Article 53

Obligations for providers of general-purpose AI models

Author: Maarten Herbosch

Date completed: November 2025

ARTICLE 53: Obligations for providers of general-purpose AI models

1. Providers of general-purpose AI models shall:

(a) draw up and keep up-to-date the technical documentation of the model, including its training and testing process and the results of its evaluation, which shall contain, at a minimum, the information set out in Annex XI for the purpose of providing it, upon request, to the AI Office and the national competent authorities;

(b) draw up, keep up-to-date and make available information and documentation to providers of AI systems who intend to integrate the general-purpose AI model into their AI systems. Without prejudice to the need to observe and protect intellectual property rights and confidential business information or trade secrets in accordance with Union and national law, the information and documentation shall:

(i) enable providers of AI systems to have a good understanding of the capabilities and limitations of the general-purpose AI model and to comply with their obligations pursuant to this Regulation; and

(ii) contain, at a minimum, the elements set out in Annex XII;

(c) put in place a policy to comply with Union law on copyright and related rights, 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.

2. The obligations set out in paragraph 1, points (a) and (b), shall not apply to providers of AI models that are released under a free and open-source licence that allows for the access, usage, modification, and distribution of the model, and whose parameters, including the weights, the information on the model architecture, and the information on model usage, are made publicly available. This exception shall not apply to general-purpose AI models with systemic risks.

3. Providers of general-purpose AI models shall cooperate as necessary with the Commission and the national competent authorities in the exercise of their competences and powers pursuant to this Regulation.

4. Providers of general-purpose AI models may rely on codes of practice within the meaning of Article 56 to demonstrate compliance with the obligations set out in paragraph 1 of this Article, until a harmonised standard is published. Compliance with European harmonised standards grants providers the presumption of conformity to the extent that those standards cover those obligations. Providers of general-purpose AI models who do not adhere to an approved code of practice or do not comply with a European harmonised standard shall demonstrate alternative adequate means of compliance for assessment by the Commission.

5. For the purpose of facilitating compliance with Annex XI, in particular points 2 (d) and (e) thereof, the Commission is empowered to adopt delegated acts in accordance with Article 97 to detail measurement and calculation methodologies with a view to allowing for comparable and verifiable documentation.

6. The Commission is empowered to adopt delegated acts in accordance with Article 97(2) to amend Annexes XI and XII in light of evolving technological developments.

7. Any information or documentation obtained pursuant to this Article, including trade secrets, shall be treated in accordance with the confidentiality obligations set out in Article 78.

RECITALS

Recital 100

When a general-purpose AI model is integrated into or forms part of an AI system, this system should be considered to be general-purpose AI system when, due to this integration, this system has the capability to serve a variety of purposes. A general-purpose AI system can be used directly, or it may be integrated into other AI systems.

Recital 101

Providers of general-purpose AI models have a particular role and responsibility along the AI value chain, as the models they provide may form the basis for a range of downstream systems, often provided by downstream providers that necessitate a good understanding of the models and their capabilities, both to enable the integration of such models into their products, and to fulfil their obligations under this or other regulations. Therefore, proportionate transparency measures should be laid down, including the drawing up and keeping up to date of documentation, and the provision of information on the general-purpose AI model for its usage by the downstream providers. Technical documentation should be prepared and kept up to date by the general-purpose AI model provider for the purpose of making it available, upon request, to the AI Office and the national competent authorities. The minimal set of elements to be included in such documentation should be set out in specific annexes to this Regulation. The Commission should be empowered to amend those annexes by means of delegated acts in light of evolving technological developments.

Recital 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.

Recital 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 license 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[1].

Recital 105

General-purpose AI models, in particular large generative AI 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. 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.

Recital 106

[1] 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 and amending Directives 96/9/EC and 2001/29/EC [2019] OJ L 130/92 ("DSM Directive").

Providers that place general-purpose AI models on the Union market should ensure compliance with the relevant obligations in this Regulation. To that end, providers of general-purpose AI models should put in place a policy to comply with Union law on copyright and related rights, in particular to identify and comply with the reservation of rights expressed by rightsholders pursuant to Article 4(3) of Directive (EU) 2019/790. Any provider placing a general-purpose AI model on the Union market should comply with this obligation, regardless of the jurisdiction in which the copyright-relevant acts underpinning the training of those 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 Union market by applying lower copyright standards than those provided in the Union.

Recital 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 AI 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.

Recital 108

With regard to the obligations imposed on providers of general-purpose AI 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 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. This Regulation does not affect the enforcement of copyright rules as provided for under Union law.

Recital 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 those 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 of general-purpose AI models 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.

Recital 117

The codes of practice should represent a central tool for the proper compliance with the obligations provided for under this Regulation for providers of general-purpose AI models. Providers should be able to rely on codes of practice to demonstrate compliance with the obligations. By means of implementing acts, the Commission may decide to approve a code of practice and give it a general validity within the Union, or, alternatively, to provide common rules for the implementation of the relevant obligations, if, by the time this Regulation becomes applicable, a code of practice cannot be finalised or is not deemed adequate by the AI Office. Once a harmonised standard is published and assessed as suitable to cover the relevant obligations by the AI Office, compliance with a European harmonised standard should grant providers the presumption of conformity. Providers of general-purpose AI models should furthermore be able to demonstrate compliance using alternative adequate means, if codes of practice or harmonised standards are not available, or they choose not to rely on those.

Select bibliography

- Bernsteiner C and Schmitt T R, 'Art. 53 Pflichten für Anbieter von KI-Modellen mit allgemeinem Verwendungszweck' in Mario Martini and Christiane Wendehorst (eds), *KI-VO: Verordnung über Künstliche Intelligenz: Kommentar* (C H Beck 2024).
- de la Durantaye K, 'Nutzung urheberrechtlich geschützter Inhalte zum Training generativer künstlicher Intelligenz – ein Lagebericht' (2024) 55 AfP 9.
- de la Durantaye K, 'Akkommodation statt Assimilation. Warum die EU bei der KI-Regulierung nicht auf den Brussels Effect setzen sollte – und was stattdessen sinnvoll wäre' (2025) Zeitschrift für Urheber- und Medienrecht 165.
- Nordemann J B and Arman R, 'Die Regelungen der KI-Verordnung mit Urheberrechtsbezug – Möglichkeit der privaten Rechtsdurchsetzung?' (2024) Zeitschrift für Urheber- und Medienrecht 780.
- Peukert A, 'Copyright in the AI Act – A Primer' (2024) 73 GRUR International 497.
- Schneider, A, 'Art. 53 Pflichten für Anbieter von KI-Modellen mit allgemeinem Verwendungszweck' in Jens Schefzig and Robert Kilian (eds), *Beck'scher Online-Kommentar KI-Recht* (3rd edn, C H Beck 2025).

Commentary

1. General remarks

1.1. Introduction

1. Article 53 AI Act[2] sets out the key obligations for providers of general-purpose AI ("GPAI") models. These are models that display significant generality, can perform a wide range of distinct tasks, and can be integrated into a variety of downstream systems and applications (Article 3(63)).[3] Like the other provisions in Chapter V of the AI Act, Article 53 was introduced at a later stage of the drafting process in response to the growing prominence of large language models.[4] As the only provision that contains substantive obligations for GPAI model providers whose models do not present systemic risk,[5] it serves several distinct goals. First, it seeks to promote safe and trustworthy AI innovation in the European Union,[6] for example by empowering regulators to request information on a GPAI model's performance.[7] These provisions are of key importance in determining what information the

[2] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) [2024] OJ L 1689/1 ("AI Act").
[3] See, more extensively, the analysis on Article 3(63) discussed in forthcoming commentary on Article 3(63) in this work.
[4] Also see Adrian Schneider, 'Art. 53 Pflichten für Anbieter von KI-Modellen mit allgemeinem Verwendungszweck' in Jens Schefzig and Robert Kilian (eds), *Beck'scher Online-Kommentar KI-Recht* (3rd edn, C.H. Beck 2025) para 5.
[5] If a GPAI model is classified as presenting systemic risk, its provider must additionally comply with Article 55 (also see forthcoming commentary on Article 55 in this work).
[6] European Commission, 'General-Purpose AI Models in the AI Act – Questions & Answers' (2025) <https://digital-strategy.ec.europa.eu/en/faqs/general-purpose-ai-models-ai-act-questions-answers> accessed 1 October 2025. See similarly Schneider (n 4) para 3.
[7] See Section 2.1.1. Also see Schneider (n 4) para 3.

Commission, AI Office and national competent authorities may request from GPAI model providers – and, conversely, which information the latter may elect to retain for themselves.

2. Second, it aims to ensure that downstream AI system providers who intend to integrate the GPAI model in their AI system have access to relevant information concerning its performance and compatibility.[8] Third, it seeks to secure compliance with Union copyright law by requiring GPAI model providers to make publicly available the information used to train their models, thereby enabling affected individuals and rights holders to better safeguard their legal interests,[9] and by arguably extending the scope of Union copyright law.[10]

3. A preliminary general observation is that Recital 109 suggests that the relevant obligations should be applied more stringently to larger providers. Particularly for obligations not explicitly scaled by provider size, this implies that Article 53's substantive provisions ought to be enforced more rigorously for such providers. In practical terms, they may therefore be expected to produce the relevant documentation with greater detail and thoroughness.

1.2. Structure & overview

4. This contribution broadly follows the structure of Article 53. It begins by situating the Article within the broader context of the AI Act before proceeding to a paragraph-by-paragraph analysis. Article 53(1) arguably contains the most substantive provision, outlining various documentation obligations for GPAI model providers. These include internal documentation – to be submitted to the AI Office and national competent authorities upon request – concerning the model's training and testing (Article 53(1)(a)). In addition, Article 53(1)(b) contains requirements regarding information to be shared with downstream system providers intending to incorporate the model into their systems.

5. Further, Article 53(1) addresses significant aspects related to copyright, requiring GPAI model providers to develop a copyright policy that underscores compliance with Union copyright law (Article 53(1)(c)) and disclosure of information on the content used for model training (Article 53(1)(d)). The analysis then turns to Article 53(2), which introduces a partial exception from certain documentation and information requirements for providers of GPAI models that do not pose systemic risks. The Article 53(3) discussion considers the obligation to cooperate with the Commission and national competent authorities.

6. The next three sections examine compliance mechanisms, including codes of practice (Article 53(4)) and the use of delegated acts (Articles 53(5) and (6)). The discussion concludes with Article 53(7), which refers to the confidentiality obligations set out in Article 78 in relation to any information communicated pursuant to Article 53.

[8] See Section 2.1.2. Also see European Commission, 'General-Purpose AI Models in the AI Act – Questions & Answers' (n 6).
[9] See Section 2.1.4.
[10] See Section 2.1.3.

2. Substance

2.1. Article 53(1): Information & documentation

7. Article 53(1) sets out the substantive obligations for GPAI model providers, distinguishing four main categories of information that must be provided or made accessible. First, GPAI model providers must, upon request, supply certain documentation to the AI Office and national competent authorities (Article 53 (1)(a)). Second, they are required to provide specific information and documentation to AI system providers intending to integrate the GPAI model into their systems ((1)(b)). Third, they must establish a copyright compliance policy ((1)(c)). Fourth and finally, they are obliged to make publicly available a summary of the data used to train the model ((1)(d)). As discussed further below, all such information must be kept up to date.

8. Recital 109 indicates that persons who develop or use GPAI models for non-professional or scientific research purposes are encouraged rather than obliged to comply with the relevant documentation and transparency rules for GPAI models.[11] This exclusion is, however, not reiterated in the text of Article 53, but it could be read to follow from Article 2.[12] In any case, it is key to underscore that this exception – as well as the open-source exception (Article 53(2)) discussed below[13] – do not go as far as to exempt the provider from some of the underlying requirements, including the need to respect Union copyright law, which features its own independent exceptions.[14]

2.1.1. Article 53(1)(a): Documentation for AI Office and national competent authorities

2.1.1.1. All GPAI models

2.1.1.1.1. Required information

9. Article 53(1)(a) requires GPAI model providers to document and maintain up-to-date technical information concerning the model, its training and testing processes, and model evaluation. This documentation must, at a minimum, include the information specified in Annex XI and must be made available, upon request, to the AI Office and/or the national competent authorities. The latter refers to the authorities designated by Member States pursuant to Article 70 of the AI Act. We will briefly reconsider the procedural implications for national competent authorities, the AI Office and the Commission below.[15]

[11] Also see Schneider (n 4) para 30.
[12] See forthcoming commentary on Article 2 in this work.
[13] See Section 2.2.
[14] Notably, the research exception found in copyright law is restricted to research organisations, see art 3 DSM Directive.
[15] See paras 38–41.

10. Annex XI provides further detail on the required information, dividing it into two sections. The first outlines the information applicable to all GPAI models, while the second sets out additional requirements for models presenting systemic risk.

11. In examining these obligations, we will frequently reference the corresponding provisions of the relevant code of practice.[16] Although adherence to such codes does not, in and of itself, constitute conclusive evidence of compliance with the AI Act in general,[17] such codes do provide useful guidance as to how the AI Office may interpret the relevant provisions.[18] Furthermore, once the AI Office and the Board deem a code of practice adequate (Article 56(6) AI Act), such a code of practice can be used to demonstrate compliance with the relevant obligations.[19] It is thus key to note that the European Commission[20] and AI Board[21] have confirmed the adequacy of the 2025 Code of Practice on 1 August 2025.

