EUROPEAN COMMISSION
DIRECTORATE-GENERAL FOR COMMUNICATIONS NETWORKS, CONTENT AND TECHNOLOGY

Artificial Intelligence Office
Artificial Intelligence Safety

European Commission

Call for tenders EC-CNECT/2025/OP/0032 - ARTIFICIAL INTELLIGENCE ACT: TECHNICAL ASSISTANCE FOR AI SAFETY

Open procedure

TENDER SPECIFICATIONS

TABLE OF CONTENTS

1. SCOPE AND DESCRIPTION OF THE PROCUREMENT

1.1. Contracting authority: who is the buyer?

This call for tenders is launched and managed by the European Commission, DG CNECT - Communications Networks, Content and Technology, referred to as the contracting authority for the purposes of this call for tenders.

1.2. Subject: what is this call for tenders about?

The subject of this call for tenders is "ARTIFICIAL INTELLIGENCE ACT: TECHNICAL ASSISTANCE FOR AI SAFETY".

1.3. Lots: is this call for tenders divided into lots?

This call for tenders is divided into 6 lots:

Lot number	Lot title
Lot 1	CBRN Risk Modelling and Evaluation
Lot 2	Cyber Offence Risk Modelling and Evaluation
Lot 3	Loss of Control Risk Modelling and Evaluation
Lot 4	Harmful Manipulation Risk Modelling and Evaluation
Lot 5	Sociotechnical Risk Modelling and Evaluation
Lot 6	Agentic Evaluation Interface

Tenders may be submitted for any lot. Each lot will be assessed independently of any other lot. Tenders, which cover only part of one lot or are declared as being conditional on the award of any other lots, are not permitted.

1.4. Description: what do we want to buy through this call for tenders?

The purchases that are the subject of this call for tenders, including any minimum requirements, are described in detail below.

Variants (alternatives to the model solution described in the tender specifications) are not allowed for any lot. The contracting authority will disregard any variants described in a tender.

1.4.1. Background and objectives

1.4.1.1. Context of the procurement

Regulation (EU) 2024/1689 (Artificial Intelligence Act) ([1]), which has entered into force on the 1st of August 2024, and for which certain rules will become applicable on the 2nd of August 2025 establishes a comprehensive legal framework governing the area of Artificial Intelligence, and in particular relevance to this procurement, establishes responsibilities for providers of General-Purpose Artificial Intelligence (GPAI) models, aiming to ensure transparency, safety, and accountability in the deployment and use of AI technologies.

The Artificial Intelligence Act (AI Act) applies to both public and private actors inside and outside the EU as long as the AI system is placed on the Union market, or its use has an impact on people located in the EU. With respect to GPAI models, Article 53 of the AI Act imposes obligations to draw up and keep up-to-date technical documentation of the model, making it available to downstream providers, or upon request, to the AI Office and the national competent authorities. Certain substantive obligations are limited to GPAI models with systemic risk, and Article 55 of the AI Act imposes obligations for providers of GPAI models with systemic risks to, among others, perform model evaluations, perform adversarial testing, and assess and mitigate possible systemic risks at Union level.

Articles 89 and 92 of the AI Act respectively allow the AI Office to monitor the compliance with regards to the AI Act and perform evaluations of GPAI models to assess the compliance and investigate systemic risks at Union level. Article 93 of the AI Act allows the Commission to request providers to take appropriate measures to comply with the obligations, to implement mitigation measures, or to restrict the availability of the model.

In order to carry out these enforcement tasks in relation the AI Act, the AI Office is seeking support of suitable third-party contractors to provide technical assistance aimed at supporting the monitoring of compliance, in particular in its assessment of risks posed by GPAI models at Union level.

1.4.1.2. Objectives of the procurement

The services under this procurement will support the AI Office of the European Commission to carry out its tasks laid down in Regulation (EU) 2024/1689 as regards to GPAI models and GPAI models with systemic risk.

([1]) Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (Text with EEA relevance), https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32024R1689

1.4.2. Detailed characteristics of the purchase

1.4.2.1. Overview of the procurement

This tender has been split into six lots. Five of those lots are dedicated to specific systemic risks, namely (i) CBRN risks, (ii) cyber offence risks, (iii) loss of control risks, (iv) harmful manipulation risks and (v) sociotechnical risks. These lots all share a similar structure.

Furthermore, there is one lot that is cross-cutting across various risk.

	Lot number	Lot title
Risk-specific	Lot 1	CBRN Risk Modelling and Evaluation
	Lot 2	Cyber Offence Risk Modelling and Evaluation
	Lot 3	Loss of Control Risk Modelling and Evaluation
	Lot 4	Harmful Manipulation Risk Modelling and Evaluation
	Lot 5	Sociotechnical Risk Modelling and Evaluation
Cross-cutting	Lot 6	Agentic Evaluation Interface

1.4.2.1.1. Risk-specific lots: risk modelling and evaluation suites

Every risk-specific lot has a similar structure, containing the following services:

- The organisation of multiple risk modelling workshops together with the AI Office, producing a risk model, risk scenarios, as well as corresponding thresholds, informing further work on evaluations.
- The development of new evaluation tools that are private to the Commission.
- The onboarding of existing evaluations that are publicly available into the developed risk framework.
- The creation of a reference procedure and reporting template for doing risk assessment with respect to the developed evaluations.
- Ad hoc and on demand services for complex evaluations, including for example red-teaming and uplift evaluations.
- A risk monitoring service in the form of regular briefings on new developments (e.g. models, risk sources, elicitation techniques, mitigation, incidents) to complement existing AI Office monitoring actions.

The scope of these tasks varies by lot.

1.4.2.1.2. Cross-cutting lots relevant to multiple risks

- Lot 6: agentic evaluation interface

Software, Cloud infrastructure support, and methodology to evaluate GPAI on diverse types of benchmarks, focusing on agentic interaction and complex agentic benchmarks.

1.4.2.2. Description of Lot 1: CBRN Risk Modelling and Evaluation

All tasks are subject to the provisions outlined in section 1.4.2.8 Description: Shared Provisions.

1.4.2.2.1. Objective

The objective of this lot is to support the AI Office in its enforcement of the AI Act in relation to General-Purpose AI Models (GPAI) and General-Purpose AI Models with systemic risk (GPAISR), specifically where such models may pose risks of enabling chemical, biological, radiological, and nuclear attacks or accidents. This includes significantly lowering the barriers to entry for malicious actors, or increasing the potential impact achieved, in the design, development, acquisition, and use of such weapon. The AI Office seeks to improve its capacity in identifying, monitoring, and assessing these risks, including the capacity to measure relevant capabilities of GPAI models and assess the effectiveness of mitigations aimed at preventing malicious actors from gaining access to these capabilities.

This lot aims to allow for flexible and close collaboration in developing risk models, prioritising risk scenarios, setting risk levels including a level of unacceptable systemic risk at Union level, identifying key model capabilities that could be linked to risk levels, and tailoring both new and existing risk measurement instruments to the context of the AI Office. Flexible collaboration between the AI Office and the contractor will be needed to ensure robust risk assessment as risks and our understanding of them evolve.

1.4.2.2.2. Deliverables

(a) Inception meeting and report

The contractor will prepare an inception report, covering all tasks in this lot, providing a preliminary overview of the proposed approach, including staff capacities as well as measures for compliance with personal data protection legislation. Sufficient attention should be given to the planned approach for the risk modelling workshop. For avoidance of doubt, the draft inception report can be substantively the same as the work plan submitted with the tender, if such a plan is sufficiently detailed and up to date to reflect the contract as signed. The inception report will be discussed at the inception meeting, which will be a digital meeting happening at the latest 1 month after the signing of the contract. The inception report draft will be provided at least 5 working days before the inception meeting, and the final version will be provided at most 5 working days after the inception meeting, integrating any discussion and feedback for such meeting.

The report will serve as the initial foundation for the execution of the tasks and any collaboration.

(b) Risk modelling workshops and reports

A risk modelling workshop early in the contract, at most two months after contract signing, serves to align contractor and AI Office perspectives on enforcement priorities and the desired roadmap to relevant risk measurement instruments. This includes for example discussion on which domains (chemical, biological, radiological, nuclear) to prioritise, or on the importance of accessibility risks, i.e. the risks in relation to the amplification of capabilities of non-expert threats, versus ceiling risks, i.e. those risks related to amplification of capabilities from actors that can already cause significant harm.

The contractor shall develop a coherent risk model that can serve as a framework for the development and integration of risk measurement instruments through the other deliverables in this lot. Although details are to be agreed with AI Office during the workshop and further communication, such risk model could include, but are by no means limited to:

- Development of several risk scenarios
- Execution of specific risk analysis methods established in literature
- Proposals for concrete risk levels for which to develop measurement instruments
- Surveying existing risk modelling literature and its adaptation to the context of the AI Office

Risk models should consider mitigations, technical or otherwise, and their weaknesses, applying both to private GPAI models as well as GPAI models made available to the public through the public release of their weights. Risk models should also integrate with existing EU policies on risk management.

The risk modelling workshop will be a physical meeting taking place at the premises of Directorate-General for Communications Networks, Content and Technology (DG CNECT) in Brussels. A draft report of the approach to risk modelling and for the risk modelling workshop will be provided one month before the workshop, while one month after the workshop a draft risk model must be provided, to be finalised three months after the workshop. Any reports will include a summary oriented to policymakers. Regular calls can be scheduled by the contractor or the AI Office in the case more alignment on the risk modelling is needed.

This risk modelling workshop, including the pre- and post-workshop drafts reports and final reports, will be repeated one year after the first one.

Either or both workshops may be conducted digitally upon mutual agreement between the contractor and the AI Office, with corresponding adjustments to resource allocation across deliverables.

Data protection: where the contractor processes personal data in the context of this Lot, the contractor must respect Regulation (EU) 2018/1725, including following instructions (in the light of Article 29 Regulation (EU) 2018/1725) from the Commission. In particular, the contractor will ensure that the processing activities will comply with the relevant data protection record (and privacy statement), namely DPR-EC-01063 "Processing of personal data linked to meetings and events".

Where they are fit for purpose, the contractor will use Commission corporate tools, such as EU Survey or EventWorks for registration purposes, Slido for interactivity, etc.

If the Commission deems it necessary, the contractor will support the Commission in any other relevant data protection aspect, and in drafting any needed data protection documentation (data protection record, data protection impact assessment, etc.). The Commission will inform the contractor of the need for such documentation during the implementation of the contract.

Intellectual property rights: As an exception to the modes of exploitation set out in Article I.8 of the service contract, the Commission may NOT make the results corresponding with these deliverables publicly available, including the distribution to the public in hard copies, in electronic or digital format, on the internet including social networks as a downloadable or non-downloadable file or the communication through press information services, unless with written agreement from the contractor. The Commission may still use the results for all other listed modes of exploitation.

(c) Onboarded public evaluation tools

Based on the developed risk scenarios and risk levels, the contractor shall analyse and assess for use any relevant existing, publicly available, and risk related evaluations tools, for example benchmark datasets. Evaluation tools that are deemed suitable or close to suitable shall be adopted, amended, or improved for use within the AI Office context.

This process can include, but is no way limited to, the filtering and cleaning of data points from benchmarks, the setting of relevant risk levels of the corresponding measurement instrument, the integration with tailored elicitation techniques, instance-level analysis of the cognitive demands required, or the measurement of human expert performance on existing benchmarks to compare GPAI model performance against.

The onboarding of public evaluation tools may also include a contamination study of the respective datasets, i.e. a quantitative or qualitative analysis of the degree to which the benchmark data has been present in the training material for existing GPAI models or their safeguards, using e.g. the presence of "canary strings" or other benchmark specific information in model responses, acknowledging the speculative nature of such a study.

There will be three corresponding deliverables. Firstly, a report detailing which public benchmarks to onboard, including a motivation for this selection, as well as a short, non-exhaustive list of example candidate evaluation tools that did not make the selection, and a motivation for their exclusion.

Secondly, an intermediary handover dedicated to the onboarding of a subset of the evaluation tools to be onboarded. Such a handover will contain:

- A report detailing any changes made to the evaluation tools included and any difficulties encountered.
- A report on the methodology and results for any substantive work conducted, e.g., to obtain human responses.
- Any modified or newly created resources such as benchmarks, datasets of human responses and their grading, evaluation software, etc.

Thirdly, a complete handover, similar to the intermediary handover, but now covering the full set of evaluation tools to onboard.

Both handovers should make the relevant onboarded evaluation tools ready-to-use in the context of the AI Office.

The initial proposal will be delivered at the latest 3 months after signing of the contract, while the intermediate handover will be delivered at the latest 6 months after signing of the contract, and the final handover 12 months after signing of the contract.

When doing so would be deemed beneficial for the overall procurement objective, for example in the case that no public resources fit for onboarding can be identified in the timeline currently set out, the contractor and the AI Office may decide jointly to postpone execution of this task, otherwise arrange the execution timeline in a more flexible manner, or reorganise team capacity to improve the scope or quality of other tasks. This flexibility must in all case respect the maximum duration of the contract.

Work on these deliverables should be informed by the outcomes of the risk modelling deliverables, and in particular by the first risk modelling workshop.

Data protection: For all processing activities, the contractor will support the Commission in all relevant data protection aspects, and in particular in drafting any needed data protection documentation (data protection record, data protection impact assessment, etc.). The Commission will inform the contractor of the need for such documentation during the implementation of the contract. The contractor will also ensure that the deliverables do not contain any personal data within the meaning of Regulation (EU) 2018/1725 (the EUDPR) or Regulation (EU) 2016/679 (GDPR).

(d) Private evaluation tools

Based on the developed risk model and the established utility of public evaluations tools (including the plans for their adaptation), the contractor will develop new evaluation tools for private use by the AI Office. These tools are expected to be one or multiple evaluation datasets that are fit for cheap automatic evaluation procedures and executable by AI Office technical experts. The evaluation methodology may extend beyond simple datasets of question-answer pairs, but the evaluation must be able to be executed repeatedly by the AI Office at acceptable costs and operational complexity.

Where relevant, these private evaluations will be complemented by the collection of human expert and non-expert answers as to compare GPAI model performance to.

The focus should be on a small but highly relevant selection of data, that is unique to the AI Office, can act as a verification of public test results with uncontaminated data, and serves as a proxy for expensive evaluations such as uplift-oriented evaluations.

There will be three corresponding deliverables. Firstly, a report detailing which evaluation tools are to be developed including the approach planned to develop them, to be delivered at most 1 month after the conclusion of the first risk modelling workshop.

Secondly, an intermediary handover dedicated to the development of a subset of the evaluation tools to be developed. Such a handover will contain:

- A report detailing the structure of the evaluation tool, its strengths and limitations, the method used for constructing it, and any problems encountered during construction.
- Any modified or newly created resources such as benchmarks, datasets of human responses and their grading, evaluation software, etc.
- A report on a sample evaluation, where an existing GPAI model available on the market is evaluated using the newly developed tools, insofar as it is reasonably feasible for the contractor to do so while preserving the privacy of the dataset. For instance, a privately hosted instance of a state-of-the-art open-source model could be used for the sample evaluation if non-logging assurances cannot be obtained for commercial APIs.
- Information on the measures taken to preserve the confidentiality of the evaluation data, for example, the presence of agreements with the provider of the GPAI model used in the sample evaluations regarding data confidentiality.

Thirdly, a complete handover, similar to the intermediary handover, but now covering the full set of private evaluations tools.

Both handovers should make the developed evaluation tools ready-to-use in the context of the AI Office.

The initial proposal will be delivered at the latest 3 months after signing of the contract, while the intermediate handover will be delivered at the latest 6 months after signing of the contract, and the final handover 12 months after signing of the contract.

Work on these deliverables should be informed by the first risk modelling workshop.

Data protection: For all processing activities, the contractor will support the Commission in all relevant data protection aspects, and in particular in drafting any needed data protection documentation (data protection record, data protection impact assessment, etc.). The Commission will inform the contractor of the need for such documentation during the implementation of the contract. The contractor will also ensure that the deliverables do not contain any personal data within the meaning of Regulation (EU) 2018/1725 (the EUDPR) or Regulation (EU) 2016/679 (GDPR).

(e) Technical compliance toolkit

Based on the developed risk model, the onboarded public evaluations, and newly created private evaluations tools, the contractor shall design a comprehensive draft evaluation process and draft reporting template that allows the AI Office to assess the systemic risk at Union level from a given model in a consistent, unified, and legible way.

The contractor will also design a redacted version of this technical compliance toolkit for potential publication by the AI Office, which could, if published, act as minimum standard for risk analysis to be referred to by GPAI providers and developers. The publication decision will be taken in due course by the AI Office.

This toolkit shall be delivered to the AI Office one month after the completion of both the onboarding of public evaluation tools and the development of private evaluation tools, and shall include:

- A report detailing the evaluation process and its methods, referring the onboarded public evaluation tools and developed private evaluation tools.

- A reporting template, tailored to policymakers in language and structure, which could be filled in and completed based on the results of the aforementioned evaluation process.

- A redacted version of both the evaluation process report and the reporting template (see above) containing no references to confidential information or resources, for potential publication.

(f) Risk monitoring

To complement the AI Office risk monitoring capacities, the contractor will provide the AI Office with regular update briefings on any events that are relevant for the developed risk models, for example major new GPAI models or new GPAI providers, algorithmic improvements, new insights into the specific risks, updated safety frameworks, policy changes, incidents, new mitigations, or the documented breaking of a mitigation.

These briefings should be tailored to the specific risks considered in this lot, can be informal in nature, and are not expected to be exhaustive. They should integrate with contractors existing effort to stay up to date with the field of AI and the relevant risks in light of their other tasks, providing a lightweight instrument for sharing relevant updates with the AI Office.

The AI Office expects there to be events potentially worthy of briefing roughly every two weeks. This expectation should not be seen as prescriptive, and the sensitivity of briefing regime is expected to be calibrated in agreement with the AI Office, considering the intention to avoid unnecessary burden on the contractor.

(g) Ad hoc uplift evaluations

To quickly respond to changes in the AI risk landscape, the AI Office will need to evaluate particular GPAI models on short notice. Therefore, the AI Office expects to request the contractor to conduct ad hoc evaluations at certain points in time. A maximum of 3 such evaluations may be commissioned.

For one such evaluation, the expectation regarding the contractor, is to make available a fixed budget of 200 expert hours with 2 weeks' notice to coordinate "uplift evaluations" for a GPAI model of the AI Office's choosing, to be spend over at most 4 weeks, where "uplift evaluations" are investigations into the degree to which GPAI systems can help humans in completing certain tasks, and any conclusions that can be drawn about systemic risks.

This specific method for the evaluation shall be tailored to the developed risk models and evaluation tools, to the degree these have already been finished. The AI Office may request regular updates from the moment of notice until the moment the evaluation report has been approved.

An evaluation report shall be delivered, at most 6 weeks after notice, containing

- A general description of the uplift evaluation, its methodology, results, and conclusions.
- A policymaker-oriented summary.
- A dataset of the anonymized interaction transcripts, containing no personal data within the meaning of Regulation (EU) 2018/1725 (EUDPR) or Regulation (EU) 2016/679 (GDPR), including e.g. requests send to the GPAI model, GPAI model outputs, any assigned labels or grades, and other information that would allow for the reconstruction of the evaluation results.

The contractor is responsible for any costs associated with the request evaluation, including but not limited to, the payment of study participants, or the API costs associated with the GPAI model.

Informed by the first risk modelling workshop, AI Office and contractor may begin to consider a potential experimental design for upcoming ad hoc uplift evaluations, to facilitate rapid execution when the evaluation is required.

Data protection: the contractor will support the Commission in all relevant data protection aspects, and in particular in drafting any needed data protection documentation (data protection record, data protection impact assessment, etc.). The Commission will inform the contractor of the need for such documentation during the implementation of the contract.

(h) Ad hoc red teaming

To quickly respond to changes in the AI risk landscape, the AI Office will need to evaluate particular GPAI models on short notice. Therefore, the AI Office expects to request the contractor to conduct ad hoc evaluations at certain points in time. A maximum of 3 such evaluations may be commissioned.

For one such evaluation, the expectation regarding the contractor, is to make available a fixed budget of 60 expert hours with 2 weeks' notice, to be spend in a maximum of 2 weeks, to execute red teaming exercises with a GPAI model of the AI Office's choosing. These red teaming exercises may aim to

stress test mitigations put in place my model providers to prevent malicious use, or to explore certain types of usage patterns.

This specific method of red teaming shall be tailored to the developed risk models and evaluation tools, to the degree these have already been finished, and aligned with the AI Office ahead of starting the red teaming.

The AI Office may request regular updates from the moment of notice until the moment the evaluation report has been approved.

