GPAI Considerations for EU AI Act High-Risk Requirements

This document lists certain obligations imposed by Articles 8-15 and Annex IV of the Council's text on the proposed EU AI Act—obligations that currently apply only to AI systems designated as "high risk" in the Act—and identifies those that would be difficult or even impossible for developers and deployers of general-purpose AI systems ("GPAI systems") to meet.[1] The analysis of these compliance challenges in the chart below raises the following major themes:

- **GPAI systems should not be generally designated as "high-risk."** GPAI systems have many low-risk, socially beneficial uses. For example, GPAI can be used to create subtitles on live video feeds to improve accessibility for hearing-impaired listeners, or to read out text or describe video feeds for visually-impaired users. Subjecting these systems to the Act's high-risk requirements places is therefore unnecessary and disproportionate, and contrary to long-standing principles of EU law. This approach could prevent the placing on the EU market of GPAI for low-risk AI systems that integrate or otherwise are based on GPAI systems, even where these downstream systems have the potential to improve health or safety, or better safeguard fundamental rights, than current technologies and practices, e.g. AI systems that include waste-sorting technology that leverages a GPAI object recognition system. Imposing the AI Act's high-risk requirements on all GPAI systems could thus have the effect of depriving the EU of access to GPAI which is essential for low-risk AI systems that could improve people's lives.

- **GPAI systems have generalized, adaptable purposes.** A defining quality of GPAI systems is that they can be adapted for multiple uses, deployed used in a range of contexts by the end users, and are often integrated into third-party products and services. In most cases, whether a GPAI system could present unexpected high risks to health, safety, or fundamental rights will depend entirely on its context of use and the specific use cases. As a result, it can be difficult or even impossible for both developers and deployers of these systems to anticipate the full range of uses for which they can be used by end users. Imposing the AI Act's high-risk requirements on developers and deployers of all GPAI systems, merely based on the chance that it could be used in a high-risk setting by end users, will make it more difficult for developers and deployers to supply such systems in the EU, even for low- and no-risk uses. Additionally, because many of the requirements for high-risk AI systems set out in the Act are grounded in the notion of the system's intended purpose or use, they are incompatible with the inherently multi-purpose nature of GPAI systems. A more coherent approach, in line with the methodology of the AI Act, would be to impose these compliance obligations only on GPAI that is part of an AI system deployed in a high-risk area.

- **Direct legal obligations for use of GPAI in high risk applications should be on deployers**. In most cases, whether a GPAI system could present unexpected high risks to health, safety, or fundamental rights will depend entirely on its context of use and the specific use cases. Imposing the AI Act's high-risk requirements on developers and deployers of GPAI systems, merely based on the chance that it could be used in a high-risk setting by end users, will make it more difficult for developers and deployers to supply such systems in the EU, even for low- and no-risk uses. Where a GPAI system is used a high-risk system, the obligations for high-risk AI should be placed on the deployer of that system.

- **GPAI providers should be obliged to support deployers where required**. The deployer of GPAI is not always in the position to comply with all proposed high-risk requirements alone. To the extent a deployer of a high-risk AI system requires an AI developer's assistance to comply with the high-risk requirements, and the developer has agreed to provide the AI system for a high-risk use case, the deployer should have the right to enter into a contract that is binding on the developer and sets out the obligations for the latter.

[1] The Council text defines "general purpose AI system" as "an AI system that - irrespective of how it is placed on the market or put into service, including as open source software - is intended by the provider to perform generally applicable functions such as image and speech recognition, audio and video generation, pattern detection, question answering, translation and others; a general purpose AI system may be used in a plurality of contexts and be integrated in a plurality of other AI systems" (Art. 3(1b)).