12. As per Section 1 of Annex XI, all GPAI model providers are required to document at least 'A general description of the general-purpose AI model including:'

- '(a) the tasks that the model is intended to perform and the type and nature of AI systems in which it can be integrated;'

13. The tasks the model is intended to perform – 'intended uses' in the Code of Practice Transparency form[22] – refers to the specific functions the model is designed to carry out, such as 'productivity enhancement, translation, creative content generation, data analysis, data visualisation, programming assistance, scheduling, customer support, variety of natural language tasks'.[23] The type and nature of AI systems in which the model can be integrated refers to the category and characteristics of those systems – such as 'autonomous systems, conversational assistants, decision support systems, creative AI systems, predictive systems, cybersecurity, surveillance, or human-AI collaboration'.[24] While it has been argued that the tasks of a model will generally correspond to the types of systems into which it

[16] For an overview of the Code of Practice and its various chapters, see European Commission, 'The General-Purpose AI Code of Practice' (2025), <https://digital-strategy.ec.europa.eu/en/policies/contents-code-gpai> accessed 1 October 2025.
[17] E.g., European Commission, 'Code of Practice for General-Purpose AI Models – Transparency Chapter' (2025) <https://ec.europa.eu/newsroom/dae/redirection/document/118120> accessed 1 October 2025, 3.
[18] See commentary on Article 56 in this work.
[19] See commentary on Article 56 in this work; Annex to the Communication to the Commission – Approval of the content of the draft Communication from the Commission – Guidelines on the scope of the obligations for general-purpose AI models established by Regulation (EU) 2024/1689 (AI Act)' C(2025) 5045 final para 94.
[20] European Commission, 'Commission Opinion of 1 August 2025 on the assessment of the General-Purpose AI Code of Practice within the meaning of Article 56 of Regulation (EU) 2024/1689' COM (2025) 5361 final.
[21] European Commission, 'Conclusion of the Artificial Intelligence Board on the Assessment of the General-Purpose AI Code of Practice pursuant to Article 56 of Regulation 2024/1689 (Artificial Intelligence Act)' (2025) <https://ec.europa.eu/newsroom/dae/redirection/document/118687> accessed 1 October 2025.
[22] European Commission, 'Model Documentation Form' (2025) <https://ec.europa.eu/newsroom/dae/redirection/document/118118> accessed 1 October 2025.
[23] See more extensively on model tasks: forthcoming commentary on Article 3(63) in this work.
[24] For these examples, see the Code of Practice Model Documentation Form (n 22) 2.

may be integrated,[25] the Code of Practice form does distinguish between the two.[26] While the wording of Article 53(1)(a) appears to require only a relatively abstract description,[27] the Code of Practice calls for a more detailed approach, suggesting a recommended length of 200 words for intended uses and 300 words for the type and nature of AI systems in which the GPAI model can be integrated.[28] Interestingly, it suggests that those elements do not need to be described positively but may also be adhered to negatively by describing the restricted or prohibited uses or the type of nature of AI systems in which the GPAI model should not be integrated, respectively.[29]

- '(b) the acceptable use policies applicable'

14. The relevant use policies are likely to concern restrictions on the use of the GPAI model, for example to prevent its deployment in the commission of criminal offences[30] or in breach of copyright law. Such policies may also impose limitations on the use of the model's outputs. GPAI model providers may, for instance, restrict use to non-commercial purposes unless a specific licence is obtained. The model form annexed to the Code of Practice suggests that providers should indicate whether such a policy exists. This, together with the use of the term 'applicable', may be taken to imply that the absence of an acceptable use policy is permissible.

15. The question of whether the modification of the model can be restricted in such use policies is discussed elsewhere in this commentary.[31]

- '(c) the date of release and methods of distribution'

16. The notion of 'release' differs from the concept of 'placing on the market' as defined in Article 3(9)[32] in that release does not require the model to be made available on the Union market.[33] In line with arguments by some authors that model providers should document the release date of the model following each modification or technical change,[34] the form provided in the Code of Practice requires GPAI model providers to indicate the release date of the current model, the release date on the Union market,[35] and any 'model dependencies' – meaning an overview of previous versions of the model and their respective release dates. The term 'Union market release date' arguably corresponds more closely to the notion of 'placing on the market' in Article 3(9).

[25] Clemens Bernsteiner and Thomas Rainer Schmitt, 'Art. 53 Pflichten für Anbieter von KI-Modellen mit allgemeinem Verwendungszweck' in Mario Martini and Christiane Wendehorst (eds), *KI-VO: Verordnung über Künstliche Intelligenz: Kommentar* (C.H. Beck 2024) para 18.
[26] Code of Practice Model Documentation Form (n 22) 2.
[27] Bernsteiner and Schmitt (n 25) para 18.
[28] Code of Practice Model Documentation Form (n 22) 2.
[29] ibid. 2.
[30] Bernsteiner and Schmitt (n 25) para 18.
[31] Also see forthcoming chapter on Modifications in this work.
[32] That provision reads '"placing on the market" means the first making available of an AI system or a general-purpose AI model on the Union market'.
[33] See also Bernsteiner and Schmitt (n 25) para 18.
[34] ibid. para 18.
[35] Code of Practice Model Documentation Form (n 22) 1.

17. Providers must also disclose the 'methods of distribution'. The model form in the Code of Practice lists several examples – 'e.g. enterprise or subscription-based access through existing software suites or enterprise-specific solutions; public or subscription-based access through an API; public or proprietary access through integrated development environments, device-specific applications or firmware, open-source repositories'.[36] For each distinct method of distribution, a link (where available) to information about how the model can be accessed should be provided, along with a brief description of the level of the access (e.g. 'weights-level access' or 'black-box access').[37]

- '(d) the architecture and number of parameters'

18. With respect to the architecture, a brief description is required – the Code of Practice form recommends a length of approximately 20 words (e.g. 'a transformer architecture').[38] The model provider must also disclose the total number of parameters. The form mandates the use of at least two significant figures (e.g. '7.3×10^{10}') and further requires the provider to indicate the range within which this number falls, selecting from a set of predefined options.[39]

- '(e) the modality (e.g. text, image) and format of inputs and outputs'

19. The information concerning the modality and format of inputs and outputs refers to the types of data a model can process or generate – for example, text, images, audio, video, or other types. It should also include the maximum input and output file sizes,[40] where such limits are defined.[41]

- '(f) the licence'

20. The information concerning the licence should, naturally, specify the licence under which access to the model is granted, or, alternatively, indicate that no such licence exists.

21. Annex XI further requires the model developer to provide 'relevant information on the development process, including the following elements:'

- '(a) the technical means (e.g. instructions of use, infrastructure, tools) required for the general-purpose AI model to be integrated in AI systems'

22. The model provider should specify the infrastructure requirements and necessary tools, along with any relevant operating instructions needed for the system provider to integrate the GPAI model successfully within the intended system. This includes a description of the required hardware (if any – none may be needed if the model is accessed via an API) and software.

---

[36] Code of Practice Model Documentation Form (n 22) 2.
[37] ibid. 1–2.
[38] ibid. 1.
[39] Code of Practice Model Documentation Form (n 22) 1, which lists the options 1–500M, 500M–5B, 5B–15B, 15B–50B, 50B–100B, 100B–500B, 500B–1T, >1T.
[40] While Annex XI's text would imply that the output size should also be shared, the Code of Practice implies that this size is only relevant for downstream system providers, see Code of Practice Model Documentation Form (n 22) 1.
[41] Also see ibid. 1.

- ‘(b) the design specifications of the model and training process, including training methodologies and techniques, the key design choices including the rationale and assumptions made; what the model is designed to optimise for and the relevance of the different parameters, as applicable’

23. The design specifications of the training process should be described in reasonable detail, covering the elements listed. The Code of Practice offers an illustrative example: ‘the model is initialized with randomly selected weights and optimised using gradient-based optimisation via the Adam optimiser in two steps. First, the model is trained to predict the next word on a large pre-training corpus using the cross-entropy loss, passing over the data for a single epoch. Second, the model is post-trained on a dataset of human preferences for 10 epochs to align the model with human values and make it more useful in responding to user prompts.’[42] This level of detail goes beyond a mere characterisation of the training method (e.g. ‘supervised learning’) and should also include a description of the key design choices made during model training, accompanied by a rationale for their adoption.

- ‘(c) information on the data used for training, testing and validation, where applicable, including the type and provenance of data and curation methodologies (e.g. cleaning, filtering, etc.), the number of data points, their scope and main characteristics; how the data was obtained and selected as well as all other measures to detect the unsuitability of data sources and methods to detect identifiable biases, where applicable’

24. The information regarding the data used for training, testing and validation should include the type of data (e.g. text, image, video or audio) and how it was obtained (e.g. via web-crawling, publicly available sources, synthetic data, user data), as well as the specific means and criteria used for its collection and selection. This includes the methods and resources employed for data annotation as well as any tools or techniques used to generate synthetic data. Where data has been sourced from third parties, the GPAI model provider should explain how the rights to use such data were acquired, unless this information has already been made publicly available under Article 53(1)(d).[43]

25. The number of data points used in training, testing and validation should be indicated, along with the relevant unit (‘e.g. tokens, documents, images, hours of video or frames’[44]). The Code of Practice requires this information to be reported with at least one or two significant figures (e.g. ‘3×10^{13} tokens’), depending on whether the data is submitted to national competent authorities or the AI Office.[45]

26. The scope and principal characteristics of the data should also be described. This includes the domain (e.g. ‘healthcare, science, law’ or scientific data), geographical origin (e.g. European, global, or US-based data), and the language(s) of the data (in the case of text, audio or video).[46] If applicable, the description should also include the modality coverage of the dataset.[47] Additionally, the information must detail the measures taken to identify potential biases – specifically methods used during data acquisition or processing – and describe how the provider assessed the potential unsuitability of the

[42] ibid. 2.
[43] See paras 97 ff.
[44] Code of Practice Model Documentation Form (n 22) 3.
[45] ibid. 3.
[46] Also see ibid. 3.
[47] ibid. 3.

data sources. This latter obligation extends beyond legal compliance (e.g. exclusion of unlawfully processed personal data or non-consensual intimate imagery) and also encompasses data that may be unsuitable for robust model training given the intended use of the model.[48]

- '(d) the computational resources used to train the model (e.g. number of floating point operations), training time, and other relevant details related to the training'

27. The information regarding the computational resources used to train the model should include the total duration of the training process. The Code of Practice distinguishes between national competent authorities and the AI Office with respect to the required level of detail – for example, indicating the number of months for the former and specifying the duration in wall-clock days for the latter.[49] The Code of Practice indicates that a similar distinction applies to the reporting of computational resources used: for national competent authorities, this should be recorded in floating point operations (FLOPs) to the correct order of magnitude, whereas for the AI Office, the figure must be provided with at least two significant figures.[50]

28. The Code of Practice further clarifies that a description of the methodology used to measure or estimate the volume of computation should be included, in the absence of a delegated act adopted under Article 53(5).[51]

29. To support compliance with these requirements, the Commission is empowered to adopt delegated acts pursuant to Article 97 of the AI Act. These acts may establish harmonised methods for calculating and measuring computational resources, training duration, and other relevant aspects of model training.[52] Such delegated acts can supplement or amend non-essential elements of the AI Act and are thus binding on those subject to its requirements, including GPAI model providers.[53]

- '(e) known or estimated energy consumption of the model'

30. Lastly, the model provider should provide information about the energy consumption of the model. Annex XI clarifies that 'where the energy consumption of the model is unknown, the energy consumption may be based on information about computational resources used.' While the past tense of that provision might be interpreted to imply that this energy use only refers to the training of the model, some authors argued that it should also include the model's energy consumption during inference.[54] In any case, such an interpretation cannot fully be ruled out as the energy consumption for training (past tense) may be indicative of that during (future) inference.[55]

[48] ibid. 3.
[49] ibid. 3.
[50] ibid. 3.
[51] See Section 2.5.
[52] AI Act, art 53(5). Also see Section 2.5.
[53] Consolidated version of the Treaty on the Functioning of the European Union [2012] OJ C 326/47 ("TFEU") art 290
[54] Nicolas Alder, Kai Ebert, Ralf Herbrich and Philip Hacker, 'AI, Climate, and Transparency: Operationalizing and Improving the AI Act' (2024) <https://arxiv.org/abs/2409.07471> s 2.
[55] Also see para 131.

31. The Code of Practice takes a clearer position, as it provides that the energy consumption of the model does not only relate to the energy used during model deployment in a potential AI system ('energy consumption during inference'), but also to the amount of energy that was used for model training ('energy used for training'),[56] as had been previously recommended by some authors. For the former, the model provider should detail the amount of computation used for inference measured in FLOP with at least two significant figures, per the Code of Practice.[57] Moreover, in the absence of a delegated act in accordance with Article 53(5),[58] the model provider should provide a description of the computational tasks and the hardware used to measure or estimate the inference-time computation, and that used for any estimations of the energy consumption of the model.[59]

32. The amount of energy used for training should be estimated and reported in megawatt-hours with at least two significant figures according to the Code of Practice, though the provider can also indicate that they do not have the necessary information (e.g. due to unknowns related to hardware provided by an external provider) to make a reasonable assessment.[60] Here too, the Code of Practice requires the provider to clarify the methodology used for their measurements.[61] In the absence of a delegated act as per Article 53(5), the provider should estimate this on the basis of computational resources required for model training.[62] As before, it is permissible to indicate that no assessment could be made due to missing information, though the provider should specify the information that they lack.[63] If the model results from the modification or fine-tuning of another GPAI model, it is reasonable to estimate its energy consumption using known estimates or information about the parent model, in the absence of more specific data.

33. As for the computational resources used to train the model,[64] the Commission is empowered to adopt delegated acts in application of Article 97 AI Act[65] to facilitate compliance.[66] These can support the calculation and measurement methods to support comparable and verifiable documentation of the energy consumption of the model. An approach might be for the Commission to establish that the environmental impact of GPAI models should be estimated using the relevant servers' or data centres' energy use and their efficiency ratings (the power usage effectiveness, PUE[67]).[68]

---

[56] Code of Practice Model Documentation Form (n 22) 3, fn 1.
[57] ibid. 3.
[58] See Section 2.5.
[59] Code of Practice Model Documentation Form (n 22) 3.
[60] ibid. 3.
[61] ibid. 3.
[62] ibid. 3.
[63] ibid. 3.
[64] See para 29.
[65] Cf. Section 2.5.
[66] AI Act, art 53(5).
[67] Also see Annex III to Commission Delegated Regulation (EU) 2024/1364 of 14 March 2024 on the first phase of the establishment of a common Union rating scheme for data centres [2024] OJ L 1364/1 (as well as Directive (EU) 2023/1791 of the European Parliament and of the Council of 13 September 2023 on energy efficiency and amending Regulation (EU) 2023/955 (recast) [2023] OJ L 231/1).
[68] Alder, Ebert, Herbrich and Hacker (n 54) s 4.