A red teaming report shall be delivered, at most 4 weeks after notice, containing

- A general description of the red teaming, its methodology, results, and conclusions.
- A policymaker-oriented summary.
- A dataset of the interaction transcripts, containing no personal data within the meaning of Regulation (EU) 2018/1725 (EUDPR) or Regulation (EU) 2016/679 (GDPR), including e.g. requests send to the GPAI model, GPAI model outputs, assigned labels or grades, and other information that would allow for the reconstruction of the results.

The contractor is responsible for any costs associated with the requested evaluation, including but not limited to, the API costs associated with the GPAI model. Where reasonably feasibly, the contractor should ensure they have permission from the GPAI model provider to temporarily break the terms of service for the purpose of the exercise. Where not feasible, the AI Office shall aim to provide the required access.

Data protection: the contractor will support the Commission in all relevant data protection aspects, and in particular in drafting any needed data protection documentation (data protection record, data protection impact assessment, etc.). The Commission will inform the contractor of the need for such documentation during the implementation of the contract.

(i) Monthly calls

The contractor shall organise monthly calls with the AI Office to provide updates on the different tasks on the basis of the methodology and timelines agreed. Meeting minutes will be provided by the contractor at most five days after the monthly call.

Any reports or materials due in the same month shall be submitted at least five working days in advance of the monthly call to allow for review by the AI Office.

In justified cases, the AI Office may request ad hoc calls to complement the monthly calls and address specific issues that might arise during the execution of the contract. This will not impact the overall effort from the side of the contractor and will not cause any change to the minimum requirements of the tender specifications

(j) Final meeting and report

The contractor shall prepare a final report, which shall provide a comprehensive and detailed overview of the tasks conducted in the previous years. The final report should include the measures adopted by the contractor to ensure data protection compliance.

The submission of the final report is expected at the latest within 36 months from the contract's start date. The said final report will be linked to the final payment associated with the end of the contract.

The final report will be presented during a final meeting, organised by the contractor and to take place digitally.

The draft final report shall be sent to the European Commission 10 days before the final meeting at the latest. This will allow the Commission to assess the final results and provide comments where necessary, and the contractor to amend the report if required.

1.4.2.2.3. Timetable

Title	Due month (at the latest)	Linked to payment
Core Deliverables		
Risk modelling draft reports	M1, M3; M13, M15	No
Risk modelling workshops	M2; M14	No
Risk modelling reports	M6; M18	Yes
Public eval. tools proposal report	M3	No
Public eval. tools intermediate handover	M6	No
Public eval. tools handover	M12	Yes
Private eval. tools proposal report	M3	No
Private eval. tools intermediate handover	M6	No
Private evaluation tools handover	M12	Yes
Technical compliance toolkit	M13	No
Continued services		
Risk monitoring	Continuous	No
Ad hoc uplift evaluations	Upon request	No
Ad hoc red teaming	Upon request	No
Coordination		
Inception meeting	M1	No
Inception report	M1	No
Monthly calls	M2-M35	No
Final meeting	M36	No
Final report	M36	Yes

1.4.2.3. Description of Lot 2: Cyber Offence Risk Modelling and Evaluation

All tasks are subject to the provisions outlined in section 1.4.2.8 Description: Shared Provisions.

1.4.2.3.1. Objective

The objective of this lot is to support the AI Office in its enforcement of the AI Act in relation to General-Purpose AI Models (GPAI) and General-Purpose AI Models with systemic risk (GPAISR), specifically where such models may pose risks related to offensive cyber capabilities that could enable large-scale or sophisticated cyber-attacks, including on critical systems (e.g. critical infrastructure). This includes automated vulnerability discovery, exploit generation, operational use, and attack scaling. The AI Office seeks to improve its capacity in identifying, monitoring, and assessing these risks, including the capacity to measure relevant capabilities of GPAI models and assess the effectiveness of mitigations aimed at preventing malicious actors from gaining access to these capabilities.

This lot aims to allow for flexible and close collaboration in developing risk models, prioritising risk scenarios, setting risk levels including a level of unacceptable systemic risk at Union level, identifying key model capabilities that could be linked to risk levels, and tailoring both new and existing risk measurement instruments to the context of the AI Office. Flexible collaboration between the AI Office and the contractor will be needed to ensure robust risk assessment as risks and our understanding of them evolve.

1.4.2.3.2. Deliverables

(a) Inception meeting and report

The contractor will prepare an inception report, covering all tasks in this lot, providing a preliminary overview of the proposed approach, including staff capacities as well as measures for compliance with personal data protection legislation. Sufficient attention should be given to the planned approach for the risk modelling workshop. For avoidance of doubt, the draft inception report can be substantively the same as the work plan submitted with the tender, if such a plan is sufficiently detailed and up to date to reflect the contract as signed. The inception report will be discussed at the inception meeting, which will be a digital meeting happening at the latest 1 month after the signing of the contract. The inception report draft will be provided at least 5 working days before the inception meeting, and the final version will be provided at most 5 working days after the inception meeting, integrating any discussion and feedback for such meeting.

The report will serve as the initial foundation for the execution of the tasks and any collaboration.

(b) Risk modelling workshops and reports

A risk modelling workshop early in the contract, at most two months after contract signing, serves to align contractor and AI Office perspectives on enforcement priorities and the desired roadmap to relevant risk measurement instruments. This includes for example discussion on which parts of the cyber kill chain to prioritise, or on the importance of accessibility risks, i.e. the risks in relation to the amplification of capabilities of non-expert threats, versus ceiling risks, i.e. those risks related to amplification of capabilities from actors that can already cause significant harm.

The contractor shall develop a coherent risk model that can serve as a framework for the development and integration of risk measurement instruments through the other deliverables in this lot. Although

details are to be agreed with AI Office during the workshop and further communication, such risk model could include, but are by no means limited to:

- Development of several risk scenarios
- Execution of specific risk analysis methods established in literature
- Proposals for concrete risk levels for which to develop measurement instruments
- Surveying existing risk modelling literature and its adaptation to the context of the AI Office

Risk models should consider mitigations, technical or otherwise, and their weaknesses, applying both to private GPAI models as well as GPAI models made available to the public through the public release of their weights. Risk models should also integrate with existing EU policies on risk management.

The risk modelling workshop will be a physical meeting taking place at the premises of Directorate-General for Communications Networks, Content and Technology (DG CNECT) in Brussels. A draft report of the approach to risk modelling and for the risk modelling workshop will be provided one month before the workshop, while one month after the workshop a draft risk model must be provided, to be finalised three months after the workshop. Any reports will include a summary oriented to policymakers. Regular calls can be scheduled by the contractor or the AI Office in the case more alignment on the risk modelling is needed.

This risk modelling workshop, including the pre- and post-workshop drafts reports and final reports, will be repeated one year after the first one.

Either or both workshops may be conducted digitally upon mutual agreement between the contractor and the AI Office, with corresponding adjustments to resource allocation across deliverables.

Data protection: where the contractor processes personal data in the context of this Lot, the contractor must respect Regulation (EU) 2018/1725, including following instructions (in the light of Article 29 Regulation (EU) 2018/1725) from the Commission. In particular, the contractor will ensure that the processing activities will comply with the relevant data protection record (and privacy statement), namely DPR-EC-01063 “Processing of personal data linked to meetings and events”.

Where they are fit for purpose, the contractor will use Commission corporate tools, such as EU Survey or EventWorks for registration purposes, Slido for interactivity, etc.

If the Commission deems it necessary, the contractor will support the Commission in any other relevant data protection aspect, and in drafting any needed data protection documentation (data protection record, data protection impact assessment, etc.). The Commission will inform the contractor of the need for such documentation during the implementation of the contract.

Intellectual property rights: As an exception to the modes of exploitation set out in Article I.8 of the service contract, the Commission may NOT make the results corresponding with these deliverables publicly available, including the distribution to the public in hard copies, in electronic or digital format, on the internet including social networks as a downloadable or non-downloadable file or the communication through press information services, unless with written agreement from the contractor. The Commission may still use the results for all other listed modes of exploitation.

(c) Onboarded public evaluation tools

Based on the developed risk scenarios and risk levels, the contractor shall analyse and assess for use any relevant existing, publicly available, and risk related evaluations tools, for example benchmark datasets. Evaluation tools that are deemed suitable or close to suitable shall be adopted, amended, or improved for use within the AI Office context.

This process can include, but is no way limited to, the filtering and cleaning of data points from benchmarks, the setting of relevant risk levels of the corresponding measurement instrument, the integration with tailored elicitation techniques, instance-level analysis of the cognitive demands required, or the measurement of human expert performance on existing benchmarks to compare GPAI model performance against.

The onboarding of public evaluation tools may also include a contamination study of the respective datasets, i.e. a quantitative or qualitative analysis of the degree to which the benchmark data has been present in the training material for existing GPAI models or their safeguards, using e.g. the presence of "canary strings" or other benchmark specific information in model responses, acknowledging the speculative nature of such a study.

There will be three corresponding deliverables. Firstly, a report detailing which public benchmarks to onboard, including a motivation for this selection, as well as a short, non-exhaustive list of example candidate evaluation tools that did not make the selection, and a motivation for their exclusion, to be delivered at most 1 month after the conclusion of the workshop.

Secondly, an intermediary handover dedicated to the onboarding of a subset of the evaluation tools to be onboarded. Such a handover will contain:

- A report detailing any changes made to the evaluation tools included and any difficulties encountered.
- A report on the methodology and results for any substantive work conducted, e.g., to obtain human responses.
- Any modified or newly created resources such as benchmarks, datasets of human responses and their grading, evaluation software, etc.

Thirdly, a complete handover, similar to the intermediary handover, but now covering the full set of evaluation tools to onboard.

Both handovers should make the relevant onboarded evaluation tools ready-to-use in the context of the AI Office.

The initial proposal will be delivered at the latest 3 months after signing of the contract, while the intermediate handover will be delivered at the latest 6 months after signing of the contract, and the final handover 12 months after signing of the contract.

When doing so would be deemed beneficial for the overall procurement objective, for example in the case that no public resources fit for onboarding can be identified in the timeline currently set out, the contractor and the AI Office may decide jointly to postpone execution of this task, otherwise arrange the execution timeline in a more flexible manner, or reorganise team capacity to improve the scope or quality of other tasks. This flexibility must in all case respect the maximum duration of the contract.

Work on these deliverables should be informed by the outcomes of the risk modelling deliverables, and in particular by the first risk modelling workshop.

Data protection: For all processing activities, the contractor will support the Commission in all relevant data protection aspects, and in particular in drafting any needed data protection documentation (data protection record, data protection impact assessment, etc.). The Commission will inform the contractor of the need for such documentation during the implementation of the contract. The contractor will also ensure that the deliverables do not contain any personal data within the meaning of Regulation (EU) 2018/1725 (the EUDPR) or Regulation (EU) 2016/679 (GDPR).

(d) Private evaluation tools

Based on the developed risk model and the established utility of public evaluations tools (including the plans for their adaptation), the contractor will develop new evaluation tools for private use by the AI Office. These tools are expected to be one or multiple evaluation datasets that are fit for cheap automatic evaluation procedures and executable by AI Office technical experts. The evaluation methodology may extend beyond simple datasets of question-answer pairs, but the evaluation must be able to be executed repeatedly by the AI Office at acceptable costs and operational complexity.

Where relevant, these private evaluations will be complemented by the collection of human expert and non-expert answers as to compare GPAI model performance to.

The focus should be on a small but highly relevant selection of data, that is unique to the AI Office, can act as a verification of public test results with uncontaminated data, and serves as a proxy for expensive evaluations such as uplift-oriented evaluations.

There will be three corresponding deliverables. Firstly, a report detailing which evaluation tools are to be developed including the approach planned to develop them, to be delivered at most 1 month after the conclusion of the first risk modelling workshop.

Secondly, an intermediary handover dedicated to the development of a subset of the evaluation tools to be developed. Such a handover will contain:

- A report detailing the structure of the evaluation tool, its strengths and limitations, the method used for constructing it, and any problems encountered during construction.
- Any modified or newly created resources such as benchmarks, datasets of human responses and their grading, evaluation software, etc.
- A report on a sample evaluation, where an existing GPAI model available on the market is evaluated using the newly developed tools, insofar as it is reasonably feasible for the contractor to do so while preserving the privacy of the dataset. For instance, a privately hosted instance of a state-of-the-art open-source model could be used for the sample evaluation if non-logging assurances cannot be obtained for commercial APIs.
- Information on the measures taken to preserve the confidentiality of the evaluation data, for example, the presence of agreements with the provider of the GPAI model used in the sample evaluations regarding data confidentiality.

Thirdly, a complete handover, similar to the intermediary handover, but now covering the full set of private evaluations tools.

Both handovers should make the developed evaluation tools ready-to-use in the context of the AI Office.

The initial proposal will be delivered at the latest 3 months after signing of the contract, while the intermediate handover will be delivered at the latest 6 months after signing of the contract, and the final handover 12 months after signing of the contract.

Work on these deliverables should be informed by the outcomes of the risk modelling deliverables, and in particular by the first risk modelling workshop.

Data protection: For all processing activities, the contractor will support the Commission in all relevant data protection aspects, and in particular in drafting any needed data protection documentation (data protection record, data protection impact assessment, etc.). The Commission will inform the contractor of the need for such documentation during the implementation of the contract. The contractor will also ensure that the deliverables do not contain any personal data within the meaning of Regulation (EU) 2018/1725 (the EUDPR) or Regulation (EU) 2016/679 (GDPR).

(e) Technical compliance toolkit

Based on the developed risk model, the onboarded public evaluations, and newly created private evaluations tools, the contractor shall design a comprehensive draft evaluation process and draft reporting template that allows the AI Office to assess the systemic risk at Union level from a given model in a consistent, unified, and legible way.

The contractor will also design a redacted version of this technical compliance toolkit for potential publication by the AI Office, which could, if published, act as minimum standard for risk analysis to be referred to by GPAI providers and developers. The publication decision will be taken in due course by the AI Office.

This toolkit shall be delivered to the AI Office one month after the completion of both the onboarding of public evaluation tools and the development of private evaluation tools, and shall include:

- A report detailing the evaluation process and its methods, referring the onboarded public evaluation tools and developed private evaluation tools.
- A reporting template, tailored to policymakers in language and structure, which could be filled in and completed based on the results of the aforementioned evaluation process.
- A redacted version of both the evaluation process report and the reporting template (see above) containing no references to confidential information or resources, for potential publication.

(f) Risk monitoring

To complement the AI Office risk monitoring capacities, the contractor will provide the AI Office with regular update briefings on any events that are relevant for the developed risk models, for example major new GPAI models or new GPAI providers, algorithmic improvements, new insights into the specific risks, updated safety frameworks, policy changes, incidents, new mitigations, or the documented breaking of a mitigation.

These briefings should be tailored to the specific risks considered in this lot, can be informal in nature, and are not expected to be exhaustive. They should integrate with contractors existing effort to stay

up to date with the field of AI and the relevant risks in light of their other tasks, providing a lightweight instrument for sharing relevant updates with the AI Office.

The AI Office expects there to be events potentially worthy of briefing roughly every two weeks. This expectation should not be seen as prescriptive, and the sensitivity of briefing regime is expected to be calibrated in agreement with the AI Office, considering the intention to avoid unnecessary burden on the contractor.

(g) Ad hoc uplift evaluations

To quickly respond to changes in the AI risk landscape, the AI Office will need to evaluate particular GPAI models on short notice. Therefore, the AI Office expects to request the contractor to conduct ad hoc evaluations at certain points in time. A maximum of 3 such evaluations may be commissioned.

For one such evaluation, the expectation regarding the contractor, is to make available a fixed budget of 200 expert hours with 2 weeks' notice to coordinate "uplift evaluations" for a GPAI model of the AI Office's choosing, to be spend over at most 4 weeks, where "uplift evaluations" are investigations into the degree to which GPAI systems can help humans in completing certain tasks, and any conclusions that can be drawn about systemic risks.

This specific method for the evaluation shall be tailored to the developed risk models and evaluation tools, to the degree these have already been finished. The AI Office may request regular updates from the moment of notice until the moment the evaluation report has been approved.

An evaluation report shall be delivered, at most 6 weeks after notice, containing

- A general description of the uplift evaluation, its methodology, results, and conclusions.
- A policymaker-oriented summary.
- A dataset of the anonymized interaction transcripts, containing no personal data within the meaning of Regulation (EU) 2018/1725 (EUDPR) or Regulation (EU) 2016/679 (GDPR), including e.g. requests send to the GPAI model, GPAI model outputs, any assigned labels or grades, and other information that would allow for the reconstruction of the evaluation results.

The contractor is responsible for any costs associated with the request evaluation, including but not limited to, the payment of study participants, or the API costs associated with the GPAI model.

Informed by the first risk modelling workshop, AI Office and contractor may begin to consider a potential experimental design for upcoming ad hoc uplift evaluations, to facilitate rapid execution when the evaluation is required.

Data protection: the contractor will support the Commission in all relevant data protection aspects, and in particular in drafting any needed data protection documentation (data protection record, data protection impact assessment, etc.). The Commission will inform the contractor of the need for such documentation during the implementation of the contract.

(h) Ad hoc red teaming

To quickly respond to changes in the AI risk landscape, the AI Office will need to evaluate particular GPAI models on short notice. Therefore, the AI Office expects to request the contractor to conduct ad hoc evaluations at certain points in time. A maximum of 3 such evaluations may be commissioned.

For one such evaluation, the expectation regarding the contractor, is to make available a fixed budget of 60 expert hours with 2 weeks' notice, to be spend in a maximum of 2 weeks, to execute red teaming exercises with a GPAI model of the AI Office's choosing. These red teaming exercises may aim to stress test mitigations put in place my model providers to prevent malicious use, or to explore certain types of usage patterns.

This specific method of red teaming shall be tailored to the developed risk models and evaluation tools, to the degree these have already been finished, and aligned with the AI Office ahead of starting the red teaming.

The AI Office may request regular updates from the moment of notice until the moment the evaluation report has been approved.

A red teaming report shall be delivered, at most 4 weeks after notice, containing

- A general description of the red teaming, its methodology, results, and conclusions.
- A policymaker-oriented summary.
- A dataset of the interaction transcripts, containing no personal data within the meaning of Regulation (EU) 2018/1725 (EUDPR) or Regulation (EU) 2016/679 (GDPR), including e.g. requests send to the GPAI model, GPAI model outputs, assigned labels or grades, and other information that would allow for the reconstruction of the results.

The contractor is responsible for any costs associated with the requested evaluation, including but not limited to, the API costs associated with the GPAI model. Where reasonably feasibly, the contractor should ensure they have permission from the GPAI model provider to temporarily break the terms of service for the purpose of the exercise. Where not feasible, the AI Office shall aim to provide the required access.

Data protection: the contractor will support the Commission in all relevant data protection aspects, and in particular in drafting any needed data protection documentation (data protection record, data protection impact assessment, etc.). The Commission will inform the contractor of the need for such documentation during the implementation of the contract.

(i) Ad hoc evaluations requiring complex infrastructure

To quickly respond to changes in the AI risk landscape, the AI Office will need to evaluate particular GPAI models on short notice. Therefore, the AI Office expects to request the contractor to conduct ad hoc evaluations at certain points in time. A maximum of 2 such evaluations may be commissioned.

For one such evaluation, the expectation regarding the contractor, is to make available a fixed budget of 200 expert hours with 2 weeks' notice, to be spend in a maximum of 4 weeks, to execute evaluations with a GPAI model of the AI Office's choosing, and which potentially require complex digital infrastructure, for example testing the capability of the GPAI model to penetrate realistic IT infrastructure. The AI Office may also specify the benchmark to be used, where feasible for the contractor to set up the required infrastructure.

This specific method of evaluation shall be tailored to the developed risk models and evaluation tools, to the degree these have already been finished, and aligned with the AI Office ahead of starting the evaluation.

The AI Office may request regular updates from the moment of notice until the moment the evaluation report has been approved.

An evaluation report shall be delivered, at most 6 weeks after notice, containing

- A general description of the evaluation, its methodology, results, and conclusions.
- A policymaker-oriented summary.
- A dataset of the interaction transcripts, containing no personal data within the meaning of Regulation (EU) 2018/1725 (EUDPR) or Regulation (EU) 2016/679 (GDPR), including e.g. requests send to the GPAI model, GPAI model outputs, assigned labels or grades, and other information that would allow for the reconstruction of the results.

The contractor is responsible for any costs associated with the requested evaluation, including but not limited to, the API costs associated with the GPAI model. Where reasonably feasibly, the contractor should ensure they have permission from the GPAI model provider to temporarily break the terms of service for the purpose of the exercise. Where not feasible, the AI Office shall aim to provide the required access.

Data protection: the contractor will support the Commission in all relevant data protection aspects, and in particular in drafting any needed data protection documentation (data protection record, data protection impact assessment, etc.). The Commission will inform the contractor of the need for such documentation during the implementation of the contract.