Article	Text	GPAI Concerns
Compliance (Article 8)		
Article 8	High-risk AI systems shall comply with the requirements established in this Chapter, taking into account the generally acknowledged state of the art. The intended purpose of the high-risk AI system and the risk management system referred to in Article 9 shall be taken into account when ensuring compliance with those requirements.	GPAI systems are by their nature general purpose, thus making it difficult to meaningfully define their "intended purpose." For example, the same image recognition system that could be used to detect potholes in images of city roads could also be used to identify broken bones in an xray. As a result, efforts to anticipate the intended purpose(s) for a GPAI system will almost inevitably be under- or over-inclusive for developers. If the "intended purpose" is defined in general terms, this will make it more difficult for developers and deployers of GPAI systems to adopt risk management systems that satisfy Article 9's requirements, particularly as they relate to identifying and mitigating risks.
Risk Management System (Article 9)		
Article 9(2)-(3)	The risk management system shall be understood as a continuous iterative process planned and run through the entire lifecycle of a high-risk AI system, requiring regular systematic updating. It shall comprise of the following steps: (a) Identification and analysis of the known and foreseeable risks most likely to occur to health, safety and fundamental rights in view of the intended purpose of the high-risk AI system; (b) [deleted] (c) Evaluation of other possibly arising risks based on the analysis of data gathered from the post-marketing monitoring system referred to in Article 61; (d) Adoption of suitable risk management measures in accordance with the provisions in the following paragraphs. The risks referred to in this paragraph shall concern only those which may be reasonably mitigated or eliminated through the development or design of the high-risk AI system, or the provision of adequate technical information. The risk management measures referred to in [2(d)] shall give due consideration to the effects and possible interaction resulting from the combined application of the requirements set out in this Chapter 2, with a view to minimizing risks more effectively while achieving the appropriate balance in implementing the measures to fulfil those requirements.	Given the many potential applications of GPAI systems and the anticipated volume of GPAI system deployments in the marketplace, requiring each developer or deployer of GPAI to implement a risk management system would likely result in a significant compliance burden and would fail to specifically target actual high risk uses. • Because GPAI systems are often developed for use by third parties, developers of GPAI systems often don't have the insights that would enable identification and analysis of the known or foreseeable risks to health, safety and fundamental rights, or to determine how best to mitigate those risks. For example, the developer of a GPAI image recognition model may not know how the model could be fine-tuned to review x-rays in the healthcare context, and therefore would be unable to analyze and address all foreseeable risks in that use case. • Similarly, in light of the broad scope of potential contexts within which a GPAI system could be used, a GPAI deployer who had no hand in designing or training a GPAI system would not be in a position to predict and manage all risks that could be associated with all of its other "intended purposes." For a deployer that integrates a third-party GPAI model into a downstream AI system, adhering to compliance obligations for both systems would be unworkable. Additionally, requirements to identify and analyze the known and foreseeable risks most likely to occur to health, safety and fundamental rights in view of the intended purpose of the high-risk AI system, or to adopt suitable mitigation measures, are superfluous if the use case for GPAI is not high risk. For example, a natural language processing-based search assistant that can provide responses to general questions about weather conditions or movie showtimes is unlikely to create any real risk for health, safety, or fundamental rights.
Data (Article 10)		

Article 10(1)-(2)	High-risk AI systems which make use of techniques involving the training of models with data shall be developed on the basis of training, validation and testing data sets that meet the quality criteria referred to in paragraphs 2 to 5. Training, validation, and testing data sets shall be subject to appropriate data governance and management practices. Those practices shall consider in particular, (a) the relevant design choices; (b) data collection processes; (c) relevant data preparation processing operations, such as annotation, labelling, cleaning, enrichment and aggregation; (d) the formulation of relevant assumptions, notably with respect to the information that the data are supposed to measure and represent; (e) a prior assessment of the availability, quantity and suitability of the data sets that are needed; (f) examination in view of possible biases that are likely to affect health and safety of natural persons or lead to discrimination prohibited by Union law; (g) the identification of any possible data gaps or shortcomings, and how those gaps and shortcomings can be addressed.	• GPAI models are often trained on large, relatively undifferentiated datasets, and are then fine-tuned to address specific use scenarios. These functions are distributed across the value chain, meaning that neither a developer or deployer of GPAI alone will have the access required to comply fully with data governance practices of the sort contemplated in Article 10. • Deployers may not have visibility into, or control over, the training, validation or testing data for the GPAI system itself. Similarly, once a GPAI system is deployed, developers often will not be in a position to validate or test that fine-tuned model in deployment. • The process of assessing a model and its data's suitability for a particular high-risk purpose will often be context-specific. As noted above, doing such an assessment for all possible uses of a GPAI system will often be impractical, overbroad (including low risk uses), and would fail to specifically target actual high risk uses.
Article 10(3)-(4)	Training, validation and testing data shall be relevant, representative, and to the best extent possible, free of errors and complete. They shall have the appropriate statistical properties, including, where applicable, as regards to the persons or groups of persons on which the high-risk AI system is intended to be used. These characteristics of the data sets may be met at the level of individual data sets or a combination thereof. Training, validation and testing 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, behavioural or functional setting within which the high-risk AI system is intended to be used.	• Given the many possible use cases for GPAI systems and the fact that these models are typically trained on large, undifferentiated datasets, ensuring that training, validation, and testing data sets are "relevant and representative" will in many cases be impossible. Such datasets will also never be "complete." The specific context also impacts what data would be "relevant and representative" (e.g. training data relevant for a healthcare context will be different than an entertainment context). • These requirements are furthermore unnecessary if the use case for the GPAI is low risk and, potentially, intended to be specific – for example, a social chat bot built on an LLM that can discuss indie movies or winning strategies for board games.
Traceability (Article 12)		