2.1.1.1.2. Extent

34. Interestingly, the extent of these obligations and the information required should be 'appropriate' to the 'size and risk profile of the model'.[69] Notably, this is the only explicit reference to the 'size' of the model in the AI Act. While 'size' most likely denotes the number of parameters of the model, it is notable that the AI Act explicitly refers to the 'number of parameters' elsewhere – for example, in Recital (98), Annex XI(1)(d), and Annex XII(1)(f) – without using the term 'size', which could imply that 'size' should be read differently, referring more broadly to, for example, the volume of training data,[70] the computational resources required,[71] the model's complexity[72] or the breadth of its capabilities[73].

35. Likewise, the AI Act neither defines the notion of a 'risk profile' nor employs the term elsewhere in its text. The Act does not clarify whether this should be understood, in the context of GPAI models, as referring to the (potential) systemic risk presented by the model or, rather, whether it relates to the likelihood that the model will be incorporated into high-risk systems within the meaning of Article 6. The latter interpretation may be defensible, given the broader relevance of various Annex XII elements for high-risk systems, discussed below.[74] However, the most natural reading is likely the more general one, which would have 'risk profile' refer to the risk definition set out in Article 3(2).[75] This definition is not confined to the categorisation requirements of high-risk systems or systemic risk models, but rather entails a contextualised assessment of the model's potential for misuse or harm, as well as the likelihood of such incidents, taking into account the scale and scope of deployment, the model's capabilities, the vulnerability of affected parties, and any available information about past incidents or vulnerabilities.[76]

2.1.1.2. GPAI Models presenting systemic risk

2.1.1.2.1. Scope

36. Specifically for providers of GPAI models with systemic risk, who are also subject to Article 55,[77] Section 2 of Annex XI lists additional information that providers of GPAI models with systemic risk must document and provide to the AI Office and national competent authorities upon request. This information largely concerns the methods used to evaluate the model's performance, the measures taken

[69] Annex XI AI Act s 1.
[70] This, too, is referred to more directly in the AI Act, e.g., the 'cumulative amount of computation used for its training' in article 51(2), Annex XI (2)(d), and Annex XIII (c).
[71] The AI Act refers to the computational resources required for the model directly in Annex XI (2)(d) (interestingly using the number of floating point operations as a proxy) and Annex IV (2)(c).
[72] The AI Act directly refers to the 'complexity' of AI *systems* (not models) in some provisions, e.g., AI Act, recital 72, recital 125, art 31(8), and art 34(2).
[73] Denoted elsewhere more directly as the 'generality' of the model, see AI Act, recitals 97, 98, and art 3(63).
[74] See paras 59 ff.
[75] This provision holds that '"risk" means the combination of the probability of an occurrence of harm and the severity of that harm'.
[76] See similarly Schneider (n 4) para 11.
[77] See forthcoming commentary on Article 55 in this work

to identify and address model vulnerabilities, and the interactions between various software components.

37. For the interpretation of these requirements, it is useful to highlight that the AI Act does not explicitly refer to Annex XI outside Articles 53 and 54, with the latter only imposing an obligation to verify the relevant documentation.[78] This is remarkable, as some of the obligations found in Article 55(1)(a), (b), and (d) could,[79] seemingly, be interpreted to closely relate to Annex XI Section 2's documentation requirements. This applies, for example, to the requirement that model developers should 'perform model evaluation in accordance with standardised protocols and tools reflecting the state-of-the-art, including conducting and documenting adversarial testing of the model with a view to identifying and mitigating systemic risks' (Article 55(1)(a)), which arguably relates to 'a detailed description of the evaluation strategies' discussed below.

38. A more extensive interpretation of Annex XI Section 2 would strengthen the link with Article 55 and thus capture Article 55(1)'s obligations, resulting in more extensive documentation under that Section of Annex XI. An important policy implication of that interpretation is that it would require providers of GPAI models with systemic risk to document the assessments made under Article 55 as well – some of which seem to be captured more explicitly by the wording of Annex XI Section 2. As a result, that information could be requested both by the AI Office and the national competent authorities on the basis of Article 53(1)(a)'s reference to Annex XI.

39. This interpretation appears most consistent with the objectives of the AI Act in relation to these documentation requirements – namely, that this documentation should inform the AI Office (and national competent authorities), upon request, regarding the potential risks and capabilities of these models. Nevertheless, the need for legal certainty could be raised as a counterargument, as the Act does not explicitly set out this link between Annex XI Section 2 and Article 55(1). That argument would, however, be largely countered by the broad language in Annex XI Section 2, together with Article 55 and the AI Office's authority under Article 91 to request relevant information, which together provide a sufficiently clear legal basis for requiring disclosure. In other words, even without an explicit textual cross-reference between Article 55 and Annex XI Section 2, the combined provisions establish a sufficiently predictable and clear framework.

40. If, however, Section 2 of Annex XI were to be given a more restrictive interpretation – which would exclude any or some elements found in Article 55 that Annex XI does not explicitly reference – the direct result could be that Article 55 obliges providers of GPAI models with systemic risk to carry out certain assessments, such as an assessment of 'systemic risks at Union level', without requiring them to document those assessments. One could argue, in such a scenario, that national competent authorities

[78] See commentary on Article 54 in this work

[79] Article 55(1) reads '1. In addition to the obligations listed in Articles 53 and 54, providers of general-purpose AI models with systemic risk shall: (a) perform model evaluation in accordance with standardised protocols and tools reflecting the state of the art, including conducting and documenting adversarial testing of the model with a view to identifying and mitigating systemic risks; (b) assess and mitigate possible systemic risks at Union level, including their sources, that may stem from the development, the placing on the market, or the use of general-purpose AI models with systemic risk; [...] (d) ensure an adequate level of cybersecurity protection for the general-purpose AI model with systemic risk and the physical infrastructure of the model.'

might not be able to request that information, as such authority is not expressly provided for in the text of Article 55. The AI Office would still be able to request information on the basis of Article 91(1), though arguably, in this interpretation, there might be little to request, since Article 55 does not explicitly require providers of GPAI models with systemic risk to document those assessments.

41. The Code of Practice seems to support the more extensive 'linked' interpretation, as it exclusively considers Annex XI Section 2's obligations together with the Article 55(1) obligations. Moreover, the Code of Practice imposes documentation for various of the obligations found in Article 55[80] – with some exceptions.[81] Interestingly, the Code of Practice also indicates that the information that should be provided on the basis of Article 53, and thus Annex XI Section 2, is more detailed than the information required under Article 55 alone.[82] While the text of the AI Act leaves open whether the evaluations and tests to be documented under Section 2 of Annex XI necessarily pertain to the systemic risk assessments mandated by Article 55(1), given that Annex XI does not mention systemic risk, that interpretation is strongly implied by the Code of Practice, which treats the Annex's obligations (to be read as part of Article 53(1)) in conjunction with those arising under Article 55(1).

2.1.1.2.2. Content

42. The information that should be documented pursuant to Annex XI Section 2 entails:

- '1. A detailed description of the evaluation strategies, including evaluation results, on the basis of available public evaluation protocols and tools or otherwise of other evaluation methodologies. Evaluation strategies shall include evaluation criteria, metrics and the methodology on the identification of limitations.'

43. The provider of a GPAI model with systemic risk should describe the benchmarks used to evaluate the model's performance. Interestingly, Annex XI does not go so far as to explicitly require testing for specific use cases or capabilities. Various public evaluation tools are available[83] that can be used to assess model performance across a range of aspects and applications. Examples include benchmarks for question answering[84], reasoning puzzles[85], coding challenges[86], and safety evaluations[87].

[80] E.g., Measures 1.1, 1.3, 7.1, 7.2 and 7.3 of the European Commission, 'Code of Practice for General-Purpose AI Models – Safety and Security Chapter' (2025) <https://ec.europa.eu/newsroom/dae/redirection/document/118119> accessed 1 October 2025. Also see forthcoming commentary on Article 55 in this work.
[81] E.g., ibid. Measure 10.1 third and fourth paragraph. The fourth paragraph adds that the information, required by the third paragraph, does not have to be collected but may be compiled upon the AI Office's request.
[82] This is evident from the reference to 'a high-level description of the model's architecture' in the Code of Practice Safety and Security Chapter (n 80) Measure 7.1(1), which is based on articles 55(1) and 56(5), and the more detailed reference to 'a detailed description of the model's architecture' in Measure 10.1, based on article 55(1) *and* article 53(1)(a) (and thus Annex XI s 2).
[83] E.g., MMLU, ARC-Challenge, PubMedQA, GSM8K, FrontierMath, MGSM, HellaSwag, WinoGrande, DROP, RACE-M/H, HumanEval, MBPP, BIG-Bench-Hard, AMC, GRE, AI2D, MMMU, DocVQA, MathVista, BBQ, and WildBench.
[84] E.g., GPQA.
[85] E.g., MGSM.
[86] E.g., HumanEval.
[87] E.g., BBQ for bias evaluation. The adequacy of specific benchmarks and evaluations is discussed in more detail in forthcoming commentary on Article 55 in this work.

44. Providers are not, however, obliged to use established or publicly available benchmarks under Section 2 of Annex XI. Arguably, though, they should document additional information regarding their evaluation procedure if unconventional or non-public benchmarks are employed. Particularly in such cases, it is important to elaborate on the evaluation criteria, metrics and methodology for identifying limitations. Evaluation criteria might include the accuracy on question answering[88], the extent of bias displayed, or the frequency of safety refusals.[89] Relevant metrics could then include accuracy percentages, bias scores, or refusal rates. The provider should also set out how specific limitations in the model's capabilities were identified through the evaluation process, and the procedures used to uncover them.

45. A key observation, based on reporting practices prior to the application and enforcement of the AI Act, is that there is little consistency in what providers disclose and that their reports tend to lack sufficient information – sometimes using vague claims such as 'above human average' – rather than reporting in-depth testing results.[90] Moreover, providers might be inclined to withhold certain information for competitive or marketing reasons, or may selectively highlight other data for similar purposes. Nevertheless, as discussed further below,[91] the requirements under Article 53 and Annex XI should arguably be interpreted as requiring extensive documentation, as such an interpretation most closely aligns with the rationale of these requirements.

- '2. Where applicable, a detailed description of the measures put in place for the purpose of conducting internal and/or external adversarial testing (e.g. red-teaming), model adaptations, including alignment and fine-tuning.'

46. There is a similar requirement to document adversarial testing and model adaptations. Here, too, the provision does not appear to go so far as to prescribe such testing – given the phrase 'where applicable' – but merely requires its documentation if it has taken place, for example on the basis of an obligation found in Article 55(1)(a).[92]

47. Adversarial testing refers to exercises where designated users deliberately attempt to elicit harmful, unsafe or otherwise undesirable outputs from the model in order to uncover risks and weaknesses.[93] For internal adversarial testing, this involves in-house teams red-teaming or attempting to break the model using challenging inputs. Depending on the level of detail required, which arguably depends on the size of the provider,[94] this documentation should include which scenarios these teams tested (such as prompts designed to elicit biased responses or to obtain confidential or harmful outputs), as well as the

[88] E.g., BBQ evaluation accuracy.
[89] Also see Tegan McCaslin and others, 'STREAM (ChemBio): A Standard for Transparently Reporting Evaluations in AI Model Reports' (2025) <https://arxiv.org/abs/2508.09853>. Various of these examples are commonly included in model cards, e.g., OpenAI, *GPT-5 System Card* (13 August 2025) <https://cdn.openai.com/gpt-5-system-card.pdf> accessed 1 October 2025.
[90] See, e.g., McCaslin and others (n 89) s 3.
[91] See para 56.
[92] See forthcoming commentary on Article 55 in this work
[93] E.g., Y Kumar and others, 'Adversarial Testing of LLMs Across Multiple Languages' (International Symposium on Networks, Computers and Communications, Washington, DC, 2024) <https://doi.org/10.1109/ISNCC62547.2024.10758949>, 1.
[94] See para 3 and AI Act, recital 109.

findings. For external adversarial testing, external experts or third parties are engaged to perform similar evaluations. Here, the documentation should arguably specify who was involved and their relevant expertise and should detail the scope and results of the testing.[95]

48. Model adaptations,[96] including alignment and fine-tuning, refer to modifications made to the model following its initial testing. It is common practice to first train a GPAI model using a large dataset before applying further guidelines or restrictions to better align the model's output with human values or improve safety.[97] Techniques that may be employed – and should be described if so – include reinforcement learning with human feedback, or additional training with specialised or curated data.[98]

49. While Annex XI Section 2(2) only mandates these descriptions 'where applicable', the implementation of such measures is widespread.[99] One could also argue that they are required by Article 55(1).[100] Furthermore, if the model is intended for deployment in a high-risk system, it could be argued that such modification is generally required (for example in the context of Article 9 AI Act) to enable a downstream system provider to meet their obligations.[101] Nevertheless, it is conceivable that model providers could restrict themselves to their obligations under Article 55 and delegate any additional fine-tuning to providers intending to implement the model in their high-risk system, thus limiting the information that needs to be documented according to Annex XI Section 2.

50. Here, too, these elements should be described in detail, rather than superficially. This is discussed in more detail below.[102]

- '3. Where applicable, a detailed description of the system architecture explaining how software components build or feed into each other and integrate into the overall processing.'

51. The third part of Section 2 similarly employs the terminology 'where applicable'. Notably, the Code of Practice subjects this requirement to an alternatively phrased condition – namely, 'insofar as the Signatory is aware of such information'.[103] Thus, it can generally be assumed that this information is required, except where the provider of the GPAI model with systemic risk is entirely unaware of it – which would, as discussed more extensively below,[104] significantly restrict their ability to market the model.