(j) Stress-testing of mitigation methods applicable to openly released GPAI models

N.B. This deliverable relates to all malicious use risks, not only those related to cyber offence. Any execution of the task may focus on cyber offence as an example but must keep in mind the cross-cutting nature of the task.

Any GPAI models can be released in an open manner, for example through the publication of their weights and associated software for interaction with the model, allowing any interested and sufficiently capable party to host the GPAI model and modify to their needs. An open model releases affects which mitigations are effective at reducing risks or preventing them from materialising. For example, content filters are no longer applicable, as there might not be a responsible intermediary in between a GPAI model and a malicious actor.

Jointly with the AI Office, the contractor shall identify relevant mitigations that can be used to protect 'open-weight' models from malicious use, taking into account for example which of such mitigations have historically been used, and which ones are considered by the scientific community to be potentially useful in this regard. The contractor will, at most 5 months after the signing of the contract, provide a report with an overview of mitigations, containing:

- A list of mitigations.
- For each such mitigation, a high-level explanation of why the mitigation was chosen and why it be a relevant target for further analysis, for example based on its use in the GPAI ecosystem.
- For each such mitigation, a basic explanation of the relevant method with the appropriate references, oriented toward technical colleagues.

- For each such mitigation, a preliminary analysis in ways in which such mitigations may be circumvented or broken.

- Any overarching insights or connection between mitigations the contractor deems worthy of note.

- A policy-maker oriented summary.

The AI Office retains the right to request, with the ample notice, the analysis of specific mitigation techniques to be included in the overview report.

The AI Office shall then create a selection of the 3 most pertinent mitigations, which the contractor shall test for robustness. Such a robustness test will stress-test the mitigation, simulating a malicious actor aiming to break or circumvent the mitigation put in place to unlock access to specific capabilities. For each such mitigation, the contractor will prepare a testing report, which will contain:

- The overall results of the stress-testing, showing whether the robustness of the mitigation was disproven, proven, or was failed to be either proven or disproven.

- A list and basic description of the methods tried to circumvent or break the mitigation and their corresponding impact on the mitigation, nothing that the analysis need not be limited to a single method, even when that method quickly proves successful in breaking the mitigation.

- A collection of examples demonstrating GPAI model behaviour the mitigation aimed to prevent, e.g. assistance with malicious use, in the case the robustness of the mitigation could be disproven.

- An assessment of the completeness of the stress-testing and any limitations of the methods used.

- A policy-maker oriented summary.

Reports may be separated in time, and are expected respectively at most at 8, 16, and 24 months after signing of the contract. Both parties may agree to update the selection of mitigations to stress-test.

In addition to the three specified mitigations, the AI Office may request up to two additional mitigation stress-test analysis and reports during the duration of the contract, taking into account the workload associated with other tasks and deadlines.

Data protection: the contractor will support the Commission in all relevant data protection aspects, and in particular in drafting any needed data protection documentation (data protection record, data protection impact assessment, etc.). The Commission will inform the contractor of the need for such documentation during the implementation of the contract.

(k) Monthly calls

The contractor shall organise monthly calls with the AI Office to provide updates on the different tasks on the basis of the methodology and timelines agreed. Meeting minutes will be provided by the contractor at most five days after the monthly call.

Any reports or materials due in the same month shall be submitted at least five working days in advance of the monthly call to allow for review by the AI Office.

In justified cases, the AI Office may request ad hoc calls to complement the monthly calls and address specific issues that might arise during the execution of the contract. This will not impact the overall effort from the side of the contractor and will not cause any change to the minimum requirements of the tender specifications

(l) Final meeting and report

The contractor shall prepare a final report, which shall provide a comprehensive and detailed overview of the tasks conducted in the previous years. The final report should include the measures adopted by the contractor to ensure data protection compliance.

The submission of the final report is expected at the latest within 36 months from the contract's start date. The said final report will be linked to the final payment associated with the end of the contract.

The final report will be presented during a final meeting, organised by the contractor and to take place digitally.

The draft final report shall be sent to the European Commission 10 days before the final meeting at the latest. This will allow the Commission to assess the final results and provide comments where necessary, and the contractor to amend the report if required.

1.4.2.3.3. Timetable

Title	Due month (at the latest)	Linked to payment
Core Deliverables		
Risk modelling draft reports	M1, M3; M13, M15	No
Risk modelling workshops	M2; M14	No
Risk modelling reports	M6; M18	Yes
Public eval. tools proposal report	M3	No
Public eval. tools intermediate handover	M6	No
Public eval. tools handover	M12	Yes
Private eval. tools proposal report	M3	No
Private eval. tools intermediate handover	M6	No
Private evaluation tools handover	M12	Yes
Technical compliance toolkit	M13	No
Stress-testing of mitigation methods: overview report	M5	No
Stress-testing of mitigation methods: Mitigation testing reports	M8, M16, M24	Yes
Continued services		
Risk monitoring	Continuous	No
Ad hoc uplift evaluations	Upon request	No
Ad hoc red teaming	Upon request	No
Ad hoc evaluations requiring complex infrastructure	Upon request	No
Ad hoc mitigation testing	Upon request	Yes
Coordination		
Inception meeting	M1	No
Inception report	M1	No
Monthly calls	M2-M35	No

Final meeting	M36	No
Final report	M36	Yes

1.4.2.4. Description of Lot 3: Loss of Control Risk Modelling and Evaluation

All tasks are subject to the provisions outlined in section 1.4.2.8 Description: Shared Provisions.

1.4.2.4.1. Objective

The objective of this lot is to support the AI Office in its enforcement of the AI Act in relation to General-Purpose AI Models (GPAI) and General-Purpose AI Models with systemic risk (GPAISR), specifically with regards to risks related to the inability to oversee and control autonomous GPAISRs that may result in large-scale safety or security threats. This includes model capabilities and propensities related to autonomy, alignment with human intent or values, self-reasoning, self-replication, self-improvement, evading human oversight, deception, or resistance to goal modification. It further includes model capabilities of conducting autonomous AI research and development that could lead to the unpredictable emergence of GPAISRs without adequate risk mitigations.

The AI Office seeks to improve its capacity in identifying, monitoring, and assessing these risks, including the capacity to measure relevant capabilities and propensities of GPAI models and assess the effectiveness of mitigations.

This lot aims to allow for flexible and close collaboration in developing risk models, prioritising risk scenarios, setting risk levels including a level of unacceptable systemic risk at Union level, identifying key model capabilities that could be linked to risk levels, and tailoring both new and existing risk measurement instruments to the context of the AI Office. Flexible collaboration between the AI Office and the contractor will be needed to ensure robust risk assessment as risks and our understanding of them evolve.

1.4.2.4.2. Deliverables

(a) Inception meeting and report

The contractor will prepare an inception report, covering all tasks in this lot, providing a preliminary overview of the proposed approach, including staff capacities as well as measures for compliance with personal data protection legislation. Sufficient attention should be given to the planned approach for the risk modelling workshop. For avoidance of doubt, the draft inception report can be substantively the same as the work plan submitted with the tender, if such a plan is sufficiently detailed and up to date to reflect the contract as signed. The inception report will be discussed at the inception meeting, which will be a digital meeting happening at the latest 1 month after the signing of the contract. The inception report draft will be provided at least 5 working days before the inception meeting, and the final version will be provided at most 5 working days after the inception meeting, integrating any discussion and feedback for such meeting.

The report will serve as the initial foundation for the execution of the tasks and any collaboration.

(b) Risk modelling workshops and reports

A risk modelling workshop early in the contract, at most two months after contract signing, serves to align contractor and AI Office perspectives on enforcement priorities and the desired roadmap to relevant risk measurement instruments. This includes for example discussion on which capabilities to prioritise, or to what degree to pay attention to propensities and misalignment.

The contractor shall develop a coherent risk model that can serve as a framework for the development and integration of risk measurement instruments through the other deliverables in this lot. Although details are to be agreed with AI Office during the workshop and further communication, such risk model could include, but are by no means limited to:

- Development of several risk scenarios
- Execution of specific risk analysis methods established in literature
- Proposals for concrete risk levels for which to develop measurement instruments
- Surveying existing risk modelling literature and its adaptation to the context of the AI Office

Risk models should consider mitigations, technical or otherwise, and their weaknesses, applying both to private GPAI models as well as GPAI models made available to the public through the public release of their weights. Risk models should also integrate with existing EU policies on risk management.

The risk modelling workshop will be a physical meeting taking place at the premises of Directorate-General for Communications Networks, Content and Technology (DG CNECT) in Brussels. A draft report of the approach to risk modelling and for the risk modelling workshop will be provided one month before the workshop, while one month after the workshop a draft risk model must be provided, to be finalised three months after the workshop. Any reports will include a summary oriented to policymakers. Regular calls can be scheduled by the contractor or the AI Office in the case more alignment on the risk modelling is needed.

This risk modelling workshop, including the pre- and post-workshop drafts reports and final reports, will be repeated one year after the first one.

Either or both workshops may be conducted digitally upon mutual agreement between the contractor and the AI Office, with corresponding adjustments to resource allocation across deliverables.

Data protection: where the contractor processes personal data in the context of this Lot, the contractor must respect Regulation (EU) 2018/1725, including following instructions (in the light of Article 29 Regulation (EU) 2018/1725) from the Commission. In particular, the contractor will ensure that the processing activities will comply with the relevant data protection record (and privacy statement), namely DPR-EC-01063 “Processing of personal data linked to meetings and events”.

Where they are fit for purpose, the contractor will use Commission corporate tools, such as EU Survey or EventWorks for registration purposes, Slido for interactivity, etc.

If the Commission deems it necessary, the contractor will support the Commission in any other relevant data protection aspect, and in drafting any needed data protection documentation (data protection record, data protection impact assessment, etc.). The Commission will inform the contractor of the need for such documentation during the implementation of the contract.

Intellectual property rights: As an exception to the modes of exploitation set out in Article I.8 of the service contract, the Commission may NOT make the results corresponding with these deliverables publicly available, including the distribution to the public in hard copies, in electronic or digital format, on the internet including social networks as a downloadable or non-downloadable file or the communication through press information services, unless with written agreement from the contractor. The Commission may still use the results for all other listed modes of exploitation.

(c) Onboarded public evaluation tools

Based on the developed risk scenarios and risk levels, the contractor shall analyse and assess for use any relevant existing, publicly available, and risk related evaluations tools, for example benchmark datasets. Evaluation tools that are deemed suitable or close to suitable shall be adopted, amended, or improved for use within the AI Office context.

This process can include, but is no way limited to, the filtering and cleaning of data points from benchmarks, the setting of relevant risk levels of the corresponding measurement instrument, the integration with tailored elicitation techniques, instance-level analysis of the cognitive demands required, or the measurement of human expert performance on existing benchmarks to compare GPAI model performance against.

The onboarding of public evaluation tools may also include a contamination study of the respective datasets, i.e. a quantitative or qualitative analysis of the degree to which the benchmark data has been present in the training material for existing GPAI models or their safeguards, using e.g. the presence of "canary strings" or other benchmark specific information in model responses, acknowledging the speculative nature of such a study.

There will be three corresponding deliverables. Firstly, a report detailing which public benchmarks to onboard, including a motivation for this selection, as well as a short, non-exhaustive list of example candidate evaluation tools that did not make the selection, and a motivation for their exclusion, to be delivered at most 1 month after the conclusion of the workshop.

Secondly, an intermediary handover dedicated to the onboarding of a subset of the evaluation tools to be onboarded. Such a handover will contain:

- A report detailing any changes made to the evaluation tools included and any difficulties encountered.
- A report on the methodology and results for any substantive work conducted, e.g., to obtain human responses.
- Any modified or newly created resources such as benchmarks, datasets of human responses and their grading, evaluation software, etc.

Thirdly, a complete handover, similar to the intermediary handover, but now covering the full set of evaluation tools to onboard.

Both handovers should make the relevant onboarded evaluation tools ready-to-use in the context of the AI Office.

The initial proposal will be delivered at the latest 3 months after signing of the contract, while the intermediate handover will be delivered at the latest 6 months after signing of the contract, and the final handover 12 months after signing of the contract.

When doing so would be deemed beneficial for the overall procurement objective, for example in the case that no public resources fit for onboarding can be identified in the timeline currently set out, the contractor and the AI Office may decide jointly to postpone execution of this task, otherwise arrange the execution timeline in a more flexible manner, or reorganise team capacity to improve the scope or quality of other tasks. This flexibility must in all case respect the maximum duration of the contract.

Work on these deliverables should be informed by the outcomes of the risk modelling deliverables, and in particular by the first risk modelling workshop..

Data protection: For all processing activities, the contractor will support the Commission in all relevant data protection aspects, and in particular in drafting any needed data protection documentation (data protection record, data protection impact assessment, etc.). The Commission will inform the contractor of the need for such documentation during the implementation of the contract. The contractor will also ensure that the deliverables do not contain any personal data within the meaning of Regulation (EU) 2018/1725 (the EUDPR) or Regulation (EU) 2016/679 (GDPR).

(d) Private evaluation tools

Based on the developed risk model and the established utility of public evaluations tools (including the plans for their adaptation), the contractor will develop new evaluation tools for private use by the AI Office. These tools are expected to be one or multiple evaluation datasets that are fit for cheap automatic evaluation procedures and executable by AI Office technical experts. The evaluation methodology may extend beyond simple datasets of question-answer pairs, but the evaluation must be able to be executed repeatedly by the AI Office at acceptable costs and operational complexity.

Where relevant, these private evaluations will be complemented by the collection of human expert and non-expert answers as to compare GPAI model performance to.

The focus should be on a small but highly relevant selection of data, that is unique to the AI Office, can act as a verification of public test results with uncontaminated data, and serves as a proxy for expensive evaluations such as uplift-oriented evaluations.

There will be three corresponding deliverables. Firstly, a report detailing which evaluation tools are to be developed including the approach planned to develop them, to be delivered at most 1 month after the conclusion of the first risk modelling workshop.

Secondly, an intermediary handover dedicated to the development of a subset of the evaluation tools to be developed. Such a handover will contain:

- A report detailing the structure of the evaluation tool, its strengths and limitations, the method used for constructing it, and any problems encountered during construction.
- Any modified or newly created resources such as benchmarks, datasets of human responses and their grading, evaluation software, etc.
- A report on a sample evaluation, where an existing GPAI model available on the market is evaluated using the newly developed tools, insofar as it is reasonably feasible for the contractor to do so while preserving the privacy of the dataset. For instance, a privately hosted instance of a state-of-the-art open-source model could be used for the sample evaluation if non-logging assurances cannot be obtained for commercial APIs.
- Information on the measures taken to preserve the confidentiality of the evaluation data, for example, the presence of agreements with the provider of the GPAI model used in the sample evaluations regarding data confidentiality.

Thirdly, a complete handover, similar to the intermediary handover, but now covering the full set of private evaluations tools.

Both handovers should make the developed evaluation tools ready-to-use in the context of the AI Office.

The initial proposal will be delivered at the latest 3 months after signing of the contract, while the intermediate handover will be delivered at the latest 6 months after signing of the contract, and the final handover 12 months after signing of the contract.

Work on these deliverables should be informed by the outcomes of the risk modelling deliverables, and in particular by the first risk modelling workshop.

Data protection: For all processing activities, the contractor will support the Commission in all relevant data protection aspects, and in particular in drafting any needed data protection documentation (data protection record, data protection impact assessment, etc.). The Commission will inform the contractor of the need for such documentation during the implementation of the contract. The contractor will also ensure that the deliverables do not contain any personal data within the meaning of Regulation (EU) 2018/1725 (the EUDPR) or Regulation (EU) 2016/679 (GDPR).

(e) Technical compliance toolkit

Based on the developed risk model, the onboarded public evaluations, and newly created private evaluations tools, the contractor shall design a comprehensive draft evaluation process and draft reporting template that allows the AI Office to assess the systemic risk at Union level from a given model in a consistent, unified, and legible way.

The contractor will also design a redacted version of this technical compliance toolkit for potential publication by the AI Office, which could, if published, act as minimum standard for risk analysis to be referred to by GPAI providers and developers. The publication decision will be taken in due course by the AI Office.

This toolkit shall be delivered to the AI Office one month after the completion of both the onboarding of public evaluation tools and the development of private evaluation tools, and shall include:

- A report detailing the evaluation process and its methods, referring the onboarded public evaluation tools and developed private evaluation tools.
- A reporting template, tailored to policymakers in language and structure, which could be filled in and completed based on the results of the aforementioned evaluation process.
- A redacted version of both the evaluation process report and the reporting template (see above) containing no references to confidential information or resources, for potential publication.

(f) Risk monitoring

To complement the AI Office risk monitoring capacities, the contractor will provide the AI Office with regular update briefings on any events that are relevant for the developed risk models, for example major new GPAI models or new GPAI providers, algorithmic improvements, new insights into the specific risks, updated safety frameworks, policy changes, incidents, new mitigations, or the documented breaking of a mitigation.

These briefings should be tailored to the specific risks considered in this lot, can be informal in nature, and are not expected to be exhaustive. They should integrate with contractors existing effort to stay

up to date with the field of AI and the relevant risks in light of their other tasks, providing a lightweight instrument for sharing relevant updates with the AI Office.

The AI Office expects there to be events potentially worthy of briefing roughly every two weeks. This expectation should not be seen as prescriptive, and the sensitivity of briefing regime is expected to be calibrated in agreement with the AI Office, considering the intention to avoid unnecessary burden on the contractor.

(g) Ad hoc evaluations requiring complex infrastructure

To quickly respond to changes in the AI risk landscape, the AI Office will need to evaluate particular GPAI models on short notice. Therefore, the AI Office expects to request the contractor to conduct ad hoc evaluations at certain points in time. A maximum of 4 such evaluations may be commissioned.

For one such evaluation, the expectation regarding the contractor, is to make available a fixed budget of 200 expert hours with 2 weeks' notice, to be spend in a maximum of 4 weeks, to execute evaluations with a GPAI model of the AI Office's choosing, and which potentially require complex digital infrastructure, for example testing the capability of the GPAI model to penetrate realistic IT infrastructure. The AI Office may also specify the benchmark to be used, where feasible for the contractor to set up the required infrastructure.

This specific method of evaluation shall be tailored to the developed risk models and evaluation tools, to the degree these have already been finished, and aligned with the AI Office ahead of starting the evaluation.

The AI Office may request regular updates from the moment of notice until the moment the evaluation report has been approved.

An evaluation report shall be delivered, at most 6 weeks after notice, containing

- A general description of the evaluation, its methodology, results, and conclusions.
- A policymaker-oriented summary.
- A dataset of the interaction transcripts, containing no personal data within the meaning of Regulation (EU) 2018/1725 (EUDPR) or Regulation (EU) 2016/679 (GDPR), including e.g. requests send to the GPAI model, GPAI model outputs, assigned labels or grades, and other information that would allow for the reconstruction of the results.

The contractor is responsible for any costs associated with the requested evaluation, including but not limited to, the API costs associated with the GPAI model. Where reasonably feasibly, the contractor should ensure they have permission from the GPAI model provider to temporarily break the terms of service for the purpose of the exercise. Where not feasible, the AI Office shall aim to provide the required access.

Data protection: the contractor will support the Commission in all relevant data protection aspects, and in particular in drafting any needed data protection documentation (data protection record, data protection impact assessment, etc.). The Commission will inform the contractor of the need for such documentation during the implementation of the contract.

(h) Monthly calls

The contractor shall organise monthly calls with the AI Office to provide updates on the different tasks on the basis of the methodology and timelines agreed. Meeting minutes will be provided by the contractor at most five days after the monthly call.

Any reports or materials due in the same month shall be submitted at least five working days in advance of the monthly call to allow for review by the AI Office.

In justified cases, the AI Office may request ad hoc calls to complement the monthly calls and address specific issues that might arise during the execution of the contract. This will not impact the overall effort from the side of the contractor and will not cause any change to the minimum requirements of the tender specifications

(i) Final meeting and report

The contractor shall prepare a final report, which shall provide a comprehensive and detailed overview of the tasks conducted in the previous years. The final report should include the measures adopted by the contractor to ensure data protection compliance.

The submission of the final report is expected at the latest within 36 months from the contract's start date. The said final report will be linked to the final payment associated with the end of the contract.

The final report will be presented during a final meeting, organised by the contractor and to take place digitally.

The draft final report shall be sent to the European Commission 10 days before the final meeting at the latest. This will allow the Commission to assess the final results and provide comments where necessary, and the contractor to amend the report if required.

