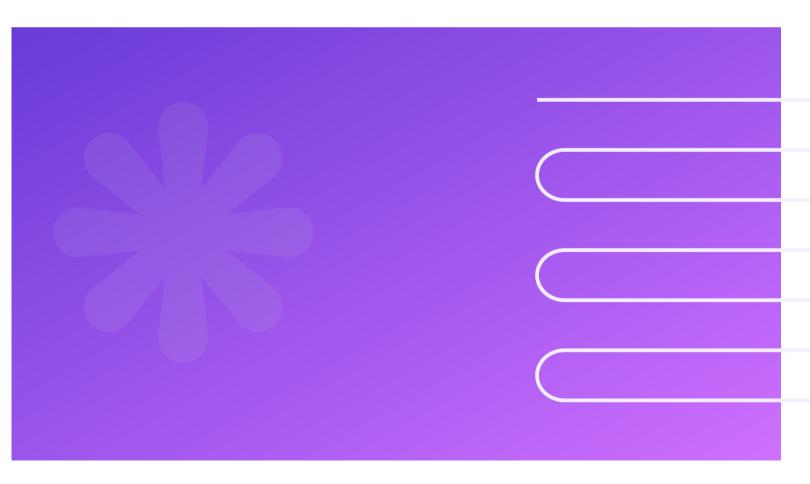
Free and open-source AI components can be provided through different channels, including their development on open repositories. For the purposes of this Regulation, AI components that are provided against a price or otherwise monetised, including through the provision of technical support or other services, including through a software platform, related to the AI component, or the use of personal data for reasons other than exclusively for improving the security, compatibility or interoperability of the software, with the exception of transactions between microenterprises, should not benefit from the exceptions provided to free and open-source AI components.

The fact of making AI components available through open repositories should not, in itself, constitute a monetisation.



(104) The providers of general-purpose AI models that are released under a free and open-source licence, and whose parameters, including the weights, the information on the model architecture, and the information on model usage, are made publicly available should be subject to exceptions as regards the transparency-related requirements imposed on general-purpose AI models, unless they can be considered to present a systemic risk, in which case the circumstance that the model is transparent and accompanied by an open-source licence should not be considered to be a sufficient reason to exclude compliance with the obligations under this Regulation.

In any case, given that the release of general-purpose AI models under free and open-source licence does not necessarily reveal substantial information on the data set used for the training or fine-tuning of the model and on how compliance of copyright law was thereby ensured, the exception provided for general-purpose AI models from compliance with the transparency-related requirements should not concern the obligation to produce a summary about the content used for model training and the obligation to put in place a policy to comply with Union copyright law, in particular to identify and comply with the reservation of rights pursuant to Article 4(3) of Directive (EU) 2019/790 of the European Parliament and of the Council.



# ANNEX II - PROVISIONS ON THE INTERPLAY BETWEEN THE AI ACT AND THE DIGITAL SERVICES ACT

## Article 2(5) Scope

This Regulation shall not affect the application of the provisions on the liability of providers of intermediary services as set out in Chapter II of Regulation (EU) 2022/2065.

## **Accompanying Recitals**

(118) This Regulation regulates AI systems and models by imposing certain requirements and obligations for relevant market actors that are placing them on the market, putting into service or use in the Union, thereby complementing obligations for providers of intermediary services that embed such systems or models into their services regulated by Regulation (EU) 2022/2065 of the European Parliament and of the Council42. To the extent that such systems or models are embedded into designated very large online platforms or very large online search engines, they are subject to the risk-management framework provided for in Regulation (EU) 2022/2065.

Consequently, the corresponding obligations of this Regulation should be presumed to be fulfilled, unless significant systemic risks not covered by Regulation (EU) 2022/2065 emerge and are identified in such models.

Within this framework, providers of very large online platforms and very large online search engines are obliged to assess potential systemic risks stemming from the design, functioning and use of their services, including how the design of algorithmic systems used in the service may contribute to such risks, as well as systemic risks stemming from potential misuses. Those providers are also obliged to take appropriate mitigating measures in observance of fundamental rights.