---

[95] See, in this sense, article 55(1)(a) as well as the Code of Practice Safety and Security Chapter (n 80) Measures 7.3 and 7.4. Also see forthcoming commentary on Article 55 in this work.
[96] Also see, on model modification, the forthcoming chapter on Modifications in this work.
[97] E.g., Humza Naveed and others, 'A Comprehensive Overview of Large Language Models' (2024) <https://arxiv.org/abs/2307.06435> ss 1 and 2.
[98] E.g., ibid. s 2 (on reinforcement learning with human feedback).
[99] E.g., Anusha Sinha and others, 'What Can Generative AI Red-Teaming Learn from Cyber Red-Teaming?' (Technical Report CMU/SEI-2025-TR-006, July 2025) <https://www.sei.cmu.edu/documents/6301/What_Can_Generative_AI_Red-Teaming_Learn_from_Cyber_Red-Teaming.pdf> accessed 1 October 2025 (on the widespread nature of red-teaming).
[100] See in more detail forthcoming commentary on Article 55 in this work.
[101] Also see para 60.
[102] See para 56.
[103] See Code of Practice Safety and Security Chapter (n 80) Measure 10.1.
[104] See paras 59 ff.

52. It is interesting that Annex XI Section 2 refers to the 'system architecture' of the GPAI model with systemic risk, rather than the 'model architecture'.[105] While the Code of Practice does discuss the need to document the 'model's architecture',[106] it subsequently shifts the focus to AI systems by requiring 'a detailed description of how the model is integrated into AI systems, explaining how software components build or feed into each other and integrate into the overall processing [...]',[107] which seems to indicate that Annex XI Section 2(3) should be understood as largely referring to systems implementing the model, rather than to the model itself. This interpretation is more consistent with the phrases 'where applicable' and 'insofar as the Signatory is aware of such information', as there may either be no system implementation (yet), or the provider of the GPAI model with systemic risk may not be aware of the specific way in which the model was implemented in a system. It is also more consistent with Annex XI, Section 1(1)(d), which requires documentation of the model architecture, as it would be illogical for both provisions to mandate identical information.

53. As such, this provision requires the provider of the GPAI model with systemic risk to provide information about the way in which the model is integrated in a specific system. That description must also address the interaction between these components, which should arguably cover what input is fed to the GPAI model with systemic risk and how the system processes that model's output.

54. Interestingly, the wording of Annex XI Section 2(3) mirrors that of Annex IV(2)(c),[108] which sets out the technical documentation required for high-risk systems (Article 11). In that sense, such information is likely to be required for most GPAI models, given the interaction with high-risk system requirements described below.[109]

2.1.1.3. Modalities and exceptions

2.1.1.3.1. Modalities and extent

55. Regarding the modalities of these obligations, a first aspect concerns how long providers should preserve the relevant information. The Code of Practice chapter on safety and security indicates that GPAI model providers should retain the relevant information for ten years.[110]

56. As discussed more extensively above regarding Annex XI Section 2 and GPAI models with systemic risk,[111] there are some arguments that would support a broad interpretation of the relevant documentation requirements. A broad and thorough approach would best enable the AI Office and national competent authorities to exercise their competences under the AI Act. This interpretation is

[105] E.g., AI Act, arts 53(2) and 54(6).
[106] See forthcoming commentary on Article 55 in this work.
[107] Measure 10.1 Code of Practice Safety and Security Chapter (n 83).
[108] '2. A detailed description of the elements of the AI system and of the process for its development, including: […] (c) the description of the system architecture explaining how software components build on or feed into each other and integrate into the overall processing; the computational resources used to develop, train, test and validate the AI system.'
[109] See paras 59 ff.
[110] Code of Practice Safety and Security Chapter (n 80) Measure 3.2 as well as app 3.
[111] See para 45.

further supported – even under a restrictive reading – by the duty to cooperate, discussed below.[112] This is especially pertinent to the level of detail required in the documentation. It is therefore insufficient, for instance, to describe a model's performance only in relative terms ('below-human capabilities') while omitting the precise test score. That said, it is equally clear that the extent of documentation and level of detail also depend on the size of the provider of the GPAI model.[113]

57. The Transparency[114] and Safety and Security[115] chapters of the Code of Practice further clarify that this information, if required, does not generally[116] need to be shared proactively or made publicly available, but should only be provided to the AI Office and national competent authorities upon request.

2.1.1.3.2. Open-source exception

58. Lastly, it is important to note that these obligations do not apply to open-source models, as discussed below.[117] This exception does not apply if the open-source GPAI model presents systemic risks.[118]

2.1.2. Article 53(1)(b): Transparency for downstream AI system providers

2.1.2.1. Context

59. Article 53(1)(b) requires GPAI model providers to make information and documentation available to system providers intending to integrate the model into their AI system. To properly understand this requirement, it is useful to briefly contextualise it: any AI system provider seeking to implement a GPAI model will themselves be subject to the AI Act's provisions on AI system providers, including – if the relevant requirements are met – the provisions regarding prohibited practices and high-risk systems. Access to key information about the integrated GPAI model is therefore essential for system providers to ensure compliance with these obligations.[119] The associated penalties (Articles 99–101 AI Act) are sufficiently severe that one could reasonably expect system providers to refrain from incorporating any GPAI model that would hinder their compliance and thereby their ability to market their system without incurring significant sanctions. Consequently, in the context of prohibited AI practices and high-risk systems, there exists a strong upstream incentive to provide the information necessary to assess whether incorporating a given GPAI model into the intended system configuration might result in such a prohibited practice or a (non-)compliant high-risk system. Specifically for those high-risk systems, the AI Act's enforcement mechanism generates a robust upstream incentive to incorporate only those (GPAI) models that allow – one could even say *facilitate* – high-risk system

[112] See Section 2.3.
[113] AI Act, recital 109.
[114] Measures 1.1 and 1.2 of the Code of Practice Transparency Chapter (n 17).
[115] Measure 10.1 Code of Practice Safety & Security Chapter (n 80).
[116] Exceptions do exist, such as in Measure 10.2 Code of Practice Safety and Security Chapter (n 80), which implements public transparency of the framework and model reports '[i]f and insofar required to assess and/or mitigate systemic risks'.
[117] See Section 2.2.
[118] AI Act, art 53(2), last sentence.
[119] See similarly: AI Act, recital 101; Schneider (n 4) para 13.

providers to comply with the AI Act's requirements. We will discuss some of the relevant high-risk system requirements in more detail below.[120]

60. As a result, the explicit requirement to inform downstream AI system providers to enable compliance (Article 53(1)(b)(i)), and more generally about the GPAI model, may appear somewhat remarkable, particularly with respect to information that would be crucial for such providers even absent a formal duty of disclosure. Admittedly, some of the information and documentation mandated by Article 53(1)(b) goes beyond what is directly relevant to assessing compliance with prohibited practices and high-risk system requirements, but the observation holds for those sections that do overlap.[121] It also applies to certain information that is essential for the downstream system provider to implement the model in the first place.[122] As such, the inclusion of these provisions appears to assume failures[123] or frictions in information exchange.[124] While such an assumption is, to some extent, justified given the complexity, unpredictability and opacity of many AI models and systems[125] – particularly those deployed in GPAI contexts – it is given a remarkable interpretation here, as is discussed more extensively below.[126]

61. An interesting and arguably desirable consequence of this approach is that some of the concerns that typically dominate high-risk AI system compliance will be brought to the attention of low-risk AI system providers as well if they decide to implement a GPAI model in their system. As such, the mechanism of requiring the GPAI model provider to share information with the integrating system provider results in some effects akin to some of the high-risk requirements – increasing awareness among low-risk system providers – without directly regulating those systems. Article 53 ensures that providers of such low-risk systems have access to more information and records than they would be required to maintain had they not incorporated a GPAI model.[127]

[120] See paras 73 ff.

[121] This is particularly clear for Annex XII (2)(c), in light of article 10 AI Act.

[122] Also see Annex XII (1)(d) and (e) concerning information that is highly relevant for commercial/marketing purposes.

[123] See for similar reservations of market-based solutions for compliance: Alexander Peukert, 'Copyright in the Artificial Intelligence Act – A Primer' (2024) 73 GRUR International 497, 507.

[124] See, in general, on the Coase theorem which describes the absence of such failures or frictions in ideal circumstances: R. H. Coase, 'The Problem of Social Cost' (1960) 3 Journal of Law and Economics 1, 1 ff; Christine Jolls, Cass R. Sunstein and Richard Thaler, 'A Behavioral Approach to Law and Economics' (1998) 50 Stanford Law Review 1471, 1483; Russell B. Korobkin and Thomas S. Ulen, 'Law and Behavioral Science: Removing the Rationality Assumption from Law and Economics' (2000) 88 California Law Review 1051, 1094-1095; Steven Shavell, *Foundations of Economic Analysis of Law* (Belknap Press of Harvard University Press 2004) 84 and 102 ff.

[125] E.g., Maarten Herbosch, 'Liability for AI Agents' (2025) 26(3) North Carolina Journal of Law & Technology 391, 412, fn 114.

[126] See paras 70 ff.

[127] One could argue, though, that this effect is not entirely due to the AI Act's requirements, as this could, more generally, follow from the proper functioning of the market as well in cases where GPAI model providers choose to model the same model similarly to integrating high- and low-risk system providers alike, also see para 77.

2.1.2.2. Listed requirements

2.1.2.2.1. General

62. Article 53(1)(b) requires GPAI model providers to make information and documentation available to AI system providers intending to integrate the relevant GPAI model in an AI system. This requirement – described as 'transparency' in Annex XII – consists of two components. To quote Article 53, the information and documentation 'shall:

> (i) enable providers of AI systems to have a good understanding of the capabilities and limitations of the general-purpose AI model and to comply with their obligations pursuant to this Regulation; and
>
> (ii) contain, at a minimum, the elements set out in Annex XII*'*

63. The first of these two components does not specify particular information to be provided or documented, but rather clarifies the objective of the documentation or information requirement. The second element refers to Annex XII and adds that the relevant information and documentation should 'at a minimum'[128] contain:

> '1. A general description of the general-purpose AI model including:
>
> (a) the tasks that the model is intended to perform and the type and nature of AI systems into which it can be integrated;
>
> (b) the acceptable use policies applicable;
>
> (c) the date of release and methods of distribution;
>
> (d) how the model interacts, or can be used to interact, with hardware or software that is not part of the model itself, where applicable;
>
> (e) the versions of relevant software related to the use of the general-purpose AI model, where applicable;
>
> (f) the architecture and number of parameters;
>
> (g) the modality (e.g. text, image) and format of inputs and outputs;
>
> (h) the licence for the model.
>
> 2. A description of the elements of the model and of the process for its development, including:
>
> (a) the technical means (e.g. instructions for use, infrastructure, tools) required for the general-purpose AI model to be integrated into AI systems;

[128] Reiterated in Annex XII as 'at least'.

(b) the modality (e.g. text, image, etc.) and format of the inputs and outputs and their maximum size (e.g. context window length, etc.);

(c) information on the data used for training, testing and validation, where applicable, including the type and provenance of data and curation methodologies.'

64. Many – though not all[129] – of these requirements largely mirror those found in Annex XI, discussed earlier,[130] albeit generally with a lower level of detail. The Code of Practice helps to illustrate some of these differences.[131]

2.1.2.2.2. Additional elements

65. In comparison to Annex XI, Annex XII(1) adds two further elements:

- '(d) how the model interacts, or can be used to interact, with hardware or software that is not part of the model itself, where applicable'

66. The model provider should supply information regarding the relevant hardware required to use the model. The Code of Practice adopts a rather minimal interpretation of this provision, requiring the GPAI model provider only to describe the technical means for model integration (such as 'instructions for use' or 'infrastructure tools'), and, where applicable, the hardware and software required (including its version).[132] As a result, this involves a rather limited level of detail compared to how this requirement is phrased in Annex XII(1). This is, in part, nuanced because certain elements discussed above,[133] such as a description of the type and nature of AI systems in which the GPAI model can be integrated, also need to be communicated to the downstream provider.

67. Nevertheless, a more extensive interpretation – more closely aligned with the literal wording of Annex XII – could be preferred by downstream providers. For example, it is highly useful for them to know whether the model can (or must) connect to external software or hardware, such as being designed to send commands to external tools or applications. It is also important for downstream providers to know whether the model can or should be integrated with third-party applications, such as email or text processing software, CRM systems or social media platforms. At the same time, they can likely be expected to obtain these details irrespective of the AI Act's obligations; that is, the frictions characterizing information exchange in the AI context[134] are unlikely to affect these elements, as they do not directly pertain to the complexity of AI models.

- '(e) the versions of relevant software related to the use of the general-purpose AI model, where applicable'

[129] See paras 65 ff.
[130] See paras 12 ff.
[131] Code of Practice Model Documentation Form (n 22) 1–3
[132] Code of Practice Model Documentation Form (n 22) 2.
[133] See paras 12 ff.
[134] See para 60.

68. Annex XII(1) also requires GPAI model providers to inform downstream system providers of the relevant software versions that might be related to the use of the GPAI model. The Code of Practice gives this provision a similarly minimal interpretation, requiring the model provider only to submit information about the software – including the relevant version – necessary to use the model.[135] Similar considerations as for the model interactions, discussed above, apply.

69. Interestingly, Annex XII does not impose explicit requirements regarding the model's energy consumption, but it does require providers to disclose the maximum size of the model's output. Some authors argue that the hardware requirements discussed above should be interpreted even more broadly to encompass the model's energy performance,[136] though this approach does not appear to be supported by the wording of the Annex or by the interpretation set out in the Code of Practice.

2.1.2.3. Understanding and compliance requirements

70. Article 53(1)(b)(ii) provides that the relevant information must include, 'at a minimum,' the elements set out in Annex XII. In addition, Article 53(1)(b)(i), in turn, points to other information that should be communicated, irrespective of Annex XII's minimal requirements. A first element is that the provider of the downstream AI system must be able to understand the capabilities and limitations of the GPAI model. This is a sensible requirement that strongly underpins the first paragraph of Annex XII. Where the peculiarities of the GPAI model, or advances in the state-of-the-art in GPAI techniques, necessitate relevant information beyond that listed in Annex XII for proper understanding, the GPAI model provider should supply it.[137]

71. A second element under Article 53(1)(b)(i) requires that the information provided enable downstream system providers 'to comply with their obligations pursuant to this Regulation.' This requirement is notable for several reasons.