1.4.2.4.3. Timetable

Title	Due month (at the latest)	Linked to payment
Core Deliverables		
Risk modelling draft reports	M1, M3; M13, M15	No
Risk modelling workshops	M2; M14	No
Risk modelling reports	M6; M18	Yes
Public eval. tools proposal report	M3	No
Public eval. tools intermediate handover	M6	No
Public eval. tools handover	M12	Yes
Private eval. tools proposal report	M3	No
Private eval. tools intermediate handover	M6	No
Private evaluation tools handover	M12	Yes
Technical compliance toolkit	M13	No
Continued services		
Risk monitoring	Continuous	No
Ad hoc evaluations with complex infra.	Upon request	No
Coordination		
Inception meeting	M1	No
Inception report	M1	No

Title	Due month (at the latest)	Linked to payment
Monthly calls	M2-M35	No
Final meeting	M36	No
Final report	M36	Yes

1.4.2.5. Description of Lot 4: Harmful Manipulation Risk Modelling and Evaluation

All tasks are subject to the provisions outlined in section 1.4.2.8 Description: Shared Provisions.

1.4.2.5.1. Objective

The objective of this lot is to support the AI Office in its enforcement of the AI Act in relation to General-Purpose AI Models (GPAI) and General-Purpose AI Models with systemic risk (GPAISR), specifically where such models may pose risks of enabling the targeted distortion of the behaviour of persons, in particular through multi-turn interactions, that causes them to take a decision that they would not have otherwise taken, in a manner that causes, or is reasonably likely to cause, significant harm on a large scale. This includes the capability to manipulate through multi-turn interaction and the propensity of models to manipulate, including manipulation of high-stakes decision makers, large-scale fraud, or exploitation of people based on protected characteristics. As a guide, risk of harmful manipulation exists if it cannot, without reasonable doubt, be ruled out that a GPAISR, when integrated into an AI system, enables the AI system, irrespective of the intention of the AI system provider or deployer, to deploy subliminal, purposefully manipulative, or deceptive techniques as outlined in the Commission Guidelines on prohibited artificial intelligence practices established by Regulation (EU) 2024/1689 (AI Act) ([2]).

The AI Office seeks to improve its capacity in identifying, monitoring, and assessing these risks, including the capacity to measure relevant capabilities of GPAI models and assess the effectiveness of mitigations aimed at preventing malicious actors from gaining access to these capabilities.

This lot aims to allow for flexible and close collaboration in developing risk models, prioritising risk scenarios, setting risk levels including a level of unacceptable systemic risk at Union level, identifying key model capabilities that could be linked to risk levels, and tailoring both new and existing risk measurement instruments to the context of the AI Office. Flexible collaboration between the AI Office and the contractor will be needed to ensure robust risk assessment as risks and our understanding of them evolve.

1.4.2.5.2. Deliverables

(a) Inception meeting and report

The contractor will prepare an inception report, covering all tasks in this lot, providing a preliminary overview of the proposed approach, including staff capacities as well as measures for compliance with personal data protection legislation. Sufficient attention should be given to the planned approach for the risk modelling workshop. For avoidance of doubt, the draft inception report can be substantively the same as the work plan submitted with the tender, if such a plan is sufficiently detailed and up to date to reflect the contract as signed. The inception report will be discussed at the inception meeting, which will be a digital meeting happening at the latest 1 month after the signing of the contract. The inception report draft will be provided at least 5 working days before the inception meeting, and the final version will be provided at most 5 working days after the inception meeting, integrating any discussion and feedback for such meeting.

([2]) https://digital-strategy.ec.europa.eu/en/library/commission-publishes-guidelines-prohibited-artificial-intelligence-ai-practices-defined-ai-act

The report will serve as the initial foundation for the execution of the tasks and any collaboration.

(b) Risk modelling workshops and reports

A risk modelling workshop early in the contract, at most two months after contract signing, serves to align contractor and AI Office perspectives on enforcement priorities and the desired roadmap to relevant risk measurement instruments. This includes for example discussion on which risks or scenarios to prioritise, for example those dealing with human-AI teams versus autonomous or unintended AI risks, or those dealing with large scale manipulation versus the targeting of key decisionmakers.

The contractor shall develop a coherent risk model that can serve as a framework for the development and integration of risk measurement instruments through the other deliverables in this lot. Although details are to be agreed with AI Office during the workshop and further communication, such risk model could include, but are by no means limited to:

- Development of several risk scenarios
- Execution of specific risk analysis methods established in literature
- Proposals for concrete risk levels for which to develop measurement instruments
- Surveying existing risk modelling literature and its adaptation to the context of the AI Office

Risk models should consider mitigations, technical or otherwise, and their weaknesses, applying both to private GPAI models as well as GPAI models made available to the public through the public release of their weights. Risk models should also integrate with existing EU policies on risk management.

Risk scenarios will be triaged in close alignment with the AI Office to ensure clear separation with other risk categories and procurement efforts.

The risk modelling workshop will be a physical meeting taking place at the premises of Directorate-General for Communications Networks, Content and Technology (DG CNECT) in Brussels. A draft report of the approach to risk modelling and for the risk modelling workshop will be provided one month before the workshop, while one month after the workshop a draft risk model must be provided, to be finalised three months after the workshop. Any reports will include a summary oriented to policymakers. Regular calls can be scheduled by the contractor or the AI Office in the case more alignment on the risk modelling is needed.

This risk modelling workshop, including the pre- and post-workshop drafts reports and final reports, will be repeated one year after the first one.

Either or both workshops may be conducted digitally upon mutual agreement between the contractor and the AI Office, with corresponding adjustments to resource allocation across deliverables.

Data protection: where the contractor processes personal data in the context of this Lot, the contractor must respect Regulation (EU) 2018/1725, including following instructions (in the light of Article 29 Regulation (EU) 2018/1725) from the Commission. In particular, the contractor will ensure that the processing activities will comply with the relevant data protection record (and privacy statement), namely DPR-EC-01063 “Processing of personal data linked to meetings and events”.

Where they are fit for purpose, the contractor will use Commission corporate tools, such as EU Survey or EventWorks for registration purposes, Slido for interactivity, etc.

If the Commission deems it necessary, the contractor will support the Commission in any other relevant data protection aspect, and in drafting any needed data protection documentation (data protection record, data protection impact assessment, etc.). The Commission will inform the contractor of the need for such documentation during the implementation of the contract.

Intellectual property rights: As an exception to the modes of exploitation set out in Article I.8 of the service contract, the Commission may NOT make the results corresponding with these deliverables publicly available, including the distribution to the public in hard copies, in electronic or digital format, on the internet including social networks as a downloadable or non-downloadable file or the communication through press information services, unless with written agreement from the contractor. The Commission may still use the results for all other listed modes of exploitation.

(c) Onboarded public evaluation tools

Based on the developed risk scenarios and risk levels, the contractor shall analyse and assess for use any relevant existing, publicly available, and risk related evaluations tools, for example benchmark datasets. Evaluation tools that are deemed suitable or close to suitable shall be adopted, amended, or improved for use within the AI Office context.

This process can include, but is no way limited to, the filtering and cleaning of data points from benchmarks, the setting of relevant risk levels of the corresponding measurement instrument, the integration with tailored elicitation techniques, instance-level analysis of the cognitive demands required, or the measurement of human expert performance on existing benchmarks to compare GPAI model performance against.

The onboarding of public evaluation tools may also include a contamination study of the respective datasets, i.e. a quantitative or qualitative analysis of the degree to which the benchmark data has been present in the training material for existing GPAI models or their safeguards, using e.g. the presence of "canary strings" or other benchmark specific information in model responses, acknowledging the speculative nature of such a study.

There will be three corresponding deliverables. Firstly, a report detailing which public benchmarks to onboard, including a motivation for this selection, as well as a short, non-exhaustive list of example candidate evaluation tools that did not make the selection, and a motivation for their exclusion, to be delivered at most 1 month after the conclusion of the workshop.

Secondly, an intermediary handover dedicated to the onboarding of a subset of the evaluation tools to be onboarded. Such a handover will contain:

- A report detailing any changes made to the evaluation tools included and any difficulties encountered.
- A report on the methodology and results for any substantive work conducted, e.g., to obtain human responses.
- Any modified or newly created resources such as benchmarks, datasets of human responses and their grading, evaluation software, etc.

Thirdly, a complete handover, similar to the intermediary handover, but now covering the full set of evaluation tools to onboard.

Both handovers should make the relevant onboarded evaluation tools ready-to-use in the context of the AI Office.

The initial proposal will be delivered at the latest 3 months after signing of the contract, while the intermediate handover will be delivered at the latest 6 months after signing of the contract, and the final handover 12 months after signing of the contract.

When doing so would be deemed beneficial for the overall procurement objective, for example in the case that no public resources fit for onboarding can be identified in the timeline currently set out, the contractor and the AI Office may decide jointly to postpone execution of this task, otherwise arrange the execution timeline in a more flexible manner, or reorganise team capacity to improve the scope or quality of other tasks. This flexibility must in all case respect the maximum duration of the contract.

Work on these deliverables should be informed by the outcomes of the risk modelling deliverables, and in particular by the first risk modelling workshop.

Data protection: For all processing activities, the contractor will support the Commission in all relevant data protection aspects, and in particular in drafting any needed data protection documentation (data protection record, data protection impact assessment, etc.). The Commission will inform the contractor of the need for such documentation during the implementation of the contract. The contractor will also ensure that the deliverables do not contain any personal data within the meaning of Regulation (EU) 2018/1725 (the EUDPR) or Regulation (EU) 2016/679 (GDPR).

(d) Private evaluation tools

Based on the developed risk model and the established utility of public evaluations tools (including the plans for their adaptation), the contractor will develop new evaluation tools for private use by the AI Office. These tools are expected to be one or multiple evaluation datasets that are fit for cheap automatic evaluation procedures and executable by AI Office technical experts. The evaluation methodology may extend beyond simple datasets of question-answer pairs, but the evaluation must be able to be executed repeatedly by the AI Office at acceptable costs and operational complexity.

Where relevant, these private evaluations will be complemented by the collection of human expert and non-expert answers as to compare GPAI model performance to.

The focus should be on a small but highly relevant selection of data, that is unique to the AI Office, can act as a verification of public test results with uncontaminated data, and serves as a proxy for expensive evaluations such as uplift-oriented evaluations.

There will be three corresponding deliverables. Firstly, a report detailing which evaluation tools are to be developed including the approach planned to develop them, to be delivered at most 1 month after the conclusion of the first risk modelling workshop.

Secondly, an intermediary handover dedicated to the development of a subset of the evaluation tools to be developed. Such a handover will contain:

- A report detailing the structure of the evaluation tool, its strengths and limitations, the method used for constructing it, and any problems encountered during construction.

- Any modified or newly created resources such as benchmarks, datasets of human responses and their grading, evaluation software, etc.

- A report on a sample evaluation, where an existing GPAI model available on the market is evaluated using the newly developed tools, insofar as it is reasonably feasible for the contractor to do so while preserving the privacy of the dataset. For instance, a privately hosted instance of a state-of-the-art open-source model could be used for the sample evaluation if non-logging assurances cannot be obtained for commercial APIs.

- Information on the measures taken to preserve the confidentiality of the evaluation data, for example, the presence of agreements with the provider of the GPAI model used in the sample evaluations regarding data confidentiality.

Thirdly, a complete handover, similar to the intermediary handover, but now covering the full set of private evaluations tools.

The initial proposal will be delivered at the latest 3 months after signing of the contract, while the intermediate handover will be delivered at the latest 6 months after signing of the contract, and the final handover 12 months after signing of the contract.

Both handovers should make the developed evaluation tools ready-to-use in the context of the AI Office.

Work on these deliverables should be informed by the outcomes of the risk modelling deliverables, and in particular by the first risk modelling workshop.

Data protection: For all processing activities, the contractor will support the Commission in all relevant data protection aspects, and in particular in drafting any needed data protection documentation (data protection record, data protection impact assessment, etc.). The Commission will inform the contractor of the need for such documentation during the implementation of the contract. The contractor will also ensure that the deliverables do not contain any personal data within the meaning of Regulation (EU) 2018/1725 (the EUDPR) or Regulation (EU) 2016/679 (GDPR).

(e) Technical compliance toolkit

Based on the developed risk model, the onboarded public evaluations, and newly created private evaluations tools, the contractor shall design a comprehensive draft evaluation process and draft reporting template that allows the AI Office to assess the systemic risk at Union level from a given model in a consistent, unified, and legible way.

The contractor will also design a redacted version of this technical compliance toolkit for potential publication by the AI Office, which could, if published, act as minimum standard for risk analysis to be referred to by GPAI providers and developers. The publication decision will be taken in due course by the AI Office.

This toolkit shall be delivered to the AI Office one month after the completion of both the onboarding of public evaluation tools and the development of private evaluation tools, and shall include:

- A report detailing the evaluation process and its methods, referring the onboarded public evaluation tools and developed private evaluation tools.

- A reporting template, tailored to policymakers in language and structure, which could be filled in and completed based on the results of the aforementioned evaluation process.

- A redacted version of both the evaluation process report and the reporting template (see above) containing no references to confidential information or resources, for potential publication.

(f) Risk monitoring

To complement the AI Office risk monitoring capacities, the contractor will provide the AI Office with regular update briefings on any events that are relevant for the developed risk models, for example major new GPAI models or new GPAI providers, algorithmic improvements, new insights into the specific risks, updated safety frameworks, policy changes, incidents, new mitigations, or the documented breaking of a mitigation.

These briefings should be tailored to the specific risks considered in this lot, can be informal in nature, and are not expected to be exhaustive. They should integrate with contractors existing effort to stay up to date with the field of AI and the relevant risks in light of their other tasks, providing a lightweight instrument for sharing relevant updates with the AI Office.

The AI Office expects there to be events potentially worthy of briefing roughly every two weeks. This expectation should not be seen as prescriptive, and the sensitivity of briefing regime is expected to be calibrated in agreement with the AI Office, considering the intention to avoid unnecessary burden on the contractor.

(g) Ad hoc evaluations with humans

To quickly respond to changes in the AI risk landscape, the AI Office will need to evaluate particular GPAI models on short notice. Therefore, the AI Office expects to request the contractor to conduct ad hoc evaluations at certain points in time. A maximum of 3 such evaluations may be commissioned.

For one such evaluation, the expectation regarding the contractor, is to make available a fixed budget of 200 expert hours with 2 weeks' notice to coordinate an evaluation of a GPAI model of the AI Office's choosing, to be spend over at most 4 weeks. These evaluations will involve human participants, for example for studying the effect of GPAI manipulation capabilities on humans, and/or in the sense of "uplift evaluations", where uplift evaluations are investigations into the degree to which GPAI systems can help humans in completing certain tasks (such as manipulating a human or AI counterpart), and any conclusions that can be drawn about systemic risks.

This specific method for the evaluation shall be tailored to the developed risk models and evaluation tools, to the degree these have already been finished. The AI Office may request regular updates from the moment of notice until the moment the evaluation report has been approved.

An evaluation report shall be delivered, at most 6 weeks after notice, containing

- A general description of the evaluation, its methodology, results, and conclusions.
- A policymaker-oriented summary.
- A dataset of the anonymized interaction transcripts, containing no personal data within the meaning of Regulation (EU) 2018/1725 (EUDPR) or Regulation (EU) 2016/679 (GDPR), including e.g. requests send to the GPAI model, GPAI model outputs, any assigned labels or grades, and other information that would allow for the reconstruction of the evaluation results.

The contractor is responsible for any costs associated with the request evaluation, including but not limited to, the payment of study participants, or the API costs associated with the GPAI model.

Informed by the first risk modelling workshop, AI Office and contractor may begin to consider potential experimental designs for upcoming evaluations, to facilitate rapid execution when the evaluation is required.

Data protection: the contractor will support the Commission in all relevant data protection aspects, and in particular in drafting any needed data protection documentation (data protection record, data protection impact assessment, etc.). The Commission will inform the contractor of the need for such documentation during the implementation of the contract.

(h) Ad hoc red teaming

To quickly respond to changes in the AI risk landscape, the AI Office will need to evaluate particular GPAI models on short notice. Therefore, the AI Office expects to request the contractor to conduct ad hoc evaluations at certain points in time. A maximum of 3 such evaluations may be commissioned.

For one such evaluation, the expectation regarding the contractor, is to make available a fixed budget of 60 expert hours with 2 weeks' notice, to be spend in a maximum of 2 weeks, to execute red teaming exercises with a GPAI model of the AI Office's choosing. These red teaming exercises may aim to stress test mitigations put in place my model providers to prevent malicious use, or to explore certain types of usage patterns.

This specific method of red teaming shall be tailored to the developed risk models and evaluation tools, to the degree these have already been finished, and aligned with the AI Office ahead of starting the red teaming.

The AI Office may request regular updates from the moment of notice until the moment the evaluation report has been approved.

A red teaming report shall be delivered, at most 4 weeks after notice, containing

- A general description of the red teaming, its methodology, results, and conclusions.
- A policymaker-oriented summary.
- A dataset of the interaction transcripts, containing no personal data within the meaning of Regulation (EU) 2018/1725 (EUDPR) or Regulation (EU) 2016/679 (GDPR), including e.g. requests send to the GPAI model, GPAI model outputs, assigned labels or grades, and other information that would allow for the reconstruction of the results.

The contractor is responsible for any costs associated with the requested evaluation, including but not limited to, the API costs associated with the GPAI model. Where reasonably feasibly, the contractor should ensure they have permission from the GPAI model provider to temporarily break the terms of service for the purpose of the exercise. Where not feasible, the AI Office shall aim to provide the required access.

Data protection: the contractor will support the Commission in all relevant data protection aspects, and in particular in drafting any needed data protection documentation (data protection record, data protection impact assessment, etc.). The Commission will inform the contractor of the need for such documentation during the implementation of the contract.

(i) Monthly calls

The contractor shall organise monthly calls with the AI Office to provide updates on the different tasks on the basis of the methodology and timelines agreed. Meeting minutes will be provided by the contractor at most five days after the monthly call.

Any reports or materials due in the same month shall be submitted at least five working days in advance of the monthly call to allow for review by the AI Office.

In justified cases, the AI Office may request ad hoc calls to complement the monthly calls and address specific issues that might arise during the execution of the contract. This will not impact the overall effort from the side of the contractor and will not cause any change to the minimum requirements of the tender specifications

(j) Final meeting and report

The contractor shall prepare a final report, which shall provide a comprehensive and detailed overview of the tasks conducted in the previous years. The final report should include the measures adopted by the contractor to ensure data protection compliance.

The submission of the final report is expected at the latest within 36 months from the contract's start date. The said final report will be linked to the final payment associated with the end of the contract.

The final report will be presented during a final meeting, organised by the contractor and to take place digitally.

The draft final report shall be sent to the European Commission 10 days before the final meeting at the latest. This will allow the Commission to assess the final results and provide comments where necessary, and the contractor to amend the report if required.

1.4.2.5.3. Timetable

Title	Due month (at the latest)	Linked to payment
Core Deliverables		
Risk modelling draft reports	M1, M3; M13, M15	No
Risk modelling workshops	M2; M14	No
Risk modelling reports	M6; M18	Yes
Public eval. tools proposal report	M3	No
Public eval. tools intermediate handover	M6	No
Public eval. tools handover	M12	Yes
Private eval. tools proposal report	M3	No
Private eval. tools intermediate handover	M6	No
Private evaluation tools handover	M12	Yes
Technical compliance toolkit	M13	No
Continued services		
Risk monitoring	Continuous	No
Ad hoc evaluations with humans	Upon request	No
Ad hoc red teaming	Upon request	No
Coordination		
Inception meeting	M1	No

Title	Due month (at the latest)	Linked to payment
Inception report	M1	No
Monthly calls	M2-M35	No
Final meeting	M36	No
Final report	M36	Yes

1.4.2.6. Description of Lot 5: Sociotechnical Risk Modelling and Evaluation

All tasks are subject to the provisions outlined in section 1.4.2.8 Description: Shared Provisions.

1.4.2.6.1. Objective

The objective of this lot is to support the AI Office in its enforcement of the AI Act in relation to General-Purpose AI Models (GPAI) and General-Purpose AI Models with systemic risk (GPAISR), specifically where such models may pose large-scale sociotechnical risks, including those risks stemming from harmful bias or discrimination, or from the endangering of fundamental human rights such as freedom of expression or health protection—irrespective of the downstream systems and contexts in which these models are deployed.

The AI Office aims to enhance its ability to identify, monitor, and evaluate these risks, including the capacity to measure relevant capabilities and propensities of GPAI models and asses the effectiveness of mitigations. This includes the monitoring of indicator variables and the creation of early warning indicators for tracking accumulative risks that materialize through reach, well-meaning behaviour, and adoption rather than through novel capabilities or malicious actors.

This lot aims to allow for flexible and close collaboration in developing risk models, prioritising risk scenarios, setting risk levels including a level of unacceptable systemic risk at Union level, identifying key model capabilities that could be linked to risk levels, and tailoring both new and existing risk measurement instruments to the context of the AI Office. Flexible collaboration between the AI Office and the contractor will be needed to ensure robust risk assessment as risks and our understanding of them evolve.