(119) Considering the quick pace of innovation and the technological evolution of digital services in scope of different instruments of Union law in particular having in mind the usage and the perception of their recipients, the AI systems subject to this Regulation may be provided as intermediary services or parts thereof within the meaning of Regulation (EU) 2022/2065, which should be interpreted in a technology-neutral manner.



#### con't

For example, AI systems may be used to provide online search engines, in particular, to the extent that an AI system such as an online chatbot performs searches of, in principle, all websites, then incorporates the results into its existing knowledge and uses the updated knowledge to generate a single output that combines different sources of information.

(120) Furthermore, obligations placed on providers and deployers of certain AI systems in this Regulation to enable the detection and disclosure that the outputs of those systems are artificially generated or manipulated are particularly relevant to facilitate the effective implementation of Regulation (EU) 2022/2065.

This applies in particular as regards the obligations of providers of very large online platforms or very large online search engines to identify and mitigate systemic risks that may arise from the dissemination of content that has been artificially generated or manipulated, in particular risk of the actual or foreseeable negative effects on democratic processes, civic discourse and electoral processes, including through disinformation.

(136) The obligations placed on providers and deployers of certain AI systems in this Regulation to enable the detection and disclosure that the outputs of those systems are artificially generated or manipulated are particularly relevant to facilitate the effective implementation of Regulation (EU) 2022/2065.

This applies in particular as regards the obligations of providers of very large online platforms or very large online search engines to identify and mitigate systemic risks that may arise from the dissemination of content that has been artificially generated or manipulated, in particular risk of the actual or foreseeable negative effects on democratic processes, civic discourse and electoral processes, including through disinformation.

The requirement to label content generated by AI systems under this Regulation is without prejudice to the obligation in Article 16(6) of Regulation (EU) 2022/2065 for providers of hosting services to process notices on illegal content received pursuant to Article 16(1) of that Regulation and should not influence the assessment and the decision on the illegality of the specific content.

That assessment should be performed solely with reference to the rules governing the legality of the content.

## ANNEX III - STRUCTURE OF THE AI ACT



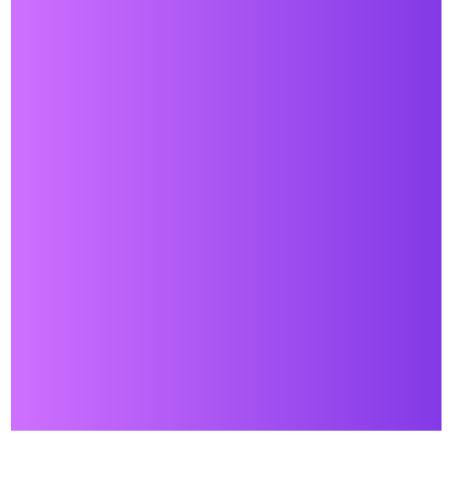
TITLE	Chapter	Recitals	Article	Title
I. Conorol		1 to 7	Article 1	Subject matter
I- General		8 to 11 21 to 27	Article 2	
Provisions		12 to 19	Article 3	Scope Definitions
		20	Article 4	Al literacy
II- Prohibited		20 +0 45	Article 5	Prohibited Al
Al practices		28 to 45	Article 5	practices
	1-CLASSIFICATION	46 to 48	Article 6	Classification rules for high-risk
	OF AI SYSTEMS AS HIGH-RISK	50-63 49	Article 7	Al systems Amendments to Annex III
		64 to 66	Article 8	Compliance with the requirements
	2-REQUIREMENTS	65	Article 9	Risk-management system
	FOR HIGH-RISK AI SYSTEMS	67 to 70	Article 10	Data & Data governance
	STOTEMO	71	Article 11	Technical documentation
			Article 12	Record-keeping
		72	Article 13	Transparency & provision of info to deployers
		73	Article 14	Human oversight
- High-risk		74 to 78	Article 15	Accuracy, robustness & cybersecurity
systems		79 to 80	Article 16	Obligations of providers of high-risk Al systems
	3-OBLIGATIONS OF	81	Article 17	Quality management system
	PROVIDERS &		Article 18	Documentation keeping
	DEPLOYERS OF		Article 19	Automatically generated logs Corrective actions & duty of information
	HIGH-RISK AI SYSTEMS & OTHER		Article 20	Cooperation with competent authorities
	PARTIES		Article 21	Authorised representatives of providers of high-risk Al systems
	-	82	Article 22	Obligations of importers
			Article 23	Obligations of distributors
			Article 24	Responsibilites along the value chain
		83 to 90	Article 25	Obligations of deployers of high-risk Al systems
		91 to 92	Article 26	Fundamental rights impact assessment for high-risk Al
		93 to 96	Article 27	systems
			Article 28	Notifying authorities
			Article 29	Application of a conformity assessment body for notificati
			Article 30	Notification procedure
	4-NOTIFYING		Article 31	Requirements related to notified bodies
	AUTHORITIES &		Article 32	Presumption of conformity with requirements relating to
	NOTIFIED BODIES		Article 32 Article 33	notified bodies
				Subsidiaries of notified bodies & subcontracting
			Article 34	Operational obligations of notified bodies
			Article 35	Identification numbers & lists of notified bodies
			Article 36	Changes to notifications
			Article 37	Challenge to the competence of notified bodies
			Article 38	Coodination of notified bodies
			Article 39	Conformity assessment bodies of third countries
		121 to 122	Article 40	Harmonised standards & standardisation deliverables
		·	Article 41	Common specifications
	5-STANDARDS,		Article 42	Presumption of conformity with certain requirements
	CONFORMITY	123 to 128		
	ASSESSEMENT,	130	Article 43	Conformity assessment
	CERTIFICATES,		Article 44	Certificates
	REGISTRATION		Article 45	Information obligations of notified bodies
			Article 46	Derogation from conformity assessment procedures
			Article 47	EU declaration of conformity
			Article 48	CE marking of conformity
		129	Article 49	Registration
V- Transparenc	су	) +o 127	Article 50	Transparency obligations for providers &
	<b>I</b> cy 132	2 to 137	Article 50	Transparency obligations for providers & users of certain Al systems