72. First, it is remarkable that such a broad obligation is not further clarified in Article 53(1)(b)(i) itself, nor is it more explicitly supported, e.g., by the transparency requirements in Annex XII. Nevertheless, this information is particularly significant for providers of downstream high-risk AI systems, who are subject to extensive obligations under Articles 8–15 of the AI Act.[138] The more specific requirements made explicit in Annex XII appear insufficient to ensure compliance with some of these obligations.

73. This is more apparent for certain requirements than others. For instance, Article 9's requirement to develop a risk management system does find some support by the obligation to provide 'risk profile'-appropriate information under Annex XII, although this raises the question of what the express requirement in Article 53(b)(i) to provide this information truly adds, given that much of this information would already be necessary for high-risk system compliance, as discussed earlier.[139]

---

[135] Code of Practice Model Documentation Form (n 22) 2.
[136] Alder, Ebert, Herbrich and Hacker (n 54) s 2.
[137] This applies in addition to the obligation to keep the model form itself up to date, see Code of Practice Transparency Chapter (n 17) Measure 1.1.
[138] Also see paras 59 ff.
[139] See paras 59 ff.

74. Article 10 may be more demanding, setting out extensive data governance requirements for high-risk system training. Relatedly, Annex XII(2)(c)'s requirement that the GPAI model provider should supply 'information on the data used for training, testing and validation, where applicable, including the type and provenance of data and curation methodologies' should be extended in such contexts, in line with the downstream compliance information requirement established by Article 53(1)(b)(i). Moreover, the likely deployment of a GPAI model in downstream high-risk systems imposes specific training data requirements[140] that are not made explicit in Chapter V of the AI Act[141] but may nonetheless prove crucial.

75. Articles 11 and 13 arguably sit at the core of the particular relationship between downstream GPAI model information requirements and the implicit assumption that the market would fail to secure these discussed earlier.[142] These provisions impose extensive technical documentation and transparency obligations on high-risk system providers. While similar in nature to the information required by Annex XII, compliance with Articles 11 and 13 requires system providers to obtain far more extensive information from the GPAI model provider whose model they implement.

76. Notably, several of Article 13's requirements, including certain explainability requirements, extend well beyond what is required under Article 53 and Annex XII. The same is true for some of the more substantive high-risk requirements set out in Articles 14 (enabling human oversight) and 15 (concerning accuracy, robustness and cybersecurity). For such provisions, the AI Act appears to implicitly rely on market functioning to ensure that GPAI model providers adapt their models in ways that enable compliance by downstream high-risk system providers, without explicitly obliging them to do so. At the same time, the Act is far more reluctant to assume that providers would share the relevant information, even though – by assumption – they would already have adapted their models due to market incentives. Consequently, the Act's approach to market functioning in this respect appears inconsistent rather than systematic.

77. Notably, these requirements for GPAI model providers whose model is implemented in high-risk systems may have an interesting spillover effect. GPAI model providers could market their models broadly – not limited to either high- or low-risk applications[143] – and thus may need to ensure that their documentation and information, or even the model itself, meets the standards set for high-risk systems. For reasons of convenience, they may choose to share this information more broadly with prospective downstream system providers,[144] even if those providers intend to use the model in a low-risk system.

---

[140] E.g., AI Act, arts 10(3) and (4) ('3. Training, validation and testing data sets shall be relevant, sufficiently representative, and to the best extent possible, free of errors and complete in view of the intended purpose. They shall have the appropriate statistical properties, including, where applicable, as regards the persons or groups of persons in relation to whom the high-risk AI system is intended to be used. Those characteristics of the data sets may be met at the level of individual data sets or at the level of a combination thereof.
4. Data sets shall take into account, to the extent required by the intended purpose, the characteristics or elements that are particular to the specific geographical, contextual, behavioural or functional setting within which the high-risk AI system is intended to be used.').

[141] E.g., Section 2.1.3.

[142] See paras 59 ff.

[143] They could, of course, also decide to develop separate models for high-risk system integration and low-risk system integration.

[144] Also see para 61.

As a result, GPAI model (self-[145]) regulation may indirectly impose certain high-risk system requirements on low-risk systems through the adoption of a shared, high-risk-compatible GPAI model.

2.1.2.4. Limitations & modalities

78. These information and documentation requirements are subject to several important limitations. First, there is the open-source exception in Article 53(2), discussed below. Secondly, Article 53(1)(b) limits the information and documentation obligations outlined above by stating that their scope is '[w]ithout prejudice to the need to observe and protect intellectual property rights and confidential business information or trade secrets in accordance with Union and national law'.

79. The 'confidential business information'[146] and 'trade secrets'[147] limitation is particularly pertinent here, as leading GPAI models often rely on proprietary techniques that providers would be reluctant to see disclosed to competitors or the wider public. Both concepts allow the GPAI model provider a degree of discretion, provided the information is not generally known.[148] By limiting transparency obligations for confidential material – as the AI Act and its GPAI provisions do elsewhere[149] – the Act seeks to balance the need for information to enable downstream compliance with the commercial interest in retaining certain information. The 'need to observe and protect' such information likely encompasses the use of non-disclosure agreements for its sharing. Notably, however, the Act addresses this tension only in relation to the explicit information obligations in Article 53(1)(b), without addressing its wider implications, for example regarding the broader compliance requirements for high-risk system providers.

80. The parallel 'intellectual property rights' reservation is likely intended to address European copyright provisions. This primarily concerns copyright,[150] which could protect aspects of the model's computer code or documentation, as well as database rights[151] that may cover databases used for training. This is particularly relevant where third parties hold the relevant rights, potentially limiting the GPAI model provider's ability to share such information. Awarded patents are likely to play a more limited role in this context,[152] as they do not typically entail confidentiality requirements, though the situation may differ in the case of unpublished patent applications. Much like with copyright, Article 53(1)(b) does not require GPAI model providers to breach their own confidentiality obligations,[153] for example when

[145] For low-risk applications, such requirements are not required by the AI Act but can, to some extent, be expected to be self-enforced as GPAI providers would likely prefer the ability to market their model widely, to both low- and high-risk system providers alike.
[146] This notion likely refers to information akin to trade secrets that does not meet the requirements imposed by article 2(1) Directive (EU) 2016/943 of the European Parliament and of the Council of 8 June 2016 on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use and disclosure [2016] OJ L 157/1.
[147] The notion likely refers to the definition found in article 2(1) of Directive (EU) 2016/943.
[148] See, in particular, art 2(1)(a) Directive (EU) 2016/943.
[149] E.g., art 25(5), art 52(6), art 53(7) (see Section 2.7), art 55(3), and art 78(1)(a), as well as Annex VII (4.5).
[150] E.g., Directive 2009/24/EC of the European Parliament and of the Council of 23 April 2009 on the legal protection of computer programs [2009] OJ L 111/16.
[151] E.g., Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases [1996] OJ L 77/20.
[152] They could, for example, cover some of the technical methods used to create the model or its training.
[153] See similarly Schneider (n 4) para 19.

they did not develop the GPAI model alone and/or are otherwise bound by non-disclosure agreements, or jeopardise the confidentiality of their know-how.

81. Both the reservation regarding confidential business information/trade secrets and the reservation regarding intellectual property rights arguably restrict the level of detail and types of information that GPAI model providers must disclose. This is reinforced by the less detailed requirements in Annex XII compared to Annex XI.[154] These limitations also suggest that GPAI model providers may be incentivised to interpret Article 53(1)(b) narrowly to maximise the protection of their trade secrets, confidential business information, and intellectual property interests. There is a direct risk, therefore, that the balance between downstream providers' need for transparency and the confidentiality interests of GPAI model providers could be determined solely at the discretion of the latter. In this context, it is worth highlighting two important safeguards for downstream providers. First, Article 53(1)(b)(ii) introduces a minimum safeguard by requiring communication of certain key elements.[155] Second, GPAI model providers must ensure that the information they do share 'enable[s] providers of AI systems to have a good understanding of the capabilities and limitations of the general-purpose AI model and to comply with their obligations pursuant to this Regulation'.[156] Both of these 'safeguard provisions' can be enforced effectively,[157] as similar information must be provided in more detail to the AI Office and national competent authorities upon request (Article 53(1)(a)). In other words, the discretion granted to GPAI model providers under Article 53(1)(b)'s confidentiality and IP provisions is, arguably, sufficiently constrained, with adequate incentives to implement their obligations in a manner that safeguards the interests of downstream system providers.

82. As the model developer is required to inform system providers who 'intend' to integrate the model into their systems, an implication is that these providers should generally have access to the relevant information before incorporating the model.[158] The text is less clear, however, as to whether this means they should receive the information prior to making a decision or before, for example, becoming contractually bound to implement the model. Both the level of detail in Annex XII and the wording of Article 53(1)(b)(i) suggest that the primary aim of the information is to ensure compliance, rather than to inform prospective downstream providers before they select a model – though it should also ensure that they have a clear understanding of the model's limitations and capabilities.

83. In any case, this provision does not preclude prior non-disclosure agreements.[159] While the Code of Practice states that GPAI model providers should provide this information 'proactively',[160] the need to preserve confidentiality may in practice require prospective system implementers to contact the provider before receiving the relevant documentation. Moreover, given the express reference to 'confidential business information' and 'trade secrets', the GPAI model provider should be assumed to be able to refuse abusive requests, for example by system providers that clearly would not adhere to

[154] See para 64.
[155] See para 70.
[156] See paras 62 ff.
[157] Also see AI Act, art 101.
[158] See similarly Schneider (n 4) para 15.
[159] See para 79.
[160] Code of Practice Model Documentation Form (n 22) 1.

the model's use policy or by competitors or their proxies without an actual intent to integrate the model.[161]

84. Interestingly, unlike Annex XI, neither Annex XII nor Article 53(1)(b) more generally refer to the risk profile of the model.[162] This is noteworthy, as Recital 101 emphasises the importance of proportionality in this context.[163] Such a requirement would also be sensible: depending on whether the downstream system qualifies as a high-risk system or one liable to be used in a prohibited AI practice, its provider is subject to more stringent obligations.[164] While GPAI model providers may be inclined to adopt a 'one size fits all' approach by supplying the same information to different downstream providers,[165] this is arguably addressed by the requirement that the information provided must enable downstream providers 'to comply with their obligations pursuant to this Regulation', which implies that less information should be required for low-risk systems. Some authors nevertheless argue that Recital 101's principle of proportionality should be read more fully into Article 53(1)(b) and Annex XII,[166] limiting the extent of information to be provided according to the size and risk profile of the GPAI model. At the same time, Recital 109 introduces a countervailing consideration,[167] imposing more stringent requirements on larger model providers, which would appear to restrict this interpretation to relatively small GPAI providers offering smaller models with more nuanced risk profiles. In any event, the impact of a proportional approach is limited, given that Article 53(1)(b) and Annex XII are already less demanding than Article 53(1)(a) and Annex XI – meaning that any obligation to share more information with downstream providers remains comparatively lighter than under the latter provisions.

2.1.3. Article 53(1)(c): Copyright compliance policy

2.1.3.1. Positioning and scope

2.1.3.1.1. Union copyright

85. Article 53(1)(c) requires GPAI model providers to establish a policy to ensure compliance with Union law and copyright and related rights regulation.[168] In particular, this policy should ensure – including through state-of-the-art technologies – that the model provider processes the text and data mining (TDM) opt-out specified in Article 4(3) of Directive (EU) 2019/790. TDM refers to the automated analysis – including for AI model training[169] – of text and data to identify patterns and extract useful information, often using web crawlers.[170] Rightsholders must permit TDM by research organisations

[161] See similarly Schneider (n 4) para 15.
[162] See para 35.
[163] Schneider (n 4) para 18.
[164] See para 59.
[165] See para 77.
[166] Schneider (n 4) para 18.
[167] See para 3.
[168] Also see Katharina de la Durantaye, 'Nutzung urheberrechtlich geschützter Inhalte zum Training generativer künstlicher Intelligenz – ein Lagebericht' (2024) 55 AfP 9, 16-17.
[169] E.g., Jan Bernd Nordemann and Arman Rasouli, 'Die Regelungen der KI-Verordnung mit Urheberrechtsbezug – Möglichkeit der privaten Rechtsdurchsetzung?' (2024) Zeitschrift für Urheber- und Medienrecht 780.
[170] See on those web-crawlers also European Commission, 'Code of Practice for General-Purpose AI Models – Copyright Chapter' (2025) <https://ec.europa.eu/newsroom/dae/redirection/document/118115> Measure 1.2(1).

and cultural heritage institutions for scientific research purposes.[171] For commercial purposes, TDM is permitted by default unless the rightsholder opts out.[172] If an opt-out is exercised, it must be in machine-readable form.[173] In the absence of an opt-out, TDM is permissible.[174] This obligation is reiterated by various measures in the Code of Practice.[175]

86. Where there is no opt-out, the party wishing to conduct TDM must have lawful access to the data and text,[176] meaning, for example, that they must not circumvent subscription models or paywalls.[177] Additionally, the Code of Practice states that GPAI model providers should 'exclude from their web-crawling websites that make available to the public content and which are, at the time of web-crawling, recognised as persistently and repeatedly infringing copyright and related rights on a commercial scale by courts or public authorities in the European Union and the European Economic Area'.[178] The Code of Practice further emphasises the efforts – including deploying state-of-the-art technologies[179] – that GPAI model providers should undertake to identify opt-outs.[180] It also provides that GPAI model providers who also operate a search engine should not interpret a TDM opt-out as also constituting an opt-out from search indexing.[181] If a TDM opt-out were interpreted that broadly, exercising it for copyright reasons could prove punitive for the rightsholder, as it would cut off an important stream of traffic to their content.