1.4.2.6.2. Deliverables

(a) Inception meeting and report

The contractor will prepare an inception report, covering all tasks in this lot, providing a preliminary overview of the proposed approach, including staff capacities as well as measures for compliance with personal data protection legislation. Sufficient attention should be given to the planned approach for the risk modelling workshop. For avoidance of doubt, the draft inception report can be substantively the same as the work plan submitted with the tender, if such a plan is sufficiently detailed and up to date to reflect the contract as signed. The inception report will be discussed at the inception meeting, which will be a digital meeting happening at the latest 1 month after the signing of the contract. The inception report draft will be provided at least 5 working days before the inception meeting, and the final version will be provided at most 5 working days after the inception meeting, integrating any discussion and feedback for such meeting.

The report will serve as the initial foundation for the execution of the tasks and any collaboration.

(b) Risk modelling workshops and reports

A risk modelling workshop early in the contract, at most two months after contract signing, serves to align contractor and AI Office perspectives on enforcement priorities and the desired roadmap to relevant risk measurement instruments. This includes discussions on prioritizing specific types of bias and risks to fundamental rights, while also taking into account how the risks covered by the GPAI provisions of the AI Act are addressed through other legal frameworks.

The contractor shall develop a coherent risk model that can serve as a framework for the development and integration of risk measurement instruments through the other deliverables in this lot. Although details are to be agreed with AI Office during the workshop and further communication, such a risk model could include, but are by no means limited to:

- Development of several risk scenarios, giving sufficient attention to 'tipping points' of different societal systems and their early indicators
- Execution of specific risk analysis methods established in literature
- Proposals for concrete risk levels linked to tipping points and early indicators for which to develop measurement instruments
- Surveying existing risk modelling literature and its adaptation to the context of the AI Office

Risk models should consider mitigations, technical or otherwise, and their weaknesses, applying both to private GPAI models as well as GPAI models made available to the public through the public release of their weights. Risk models should also integrate with existing EU policies on risk management.

Risk scenarios will be triaged in close alignment with the AI Office to ensure clear separation with other risk categories and procurement efforts.

The risk modelling workshop will be a physical meeting taking place at the premises of Directorate-General for Communications Networks, Content and Technology (DG CNECT) in Brussels. A draft report of the approach to risk modelling and for the risk modelling workshop will be provided one month before the workshop, while one month after the workshop a draft risk model must be provided, to be finalised three months after the workshop. Any reports will include a summary oriented to policymakers. Regular calls can be scheduled by the contractor or the AI Office in the case more alignment on the risk modelling is needed.

This risk modelling workshop, including the pre- and post-workshop drafts reports and final reports, will be repeated one year after the first one.

Either or both workshops may be conducted digitally upon mutual agreement between the contractor and the AI Office, with corresponding adjustments to resource allocation across deliverables.

Data protection: where the contractor processes personal data in the context of this Lot, the contractor must respect Regulation (EU) 2018/1725, including following instructions (in the light of Article 29 Regulation (EU) 2018/1725) from the Commission. In particular, the contractor will ensure that the processing activities will comply with the relevant data protection record (and privacy statement), namely DPR-EC-01063 "Processing of personal data linked to meetings and events".

Where they are fit for purpose, the contractor will use Commission corporate tools, such as EU Survey or EventWorks for registration purposes, Slido for interactivity, etc.

If the Commission deems it necessary, the contractor will support the Commission in any other relevant data protection aspect, and in drafting any needed data protection documentation (data protection record, data protection impact assessment, etc.). The Commission will inform the contractor of the need for such documentation during the implementation of the contract.

Intellectual property rights: As an exception to the modes of exploitation set out in Article I.8 of the service contract, the Commission may NOT make the results corresponding with these deliverables publicly available, including the distribution to the public in hard copies, in electronic or digital format, on the internet including social networks as a downloadable or non-downloadable file or the communication through press information services, unless with written agreement from the contractor. The Commission may still use the results for all other listed modes of exploitation.

(c) Onboarded public evaluation tools

Based on the developed risk scenarios and risk levels, the contractor shall analyse and assess for use any relevant existing, publicly available, and risk related evaluations tools, for example benchmark datasets. Evaluation tools that are deemed suitable or close to suitable shall be adopted, amended, or improved for use within the AI Office context.

This process can include, but is in no way limited to, the filtering and cleaning of data points from benchmarks, the setting of relevant risk levels of the corresponding measurement instrument, analysis of publicly available GPAI model usage data, the instance-level analysis of the propensities and dimensions of bias or impacted human rights, or the measurement of human performance on existing benchmarks to compare GPAI model performance against.

The onboarding of public evaluation tools may also include a contamination study of the respective datasets, i.e. a quantitative or qualitative analysis of the degree to which the benchmark data has been present in the training material for existing GPAI models or their safeguards, using e.g. the presence of "canary strings" or other benchmark specific information in model responses, acknowledging the speculative nature of such a study.

There will be three corresponding deliverables. Firstly, a report detailing which public benchmarks to onboard, including a motivation for this selection and discussion of their construct validity, as well as a short, non-exhaustive list of example candidate evaluation tools that did not make the selection, and a motivation for their exclusion, to be delivered at most 1 month after the conclusion of the workshop.

Secondly, an intermediary handover dedicated to the onboarding of a subset of the evaluation tools to be onboarded. Such a handover will contain:

- A report detailing any changes made to the evaluation tools included and any difficulties encountered.
- A report on the methodology and results for any substantive work conducted, e.g., to obtain human responses.
- Any modified or newly created resources such as benchmarks, datasets of human responses and their grading, evaluation software, etc.

Thirdly, a complete handover, similar to the intermediary handover, but now covering the full set of evaluation tools to onboard.

Both handovers should make the relevant onboarded evaluation tools ready-to-use in the context of the AI Office.

The initial proposal will be delivered at the latest 3 months after signing of the contract, while the intermediate handover will be delivered at the latest 6 months after signing of the contract, and the final handover 12 months after signing of the contract.

When doing so would be deemed beneficial for the overall procurement objective, for example in the case that no public resources fit for onboarding can be identified in the timeline currently set out, the contractor and the AI Office may decide jointly to postpone execution of this task, otherwise arrange the execution timeline in a more flexible manner, or reorganise team capacity to improve the scope or quality of other tasks. This flexibility must in all case respect the maximum duration of the contract.

Work on these deliverables should be informed by the outcomes of the risk modelling deliverables, and in particular by the first risk modelling workshop.

Data protection: For all processing activities, the contractor will support the Commission in all relevant data protection aspects, and in particular in drafting any needed data protection documentation (data protection record, data protection impact assessment, etc.). The Commission will inform the contractor of the need for such documentation during the implementation of the contract. The contractor will also ensure that the deliverables do not contain any personal data within the meaning of Regulation (EU) 2018/1725 (the EUDPR) or Regulation (EU) 2016/679 (GDPR).

(d) Private evaluation tools

Based on the developed risk model and the established utility of public evaluations tools (including the plans for their adaptation), the contractor will develop new evaluation tools for private use by the AI Office. These tools are expected to be one or multiple evaluation datasets that are fit for cheap automatic evaluation procedures and executable by AI Office technical experts. The evaluation methodology may extend beyond simple datasets of question-answer pairs, but the evaluation must be able to be executed repeatedly by the AI Office at acceptable costs and operational complexity.

Where relevant, these private evaluations will be complemented by the collection of human answers or human annotation and judgement of AI generated answers.

The focus should be on a small but highly relevant selection of data, that is unique to the AI Office, can act as a verification of public test results with uncontaminated data, and serves as a proxy for expensive evaluations such as those involving a large quantity of human participants.

There will be three corresponding deliverables. Firstly, a report detailing which evaluation tools are to be developed including the approach planned to develop them, to be delivered at most 1 month after the conclusion of the first risk modelling workshop.

Secondly, an intermediary handover dedicated to the development of a subset of the evaluation tools to be developed. Such a handover will contain:

- A report detailing the structure of the evaluation tool, its strengths and limitations, the method used for constructing it, and any problems encountered during construction.

- Any modified or newly created resources such as benchmarks, datasets of human responses and their grading, evaluation software, etc.

- A report on a sample evaluation, where an existing GPAI model available on the market is evaluated using the newly developed tools, insofar as it is reasonably feasible for the contractor to do so while preserving the privacy of the dataset. For instance, a privately hosted instance of a state-of-the-art open-source model could be used for the sample evaluation if non-logging assurances cannot be obtained for commercial APIs.

- Information on the measures taken to preserve the confidentiality of the evaluation data, for example, the presence of agreements with the provider of the GPAI model used in the sample evaluations regarding data confidentiality.

Thirdly, a complete handover, similar to the intermediary handover, but now covering the full set of private evaluations tools.

Both handovers should make the developed evaluation tools ready-to-use in the context of the AI Office.

The initial proposal will be delivered at the latest 3 months after signing of the contract, while the intermediate handover will be delivered at the latest 6 months after signing of the contract, and the final handover 12 months after signing of the contract.

Work on these deliverables should be informed by the outcomes of the risk modelling deliverables, and in particular by the first risk modelling workshop.

Data protection: For all processing activities, the contractor will support the Commission in all relevant data protection aspects, and in particular in drafting any needed data protection documentation (data protection record, data protection impact assessment, etc.). The Commission will inform the contractor of the need for such documentation during the implementation of the contract. The contractor will also ensure that the deliverables do not contain any personal data within the meaning of Regulation (EU) 2018/1725 (the EUDPR) or Regulation (EU) 2016/679 (GDPR).

(e) Technical compliance toolkit

Based on the developed risk model, the onboarded public evaluations, and newly created private evaluations tools, the contractor shall design a comprehensive draft evaluation process and draft reporting template that allows the AI Office to assess the systemic risk at Union level from a given model in a consistent, unified, and legible way.

The contractor will also design a redacted version of the technical compliance toolkit for potential publication by the AI Office, which could, if published, act as a minimum standard for risk analysis to be referred to by GPAI providers and developers. The publication decision will be taken in due course by the AI Office.

This toolkit shall be delivered to the AI Office one month after the completion of both the onboarding of public evaluation tools and the development of private evaluation tools, and shall include:

- A report detailing the evaluation process and its methods, referring the onboarded public evaluation tools and developed private evaluation tools.

- A reporting template, tailored to policymakers in language and structure, which could be filled in and completed based on the results of the aforementioned evaluation process.

- A redacted version of both the evaluation process report and the reporting template (see above) containing no references to confidential information or resources, for potential publication.

(f) Risk monitoring framework

Sociotechnical risks identified in the risk modelling workshops might evolve and build up over time before they reach a critical tipping point. In order to complement and inform the AI Office risk monitoring capacities, the contractor shall design a suitable risk monitoring framework to keep track of emerging risks and their early indicators, including complex risks stemming from societal dynamics and increased adoption of GPAI models, not necessarily attributable to a single model or novel capabilities. Such a framework will:

- Identify suitable variables that can act as early warning indicators for the identified risks, and which can be efficiently tracked. Any variable will be accompanied by a detailed tracking methodology and a description about how it is connected to the identified risks.

- Where they are not already covered by tracked variables (e.g. in incident counts), the monitoring framework will set out a strategy for the monitoring of incidents related to sociotechnical risks.

- Present an overview of risk blind spots, i.e. discussing those risks that might materialize with significant impact, but are not easily tracked unless more intensive tracking methods are employed. The overview will provide high-level suggestions as to how these risks could be tracked, including potential proxy variables.

A report detailing the items above will be presented to the AI Office at most 8 months after the signing of the contracts.

(g) Risk monitoring reports and briefings

The contractor shall execute the monitoring framework set-out above, providing update reports every 3 months until the end of contract. Such a report will contain:

- The status of all tracked variables, including hypotheses explaining any significant changes.

- A detailed collection of new incidents.

- Any problems with the tracking methodology and suggestions for changes, including in which variables to track.

- A policy-maker oriented summary.

The fourth update report will be linked to payment.

To complement the periodic monitoring reports above the contractor will provide the AI Office with regular update briefings on any events that are relevant for the developed risk models. Examples of such events are major new GPAI models or new GPAI providers, algorithmic improvements, new insights into the specific risks, updated safety frameworks, policy changes, incidents, new mitigations, relevant social or cultural trends and events, the activation of early warning indicators, or the documented breaking of a mitigation.

These briefings should be tailored to the specific risks considered in this lot, can be informal in nature, and are not expected to be exhaustive. They should integrate with the contractor's existing effort to stay up to date with the field of AI and the relevant risks in light of their other tasks, providing a lightweight instrument for sharing relevant updates with the AI Office.

The AI Office expects there to be events potentially worthy of briefing roughly every two weeks. This expectation should not be seen as prescriptive, and the sensitivity of briefing regime is expected to be calibrated in agreement with the AI Office, considering the intention to avoid unnecessary burden on the contractor.

(h) Ad hoc evaluations

To quickly respond to changes in the AI risk landscape, the AI Office will need to evaluate particular GPAI models on short notice. Therefore, the AI Office expects to request the contractor to conduct ad hoc evaluations at certain points in time. A maximum of 3 such evaluations may be commissioned.

For one such evaluation, the expectation regarding the contractor, is to make available a fixed budget of 200 expert hours with 2 weeks' notice to coordinate an evaluation of a GPAI model of the AI Office's choosing, to be spend over at most 4 weeks. These evaluations might involve human participants, for example for capturing realistic usage patterns in multi-turn interactions, or for assessing and labelling conversation traces and model outputs.

This specific method for the evaluation shall be tailored to the developed risk models and evaluation tools, to the degree these have already been finished. The AI Office may request regular updates from the moment of notice until the moment the evaluation report has been approved.

An evaluation report shall be delivered, at most 6 weeks after notice, containing

- A general description of the evaluation, its methodology, results, and conclusions.
- A policymaker-oriented summary.
- A dataset of the anonymized interaction transcripts, containing no personal data within the meaning of Regulation (EU) 2018/1725 (EUDPR) or Regulation (EU) 2016/679 (GDPR), including e.g. requests send to the GPAI model, GPAI model outputs, any assigned labels or grades, and other information that would allow for the reconstruction of the evaluation results.

The contractor is responsible for any costs associated with the request evaluation, including but not limited to, the payment of study participants, or the API costs associated with the GPAI model.

Informed by the first risk modelling workshop, AI Office and contractor may begin to consider potential study designs for upcoming evaluations, to facilitate rapid execution when the evaluation is required.

(i) Monthly calls

The contractor shall organise monthly calls with the AI Office to provide updates on the different tasks on the basis of the methodology and timelines agreed. Meeting minutes will be provided by the contractor at most five days after the monthly call.

Any reports or materials due in the same month shall be submitted at least five working days in advance of the monthly call to allow for review by the AI Office.

In justified cases, the AI Office may request ad hoc calls to complement the monthly calls and address specific issues that might arise during the execution of the contract. This will not impact the overall effort from the side of the contractor and will not cause any change to the minimum requirements of the tender specifications

(j) Final meeting and report

The contractor shall prepare a final report, which shall provide a comprehensive and detailed overview of the tasks conducted in the previous years. The final report should include the measures adopted by the contractor to ensure data protection compliance.

The submission of the final report is expected at the latest within 36 months from the contract's start date. The said final report will be linked to the final payment associated with the end of the contract.

The final report will be presented during a final meeting, organised by the contractor and to take place digitally.

The draft final report shall be sent to the European Commission 10 days before the final meeting at the latest. This will allow the Commission to assess the final results and provide comments where necessary, and the contractor to amend the report if required.

1.4.2.6.3. Timetable

Title	Due month (at the latest)	Linked to payment
Core Deliverables		
Risk modelling draft reports	M1, M3; M13, M15	No
Risk modelling workshops	M2; M14	No
Risk modelling reports	M6; M18	Yes
Public eval. tools proposal report	M3	No
Public eval. tools intermediate handover	M6	No
Public eval. tools handover	M12	Yes
Private eval. tools proposal report	M3	No
Private eval. tools intermediate handover	M6	No
Private evaluation tools handover	M12	Yes
Risk monitoring framework report	M8	Yes
Technical compliance toolkit	M13	No
Continued services		
Risk monitoring (briefings)	Continuous	No
Risk monitoring (update reports)	M11, M14, M17, M20, M23, M26, M29, M32, M35;	Yes (M20)
Ad hoc evaluations	Upon request	No
Coordination		
Inception meeting	M1	No
Inception report	M1	No
Monthly calls	M2-M35	No
Final meeting	M36	No
Final report	M36	Yes

1.4.2.7. Description of Lot 6: Agentic Evaluation Interface

All tasks are subject the provisions outlined in section 1.4.2.8 Description: Shared Provisions.

1.4.2.7.1. Objective

The objective of this lot is to supply the AI Office with a programmatic interface for the evaluation of General-Purpose Artificial Intelligence (GPAI) models through agentic interaction patterns, i.e. the evaluation of GPAI models for use in tasks requiring multiple decisions and interaction with digital environments such as a browser, command line, or a full operating system.

The interface is taken to mean software, methodology, configuration, cloud orchestration and other digital technical components that would allow the AI Office to execute evaluations of GPAI model and their agentic capabilities, without further dependency on the contractor, supplementing and working in tandem with AI Office digital infrastructure.

The AI Office is seeking such an interface in the form of a GPAI model evaluation workflow or 'harness', and the necessary components thereof, that allow the AI Office and its technical experts to rapidly integrate into the evaluation pipeline new GPAI models and benchmarks across all digital modalities (text, image, video, audio) and from the full spectrum of digital tasks. This harness should provide support for relevant and state of the art elicitation methods and agent scaffolding, including fallback mechanisms that connect text only models to multi-modal benchmarks.

It is expected that a significant part of the service will be dedicated to supporting the AI Office in the deployment on Commission infrastructure of existing or developed components of the evaluation interface.

Given the fast pace of development in AI, contractors should prepare for significant flexibility in execution and should structure the project to facilitate close communication for creating short feedback cycles and the purpose of alignment on design and needs.

1.4.2.7.2. Detailed characteristics and functional requirements

This section is focused on the detailed characteristics of the evaluation interfaces and discusses functional requirements, design restrictions, and the AI Office vision on the technical product. For details on coordination, support, logistics, and deliverables, see section 1.4.2.7.3. *Deliverables* below.

The primary goal for the developed work is to serve as an 'evaluation harness': a digital tool in the form of a programming library and corresponding command-line interface, together with corresponding supporting infrastructure setup, that allows the configuration of a particular evaluation, i.e. the selection of a specific GPAI model, benchmark, scaffolding techniques, and the setting of other relevant parameters, and then executing said evaluation configuration to produce evaluation results such as benchmark scores. While various such tools exist, this deliverable shall focus on complex and agentic tasks, the infrastructure to run those, and the tailoring thereof to existing AI

Office infrastructure. Some public example benchmarks are SWE Bench ([3]), OS World ([4]), Cybench ([5]), VisualWebArena ([6]), GAIA ([7]), and ToolACE ([8]).

The requirements for this harness and the corresponding infrastructure, referred to as 'the project', include the following:

- The project should be designed to be extendible by AI Office technology specialists and IT professionals, allowing for the integration of new GPAI models, new benchmarks, new model scaffolding techniques and other improvements and updates. The balance between being an optimal and easy to use product should thus always be balanced against the flexibility of the interface and the library primitives.

- The project should be deployable to Commission premises, which means it must be possible to integrate it with existing Commission computing infrastructure, including public and private cloud infrastructure. This will entail making available, among other deliverables, relevant software, configurations, or configuration templates for orchestrating networks, orchestrating containers, and building containers.

- The project can, and is strongly encouraged to, make use of existing open-source tools, given they have permissive licenses (e.g. MIT License). The contractor is expected to maintain a list of all direct and indirect dependencies of the produced code and any licenses related to these dependencies. This shall be accompanied where necessary by the indication of whether such dependencies have been modified via a software bill of materials. This list will be included in the source code and make explicit references to libraries and specific functions. The contractor will ensure compatibility between the licenses of the dependencies used.

- The project should assume black-box access to the GPAI model through an HTTP(S) API, which will be different between various GPAI model providers.

- The project should support the infrastructural needs of various advanced benchmarks, including, but not limited to,

    - Evaluations centred around measuring the capability of the GPAI model to support autonomous AI research and development (R&D), including the writing and executing of code to train or finetune other GPAI models or otherwise run AI experiments.

    - Evaluations centred around measuring the capability of GPAI models to replicate themselves to places different from their intended location, including the circumvention of any measures designed to avoid such replication.

    - Agentic (offensive) cyber evaluations, including Capture the Flag challenges, and the finding and exploiting vulnerabilities in realistic settings, etc.