TITLE	Chapter	Recitals	Article	Title
V- General Purpose Al models	1-CLASSIFICATION RULES		Article 51	Classification of general purpose Al models as general purpose Al models with systemic risk
			Article 52	Procedure
	2- OBLIGATIONS FOR PROVIDERS OF GENERAL	07 +- 115		
	PURPOSE AI MODELS	97 to 115	Article 53 Article 54	Obligations for providers of of general purpose AI models Authorised representatives of providers of general-purpose AI models
	3-OBLIGATIONS FOR PROVIDERS OF GENERAL PURPOSE AI MODELS WITH			
	SYSTEMIC RISK	116 to 117	Article 55 Article 56	Obligations for providers of general purpose Al models with systemic risk Codes of Practice
		138 to 141	Article 57	Al regulatory sandboxes
VI-Measures			Article 58	Detailed arrangements for and functioning of Al regulatory sandboxes
in support of innovation			Article 59	Further processing of personal data for developing certain Al systems in the public interest in the Al regulatory sandbox
			Article 60	Testing of high-risk AI systems in real world conditions
			Article 61	outside AI regulatory sandboxes Informed consent to participate in testing in real world
		143 to 145	Article 62	conditions outside Al regulatory sandboxes  Measures for providers and deployers, in particular SMEs, including start-ups
		146	Article 63	Derogations for specific operators
\/II	1 - GOVERNANCE AT	148	Article 64	Al Office Establishment and structure of the European Artificial Intelligence Board
VII- Governance	UNION LEVEL	149	Article 65 Article 66	Tasks of the Board
		150	Article 67	Advisory Forum Scientific panel of independent experts
		151	Article 68	Access to the pool of experts by the Member States
			Article 69	
	2 - NATIONAL COMPETENT AUTHORITIES			Designation of national competent authorities and single point of contact
		153 to 154	Article 70	
VIII - EU Databas stand-alone hig Al systems			Article 71	EU database for stand-alone high-risk AI systems
IX - Post-	1 - POST-MARKET MONITORING	156 to 160	Article 74 Article 75	Market surveillance and control of AI systems in the Union market Mutual Assistance, market surveillance and control of general purpose AI
market monitoring,	2 - SHARING OF	161 to 164		systems Supervision of testing in real world conditions by market surveillance
information	INFORMATION ON		Article 76 Article 77	authorities Powers of authorities protecting fundamental rights
sharing,	INCIDENTS AND MALFUNCTIONING		Article 78 Article 79	Procedure for dealing with AI systems presenting a risk at national level
market surveillance		1	Article 79	Procedure at national level for dealing with Al systems presenting a risk Procedure for dealing with Al systems classified by the provider as non-
surveillarice	3 - ENFORCEMENT		Article 80 Article 81	high-risk in application of Annex II
			Article 82	Union safeguard procedure
		1	Article 83 Article 84	Compliant AI systems which present a risk Formal non-compliance
	3 - REMEDIES	152		Union Al testing support structures
		170	Article 85 Article 86	Right to lodge a complaint with a market surveillance authority Right to explanation of individual decision-making
		171	Article 87	Reporting of infringements and protection of reporting persons
	5 - SUPERVISION, INVESTIGATION,			Enforcement of obligations on providers of general purpose AI models
	ENFORCEMENT AND		Article 88 Article 89	Monitoring actions
	MONITORING IN RESPECT OF		Article 90	Alerts of systemic risks by the scientific panel  Power to request documentation and information
	PROVIDERS OF GENERAL		Article 91 Article 92	Power to conduct evaluations
	PURPOSE AI		Article 92 Article 93	Power to request measures  Procedural rights of economic operators of the general purpose AI model
	MODELS		Article 94	Procedural rights of economic operators of the general purpose Al model
L				-