2.1.3.1.2. Criticism and scope

87. The inclusion of copyright in the AI Act has attracted some criticism,[182] as it conflates the private rights nature of copyright protection with the public interest objectives of the AI Act.[183] It is also notable that

[171] Article 3 Directive (EU) 2019/790.
[172] Article 4 Directive (EU) 2019/790.
[173] Generally in the metadata, the terms of use or the robots.txt file. See article 4(3) Directive (EU) 2019/790. Also see Measure 1.3 of the Code of Practice Copyright Chapter (n 170). Over time (e.g. article 53(1)(c) or Measure 1.3 (1)(b)'s reference to state-of-the-art), this may come to include natural language opt-outs, as has already been argued (see, e.g., *Robert Kneschke v LAION e.V.* (Regional Court of Hamburg, 27 September 2024), 310 O 227/23).
[174] Article 4(1) Directive (EU) 2019/790.
[175] Code of Practice Copyright Chapter (n 170) Measures 1.2 and 1.3.
[176] Article 4(3) Directive (EU) 2019/790. Also see Schneider (n 4) para 21.
[177] Also see Code of Practice Copyright Chapter (n 170) Measure 1.2(1)(a). Such paywalled data is generally more valuable for model finetuning for specific applications than publicly accessible data is, see Katharina de la Durantaye, 'Akkommodation statt Assimilation. Warum die EU bei der KI-Regulierung nicht auf den Brussels Effect setzen sollte – und was stattdessen sinnvoll wäre' (2025) Zeitschrift für Urheber- und Medienrecht 165, 173.
[178] To this end, Measure 1.2(1)(b) adds 'For the purpose of compliance with this measure, a dynamic list of hyperlinks to lists of these websites issued by the relevant bodies in the European Union and the European Economic Area will be made publicly available on an EU website'.
[179] Also see the text of art 53(1)(c) AI Act.
[180] Code of Practice Copyright Chapter (n 170) Measure 1.3(1)(b). Also see Schneider (n 4) para 20; João Pedro Quintais, 'Generative AI, Copyright and the AI Act' (2025) 56(106107) Computer Law & Security Review <https://doi.org/10.1016/j.clsr.2025.106107> 1, 9-10.
[181] Code of Practice Copyright Chapter (n 170) Measure 1.3 (5).
[182] Some would argue that the AI Act should have provided copyright exceptions rather than enforce existing copyright, e.g., David Bomhard and Jonas Siglmüller, 'AI Act – das Trilogergebnis' Recht Digital 45, 54. Also see de la Durantaye, 'Akkommodation statt Assimilation.' (n 177) 167 ff (criticising the ambitious territorial scope of the copyright provisions and their assumption of a Brussels effect).
[183] Peukert (n 126) 498-499. Also see (without criticising this) Nordemann and Rasouli (n 169) 780.

the Act imposes a specific obligation on GPAI model providers to comply with copyright law, without imposing a similar requirement – despite a generally higher level of depth and detail in the relevant provisions – on high-risk system providers or, indeed, on all AI model and system providers.[184]

88. Interestingly, a dominant reading gives this obligation a strikingly broad territorial scope.[185] Recital 106 states that GPAI model providers should comply with copyright law 'regardless of the jurisdiction in which the copyright-relevant acts underpinning the training of those general-purpose AI models take place.'[186] This extension – sometimes characterised as a 'maximalist' interpretation[187] – is not confirmed by the text of the regulation itself. Notably, it would also extend the traditional territorial scope of EU copyright law.[188] A more 'minimalist' reading would suggest that the territorial reach of EU copyright law should set an upper limit to the AI Act's effect in this area:[189] as EU copyright law and the TDM opt-out do not apply to a foreign developer training their model outside the EU, such developers could argue that they comply with EU copyright law irrespective of their training practices (i.e. even if their actions would violate EU copyright law if conducted within the EU).[190] This would mean that the EU TDM exception could become highly relevant in foreign jurisdictions, even where it is not part of the local copyright regime.[191] A less convincing intermediate approach would require compliance with the TDM opt-out only if the data is hosted on a European server.[192]

89. The maximalist interpretation is most consistent with the objectives of the AI Act, as set out in Recital 106. This 'product regulation' approach[193] is further reflected in other provisions of the Act, such as the requirement that training data for high-risk systems meet specified anti-bias standards.[194] Resultingly, even if parties develop their model or system outside the Union, their conduct during development becomes subject to the AI Act – and, in this case, to EU copyright law – if they subsequently put their system into service within the EU.[195] This interpretation best reflects the intention to create a level playing field, ensuring that developers outside the EU cannot exploit more lenient copyright regimes abroad to gain an advantage with their GPAI models within the EU.[196]

---

[184] Also see Peukert (n 126) 499-500.
[185] Also see Malte Stieper and Michael Denga, 'The International Reach of EU Copyright through the AI Act' (2024) 194 Beiträge zum Transnationalen Wirtschaftsrecht, Forschungsstelle für Transnationales Wirtschaftsrecht 1, 11 ff; de la Durantaye, 'Akkommodation statt Assimilation.' (n 177) 168.
[186] Also see, e.g., Nordemann and Rasouli (n 169) 780; Lukas 185 and Nikolaus Forgó, *KI-VO: EU-Verordnung über künstliche Intelligenz* (Verlag Österreich 2024) 371.
[187] E.g., Peukert (n 123) 506.
[188] Peukert (n 123) 506. Also see Nordemann and Rasouli (n 169) 780-781; Bernsteiner and Schmitt (n 25) para 35.
[189] Peukert (n 123) 506.
[190] Peukert (n 123) 506; Stieper and Denga (n 185) 14.
[191] Stieper and Denga (n 185) 15.
[192] Peukert (n 123) 506.
[193] Also see forthcoming chapter on Product, Model and Entity Regulation in this work. Also see Nordemann and Rasouli (n 169) 781; Quintais (n 180) 9.
[194] AI Act, art 10.
[195] Also see Peukert (n 123) 504-505; Stieper and Denga (n 185) 15. Also see forthcoming commentary on Article 2 in this work.
[196] de la Durantaye, 'Nutzung urheberrechtlich geschützter Inhalte zum Training generativer künstlicher Intelligenz' (n 168) 17; Peukert (n 126) 506; Bernsteiner and Schmitt, (n 25) para 35.

90. Some authors argue that this obligation does not mean that all AI development is covered by this provision, for example when the provider of the GPAI model (within the meaning of Article 3(3)) did not themselves develop the model but merely places on the EU market a model that was developed outside the EU by a party who does not qualify as a provider under Article 3(3) AI Act.[197] In that case, the language of Article 53(1)(c), which targets the 'provider' of the model, does not appear to apply to the model training.[198] This, however, fundamentally depends on how Article 53(1)(c) AI Act is to be understood. If it is regarded as a form of product regulation,[199] imposing requirements on how the model was developed, it could be taken to bind the provider who places the model on the market. If, instead, it is viewed as a form of entity regulation,[200] its scope would be confined to the actions of the provider of the model. While the nature of the obligation and the language of Article 53(1)(c) suggest an interpretation as entity regulation, Recital 106 could also be read as supporting a characterisation as product regulation, which the provider must ensure compliance with regardless of who trained the model beforehand.[201]

91. The obligation to comply with EU copyright law and TDM opt-outs, at least under the AI Act, is sometimes argued to be one of best efforts rather than strict liability for any (minor) breach[202] – a view supported by Recital 108[203] – because it is currently technologically impossible to perfectly filter out copyright-protected material.[204] This grants the AI Office significant authority in the copyright sphere, as it may set the relevant standard,[205] though Recital 108 makes clear that this standard should not require perfection. While the establishment of a policy is strictly required, this does thus not equate to an obligation of absolute compliance with EU copyright law. In this context, it is interesting that some German commentators have suggested that Article 53(1)(c) might constitute a so-called *Schutznorm*[206] – a notion in German law that refers to a provision intended to protect specific parties or interests from harm[207] – which would potentially allow affected individuals to derive rights from it, such as to bring liability claims. However, it is important to emphasise that a breach of copyright itself does not necessarily entail a breach of this provision – for instance, a GPAI developer could have a robust policy but still be unable, due to external factors or technical limitations, to prevent every possible copyright violation.[208] This does not, of course, preclude rightsholders from bringing claims under copyright law itself.

---

[197] de la Durantaye, 'Akkommodation statt Assimilation.' (n 177) 168.
[198] ibid. 168.
[199] Also see forthcoming chapter on Product, Model and Entity Regulation in this work.
[200] Also see forthcoming chapter on Product, Model and Entity Regulation in this work.
[201] A different reading of Article 53(1)(c) AI Act would arguably undermine the level playing field the AI Act tries to create, according to recital 106.
[202] E.g., Peukert (n 123) 505.
[203] '[T]he 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.' Also see Peukert, (n 123) 505.
[204] See Peukert (n 123) 505.
[205] Peukert (n 123) 505.
[206] Nordemann and Rasouli (n 169) 782–785; Bernsteiner and Schmitt (n 25) para 36. See similarly Schneider (n 4) para 39.
[207] See s 823(2) German Civil Code (BGB) [2002] Federal Law Gazette 1 page 42, 2909; 2003 I page 738.
[208] See Peukert (n 123) 505.

92. As a result, we would question the assertion some authors make that violations of this provision are obvious,[209] as the text does not make clear how comprehensive the required policy must be or which specific technical measures GPAI model providers are expected to implement to achieve compliance. Even if this provision were to confer an individual right of action under certain national legal systems, claimants would likely find it challenging to demonstrate that a specific copyright violation resulted from the absence of an adequate policy, particularly given the inherent difficulty in eliminating all copyright infringements, even with robust policies, and the causality issues that typically characterize cases of AI harm explored elsewhere in this work.[210]

93. Lastly, it is worth emphasising that the policy requirement in Article 53(1)(c) applies not only to the model's training data (input) but also to its output. While some maintain that GPAI providers are not obliged to verify the output of another provider's system incorporating the GPAI model[211] – nor can they be held liable for infringements arising at the level of the implementing system[212] – it is nevertheless clear that the (potential) output of the GPAI model should fall within the scope of the copyright policy.[213] The Code of Practice further states that GPAI model providers should implement safeguards to help prevent their models from infringing copyright and that they should require downstream system providers to accept terms and conditions designed to prevent copyright violations.[214]

2.1.3.2. Policy requirements

94. In general, a copyright policy requires that the GPAI model provider first assess the extent to which the training and use of the model may give rise to copyright infringements.[215] In other words, the provider must determine for which aspects of training and use the rightsholder's permission is necessary.[216] The Code of Practice addresses this most explicitly in the context of TDM opt-out detection, as discussed above, but the compliance policy requirement extends further, also covering other potential copyright infringements when collecting or processing copyrighted material.

95. At the next stage, the GPAI model provider should devise a plan to mitigate the risks of copyright infringement.[217] This could involve deploying state-of-the-art techniques – which Article 53(1)(c) refers to specifically in relation to the TDM opt-out – to prevent copyright violations more generally. For model outputs, the Code of Practice similarly requires GPAI developers to implement 'appropriate and proportionate' technical safeguards to prevent the model from reproducing copyright-protected material used during training, and to prohibit copyright-infringing uses through contractual terms with downstream providers, or, in the case of open-source models, at least to alert users to the prohibition of

[209] Bernsteiner and Schmitt (n 25) para 36 ('aus der Norm klar hervorgeht, wann diese verletzt wird').
[210] Also see forthcoming chapter on GPAI Liability in this work.
[211] Peukert (n 123) 507.
[212] Bernsteiner and Schmitt (n 25) para 43.
[213] Also see Peukert (n 123) 507; Code of Practice Copyright Chapter (n 170) Measure 1.4.
[214] ibid. Measure 1.4(1).
[215] Bernsteiner and Schmitt (n 25) para 40.
[216] ibid. para 40.
[217] ibid. para 41.

copyright-infringing use.[218] On the training side, GPAI model providers could impose similar obligations on parties assisting in the acquisition of training material.[219] It is also important to stress that the policy must not be merely theoretical but must be effectively implemented.[220]

96. According to the Code of Practice, GPAI model providers should, as part of their policy, designate a point of contact and enable stakeholders to submit complaints.[221] Such complaints should be handled diligently and in accordance with due process, without prejudicing any potential copyright-based claims by rightsholders.[222] Several other provisions in the Code of Practice also emphasise the need to make relevant information available to rightsholders who believe their copyright may have been infringed by the model.[223]

2.1.4. Article 53(1)(d): Summary of training content

2.1.4.1. Requirement and rationale

97. According to Article 53(1)(d), the GPAI model provider is also required to prepare a summary of the data used to train the model. To this end, the AI Act mandates the AI Office to provide a template for providers to complete,[224] which it has.[225] While Recital 107 emphasises the need to respect trade secrets and confidential business information, both Article 53(1)(d) and Recital 107 indicate that the summary should be reasonably detailed, though not necessarily highly technical. Recital 107 further clarifies that the purpose of the summary is 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.'[226] The level of detail – while not necessarily technical[227] – should enable those parties to assess their legal position and identify any potential concerns regarding the data used. The non-binding Explanatory Notice to the Template clarifies that this legal position extends beyond copyright concerns[228] to include all rights protected under Union law,[229] such as data protection[230] and the freedom to receive information and conduct scientific research.[231]

---

[218] Code of Practice Copyright Chapter (n 170) Measure 1.4.
[219] See in a similar sense Code of Practice Copyright Chapter (n 170) Measure 1.3 on web-crawlers used on their behalf.
[220] See similarly, Bernsteiner and Schmitt (n 25) para 42.
[221] Code of Practice Copyright Chapter (n 170) Measure 1.5; Schneider (n 4) para 24.
[222] Code of Practice Copyright Chapter (n 170) Measure 1.5.
[223] E.g., ibid. Measures 1.3(4).
[224] AI Act, art 53(1)(d).
[225] European Commission, 'Explanatory Notice and Template for the Public Summary of Training Content for general-purpose AI models required by Article 53 (1)(d) of Regulation (EU) 2024/1689 (AI Act)' C(2025) 5235 final <https://ec.europa.eu/newsroom/dae/redirection/document/118480> accessed 1 October 2025.
[226] Also see Schneider (n 4) para 27.
[227] Also see AI Act, recital 107; Schneider (n 4) para 27.
[228] European Commission, 'Template Explanatory Notice' (n 225) para 9.
[229] ibid. para 7.
[230] ibid. para 9.
[231] ibid. para 11.