([3]) https://www.swebench.com/ , https://arxiv.org/abs/2310.06770
([4]) https://os-world.github.io/ , https://arxiv.org/abs/2404.07972
([5]) https://cybench.github.io/ , https://arxiv.org/abs/2408.08926
([6]) https://github.com/web-arena-x/visualwebarena , https://arxiv.org/abs/2401.13649
([7]) https://huggingface.co/gaia-benchmark , https://arxiv.org/abs/2311.12983
([8]) https://huggingface.co/datasets/Team-ACE/ToolACE , https://arxiv.org/abs/2409.00920

    - Advanced biological capability evaluations that relate to finding, installing, and using biological design tools, writing and executing code to manipulate biological sequences, controlling lab hardware interfaces, etc.

    - Tasks that require several hours to several days for a human to complete.

- The project should encompass a collection of state of the art agentic 'scaffolding' techniques, i.e. programmatic workflows that allow GPAI models to be used for agentic tasks. Similarly, the project should encompass a collection of 'elicitation' techniques, i.e. programmatic workflows that make optimal use of GPAI model capabilities to complete a task at hand, examples of which are chain of thought prompting and tool-use.

- The project should include the possibility for interaction between GPAI model and benchmark to happen through a universal interface that resembles how humans would use a computer, where the model receives screenshots as input and outputs keyboard and mouse actions, and integrating with other peripherals such as microphone, webcam, or touchscreen through a corresponding virtual device.

- The project should be able to represent various non-agentic benchmarks, for example pertaining to an image classification task or a textual multiple-choice question answering tasks, in a web-based interface, as to integrate them into an agentic interaction pattern. Such a web-based interface should be designed as to allow a human to complete the task at hand and should thus allow for the measurement of human performance on these tasks, acting as reference to compare GPAI performance against. The measurement of human performance is not a task assigned to the contractor.

  Decision and prioritization of which classes of non-agentic benchmarks to integrate will be decided jointly with the AI Office.

- In relation to the agentic scaffolding, the project should integrate fallback mechanisms that allow GPAI models that do possess to option to handle particular input or output modalities, e.g. text only language models, to nonetheless be used for agentic tasks. Potential examples are the use of operating system provided accessibility information, the use of web-page hypertext as opposed to screenshots, or the use of intermediary AI systems such as visual question answering models.

- The project should aim to automatically grade agent behaviours where possible and should support state-of-the-art workflows such as advanced 'AI-as-a-judge' methodology. Given the flexible interaction between agent and benchmark, the project should make sure to not leak grading labels, answer keys, or reward functions to the agent.

- The project should aim to support a diverse set of benchmarks, including benchmarks such as OS World already mentioned before, non-agentic benchmarks such as image classification or translation benchmarks, but also classic control tasks such as the mountain car problem, or game-oriented benchmarks such as chess, Atari, or Nethack. Prioritisation will be decided jointly with the AI Office.

- The project should support active evaluation strategies that evaluate a GPAI model on a given benchmark by carefully selecting benchmark instances as to minimise the number of instances, and thus the corresponding compute needed, while still providing accurate performance measurement. The contractor is not expected to develop such evaluation strategies but should facilitate their use.

These requirements, *and the priorities thereof*, should be expected to change in accordance with the fast pace of change in the AI ecosystem, and will be the topic of discussion in monthly calls and other collaboration methods between the AI Office and the contractor.

1.4.2.7.3. Development coordination

To facilitate development coordination across institutions, and as to enable short feedback cycles and reduce communication overhead, the contractor shall:

- make available at least one technical member of the contractor's development team to work on-site at AI Office premises, referred to as 'intramuros', whose role it will be to ensure integration of the developed infrastructure with existing AI Office resources;
- provide designated members of the AI Office with continuous read and write access to the relevant code repositories and version control systems ([9]), as well as to read access to any other repositories, for example those used to host containers,
- organise weekly (remote) pair programming sessions with designated members of the AI Office,
- be maximally transparent about the development process, including any designated roles (e.g. project manager, AI experts, …)

Apart from the contractor personal working in intramuros capacity, all development will take place at contractor premises using contractor infrastructure and resources. The purpose of the mechanisms set out above is to facilitate coordination and does not imply the AI Office will take up significant development responsibilities, although under mutual agreement, the AI Office may decide to do so.

1.4.2.7.4. Maintenance and support

To guarantee support, maintenance, and integration of the developed works with AI Office infrastructure, as well as onboarding of AI Office technological specialists, the contractor shall make available one technical member of their development team to provide *intramuros* assistance for a period of 12 months after the final infrastructure handover. This contractor employee shall have sufficient expertise and experience to maintain and improve the evaluation interface across all its dimensions, including

- the keeping up to date of software dependencies
- adapting the project to changes in cloud services used at the Commission
- integrating new GPAI models and benchmarks that fall within the supported functionality of current implementation of the project
- adapting, expanding, and in general improving the project with new features

([9]) Write access may be gated behind approval mechanisms, e.g. common version control system pull or merge requests, and the contractor maintains the right to refuse any changes suggested by the AI Office in light of their responsibility to guarantee the quality of the deliverables.

It is expected that major parts of the development, testing, and GPAI model evaluations related to this maintenance and support would happen on AI Office premises using AI Office infrastructure.

1.4.2.7.5. Deliverables

(a) Inception meeting and report

The contractor will prepare an inception report, covering all tasks in this lot, providing a preliminary overview of the proposed approach, including staff capacities as well as measures for compliance with personal data protection legislation. For avoidance of doubt, the inception report can be substantively the same as the work plan submitted with the tender, if such a plan is sufficiently detailed and up to date to reflect the contract as signed. The inception report will be discussed at the inception meeting, which will be a digital meeting happening at the latest 1 month after the signing of the contract.

The inception report draft will be provided at least 5 working days before the inception meeting, and the final version will be provided at most 5 working days after the inception meeting, integrating any discussion and feedback for such meeting.

The inception report and meeting will serve as the initial foundation for the execution of the tasks and any collaboration.

(b) Monthly calls

The contractor shall organise monthly calls with the AI Office to provide updates on the different tasks on the basis of the methodology and timelines agreed. These calls will be used to set the priorities for the coming month, update the development roadmap, and discuss results and next steps for sample evaluations (see below).

Meeting minutes will be provided by the contractor at most five days after the monthly call.

Any reports or materials due in the same month shall be submitted at least five working days in advance of the monthly call to allow for review by the AI Office.

In justified cases, the AI Office may request ad hoc calls to complement the monthly calls and address specific issues that might arise during the execution of the contract. This will not impact the overall effort from the side of the contractor and will not cause any change to the minimum requirements of the tender specifications

(c) Sample evaluations

To support grounding the development with realistic needs and to test the functionality of the implementation, the contractor shall use their developed solutions to run sample evaluations on a monthly basis, where the evaluation configuration, i.e. which GPAI models, benchmarks, and scaffolding techniques to use, will be decided between the contractor and the AI Office as to set feasible goals and development priorities.

These evaluations shall be executed in its entirety by the contractor using contractor resources and infrastructure, and a report detailing the evaluations results, as well as any issues encountered, shall be provided to the AI Office at most 1 week later.

If sample evaluations can be executed successfully on contractor infrastructure, the contractor employees working in intramuros capacity shall reproduce said evaluations on Commission infrastructure, aiming to reproduce them at most 1 month later.

(d) Project handover

At fixed points in time, the contractor shall prepare a comprehensive handover of the developed interface to the AI Office. These milestones should include at least a prototype handover, aimed at an initial integration into the AI office for the purpose of identifying integration and requirement issues, as well as a final handover at the end of the main development period.

Such a handover package will include:

- any relevant code, configuration, containers, and other technical artifacts that allow the AI Office to integrate the developed works in their internal processes and infrastructure, without continued dependency on the contractor;
- sufficient documentation, including
    - development documentation, including an architectural overview and its design rationale, as well as how-to guides for integrating new GPAI models, new benchmarks, and new scaffolding techniques
    - user manuals tailored to technical users, including how-to guides for running evaluations and a reference of the interface and its options, troubleshooting guides and installation guides
    - project documentation, including a consolidated view of the meeting minutes and other communications, a reference of resources created, and an overview of quality assurance processes and results
- reports on the sample evaluations that have been ran, any issues that have been encountered, and any remediations that have implemented.

Such a handover package will be tested for quality by the ability of the AI Office to autonomously run sample evaluations, specifically those evaluations that were designated as the monthly targets.

(e) Project management documents

Upon request from the Commission, the contractor shall prepare or assist the Commission in the creation of project management documentation such as a project charter, an IT security plan, a risk mitigation strategy, stakeholder reports, or status reports. This list is indicative.

Data protection: upon request from the Commission, the contractor should prepare a Data Protection Concept Note, whose objective is to prove compliance with the principle of data protection by design and by default, as set out in Article 27 EUDPR, in the development of the project. The contractor therefore has to make sure it possesses the necessary data protection expertise to perform these tasks and prepare the deliverable. The deliverable should be kept up-to-date throughout the implementation of the contract.

(f) Final meeting and report

The contractor shall prepare a final report, which shall provide a comprehensive and detailed overview of the tasks conducted in the previous years. The final report shall take care to include any changes in the project relating to the period between the project handover and final report.

The submission of the final report is expected at the latest within 24 months from the contract's start date. The said final report will be linked to the final payment associated with the end of the contract.

The final report will be presented during a final meeting, organised by the contractor and to take place digitally.

The draft final report shall be sent to the European Commission 10 days before the final meeting at the latest. This will allow the Commission to assess the final results and provide comments where necessary, and the contractor to amend the report if required.

1.4.2.7.6. Timetable

Title	Due month	Linked to payment
Core deliverables		
Project prototype handover	M6	Yes
Sample evaluations	M2-M12	No
Project handover	M12	Yes
Project management documents	Upon request	No
Continuous services		
Development Coordination (intramuros, repository access, pair coding)	M1-M12 (including)	No
Maintenance and support (intramuros)	M13-M24 (including)	No
Coordination		
Inception meeting	M1	No
Inception report	M1	No
Monthly calls	M2-23	No
Final meeting	M24	No
Final report	M24	Yes

1.4.2.8. Description: Shared Provisions

These provisions apply to all lots individually and equally.

1.4.2.8.1. Experts

Any involvement of third-party experts must be in line with the provision of section 2.4.2. (Subcontracting) of these specifications.

Any involvement of third-party experts will, unless explicitly noted otherwise, always fall entirely under the responsibility of the contractor, including but not limited to, the search and selection of experts, their payment, the coordination of their tasks, and the coordination and reimbursement of any travel. The offers shall include therefore all the costs related to arrangement of third-party experts.

1.4.2.8.2. Meetings and workshops

The contractor shall cover the costs for all the supplies related to any meeting or workshop (e.g. printed materials, only if strictly needed) and for the participation of the contractor's staff/experts (travel, accommodation, subsistence). The offers shall include therefore all the costs related the organisation of meetings.

1.4.2.8.3. Terms of approval of deliverables

Except for the reports linked to payments, the Commission shall have 30 days from receipt to approve or reject the deliverable(s), and the contractor shall have 30 days in which to submit additional information or a new deliverable. The Commission may allow the contractor additional time to rework complex deliverables.

For the terms of approval of the other deliverables linked to payments, please refer to Article I.5 of the contract.

1.4.2.8.4. Reports

All reports shall be written in English and submitted electronically. Additionally, they should be consistent in style (headings, margins, citations, bibliography etc.) and contain a short executive summary. The contractor is required to apply properly quotation techniques and particular care will be taken to verify improper re-use of existing material (see also below on Intellectual Property Rights). The contractor must also ensure that all reports are drafted in a clear and easily understandable language, that they are concise, logically structured and focused on their purpose and that compliance with data protection is respected.

1.4.2.8.5. Finalisation and online dissemination of the final report

Ahead of the finalisation of the report, the contractor will ensure that the presentation of the texts, tables and graphs is clear and complete and corresponds to commonly recognised standards for publication. In particular, the contractor will make reference to and apply the Interinstitutional Style Guide (https://publications.europa.eu/code/en/en-000300.htm).

The contractor will give attention to the following requirements:

- web content accessibility guidelines
- copyright/ intellectual property rights (and see below)

1.4.2.8.6. Confidentiality of information

Contractors must guarantee the confidentiality of the information in line with the contractual provisions. In this regard, the tender must include a Declaration on confidentiality of information using the template in Annex V to the draft contract, signed by the legal representative of the group members (contractors and subcontractors) and by the proposed team members/experts that will participate in the contract's implementation. Details regarding requirements and obligations concerning confidentiality of information are set out in the draft contract.

1.4.2.8.7. Intellectual Property

Regarding the Intellectual Property Rights, the contractor's attention is brought to the contract provisions (in particular articles I.8 and II.13 of the service contract) on the ownership and/or transfer of intellectual property rights in deliverables created in the framework of this contract. This paragraph is without prejudice to specific references to intellectual property elsewhere in the tender specifications.

For the avoidance of doubt:

- The Contractor shall be responsible for the clearance of third-party intellectual property rights embedded in the deliverables.
- The Contracting Authority shall be entitled to use the deliverables for any of the modes of exploitation detailed in article I.8 of the service contract, including the publication and making publicly available online.
- Unless provided otherwise in the special conditions, the Union does not acquire ownership of pre-existing rights under this contract (Article I.8.1 and Article II.13.2 of the draft contract).
- The Contracting Authority acquires irrevocably worldwide ownership of the results and of all intellectual property rights on the newly created materials produced specifically for the Union under the contract and incorporated in the results, without prejudice however to the rules applying to pre-existing rights on pre-existing materials (Article II.13.1 of the draft contract).

☝ Please note that while the Contracting Authority acquires ownership of the results, it may decide to extend licenses to the Contractor or other parties, for use of the results or parts thereof, beyond the scope of this contract, especially when such use would benefit the objectives set out in these specifications.

The Contractor shall ensure that the deliverable already includes the necessary attributions, disclaimers and copyright notices for its publication. Any non-EU-owned content in the deliverable must be explicitly acknowledged (both in the copyright notice and where this content is used), identifying the copyright holder (and, if different and identifiable, the author of the work), the content's location in the work and any conditions/limitations on the reuse.

The Contractor shall include in the deliverable the following copyright notices:

d. Text documents free of third-party content

"© European Union, 20XX [year of first publication]

The Commission's reuse policy is implemented by Commission Decision 2011/833/EU of 12 December 2011 on the reuse of Commission documents (OJ L 330, 14.12.2011, p. 39 – https://eur-lex.europa.eu/eli/dec/2011/833/oj).

Unless otherwise noted (e.g. in individual copyright notices), the reuse of this document is authorised under the Creative Commons Attribution 4.0 International (CC BY 4.0) licence (https://creativecommons.org/licenses/by/4.0/). This means that reuse is allowed, provided appropriate credit is given and any changes are indicated."

- Text documents including third-party content

"© European Union, 20XX [year of first publication]

The Commission's reuse policy is implemented by Commission Decision 2011/833/EU of 12 December 2011 on the reuse of Commission documents (OJ L 330, 14.12.2011, p. 39 – https://eur-lex.europa.eu/eli/dec/2011/833/oj).

Unless otherwise noted (e.g. in individual copyright notices), the reuse of this document is authorised under the Creative Commons Attribution 4.0 International (CC BY 4.0) licence (https://creativecommons.org/licenses/by/4.0/). This means that reuse is allowed, provided appropriate credit is given and any changes are indicated.

For any use or reproduction of elements that are not owned by the EU, permission may need to be sought directly from the respective rightholders. The EU does not own the copyright in relation to the following elements:

- *Cover page illustration, © Author name / stock.adobe.com*
- *[page XX, element concerned], source: [e.g. Fotolia.com]*
- *…"*

The copyright notice of a text document should normally be included in the title page, on the reverse side of the title page or on either side of the front or back cover.

Concerning the possible use of AI-generated content, if a document containing AI generated output is published or is used for the purpose of the present study, the contractor should insert an explicit acknowledgement as part of the output's copyright notice, e.g., if the EU is the output's copyright owner, for instance as follows: *"© European Union 2023. Some content was created using [name of AI tool]"*.

In addition to the abovementioned copyright notices, the final report should include the following disclaimer:

"The information and views set out in this [report/study/article/publication…] are those of the author(s) and do not necessarily reflect the official opinion of the Commission. The Commission does not guarantee the accuracy of the data included in this study. Neither the Commission nor any person acting on the Commission's behalf may be held responsible for the use which may be made of the information contained therein."

1.4.2.8.8. Data Protection Compliance

Where the contractor processes personal data in the context of the present contract, the contractor has to respect Regulation (EU) 2018/1725. In this sense, the contractor will follow the instructions (in the light of Article 29 Regulation (EU) 2018/1725) of the Commission to achieve data protection compliance. This paragraph is without prejudice to specific references to data protection elsewhere in the tender specifications.

1.5. Place of performance: where will the contract be performed?

The services will be performed at the following locations:

- the contractor's premises
- the location(s) indicated in Section 1.4 of these specifications, including the European Commission's premises, applying to meetings and workshops for all Lots, and the intramuros support specified in Lot 6.

1.6. Nature of the contract: how will the contract be implemented?

The procedure will result in the conclusion of the following contract types per lot:

Lot	Contract type
Lot 1 - CBRN Risk Modelling and Evaluation	a direct contract
Lot 2 - Cyber Offence Risk Modelling and Evaluation	a direct contract
Lot 3 - Loss of Control Risk Modelling and Evaluation	a direct contract
Lot 4 - Harmful Manipulation Risk Modelling and Evaluation	a direct contract
Lot 5 - Sociotechnical Risk Modelling and Evaluation	a direct contract
Lot 6 - Agentic Evaluation Interface	a direct contract

In direct contracts all the terms governing the provision of the services, supplies or works are defined at the outset. Once signed, they can be implemented directly without any further contract procedures.

Tenderers need to take full account of the full set of procurement documents, including the provisions of the draft contract as the latter will define and govern the contractual relationships to be established between the contracting authority and the successful tenderers. Special attention is to be paid to the provisions specifying the rights and obligations of the contractor, in particular those on payments, performance of the contract, confidentiality, and checks and audits.

☝ Please be aware that if a tenderer to whom the contract is awarded (any of the group members in case of a joint tender) has established debt(s) owed to the Union, the European Atomic Energy Community or an executive agency when the latter implements the Union budget, such debt(s) may be offset, in line with Articles 101(1) and 102 of Regulation (EU, Euratom) 2024/2509 of the European Parliament and of the Council of 23 September 2024 on the financial rules applicable to

the general budget of the Union (Financial Regulation)[10] and the conditions set out in the draft contract, against any payment due under the contract. The contracting authority will verify the existence of overdue debts of the successful tenderers (any of the group members in case of a joint tender), and, if any such debt is found, will inform the tenderer (the group leader in case of a joint tender who will then have the obligation to inform all other group members before signing the contract) that the debt(s) may be offset against any payment under due the contract.

1.7. Volume and value of the contract: how much do we plan to buy?

The maximum total amount of all purchases under this call for tenders is indicated under Section 2.1.3 of the contract notice. The volumes/values of the purchases for each lot over the total duration of the contract are:

Lot	Maximum value
Lot 1 - CBRN Risk Modelling and Evaluation	1,850,000€
Lot 2 - Cyber Offence Risk Modelling and Evaluation	2,000,000€
Lot 3 - Loss of Control Risk Modelling and Evaluation	1,800,000€
Lot 4 - Harmful Manipulation Risk Modelling and Evaluation	1,300,000€
Lot 5 - Sociotechnical Risk Modelling and Evaluation	1,300,000€
Lot 6 - Agentic Evaluation Interface	1,080,000€

Within three years following the signature of the contracts resulting from the lots specified below, the contracting authority may use the negotiated procedure under point 11.1.e of Annex 1 to the Financial Regulation to procure new services from the contractors up to the following maximum percentages for the respective lot:

Lot	New services
Lot 1 - CBRN Risk Modelling and Evaluation	Max. 50% of the initial contract value
Lot 2 - Cyber Offence Risk Modelling and Evaluation	Max. 50% of the initial contract value
Lot 3 - Loss of Control Risk Modelling and Evaluation	Max. 50% of the initial contract value
Lot 4 - Harmful Manipulation Risk Modelling and Evaluation	Max. 50% of the initial contract value
Lot 5 - Sociotechnical Risk Modelling and Evaluation	Max. 50% of the initial contract value
Lot 6 - Agentic Evaluation Interface	Max. 50% of the initial contract value

These services would consist in the repetition of similar services entrusted to the contractors and

[10] Regulation (EU, Euratom) 2024/2509 of the European Parliament and of the Council of 23 September 2024 on the financial rules applicable to the general budget of the Union, amending Regulations (EU, Euratom) 2018/1046, No 1296/2013, (EU) No 1301/2013, (EU) No 1303/2013, (EU) No 1304/2013, (EU) No 1309/2013, (EU) No 1316/2013, (EU) No 223/2014, (EU) No 283/2014, and Decision No 541/2014/EU and repealing Regulation (EU, Euratom) No 966/2012 (OJ L 193 of 30.07.2018, p.1).

would be awarded under the following conditions: same conditions as the current call for tenders.