Article 12(1)-(2)	High-risk AI systems shall technically allow for the automatic recording of events (‘logs’) over the duration of the life cycle of the system. In order to ensure a level of traceability of the AI system’s functioning that is appropriate to the intended purpose of the system, logging capabilities shall enable the recording of events relevant for: i. identification of situations that may result in the AI system presenting a risk within the meaning of Article 65(1) or in a substantial modification; ii. facilitation of the post-marketing monitoring referred to in Article 61; and iii. monitoring for the operation of high-risk AI systems referred to in Article 29(4).	• Although enabling auditability and oversight of an AI system through data logging might make some sense for systems intended for high-risk uses, imposing such logging capability requirements on GPAI systems when used in low-risk scenarios would serve no justifiable purpose and could make the operation of the system cost-prohibitive. • Logging data usage with respect to low- or no-risk systems in many cases might be inconsistent with these principles.
Instructions (Article 13)		
Article 13 (2) – (3)	High-risk AI systems shall be accompanied by instructions for use in an appropriate digital format or otherwise that include concise, complete, correct and clear information that is relevant, accessible and comprehensible to users. The information referred to in paragraph 2 shall specify: (a) the identity and contact details of the provider and, where applicable of its authorised representative; (b) the characteristics, capabilities and limitations of performance of the high-risk AI system, including: i. its intended purpose, inclusive of the specific geographical, behavioural, or functional setting within which the high-risk AI system is intended to be used; ii. the level of accuracy, including its metrics, robustness and cybersecurity referred to in Article 15 against which the high-risk AI system has been tested and validated and which can be expected, and any known and foreseeable circumstances that may have an impact on that expected level of accuracy, robustness and cybersecurity; iii. any known or foreseeable circumstance, related to the use of the high-risk AI system in accordance with its intended purpose, which may lead to risks to the health and safety or fundamental rights referred to in Article 9(2); […]	Because it is difficult to anticipate all of the various ways in a GPAI system may be deployed, these requirements will often be impossible to comply with. For instance: • GPAI systems do not have a single “intended purpose,” nor do they have a “specific geographical, behavioural, or functional setting within which” they are intended to be used. • The “level of accuracy” of a GPAI system often will depend on the context of its use, as will the “foreseeability” of circumstances that could impact its “level of accuracy, robustness and cybersecurity.” For example, the accuracy of an AI translation tool may vary significantly depending on whether it is used in a retail customer service setting or a medical research setting.
Human Oversight (Article 14)		