98. Additionally, this information can be valuable for downstream system providers considering integration of the model, allowing them to assess the diversity of the data used.[232] It is also argued that such transparency may lead to more competitive markets, as it enables downstream actors to better evaluate how their data and models have been used, thus reducing lock-in effects.[233] At the same time, this transparency requirement has been criticised for imposing a significant burden on providers.[234]

99. The wording of Article 53(1)(d) strongly suggests that the information must be shared using the template provided by the AI Office. Recital 107 takes a softer approach, stating that '[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' without expressly indicating that this form should be used by providers. Nevertheless, the more literal reading of Article 53(1)(d) – that the AI Office template must be used – is reaffirmed by the Explanatory Notice to the Template[235] and, arguably, aligns more closely with the core objective of this obligation: to enable parties with legitimate interests to assess their legal position, with uniformity supporting both this goal and the document's accessibility. Even if Article 53(1)(d) were interpreted as not mandating use of the template, employing the template would remain the most straightforward way to comply with the duty to disclose the summary of training content, compared to communicating that information by other means. However, if a GPAI model provider can communicate the same information as effectively through another format, one might question whether anyone would be disadvantaged and whether the aims of the provision are not still met.[236]

100. It is also important to note that the GPAI model provider must make this document 'publicly available'. This availability should go beyond mere theoretical access.[237] Given the wide range of potentially interested parties, it is reasonable to require that the document be accessible online in a digital format. This is also what the Explanatory Notice suggests: '[the Summary] should be published on the provider's official website in a clearly visible and accessible manner, clearly indicating which model(s) (and possibly model version(s)) the Summary covers [...]. The Summary should also be made publicly available together with the model across all its public distribution channels (e.g. online platforms).'[238] While the wording of Article 53(1)(d) does not entirely preclude a provider from meeting its obligations by making the document 'publicly available' on-site, such in-person availability would run counter to the transparency that the provision and Recital 107 seek to establish. In any case, the document must be made available no later than the date on which the model is placed on the Union market.[239]

[232] ibid. para 10.
[233] ibid. para 12.
[234] The lack of copyright harmonisation in the EU is said to make it difficult for providers to assess what material is protected, see e.g., Philipp Hacker and Amelie Berz, 'Der AI Act der Europäischen Union – Überblick, Kritik und Ausblick' (2023) Zeitschrift für Rechtspolitik 226, 228; Schneider (n 4) para 28.
[235] European Commission, 'Template Explanatory Notice' (n 225) para 4.
[236] Also see Bernsteiner and Schmitt (n 25) para 45.
[237] Also see Schneider (n 4) para 26.
[238] European Commission, 'Template Explanatory Notice' (n 225) para 32.
[239] ibid. para 32.

2.1.4.2. Detail and modalities

101. The Template that was published by the European Commission provides more insight into the information that should be shared. It clarifies that the information should be provided with sufficient detail,[240] covering all types of data used throughout the model's lifecycle – from pre-training to post-training – including model alignment and fine-tuning.[241] Data used during model operation, such as in retrieval-augmented generation, falls outside the scope unless it contributes to model training.[242] The information should be comprehensive,[243] presented in a narrative, simple, and effective format to ensure it is understandable to the relevant parties.[244] It must be accurate, comprehensive, and provided in good faith.[245] The AI Office will assess compliance and, if necessary, request corrective measures or seek enforcement.[246]

102. At the same time, it is important to emphasise that the template aims to balance transparency requirements with the need to protect confidential business information and trade secrets.[247] This has resulted in significant limitations on the data that must be disclosed. The disclosure obligation for licensed data is limited, which makes sense as the relevant rightsholders are already parties to those licence agreements.[248] Private datasets that are not commercially licensed only need to be listed if they are publicly known or if the provider chooses to make them public.[249] Commercially sensitive details concerning data sources, model curation, or training methods do not need to be disclosed.[250] Only minimal information is required regarding user data from interactions, explicitly excluding data licensed via commercial agreements or customer fine-tuning data.[251] For synthetic data, disclosure is limited to the names of the models used if they are placed on the market, or a general description of model training data where necessary.[252]

103. Furthermore, only a high-level aggregated overview of training data size per modality, presented in broad ranges, is required. For publicly available datasets, more detail must be provided.[253] For scraped data, only a summarised narrative list of the most relevant domain names must be provided – not a full list of URLs.[254] In accordance with Recital 107, the Explanatory Notice also stresses that disclosures should be non-technical and presented in a summarised narrative form, so as to avoid revealing sensitive

[240] Also see AI Act, recital 107.
[241] European Commission, 'Template Explanatory Notice' (n 225) para 13. Also see forthcoming chapter on Modifications in this work.
[242] European Commission, 'Template Explanatory Notice' (n 225) para 13.
[243] AI Act, recital 107; European Commission, 'Template Explanatory Notice' (n 225) para 14.
[244] European Commission, 'Template Explanatory Notice' (n 225) para 23.
[245] ibid. paras 24-25.
[246] ibid. para 26.
[247] ibid. paras 17-22. Also see AI Act, recital 107.
[248] ibid. para 19.
[249] ibid. para 19.
[250] ibid. para 18.
[251] ibid. para 21.
[252] ibid. para 21.
[253] ibid. para 19.
[254] ibid. para 20.

information.[255] Only high-level aggregates regarding the mix and composition of data sources must be disclosed, without specifying the exact mix.[256] With respect to crawlers, providers must disclose their purpose and collection periods, but not their precise technical implementation.[257]

104. In any case, providers may go beyond these minimum requirements and disclose more information than is required by the template.[258] Providers are also encouraged to respond to requests from relevant stakeholders with legitimate interests who wish to better assess their own legal position regarding the data used, provided this does not breach the provider's own obligations. This is particularly encouraged for information scraped or crawled from the internet.[259]

2.1.4.3. Template content

105. Without seeking to reiterate the full content of the template itself,[260] it consists of three main sections.[261] First, it collects general information to identify the provider and the model, as well as the characteristics of the training data, broken down by modality (text, image, audio, video, etc.), estimated data size, and general content types. Second, it requires a detailed list and categorisation of all data sources used for training, including public datasets, private or licensed datasets, data scraped from online sources, user data, synthetic data, and any other sources, together with descriptions and, where relevant, lists of domain names. Third, it addresses data processing aspects, including measures to respect copyright opt-outs, the removal of illegal content, and other relevant steps taken to ensure compliance with Union law.[262]

2.1.4.4. Adjusted content for pre-existing GPAI models/updates

106. The requirements outlined above are affected if a downstream entity sufficiently[263] modifies a model already placed on the Union market so that they themselves become the provider of the modified model.[264] In such cases, the modifying entity should report only the training data used for the modification in the template, as well as the name of the original model.[265] If a provider continues to train their own model, they should also update the information provided at six-month intervals, or sooner if the training results in a materially significant update to the content of the summary.[266] The updated summary should then be made available alongside the modified model.[267]

---

[255] ibid. para 20.
[256] ibid. para 22.
[257] ibid. para 20.
[258] ibid. para 16.
[259] ibid. para 16.
[260] European Commission, 'Template Explanatory Notice' (n 225) 9-14.
[261] ibid. para 16.
[262] ibid. para 15.
[263] See, in more detail: Commission Guidelines on the Scope of the Obligations for General-Purpose AI Models (n 19) para 61. Also see forthcoming chapter on Modifications in this work.
[264] European Commission, 'Template Explanatory Notice' (n 225) para 28. Also see Commission Guidelines on the Scope of the Obligations for General-Purpose AI Models (n 19) paras 60 ff.
[265] AI Act, recital 107; European Commission, 'Template Explanatory Notice' (n 225) para 28.
[266] European Commission, 'Template Explanatory Notice' (n 225) para 29.
[267] ibid. para 29.

107. The Explanatory Notice does not specify the criteria for determining whether changes to the training data constitute a 'significant update'. However, minor changes – such as correcting labelling errors, adding relatively small amounts of data, or retraining on the same domains without significant alteration – are unlikely to qualify. By contrast, a significant update is more likely where entirely new data sources or domains are added (for example, the inclusion of medical or financial data), where there is a substantial expansion of the dataset, or where the dataset comes to include non-copyrighted material in addition to copyrighted material, or vice versa. The same applies in cases of substantial reweighting of the various data types used.

108. In any case, this assessment involves a degree of judgement on the part of the provider. Crucially, the purpose of the information being shared must be taken into account. Accordingly, an update is required sooner than the six-month interval if the changes made could have potentially important consequences for stakeholders such as rights managers and copyright holders. Bearing this purpose in mind, it is unlikely that changes to the training data which do not meet the described thresholds but which result in significantly altered model behaviour or performance are relevant here. While such updates may arguably be considered 'significant', their significance does not necessarily extend to stakeholders concerned with the data used.

109. A single summary may be used for multiple versions of the same model if their summaries are identical, provided that those models and versions are clearly identified. If one of these versions has already been placed on the Union market, requiring an earlier summary, the summaries for subsequent versions need only cover the training data specifically used to modify the original version, along with a clear reference to the original summary.[268]

2.2. Article 53(2): Open-source exception

110. Article 53(2) provides a partial exception from the preceding obligations for open-source GPAI models.[269] Where a GPAI model provider meets certain requirements, they are not obliged to provide documentation to the AI Office and national competent authorities nor to supply information to downstream AI system providers, provided those models do not present a systemic risk. If the model does present systemic risks, however, these information requirements do apply[270] – and are, in part, extended by Section 2 of Annex XI. This open-source exception is motivated by the need to foster innovation and the growth opportunities that such models could offer for the European Union.[271]

111. To qualify for the exception, the model must be released under a free and open-source licence. The Commission Guidelines clarify that the term 'licence' should be interpreted broadly, referring to the granting of permissions such that the original provider does not use their intellectual property rights to restrict the use of the model or charge for its use.[272] Furthermore, they indicate that the licence should provide for access, use, modification and distribution; modifying, using, or distributing the model

[268] ibid. para 30.
[269] See, generally, on this exception: Commission Guidelines on the Scope of the Obligations for General-Purpose AI Models (n 21) paras 76-92.
[270] Also see Schneider (n 4) para 32.
[271] See AI Act, recital 103.
[272] Commission Guidelines on the Scope of the Obligations for General-Purpose AI Models (n 19) para 80.

should therefore be possible without restriction, although limited conditions are permissible.[273] These 'limited conditions often consist only in crediting the author(s) and retaining their copyright notice, i.e. attribution.'[274]

112. The Commission Guidelines also identify certain restrictions that disqualify a licence from meeting these criteria.[275] These include limitations to non-commercial or research use only, prohibitions on distributing the model or its components, usage restrictions relating to user scale thresholds (which require additional licensing), and requirements to obtain a specific licence for certain use cases.[276]

113. Moreover, a model is not considered to be 'free' if it is monetised indirectly, including through additional services (for example, for necessary technical support or security),[277] or if access to the model requires the purchase of support or training.[278] This exclusion also covers instances where the model provider collects data – other than for the purpose of improving security, compatibility, or interoperability – for monetisation purposes.[279] Furthermore, it precludes models offered under licences that permit free academic use but require payment for commercial or scaled use, or licences under which use of the model necessarily requires purchasing access to a platform or server hosted by the provider.[280] Permissible practices, however, include situations where the model is offered alongside paid services that are purely optional or where such paid services or support are made available in the form of premium versions or extensions of the model.[281]

114. In addition, there is an information or transparency requirement. The provider should make publicly available the model's parameters (including weights) and the model architecture, as well as relevant information on model usage. This should arguably be interpreted broadly, given the purpose of the exception[282] – as wider access would better enable the model's open-access status to foster innovation, facilitate further development, and allow downstream providers to integrate the model. To this end, the Commission Guidelines similarly stress that this information should at least include '[i]nformation about the model's input and output modalities, capabilities, and limitations [including] the technical means (e.g. instructions for use, infrastructure, tools) required for the model to be integrated into AI systems, which may include the appropriate configuration for the intended use cases, where applicable.'[283]

---

[273] ibid. para 82.
[274] ibid. para 82.
[275] ibid. para 83.
[276] ibid. para 83.
[277] See AI Act, recital 103; European Commission, Guidelines on the Scope of the Obligations for General-Purpose AI Models (n 19) para 85.
[278] European Commission, Guidelines on the Scope of the Obligations for General-Purpose AI Models (n 19) para 86.
[279] See AI Act, recital 103; Commission Guidelines on the Scope of the Obligations for General-Purpose AI Models (n 19) para 87. See similarly in article 3(5)(f) of Directive (EU) 2019/770 of the European Parliament and of the Council of 20 May 2019 on certain aspects concerning contracts for the supply of digital content and digital services ("Digital Content Directive") [2019] OJ L 136/1; Schneider (n 4) para 31.
[280] Commission Guidelines on the Scope of the Obligations for General-Purpose AI Models (n 19) para 86.
[281] ibid. para 88.
[282] See also Bernsteiner and Schmitt (n 25) para 50.
[283] Commission Guidelines on the Scope of the Obligations for General-Purpose AI Models (n 19) para 92.

2.3. Article 53(3): Duty of cooperation

115. Article 53(3) imposes a duty on GPAI model providers to cooperate with the Commission and national competent authorities in the exercise of their AI Act competencies and powers. This obligation should also be understood as extending to the AI Office.[284] In some respects, Article 53(3) constitutes the GPAI provider equivalent of Articles 21, 23(7), 24(6) and 26(12), which impose cooperation duties on various other actors – namely, high-risk system providers, importers, distributors, and deployers, respectively.

116. Article 53(3) states that the duty to cooperate applies 'in the exercise of their competences and powers pursuant to this Regulation'. The material scope of this obligation thus extends beyond the other provisions of Article 53 and covers the entirety of the AI Act, with particular reference to Articles 88 to 94.[285] It should also be read in conjunction with Article 101(1)(b), which imposes a fine on GPAI providers who intentionally or negligently '[fail] to comply with a request for a document or for information pursuant to Article 91, or supplied incorrect, incomplete or misleading information'. The relationship between some of these provisions and Article 53 is, however, not always clear. For example, the Commission[286] can request the documentation required by Article 53(1)(a) both under that provision itself and similarly under Article 91, whereas national competent authorities can seemingly only do so on the basis of Article 53(1)(a), raising the question as to the additional function of Article 91. Its role appears clearer for other parts of Article 53(1), such as (1)(b) and (c), which do not grant a separate competence to request information. A different interpretation, suggested by the Article's title ('Obligations for providers of general-purpose AI models') could see Article 53(1)(a) as only binding on the providers without providing a power for the AI Office the Commission or national competent authorities to request the information, though that would leave the AI Office and national competent authorities without such a power under the AI Act.[287]

117. The Article 53(3) duty to cooperate is not further clarified in the recitals, leaving its interpretation largely open. This is particularly relevant to the extent of the obligation. Given the phrase 'as necessary', one could interpret this obligation as requiring GPAI model providers only to respond to requests for information and documentation made by the AI Office, the Commission, or national competent authorities, provided such requests are made within the exercise of their AI Act powers and competences.