1.8. Duration of the contract: how long do we plan to use the contract?

The contracts resulting from the award of this call for tenders will be concluded for the duration specified per lot below. The details of the initial contract duration and possible renewals per lot are set out in the draft contract for the respective lot.

Lot	Lot duration
Lot 1 - CBRN Risk Modelling and Evaluation	36 months
Lot 2 - Cyber Offence Risk Modelling and Evaluation	36 months
Lot 3 - Loss of Control Risk Modelling and Evaluation	36 months
Lot 4 - Harmful Manipulation Risk Modelling and Evaluation	36 months
Lot 5 - Sociotechnical Risk Modelling and Evaluation	36 months
Lot 6 - Agentic Evaluation Interface	24 months

1.9. Electronic exchange system: can exchanges under the contract be automated?

For all exchanges with the contractors during the implementation of the contracts as well as for future possible subsequent proceedings, including, but not limited to, for the purposes of EDES (European Union's Early Detection and Exclusion System), the contracting authority may use an electronic exchange system meeting the requirements of Article 151 of the Financial Regulation. At the request of the contracting authority, the use of such a system shall become mandatory for the contractors at no additional cost for the contracting authority. Details on specifications, access, terms and conditions of use will be provided in advance.

1.10. Security

When performing tasks for the contracting authority in execution of the contract, the contractor and its personnel shall comply with the contracting authority's applicable security requirements.

For the Commission (and, when relevant - for the Executive Agencies), the applicable security requirements include:

- ✓ For tenderers to Lot 5, “Agentic Evaluation Interface”, Commission Decision (EU, Euratom) 2017/46 of 10 January 2017 on the security of communication and information systems in the European Commission, its subsequent versions, its implementing rules (as adapted from time to time) and the corresponding security standards, guidelines and notices;

- ✓ Commission Decision (EU, Euratom) 2015/444 of 13 March 2015 on the security rules for protecting EU classified information, its subsequent versions, its implementing rules (as adapted from time to time) and the corresponding security standards, guidelines and notices;

- ✓ Commission Decision (EU, Euratom) 2015/443 of 13 March 2015 on Security in the Commission, as well as all its subsequent versions.

- ✓ https://ec.europa.eu/info/files/security-standards-information-systems_en

For tenderers to Lot 5, “Agentic Evaluation Interface”, Specific security rules for the contractor’s personnel are set out in Article I.11 of the draft contract.

Any financial burden for complying with the security measures (e.g. security background checks, security clearance etc.) will be entirely at the expense of the contractor and not of the contracting authority.

The contracting authority reserves the right to require any person involved in the provision of the services under a given project to attend security briefings or training given by the contracting authority, and/or to sign a security statement.

In exceptional cases, when required for security reasons, the contracting authority may ask the contractor to provide security vetted personnel for the provision of certain services. A positive outcome of the national vetting process leads to the status “security clearance”. This will be considered as a specific requirement for a specific project, without influencing the other conditions.

Should the contractor, during the performance of the tasks, which are the subject of the contract, need remote access to any communication and information system of Commission or data sets processed therein, one of the two following approaches should be observed:

1) Contractor’s personnel is granted remote access to any communication and information system of the Commission or data sets processed therein, without being provided with Commission IT equipment. In this case the Contractor shall be requested to comply with security rules referred to in Article 6(5) of the Commission Decision (EU, Euratom) 2017/46 of 10 January 2017. This entails prior authorisation, which shall be granted on the basis of a formal request for network access service “Remote Access for Companies”, and approval process, which takes on average 4-6 weeks. The outcome of the approval, i.e. the Interconnection Security Agreement, shall be valid for a specified duration linked to the contract and shall be obtained before the connection is activated. The formal request is initiated by the concerned Directorate-General or service of the Commission and based on the risk assessment with the focus on nature and sensitivity of the tasks to be performed remotely and the security needs of each accessed communication and information system.

   During the authorisation process the contractor is asked to describe relevant organisational, physical, logical and network security measures in order to provide reasonable assurance that the risks are adequately and systematically covered at a level equivalent to the Commission Decision (EU, Euratom) 2017/46 of 10 January 2017, its implementing rules and corresponding security standards. The authorisation process may impose additional security requirements as a prerequisite for approval, in order to protect the Commission’s communication and information systems and networks from the risks of unauthorised access or other security breaches. No remote access will be possible in this context without having in place an approved Interconnection Security Agreement (formerly called a security convention).

   Contractors and service providers may be required to comply with the baseline security measures published by the Commission at Standards & Procedures (https://ec.europa.eu/info/files/security-standards-information-systems_en).

2) Contractor's personnel use Commission IT equipment (normally a laptop PC) and connects to the Commission’s internal network via the remote access service for Commission staff. In this case, contractors are required to put in place minimum security measures in order to mitigate risks to the security of Commission information during the fulfilment of the contracted services. These measures focus mainly on the confidentiality and integrity of Commission

equipment and information. The baseline security measures for contractors in the context of remote service delivery are available for consultation at the internet address: https://ec.europa.eu/info/files/security-standards-information-systems_en. These rules apply to service providers working on contractor's premises or in home offices, where permitted by the specific contract. This baseline does not cover service providers accessing non-Commission systems, such as contractors' development environments. When the contractor undertakes to follow these controls in the contract, access is permitted without an additional Interconnection Security Agreement (security convention).

2. GENERAL INFORMATION ON TENDERING

2.1. Legal basis: what are the rules?

This call for tenders is governed by the provisions of the Financial Regulation.

The contracting authority has chosen to award the contracts resulting from this call for tenders through an open procedure pursuant to Article 167(1)(a) of the Financial Regulation.

In this procedure any interested economic operator (any natural or legal person who offers to supply products, provide services or execute works) may submit a tender.

2.2. Entities subject to restrictive measures and rules on access to procurement: who may submit a tender?

Tenderers must ensure that no involved entities (see Section 2.4) nor any subcontractors, including those which do not need to be identified in the tender (see Section 2.4.2), are subject to EU restrictive measures adopted under Article 29 of the Treaty on the European Union (TEU) or Article 215 of the Treaty on the Functioning of the EU (TFEU)[2], consisting of a prohibition to make available or transfer funds or economic resources or to provide financing or financial assistance to them directly or indirectly, or of an asset freeze. The prohibition applies throughout the whole performance of the contract.

Participation in this call for tenders is open on equal terms to all natural and legal persons coming within the scope of the Treaties, as well as to international organisations.

It is also open to all natural and legal persons established in a third country provided that it has a special agreement with the European Union in the field of public procurement on the conditions laid down in that agreement.

As the Agreement on Government Procurement[11] concluded within the World Trade Organisation applies, the participation to this call for tenders is also open to all natural and legal persons established in the countries that have ratified this Agreement, on the conditions laid down therein.

The rules on access to procurement do not apply to entities on whose capacity tenderers rely to fulfil the selection criteria nor to subcontractors. Subcontracting may not be used with the intent or effect to circumvent the rules on access to procurement.

Participation in this call for tenders is also open on equal terms to natural and legal persons established in a third country eligible for funding under the Digital Europe Programme[3].

To enable the contracting authority to verify the access, each tenderer must indicate its country of establishment (in case of a joint tender – the country of establishment of each group member) and must present the supporting evidence normally acceptable under the law of that country. The same document(s) could be used to prove country/-ies of establishment and the delegation(s) of the authorisation to sign, as described in Section 4.3.

---

[2] Please note that the EU Official Journal contains the official list and, in case of conflict, its content prevails over that of the EU Sanctions Map.

[3] https://www.wto.org/english/tratop_E/gproc_e/gp_gpa_e.htm

[4] Regulation (EU) 2021/694 of the European Parliament and of the Council of 29 April 2021 establishing the Digital Europe Programme and repealing Decision (EU) 2015/2240, OJ L 166, 11.5.2021, p. 1–34

2.3. Registration in the Participant Register: why register?

Any economic operator willing to participate in this call for tenders must be registered in the Participant Register - an online register of organisations and natural persons (participants) participating in calls for tenders or proposals of the European Commission and other EU institutions/bodies.

On registering each participant obtains a Participant Identification Code (PIC, 9-digit number), which acts as its unique identifier in the Participant Register. A participant needs to register only once – the information provided can be further updated or re-used by the participant in other calls for tenders or calls for proposals of the European Commission and other EU institutions/bodies.

☝ **Each participant needs to ensure that its SME status in the Participant Register is registered and kept up to date.**

At any moment during the procurement procedure, the Research Executive Agency Validation Services (hereafter *the EU Validation Services*) may contact the participant and ask for supporting documents on legal existence and status and financial capacity. The requests will be made through the register's messaging system to the e-mail address of the participant's contact person indicated in the register. It is the responsibility of the participant to provide a valid e-mail address and to check it regularly. The documents that may be requested by *the EU Validation Services* are listed in the EU Grants and Tenders Rules on Legal Entity Validation, LEAR appointment and Financial Capacity assessment.

☝ **Please note that a request for supporting documents by the *EU Validation Services* in no way implies that the tenderer has been successful.**

2.4. Ways to submit a tender: how can economic operators organise themselves to submit a tender?

Economic operators can submit a tender, either as a sole economic operator (sole tenderer) or as a group of economic operators (joint tender)[4]. In either case subcontracting is permitted.

Tenders must be drawn and submitted in complete independence and autonomously from the other tenders. A declaration in this regard by each tenderer (in case of a joint tender, by the group leader) shall be requested ***(Annex 2)***.

A natural or legal person cannot participate at the same time and for the same lot within the same procedure either as member of two or more groups of economic operators or as a sole tenderer and member of another group of economic operators. In such case, all tenders in which that person has participated, either as sole tenderer or as member of a group of economic operators, will be rejected.

Economic operators linked by a relationship of control or of association (e.g. belonging to the same economic/corporate group) are allowed to submit different and separate tenders, provided that each tenderer is able to demonstrate that its tender was drawn independently and autonomously.

A natural or legal person may act as subcontractor for several tenderers as long as the tenders are drawn and submitted in complete independence and autonomously from each other. However, cross subcontracting among tenderers is forbidden, more precisely an entity "A" may participate as tenderer

[4] Each economic operator participating in the joint tender is referred to as "group member".

(either as sole tenderer or as member of a group of economic operators) and as subcontractor to another tenderer "B" for the same lot within the same procurement procedure. However, in this case it is forbidden that tenderer "B" (or any of its participating members in case of a group of economic operators) is at the same time subcontractor for tenderer "A" (or for the group of economic operators in which "A" participates) for the same lot within the same procurement procedure. In this case, both tenders A and B shall be rejected.

In order to fulfil the selection criteria set out in Section 3.2 the tenderer can rely on the capacities of subcontractors (see Section 2.4.2) or other entities that are not subcontractors (see Section 2.4.3).

An **"involved entity"** is any economic operator involved in the tender. This includes the following four categories of economic operators:

- sole tenderer,
- group members (including group leader),
- identified subcontractors (see Section 2.4.2), and
- other entities (that are not subcontractors) on whose capacity the tenderer relies to fulfil the selection criteria.

The role of each entity involved in a tender must be clearly specified in the eSubmission application: i) sole tenderer, ii) group leader (in case of a joint tender), iii) group member (in case of a joint tender), or iv) subcontractor[12].

For an entity on whose capacities the tenderer relies to fulfil the selection criteria (that is not a subcontractor), this role is defined in the commitment letter (***Annex 5.2***)

2.4.1. Joint tenders

A joint tender is a situation where a tender is submitted by a group (with or without legal form) of economic operators regardless of the link they have between them in the group. The group as a whole is considered a tenderer[13].

All group members assume joint and several liability towards the contracting authority for the performance of the contract as a whole.

Group members must appoint from among themselves a group leader (the group leader) as a single point of contact authorised to act on their behalf in connection with the submission of the tender and all relevant questions, clarification requests, notifications, etc., that may be received during the evaluation, award and until the contract signature. All group members (including the group leader) must sign an Agreement/Power of attorney drawn up in the model attached in **Annex 3**.

The joint tender must clearly indicate the role and tasks of each group member, including those of the group leader who will act as the contracting authority's contact point for the contract's administrative or financial aspects and operational management. The group leader will have full authority to bind the group and each of its members during contract execution.

If the joint tender is successful, the contracting authority shall sign the contract with the group leader, authorised by the other members to sign the contract also on their behalf via the Agreement/Power of

[5] Only identified subcontractors (see Section 2.4.2) must be specified in the eSubmission application.

[6] References to *tenderer* or *tenderers* in this document shall be understood as covering both sole tenderers and groups of economic operators submitting a joint tender.

attorney drawn up in the model attached in ***Annex 3***.

Changes in the composition of the group during the procurement procedure (after the deadline for submission of tenders and before contract signature) shall lead to rejection of the tender, with the exception of the following cases:

- case of a merger or takeover of a group member (universal succession), provided that the following cumulative conditions are fulfilled:

  - the new entity is not subject to restrictive measures, has access to procurement (see Section 2.2) and is not in an exclusion situation (see Section 3.1),
  - all the tasks assigned to the former entity are taken over by the new entity member of the group,
  - the group meets the selection criteria (see Section 3.2),
  - the change must not make the tender non-compliant with the procurement documents,
  - the terms of the originally submitted tender are not altered substantially and the evaluation of award criteria of the originally submitted tender are not modified,
  - the new entity undertakes to replace the former entity for the implementation of the contract, in case of an award.

- case where a group member is subject to restrictive measures or does not have access to procurement (see Section 2.2) or is in an exclusion situation (see Section 3.1), provided the following cumulative conditions are fulfilled:

  - none of the remaining group members is subject to restrictive measures (see Section 2.2),
  - all the remaining group members have access to procurement (see Section 2.2),
  - the remaining group members meet the selection criteria (see Section 3.2),
  - the change must not make the tender non-compliant with the procurement documents,
  - the terms of the originally submitted tender are not altered substantially and the evaluation of award criteria of the originally submitted tender are not modified,
  - the continuation of the participation of the remaining group members in the procurement procedure does not put the other tenderers in a competitive disadvantage,
  - the remaining group members undertake to implement the contract, in case of an award, without the excluded group member.

The replacement of the group member not having access to procurement or in a situation of exclusion is not allowed.

2.4.2. Subcontracting

Subcontracting is the situation where the contractor enters into legal commitments with other economic operators, which will perform part of the contract on its behalf. The contractor retains full liability towards the contracting authority for performance of the contract as a whole.

The following shall not be considered subcontracting:

a) Use of workers posted to the contractor by another company owned by the same group and established in a Member State ("intra-group posting" as defined by Article 1, 3, (b) of Directive 96/71/EC concerning the posting of workers in the framework of the provision of services).
b) Use of workers hired out to the contractor by a temporary employment undertaking or placement agency established in a Member State ("hiring out of workers" as defined by Article 1, 3, (c) of Directive 96/71/EC concerning the posting of workers in the framework of the

provision of services).

c) Use of workers temporarily transferred to the contractor from an undertaking established outside the territory of a Member State and that belongs to the same group ("intra-corporate transfer" as defined by Article 3, (b) of Directive 2014/66/EU on the conditions of entry and residence of third-country nationals in the framework of an intra-corporate transfer).
d) Use of staff without employment contract ("self-employed persons working for the contractor"), without the tasks of the self-employed persons being particular well-defined parts of the contract.
e) Use of suppliers and/or transporters by the contractor, in order to perform the contract at the place of performance, unless the economic activities of the suppliers and/or the transporting services are within the subject of this call for tenders (see Section 1.4).
f) Performance of part of the contract by members of an EEIG (European Economic Interest Grouping), when the EEIG is itself a contractor or a group member.

The persons mentioned in points a), b), c) and d) above will be considered as "personnel" of the contractor as defined in the contract.

All contractual tasks may be subcontracted unless the procurement documents expressly reserve the execution of certain critical tasks to the sole tenderer itself, or in case of a joint tender, to a group member.

By filling in the form available in ***Annex 4*** (List of identified subcontractors), tenderers are required to give an indication of the proportion of the contract that they intend to subcontract, as well as to identify and describe briefly the envisaged contractual roles/tasks of subcontractors meeting any of these conditions (hereafter referred to as *identified subcontractors*):

- subcontractors on whose capacities the tenderer relies upon to fulfil the selection criteria as described under Section 3.2;
- subcontractors whose intended individual share of the contract, known at the time of submission, is above 20% .

Any such subcontractor must provide the tenderer with a commitment letter drawn up in the model attached in ***Annex 5.1*** and signed by its authorised representative.

☝ Each tenderer shall identify such subcontractors and provide the commitment letters with its tender. The information must be true and correct at the time of submitting the tender. Any changes or additions regarding the envisaged subcontractors after the deadline for submission of tenders must be justified to the contracting authority.

The above rules apply also where the economic operators, which will perform part of the contract on behalf of a successful tenderer, belong to the same economic/corporate group as the sole tenderer or a member of the group submitting the joint tender.

Changes concerning subcontractors identified in the tender (withdrawal/replacement of a subcontractor, additional subcontracting) during the procurement procedure (after the deadline for submission of tenders and before contract signature) require the prior written approval of the contracting authority subject to the following verifications:

- any new subcontractor is not subject to restrictive measures, has access to procurement if the rules on access to procurement apply also to subcontractors (see Section 2.2) and is not in an exclusion situation (see Section 3.1),
- the tenderer still fulfils the selection criteria and the new subcontractor fulfils the selection

criteria applicable to it individually, if any;

- the terms of the originally submitted tender are not altered substantially, i.e. all the tasks assigned to the former subcontractor are taken over by another involved entity, the change does not make the tender non-compliant with the tender specifications, and the evaluation of award criteria of the originally submitted tender is not modified.

Subcontracting to subcontractors identified in a tender that was accepted by the contracting authority and resulted in a signed contract, is considered authorised.

2.4.3. Entities (not subcontractors) on whose capacities the tenderer relies to fulfil the selection criteria

In order to fulfil the selection criteria a tenderer may also rely on the capacities of other entities (that are not subcontractors), regardless of the legal nature of the links it has with them. It must in that case prove that it will have at its disposal the resources necessary for the performance of the contract by producing a commitment letter in the model attached in ***Annex 5.2***, signed by the authorised representative of such an entity, and the supporting evidence that those other entities have the respective resources[14].

☝ The above rules apply also where the economic operators on whose capacities the tenderer relies to fulfil the selection criteria (that are not subcontractors) belong to the same economic/corporate group as the sole tenderer or a member of the group submitting the joint tender.

2.4.4. Rules common to subcontractors and entities (not subcontractors) on whose capacities the tenderer relies to fulfil the selection criteria

If a successful tenderer intends to rely on another entity to meet the minimum levels of economic and financial capacity, the contracting authority may require the entity to sign the contract or, alternatively, to provide a joint and several first-call financial guarantee for the performance of the contract.

With regard to technical and professional selection criteria, a tenderer may only rely on the capacities of other entities where the latter will perform the works or services for which these capacities are required, i.e. the latter will either assume the role of subcontractors or will fall within the exceptions listed in Section 2.4.2 and will then assume the role of entities (not subcontractors) on whose capacities the tenderer relies to fulfil the selection criteria.

☝ Relying on the capacities of other entities is only necessary when the capacity of the tenderer is not sufficient to fulfil the required minimum levels of capacity. Abstract commitments that other entities will put resources at the disposal of the tenderer will be disregarded.

[7] This does not apply to subcontractors on whose capacity the tenderer relies to fulfil the selection criteria – for these the documentation required for subcontractors must be provided.

3. EVALUATION AND AWARD

The evaluation of the tenders that comply with the submission conditions will consist of the following elements:

- Check if the tenderer is not subject to restrictive measures and has access to procurement (see Section 2.2);
- Verification of administrative compliance (if the tender is drawn up in one of the official EU languages and the required documents signed by duly authorised representative(s) of the tenderer);
- Verification of non-exclusion of tenderers on the basis of the exclusion criteria;
- Selection of tenderers on the basis of selection criteria;
- Verification of compliance with the minimum requirements specified in the procurement documents;
- Evaluation of tenders on the basis of the award criteria.

The contracting authority will evaluate the above mentioned elements in the order that it considers to be the most appropriate.

If the evaluation of one or more elements demonstrates that there are grounds for rejection, the tender will be rejected and will not be subjected to further full evaluation. The unsuccessful tenderers will be informed of the ground for rejection without being given feedback on the non-assessed content of their tenders. Only the tenderers for whom the verification of all elements did not reveal grounds for rejection can be awarded the contracts resulting from this call for tenders.