Article 14(1)-(3)	High-risk AI system shall be designed and developed in such a way, including with appropriate human-machine interface tools, that they can be effectively overseen by natural persons during the period in which the AI system is in use. Human oversight shall aim at preventing or minimising the risks to health, safety or fundamental rights that may emerge when a high-risk AI system is used in accordance with its intended purpose or under conditions of reasonably foreseeable misuse, in particular when such risks persist notwithstanding the application of the other requirements set out in this Chapter. Human oversight shall be ensured through either one or all of the following types of measures: a) Measures identified and built, when technically feasible, into the high-risk AI system by the provider before it is placed on the market or put into service; b) Measures identified by the provider before placing the high-risk AI system on the market or putting it into service and that are appropriate to be implemented by the user.	• To the extent a GPAI system simply generates an output in response to a given input by the end user, it is unclear what would constitute "effective overs[ight]" of such use, or why it would be necessary. For example, if an end user asks an image generation system to draw a sailboat in the style of Rubens, the end user presumably will be fully capable of determining whether the system worked as intended; separate human oversight, in these and countless other examples, would be unrealistic and provide no benefit. • Similarly, in low- and no-risk settings, such human oversight of a GPAI system's use would serve no useful purpose (e.g., using a GPAI system to organize photos from a corporate event). • Furthermore, because developers of GPAI systems often will have little visibility into how the system is deployed, they will be unable to evaluate the risks to health, safety or fundamental rights that may emerge when the system is used "in accordance with its intended purpose or under conditions of reasonably foreseeable misuse," or determine which human oversight measures are appropriate for a deployer's use case.
Accuracy, Robustness, Cybersecurity (Article 15)		
Article 15(1)	High-risk AI systems shall be designed and developed in such a way that they achieve, in light of their intended purpose, an appropriate level of accuracy, robustness and cybersecurity, and perform consistently in those respects throughout their lifecycle.	GPAI systems are capable of performing a wide array of tasks. It would not be possible for developers to determine a single "appropriate level of accuracy, robustness and cybersecurity" that applies to all possible deployments, or monitor those elements throughout the GPAI system's lifecycle. As elaborated above, appropriate levels of accuracy, robustness, and cybersecurity are best assessed on a context-specific basis for actual high risk uses.
Article 15(4)	High-risk AI systems shall be resilient as regards attempts by unauthorised third parties to alter their use or performance by exploiting the system vulnerabilities. The technical solutions aimed at ensuring the cybersecurity of high-risk AI systems shall be appropriate to the relevant circumstances and the risks. The technical solutions to address AI specific vulnerabilities shall include, where appropriate, measures to prevent and control for attacks trying to manipulate the training dataset ('data poisoning'), inputs designed to cause the model to make a mistake ('adversarial examples'), or model flaws.	• Cybersecurity standards should take into account the sensitivity of the data and use case of the GPAI system. Imposing a one-size-fits-all approach to cybersecurity would impose obligations on GPAI systems that go far beyond what is appropriate in light of the risk presented. • While developers of GPAI systems can take steps to minimize the risk of unauthorized access to their own system by third parties, including through appropriate design solutions, they normally will have no ability to control access to the system in its use, as this will be under the control of the deployer. Similarly, although deployers could implement access controls and similar measures in their own deployments, they would not be able to implement measures related to the design of the GPAI system itself—e.g., to address data poisoning vulnerabilities and attempts to exploit problems in the model itself.
Technical Documentation Requirements (Annex IV)		

Annex IV (1)-(2)	The technical documentation referred to in Article 11(1) shall contain at least the following information, as applicable to the relevant AI system: A general description of the AI system including: a. its intended purpose, the person/s developing the system the date and the version of the system b. how the AI system interacts or can be used to interact with hardware or software that is not part of the AI system itself, where applicable; c. the versions of relevant software or firmware and any requirement related to version update; d. the description of all forms in which the AI system is placed on the market or put into service (e.g., software package embedded into hardware, downloadable, API, etc.); e. the description of hardware on which the AI system is intended to run; f. where the AI system is a component of products, photographs or illustrations showing external features, marking and internal layout of those products; instructions of use for the user and, where applicable installation instructions; […]	The technical documentation required by these provisions are intended for specific, single-purpose AI systems, and even then, require a mix of developer and deployer input. Applying these obligations to providers of GPAI systems will present major compliance challenges, given the multipurpose nature of GPAI systems, and the GPAI developer's practical inability to know about or control the potentially thousands of deployments of its system.
Annex IV(8)	A detailed description of the system in place to evaluate the AI system performance in the post-market phase in accordance with Article 61, including the post-market monitoring plan referred to in Article 61(3).	• Because providers normally will not be in a position to monitor performance of the system in the post-market phase with respect to all deployments, they will be unable to provide a "detailed description" of the system in place. • As acknowledged in the Council text, any post-market monitoring system should be "proportionate to the risks" of the system. Given that many GPAI systems will be low- or no-risk in nature, it would be disproportionate to require post-market monitoring for all GPAI systems and uses.

KL/128/50/ET/2023