[284] Also see article 3(47) which implies that references to the AI Office 'shall be construed as references to the Commission'. It is sensible that the duty to cooperate would extend to the AI Office as it is also authorized (and tasked) with requesting information from GPAI model providers, see e.g. article 53(1)(a) discussed above. See similarly Code of Practice Safety and Security Chapter (n 80) recital (e); Bernsteiner and Schmitt (n 25) para 48. Also see forthcoming commentary on Article 3(47) in this work.
[285] See similarly Schneider (n 4) para 33.
[286] Based on a reading of art 3(47) AI Act that equates the AI Office (in article 53(1)(a)) with the Commission. Also see forthcoming commentary on Article 3(47) in this work.
[287] Also see AI Act, art 91.

118. While somewhat less convincing,[288] one could also interpret this obligation more broadly.[289] A broader reading might require GPAI model providers to proactively submit information that has not been expressly requested by the AI Office, the Commission, or national competent authorities, insofar as doing so could facilitate their enforcement of the AI Act. Such an interpretation appears less consistent with the literal wording of Article 53(3) and is also implicitly contradicted by the explicitly phrased duty to notify – a form of proactive cooperation – the AI Office and, where relevant, national competent authorities of serious incidents involving GPAI models with systemic risk, as set out in Article 55(1)(c).[290]

119. Regardless of these potential diverging interpretations, Article 53, taken together with the wording of Article 101(1)(b) – which requires that supplied information be correct, complete and not misleading – means that this obligation should be interpreted as requiring GPAI model providers to supply detailed, accurate and reliable information. These elements should be assessed in light of the enforcement objectives for which they are provided.

2.4. Article 53(4): Compliance pathways

120. Article 53(4)[291] sets out how GPAI model providers can demonstrate compliance with the requirements outlined in the previous sections. It offers three distinct pathways to this end. First, GPAI model providers may demonstrate compliance by adhering to harmonised standards, which creates a presumption of conformity with the AI Act insofar as the relevant obligations are addressed by those standards.[292] In the absence of published harmonised standards, model providers may rely on codes of practice.[293] Finally, GPAI model providers may demonstrate compliance by any other 'adequate means'.

2.4.1. Harmonised standards

121. Over time, compliance with harmonised standards can be expected to become the principal mechanism for fulfilling GPAI model provider obligations. This notion refers to the definition found in Article 2(1)(c) of Regulation (EU) No 1025/2012,[294] which describes a harmonised standard as 'a European standard adopted on the basis of a request made by the Commission for the application of Union harmonisation legislation'. These standards, developed by the European Committee for Standardisation (CEN), the European Committee for Electrotechnical Standardisation (Cenelec) or, the European

[288] Although the Commission Guidelines offer very limited support for a more far-reaching duty of cooperation (Commission Guidelines on the Scope of the Obligations for General-Purpose AI Models (n 19) para 102), that support is limited to the context of formal proceedings.

[289] Also see Code of Practice Safety and Security Chapter (n 80) recital (e) ('The Signatories further recognise the importance of cooperating with the AI Office (Article 53(3) AI Act) to foster collaboration between providers of general-purpose AI models with systemic risk, researchers, and regulatory bodies to address emerging challenges and opportunities in the AI landscape').

[290] Also see forthcoming commentary on Article 55 in this work.

[291] Also see the similarly phrased article 55(2), and also see forthcoming commentary on Article 55 in this work.

[292] Also see commentary on Article 56 in this work.

[293] See in particular Code of Practice Transparency Chapter (n 17); Code of Practice Copyright Chapter (n 170); Code of Practice Safety and Security Chapter (n 80).

[294] See recital 121.

Telecommunications Standards Institute (ETSI),[295] can be expected to reflect the state-of-the-art[296] and will be formulated with a 'balanced representation of interests involving all relevant stakeholders',[297] following a request by the Commission.[298] Article 10(6) Regulation (EU) No 1025/2012 indicates that a reference of the standard will be published in the Official Journal of the European Union.

122. While Recital 117 indicates that the AI Office will assess whether such a standard constitutes a 'suitable' instrument to cover the relevant obligations,[299] this idea is not directly reflected in the enacting terms. Instead, this likely refers to the Article 10(5) Regulation (EU) No 1025/2012 provision that the Commission 'shall assess the compliance of the documents drafted by the European standardisation organisations with its initial request'.

123. Once it is established that an obligation is covered by a relevant harmonised standard, and a reference of that standard has been published in the Official Journal of the European Union, Article 53(4) introduces a presumption that compliance with the standard entails compliance with Article 53 vis-à-vis that obligation. It is key to note, however, that that presumption is likely rebuttable, as it would otherwise not be meaningful to call it a presumption.[300]

2.4.2. Codes of practice

124. In the absence of a harmonised standard, Article 53(4) indicates that GPAI model providers may comply with codes of practice within the meaning of Article 56.[301] Its last sentence indicates that such a code of practice must have been assessed adequate ('approved') within the meaning of Article 56(6) in order to serve its compliance function.[302] This approach differs from that adopted for high-risk systems, where the fall-back in the absence of harmonised standards is the adoption of 'common specifications' (Article 41 AI Act).

125. While codes of practice are discussed more extensively in the chapter on Article 56,[303] it is worth noting that Article 53(4) does not extend the presumption applicable to harmonised standards to codes of practice. As a result, compliance with a code of practice does not amount to automatic compliance with the AI Act, nor does it generally give rise to a presumption of such compliance.[304]

---

[295] See article 10 Regulation (EU) No 1025/2012 of the European Parliament and of the Council of 25 October 2012 on European standardisation, amending Council Directives 89/686/EEC and 93/15/EEC and Directives 94/9/EC, 94/25/EC, 95/16/EC, 97/23/EC, 98/34/EC, 2004/22/EC, 2007/23/EC, 2009/23/EC and 2009/105/EC of the European Parliament and of the Council and repealing Council Decision 87/95/EEC and Decision No 1673/2006/EC of the European Parliament and of the Council [2012] OJ L 316/12 as well as its Annex I.

[296] Recital 121.

[297] Recital 121.

[298] See the definition found in article 2(1)(c), Regulation (EU) No 1025/2012. Also see recital 121.

[299] 'Once a harmonised standard is published and assessed as suitable to cover the relevant obligations by the AI Office, compliance with a European harmonised standard should grant providers the presumption of conformity.'

[300] Also see commentary on Article 56 in this work.

[301] Also see commentary on Article 56 in this work.

[302] See more extensively commentary on Article 56 in this work.

[303] See commentary on Article 56 in this work.

[304] Also see Bernsteiner and Schmitt (n 25) para 56.

126. Even though codes of practice do thus not directly confer (a presumption of) compliance,[305] they remain a valuable tool for interpreting the Act's provisions. Given their assessment as adequate by the AI Office, they also give rise to legitimate expectations.[306] Nevertheless, the absence of an explicit presumption of conformity underscores the – at least theoretical[307] – possibility that the Commission, the AI Office, or national competent authorities could still find a provider in violation of the AI Act despite adherence to a code of practice deemed adequate.[308] Nevertheless, the Commission Guidelines seem to equate compliance with an approved code of practice with compliance with the AI Act.[309]

2.4.3. Alternative adequate means

127. There is no obligation for GPAI model providers to adhere to codes of practice or harmonised standards, even when these are available. Irrespective of the availability of such measures, Article 53(4) indicates that providers may also demonstrate compliance with the AI Act through 'alternative adequate means'. This remains subject to assessment by the Commission.

128. Even where providers choose not to adhere to codes of practice or harmonised standards, these instruments offer valuable guidance for alternative compliance routes – particularly regarding the types and extent of information to be documented, as well as the modalities (such as duration) of documentation. For instance, compliance might be achieved by documenting the same or similar information, for a comparable duration, but in a different format to that proposed by the codes of practice or harmonised standards. The Commission Guidelines suggest conducting a 'gap analysis' to compare the adopted measures with those set out in the codes of practice. They also note that providers choosing this path may face increased information requests, as it will generally be more challenging for the AI Office to assess their compliance.[310]

129. A notable exception appears to be Article 53(1)(d), which seems to mandate use of the template provided by the AI Office, as discussed earlier.[311]

2.5. Article 53(5): Delegated acts on Annex XI methodologies

130. Articles 53(5) and 97 empower the Commission, pursuant to Article 290 TFEU, to adopt delegated acts providing further detail on measurement and calculation methodologies, 'with a view to allowing for comparable and verifiable documentation' to assess compliance with the provisions in Annex XI relating to the computational resources used to train the model, the model's training time, and other relevant details of the training process (Annex XI(2)(d)), as well as the estimated or known energy consumption of the model (Annex XI(2)(e)), as discussed above.[312] Such delegated acts are binding

[305] Also see the objectives stated at the start of those codes of practice themselves.
[306] Also see commentary on Article 56 in this work.
[307] See more extensively commentary on Article 56 in this work (on legitimate expectations).
[308] Admittedly, even a presumption would not rule out this (theoretical) possibility fully, as the presumption could be rebutted. See more extensively commentary on Article 56 in this work.
[309] Commission Guidelines on the Scope of the Obligations for General-Purpose AI Models (n 19) para 94.
[310] ibid. para 95.
[311] See para 99.
[312] See paras 30 ff.

non-legislative acts and constitute secondary legislation under Article 290 TFEU, allowing the Commission to specify these technical elements. This mechanism is particularly pertinent given the absence of agreed technical standards for certain aspects.[313]

131. Interestingly, while Article 53(5) explicitly refers to delegated acts in relation to Annex XI Section 1(e) (and (d)), the Act – and Annex XI in particular – does not clarify the relationship with the phrase at the end of Annex XI Section 1: 'With regard to point (e), where the energy consumption of the model is unknown, the energy consumption may be based on information about computational resources used.' This could give rise to a contradiction if the Commission were to adopt a delegated act specifying an estimation method not based on computational resources used. Should this occur, it is relevant to note that the Commission is also empowered to adopt delegated acts to amend Annexes XI and XII in light of technological developments (Article 53(6)).[314] However, this estimation method is arguably not encompassed by that provision, leaving it as a valid alternative to the method identified in any delegated act adopted by the Commission.

132. Procedurally, it is important to note that the European Parliament and the Council may object to a delegated act adopted by the Commission within three months of its notification.[315] This objection prevents the act from becoming binding, although it may become binding within that three-month period if both the European Parliament and the Council have informed the Commission that they do not intend to object (Article 97(6)).

2.6. Article 53(6): Delegated acts to amend Annexes XI and XII

133. In contrast to Article 53(5), Article 53(6) confers a much broader mandate on the Commission to adopt delegated acts amending Annexes XI and XII. As discussed earlier, Annex XI sets out the technical documentation that GPAI model providers must supply, upon request, to the AI Office and national competent authorities, while Annex XII specifies the information that should be communicated to downstream AI system providers intending to integrate the GPAI model into their systems.

134. Article 53(6) empowers the Commission to amend Annexes XI and XII 'in light of evolving technological developments'. This should be seen as an attempt to 'future-proof' the AI Act,[316] and in particular, the regulatory oversight it establishes. In practical terms, this means the Commission may add elements to the documentation requirements for GPAI model providers (Annex XI) – for example, novel evaluation methods or new risk assessment criteria – or require additional information to be shared with downstream providers (Annex XII), such as emerging integration challenges or limitations. As a result, GPAI model providers cannot rely on a single documentation exercise but must remain vigilant and ensure their documentation remains up to date to avoid omitting any amendments.

[313] E.g. on the lack of consensus on the environmental impact of AI models: Ian R Hodgkinson, Nick Jennings and Tom Jackson, 'Everyone Must Understand the Environmental Costs of AI' (OECD.AI, 2024) <https://oecd.ai/en/wonk/understand-environmental-costs>.
[314] See Section 2.6.
[315] See AI Act, art 97(5).
[316] Also see recital 173.

135. The procedure mirrors that described above (see also Article 97(2) AI Act). This power for the Commission to amend Annexes XII and, in particular, XI means that GPAI model providers cannot rely on a single documentation exercise; rather, they must remain vigilant and ensure their documentation remains up to date to avoid omitting any amendments.

2.7. Article 53(7): Confidentiality

136. Article 53(7) provides that the AI Office, the Commission, and national competent authorities must treat any information received pursuant to Article 53 in accordance with the confidentiality obligations set out in Article 78. Notably, Article 78(2) requires that authorities generally exercise restraint in requesting sensitive data, ensuring that such requests are 'strictly necessary' for the exercise of their powers under the AI Act or their obligations under Regulation 2019/1020 on market surveillance and compliance of products. This requirement for restraint, however, is not expressly reiterated in Article 53 itself.

137. Unlike Article 78, which applies only to 'the Commission, market surveillance authorities and notified bodies and any other natural or legal person involved in the application of this Regulation', Article 53(7) adopts a broader scope, seemingly extending to downstream system providers who have received information under Article 53(1)(b), insofar as the information is sensitive. However, as penalties under the AI Act are directed solely at GPAI model providers (Article 101), it appears that breaches by downstream system providers would not be penalised under the AI Act. This further reinforces the earlier point that it is permissible, and arguably advisable, for GPAI model providers to make access to relevant information conditional upon a non-disclosure agreement.[317]

138. As a direct consequence of this confidentiality requirement, the general public will not have access to any information documented pursuant to Article 53, with the notable exception of Article 53(1)(d) – the training data summary.[318]

[317] See para 79.
[318] See paras 97 ff.