TITLE	Chapter	Recitals	Article	Title
X - Codes of conduct and guidelines		165	Article 95	Codes of conduct for voluntary application of specific requirements
			Article 96	Guidelines from the Commission on the implementation of this Regulation
XI - Delegation			Article 97	Exercise of the delegation
of power and Committee		85		
procedure			Article 98	Committee procedure
XII - Penalties		168	Article 99 Article 100 Article 101	Penalties Administrative fines on Union institutions, agencies and bodies Fines for providers of general purpose Al models
			Article 102	Amendment to Regulation (EC) No 300/2008
XII - Final			Article 103	Amendment to Regulation (EU) No 167/2013
provisions			Article 104 Article 105	Amendment to Regulation (EU) No 168/2013  Amendment to Directive 2014/90/EU
			Article 106	Amendment to Directive (EU) 2016/797
				Amendment to Regulation (EU) 2018/858
			Article 107 Article 108	Amendment to Regulation (EU) 2018/1139
			Article 109	Amendment to Regulation (EU) 2019/2144
			Article 110	Amendment to Directive (EU) 2020/1828
		177	Article 111	Al systems already placed on the market or put into service
		174 178 to 179	Article 112 Article 113	Evaluation and review Entry into force and application
Annexes			ANNEX I	List of Union harmonisation legislation
Alliexes			ANNEX III	List of criminal offences referred to in Article 5(1), point (e) (iii)
			ANNEX III	High-risk Al systems referred to in article 6(2)
			ANNEX IV	Technical documentation referred to in article 11(1)
			ANNEX V	EU declaration of conformity
			ANNEX VI	Conformity assessment procedure based on internal control
			ANNEX VII	Conformity based on assessment of quality management system and assessment of technical documentation
			ANNEX VIII	Information to be submitted upon the registration of high-risk Al systems in accordance with Article 49
			ANNEX IX	Information to be submitted upon the registration of high-risk Al systems
			ANNEX X	listed in Annex III in relation to testing in real world conditions in accordance with Article 60
			ANNEX XI	Union legislative acts on large-scale IT systems in the area of Freedom, Security and Justice
				Technical documentation referred to in Article 53(1), point (a) - technical documentation for providers of general-purpose AI models
			ANNEX XII	Transparency information referred to in Article 53(1), point (b) - technical documentation for providers of general-purpose AI models to
			ANNIEV VIII	downstream providers that integrate the model into their Al system
			ANNEX XIII	Criteria for the designation of general-purpose AI models with systemic risk referred to in Article 51





## For more information about this publication, please contact:

Molly Stech General Counsel, STM

molly@stm-assoc.org

Claudia Russo

Regional Manager, EU Public Affairs claudia@stm-assoc.org