The evaluation will be based on the information and evidence contained in the tenders and, if applicable, on additional information and evidence provided at the request of the contracting authority during the procedure. If any of the declarations or information provided proves to be false, the contracting authority may impose administrative sanctions (exclusion or financial penalties) on the entity providing the false declarations/information.

For the purposes of the evaluation related to exclusion and selection criteria the contracting authority may also refer to publicly available information, in particular evidence that it can access on a national database free of charge.

3.1. Exclusion criteria

The objective of the exclusion criteria is to assess whether the tenderer is in any of the exclusion situations listed in Article 138(1) of the Financial Regulation.

Tenderers found to be in an exclusion situation will be rejected.

As evidence of non-exclusion, each tenderer[8] needs to submit with its tender a Declaration on Honour[9] in the model available in ***Annex 2***.[10] The declaration must be signed by an authorised representative of the entity providing the declaration. Where the declaration has been signed by hand,

[8] See Annex 1 which of the involved entities participating in a tender need to provide the Declaration on Honour.

[9] The European Single Procurement Document (ESPD) may not be used yet in calls for tenders of the European Commission.

[10] Unless the same declaration has already been submitted for the purposes of another award procedure of the European Commission, the situation has not changed, and the time elapsed since the issuing date of the declaration does not exceed one year.

the original does not need to be submitted to the contracting authority, but the latter reserves the right to request it from the tenderer at any time during the record-keeping period specified in Section 4.3.

The initial verification of non-exclusion of tenderers will be done on the basis of the submitted declarations and consultation of the European Union's Early Detection and Exclusion System.

At any time during the procurement procedure[12], the contracting authority may request the documents mentioned in the Declaration on Honour as supporting evidence on non-exclusion (the documentary evidence). It may also request information on natural or legal persons that are members of the administrative, management or supervisory body or that have powers of representation, decision or control, including legal and natural persons within the ownership and control structure and beneficial owners, and appropriate evidence that none of those persons are in one of the exclusion situations referred to in Section A point (1) (c) to (f) of the Declaration on Honour.

All tenderers are **invited to prepare in advance the documentary evidence**, since they may be requested to provide such evidence within a short deadline. In any event, the tenderers proposed by the evaluation committee for the award of the contracts will be requested to provide such evidence.

☝ If the tenderer does not provide valid documentary evidence within the deadlines set by the contracting authority, the latter reserves the right to reject the tender. In any event, in case a tenderer proposed for the award of the contract fails to comply with the above evidence requirement, its tender will be rejected, unless the tenderer can justify the failure on the grounds of material impossibility to provide such evidence.

Annex 1 specifies which of the involved entities participating in a tender need to provide the Declaration on Honour and, when requested by the contracting authority, the supporting evidence.

Please note that a request for evidence in no way implies that the tenderer has been successful.

3.2. Selection criteria

The objective of the selection criteria is to assess whether the tenderer has the legal, regulatory, economic, financial, technical and professional capacity to perform the contract.

The selection criteria for this call for tenders, including the minimum levels of capacity, the basis for assessment and the evidence required, are specified in the following subsections.

Tenders submitted by tenderers not meeting the minimum levels of capacity will be rejected.

When submitting its tender each tenderer shall declare on honour that it fulfils the selection criteria for the lot(s) for which it applies in this call for tenders. The model Declaration on Honour available in ***Annex 2*** shall be used.

The initial assessment of whether a tenderer fulfils the selection criteria will be done on the basis of

[12] The obligation to provide the supporting evidence will be waived in the following situations:
- if the same documents have already been provided in a previous award procedure of the European Commission, have been issued no more than one year before the date of their request by the contracting authority and are still valid at that date;
- if such evidence can be accessed by the contracting authority on a national database free of charge, in which case the economic operator shall provide the contracting authority with the internet address of the database and, if needed, the necessary identification data to retrieve the document;
- if there is a material impossibility to provide such evidence.

the submitted declaration(s).

The subsections below specify which selection criteria evidence must be provided with the tender or may be requested later, at any time during the procurement procedure, within a deadline given by the contracting authority[13].

The evidence must be provided in accordance with the applicable basis for assessment of each criterion: in case of a consolidated assessment – only by the involved entities who contribute to the fulfilment of the criterion, and in case of individual assessment – by each entity to whom the criterion applies individually.

In case not all selection criteria evidence is requested with the tender, all tenderers are **invited to prepare in advance the documentary evidence**, since they may be requested to provide such evidence within a short deadline. In any event, the tenderers proposed by the evaluation committee for the award of the contracts will be requested to provide such evidence.

☝ If the tenderer does not provide valid documentary evidence within the deadlines set by the contracting authority, the contracting authority reserves the right to reject the tender. In any event, in case a tenderer proposed for the award of the contract fails to comply with the above evidence requirement, its tender will be rejected, unless there is a ground for a waiver.

Please note that a request for evidence in no way implies that the tenderer has been successful.

3.2.1. Legal and regulatory capacity

Tenderers can be natural or legal persons. Tenderers are not obliged to take a specific legal form in order to submit their tenders.

Where tenderers submit a tender through an entity, which lacks legal personality (e.g., a branch), the compliance with the exclusion criteria, selection criteria, the rules on access to procurement as well as the absence of restrictive measures shall be assessed at the level of the tenderers.

Tenderers must prove that they have legal capacity to perform the contract and the regulatory capacity to pursue the professional activity necessary to carry out the work subject to this call for tenders.

The legal and regulatory capacity shall be proven by the evidence listed below:

- Proof of enrolment in a relevant trade or professional register

☝ The criterion applies to each member of the group individually.

The evidence of legal and regulatory capacity must be provided with the tender.

In addition, involved entities (see Section 2.4) and all subcontractors, including those which do not need to be identified in the tender (see Section 2.4.2), must not be subject to EU restrictive measures

[13] The obligation to provide the supporting evidence will be waived in the following situations:
- if the same documents have already been provided in a previous award procedure of the European Commission and are still up-to-date;
- if such evidence can be accessed by the contracting authority on a national database free of charge, in which case the economic operator shall provide the contracting authority with the internet address of the database and, if needed, the necessary identification data to retrieve the document.

adopted under Article 29 of the Treaty on the European Union (TEU) or Article 215 of the Treaty on the Functioning of the EU (TFEU)[14] that constitute a legal impediment to perform the contract. This requirement will be assessed by reference to the EU restrictive measures in force. Therefore, the tenderer is not required to submit any evidence of not being subject to EU restrictive measures.

3.2.2. Economic and financial capacity

Tenderers must comply with the following selection criteria in order to prove that they have the necessary economic and financial capacity to perform the contract.

Lots 1, 2, 3, 4, 5, 6

Criterion F1	
Minimum level of capacity	Average yearly turnover of the last two financial years above • EUR 700,000 for Lot 1 • EUR 700,000 for Lot 2 • EUR 500,000 for Lot 3 • EUR 400,000 for Lot 4 • EUR 400,000 for Lot 5 • EUR 500,000 for Lot 6
Basis for assessment	This criterion applies to the tenderer as a whole, i.e. a consolidated assessment of the combined capacities of all involved entities will be carried out.
Evidence	Copy of the profit and loss accounts for the last two years for which accounts have been closed from each concerned involved entity, or, failing that, appropriate statements from banks. The most recent year must have been closed within the last 18 months.

☝ The evidence of economic and financial capacity does not need to be provided with the tender but may be requested by the contracting authority at any time during the procedure.

3.2.3. Technical and professional capacity

☝ With regard to technical and professional selection criteria, a tenderer may only rely on the capacities of other entities where the latter will perform the works or services for which these capacities are required. The entity on whose capacity the tenderer relies will either assume the role of a subcontractor or fall within the exceptions listed in Section 2.4.2.

Tenderers must comply with the following selection criteria in order to prove that they have the necessary technical and professional capacity to perform the contract:

[14] Please note that the EU Official Journal contains the official list and, in case of conflict, its content prevails over that of the EU Sanctions Map.

Lots 1, 2, 3, 4, 5, 6

Criterion T1	
The tenderer must have relevant experience in and must be able to prove their capacity for (a) conducting research targeted studies, (b) elaborating and employing rigorous technical methodologies, (c) where relevant, conducting data analysis on a large scale and developing datasets, and (d) identifying and anticipating potential future trends, developments, and risks. This experience must be related to the specific domain of each lot they apply for: • CBRN risk modelling and evaluation for Lot 1; • Cyber Offence risk modelling and evaluations for Lot 2; • Loss of Control risk modelling and evaluation for Lot 3; • Harmful Manipulation risk modelling and evaluation for Lot 4; • Sociotechnical risk modelling and evaluation for Lot 5; • Agentic evaluation, software development, and infrastructure for Lot 6.	
Minimum level of capacity	At least 2 studies / reports / benchmarks/ recommendations of at least 10 pages published, co-published or contributed to in the last three years preceding the deadline for submission of tenders, in at least one of the above-mentioned fields.
Basis for assessment	This criterion applies to the tenderer as a whole, i.e. the consolidated assessment of combined capacities of all involved entities will be carried out.
Evidence	The tender must provide a list of studies or reports meeting the above-mentioned minimum level of capacity. This list will for each report be accompanied with: • the full report where possible, potentially redacted, and if not possible due to confidentiality, an abstract and the length of the report without references • details of their start and end date, and date of publication • details about the scope • the role of the contractor in the report • the amount invoiced where applicable, and where not applicable, an estimation of the number of working hours contributed by the contractor to the report.

Criterion T2
The tenderer must prove qualifications, knowledge and experience related to the specific domain of each lot they apply for: • CBRN risk modelling and evaluation for Lot 1; • Cyber Offence risk modelling and evaluations for Lot 2; • Loss of Control risk modelling and evaluation for Lot 3; • Harmful Manipulation risk modelling and evaluation for Lot 4; • Sociotechnical risk modelling and evaluation for Lot 5;

• Agentic evaluation, software development, and infrastructure for Lot 6. The tenderer must also prove their experience in the field of AI, particularly on the most advanced General-Purpose AI Models.	
Minimum level of capacity	Demonstrated track record in the specific field of the lot and GPAI evaluations. Management of at least one project similar in scope and complexity, or at least one collection of up to three projects cumulatively similar in scope and complexity, completed in the last three years preceding the deadline for submission of tenders, or currently being performed.
Basis for assessment	This criterion applies to the tenderer as a whole, i.e. the consolidated assessment of combined capacities of all involved entities will be carried out.
Evidence	A list of projects meeting the minimum level of capacity. The list shall include: (a) A short abstract of the project, (b) Start and end dates, (c) Total project amount and scope, (d) Role and amount invoiced. For ongoing projects, only the completed portion during the reference period will be taken into consideration.

Criterion T3	
The team delivering the services must prove their experience in the fields of • CBRN risk modelling and evaluation for Lot 1; • Cyber Offence risk modelling and evaluations for Lot 2; • Loss of Control risk modelling and evaluation for Lot 3; • Harmful Manipulation risk modelling and evaluation for Lot 4; • Sociotechnical risk modelling and evaluation for Lot 5; • Agentic evaluation, software development, and infrastructure for Lot 6; The team delivering must also prove their qualifications in the fields of AI, particularly on the most advanced General-Purpose AI Models.	
Minimum level of capacity	A multidisciplinary team with proven experience relevant to the topic of the lot, as well as with GPAI models evaluations, and regulatory compliance. A minimum of two years of experience among key personnel such as project manager, technical leads, or lead domain experts in managing projects of similar scope or of projects with reasonably reduced scope but with stellar results and proportionate impact. Sufficient team capacity (in FTE) to fulfil the objectives, in line with the detailed characteristics, and complete the deliverables set out in the technical specification of the relevant lot(s).
Basis for assessment	This criterion applies to the tenderer as a whole, i.e. the consolidated assessment of combined capacities of all involved

Criterion T3	
	entities will be carried out.
Evidence	The tender must provide supporting documents proving that the project team collectively meets the above requirements, including: Concise but informative CVs of each team member involved in executing the tasks, demonstrating their professional experience within the last five years.

When applying for multiple lots, tenderers must meet the minimum level of capacity of criterion T3 cumulatively, i.e. have sufficient team capacity to fulfil the cumulative criteria of all lots applied to.

☝ All of the above-specified evidence of technical and professional capacity must be provided with the tender.

***

☝ Involved entities (see Section 2.4) and all subcontractors, including those which do not need to be identified in the tender (see Section 2.4.2), must not be subject to professional conflicting interests which may negatively affect the contract performance. Where the contracting authority has established such conflicting interests, it may conclude that the tenderer or an involved entity does not possess the required professional capacity to perform the contract to an appropriate quality standard.

The presence of conflicting interests shall be examined during the evaluation phase based on the statements made through the Declarations on Honour and, where applicable, the commitment letters (***Annex 5.1 and Annex 5.2***).

Further details and obligations concerning professional conflicting interests are set out in the draft contract.

3.3. Compliance with the conditions for participation and minimum requirements specified in the procurement documents

By submitting a tender a tenderer commits to perform the contract in full compliance with the terms and conditions of the procurement documents for this call for tenders. Particular attention is drawn to the minimum requirements specified in Section 1.4 of these specifications and to the fact that tenders must comply with applicable data protection, environmental, social and labour law obligations established by Union law, national legislation, collective agreements or the international environmental, social and labour conventions listed in Annex X to Directive 2014/24/EU.

The minimum requirements shall be observed throughout the entire duration of the contract. Compliance with these requirements is mandatory and cannot be subject to any assumptions, limitations, conditions, or reservations on the part of a tenderer.

Tenderers must declare when submitting their tenders in eSubmission whether their tenders comply with the minimum requirements specified in the procurement documents.

☝ **Tenders that are not compliant with the applicable minimum requirements shall be rejected.**

3.4. Award criteria

The objective of the award criteria is to evaluate the tenders with a view to choosing the most economically advantageous tender.

Tenders will be evaluated on the basis of the following award criteria and their weighting:

1. Price - 35%

The price considered for evaluation will be the total price of the tender, covering all the requirements set out in the tender specifications.

2. Quality - 65%

The quality of the tender for each individual lot will be evaluated based on the following criteria:

Criteria	Maximum score	Threshold
1. Functional Characteristics **Sub-criterion 1.1:** Specific features and capabilities of the tender demonstrate subject-matter expertise for each of the field of lot specific risk analysis, modelling, and evaluation. Such as • CBRN risk modelling and evaluation for Lot 1; • Cyber Offence risk modelling and evaluations for Lot 2; • Loss of Control risk modelling and evaluation for Lot 3; • Harmful Manipulation risk modelling and evaluation for Lot 4; • Sociotechnical risk modelling and evaluation for Lot 5; • Agentic evaluation, software development, and infrastructure for Lot 6; **Sub-criterion 1.2:** Specific features and capabilities of the tender demonstrate subject-matter expertise in General-Purpose AI (GPAI) evaluations, including risk assessment of GPAI models. *All the sub-criteria above are of equal relative importance.*	30	15
2. Technical quality **Sub-criterion 2.1:** Clarity, credibility, quality and completeness of the tender (i.e. project description, and functionality of the approach).	40	20

Sub-criterion 2.2: Use of state-of-the-art methodologies, tools, and best practices, ensuring innovation, efficiency, and alignment with the latest industry standards. **Sub-criterion 2.3**: Adequacy of the quality control system applied to the service foreseen in the tender specifications (i.e. after-sale service and technical assistance). *All the sub-criteria above are of equal relative importance.*		
3. Team allocation and organization of the work **Sub-criterion 3.1**: Feasibility to meet the objectives specified in the tender specifications outlined by a work plan and timetable. **Sub-criterion 3.2:** Adequacy and appropriateness of the overall allocation of time and resources and to each task and deliverable. This includes specifying clearly the identity, roles, activities and responsibilities of consortium members and/or subcontractor(s) where applicable. *All the sub-criteria above are of equal relative importance.*	30	15
Maximum score 100/minimum 60	100	60

Tenders must score minimum 50% for each criterion, and minimum 60% in total. Tenders that do not reach the minimum quality levels will be rejected and will not be ranked.

3.5. Award (ranking of tenders)

Tenders for the lots for which the award method is best price-quality ratio shall be ranked according to the best price-quality ratio in accordance with the formula below:

$$\text{score for tender X} = \frac{\text{cheapest price}}{\text{price of tender X}} * 100 * 35\% + \text{total quality score (out of 100) for all award criteria of tender X} * 65\%$$

Should the outcome of the formula lead to two or more tenders with the same result, the tenderer who has been awarded the highest marks for quality will be deemed to be the most economically advantageous tender. This approach will continue to be applied to each of the award criteria in the descending order listed in below until a most economically advantageous tender can be determined:

- Quality of the proposed methodology and tools for performing the tasks
- Organisation of work and resources
- Quality control measures

☝ The contracts shall be awarded to the tenders ranked first for each lot, which comply with the minimum requirements specified in the procurement documents and are submitted by a tenderer not subject to restrictive measures, having access to procurement, not in an exclusion situation and

fulfilling the selection criteria.

☝ **Detection of abnormally low tenders**

Tenderers must be aware of Point 23 of Annex I to the Financial Regulation on abnormally low tenders and of the possibility for rejection of the tender based on it.

4. FORM AND CONTENT OF THE TENDER

4.1. Form of the tender: how to submit the tender?

Tenders are to be submitted via the eSubmission application according to the instructions laid down in the Invitation letter and the eSubmission Quick Guide available at the link below:

https://wikis.ec.europa.eu/display/FTPortal/Open+procedures_EN

☝ Make sure you prepare and submit your tender in eSubmission early enough to ensure it is received within the deadline for receipt indicated under Section 5.1.12 of the contract notice and/or on Funding & Tenders Portal (F&T Portal)[15].

4.2. Content of the tender: what documents to submit with the tender?

The documents to be submitted with the tender in eSubmission are listed in ***Annex 1***.

☝ Tenderers willing to submit tenders for more than one lot need to upload a separate technical and financial tender for each of the lots in which they are interested.

The following requirements apply to the technical and financial tender to be uploaded in eSubmission:

- *Technical tender*

The technical tender for any lot for which the award method best price-quality ratio must provide all the information needed to assess the compliance with Section 1.4 of these specifications and the award criteria. Tenders deviating from the minimum requirements or not covering all the requirements may be rejected on the basis of non-compliance and not evaluated further.

Tenderers are free to choose where the personal data will be processed or stored as long as they comply with the contractual obligations on data processing (Art.I.9.2 and Art. II.9) and, in particular, with the requirements for transfer of personal data to third countries and international organisations laid down in Chapter V of Regulation (EU) 2018/1725[115].

Tenderers must specify in their technical tender the location where the personal data will be processed and stored only where this location is outside the territory of the European Union or the European Economic Area. If no location is specified in the tender, the contracting authority will consider that the personal data will be processed and stored only within the territory of the European Union or the European Economic Area.

- *Financial tender*

A complete financial tender, including the breakdown of the price, needs to be submitted. For this purpose, the Financial Model in **Annex 6** shall be used.

The financial tender shall be:

[15] https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/home

[16] Regulation (EU) 2018/1725 of 23 October 2018 on the protection of natural persons with regard to the processing of personal data by the Union institutions, bodies, offices and agencies and on the free movement of such data, and repealing Regulation (EC) No 45/2001 and Decision No 1247/2002/EC, OJ L 295/39, 21.11.2018

- expressed in euros. Tenderers from countries outside the euro zone have to quote their prices in euro. The price quoted may not be revised in line with exchange rate movements. It is for the tenderer to bear the risks or the benefits deriving from any variation.
- quoted free of all duties, taxes and other charges, i.e. also free of VAT.

☝ The European Union Institutions are exempt from such charges in the EU under Articles 3 and 4 of the Protocol on the Privileges and Immunities of the European Union of 8 April 1965 annexed to the Treaty on the Functioning of the European Union. Exemption is granted to the Commission by the governments of the Member States, either through refunds upon presentation of documentary evidence or by direct exemption.

In case of doubt about the applicable VAT system, it is the tenderer's responsibility to contact its national authorities to clarify the way in which the European Union is exempt from VAT.

4.3. Signature policy: how can documents be signed?

Where a document needs to be signed, the signature must be either hand-written or, preferably, a qualified electronic signature (QES) as defined in Regulation (EU) No 910/2014 on electronic identification and trust services for electronic transactions in the internal market (the *eIDAS Regulation*).

Tenderers are strongly encouraged to sign with a QES[17] all documents requiring a signature and only exceptionally to sign such documents by hand as hand-written signatures lead to an additional administrative burden for both the tenderer and the contracting authority. The originals of any hand-signed documents (other than the contract) do not need to be submitted to the contracting authority but the tenderer must keep them for a period of five years starting from the notification of the outcome of the procedure or, where the tenderer has been awarded a contract resulting from this call for tenders and the contract has been signed, the payment of the balance.

All documents must be signed by the signatories (when they are individuals) or by their duly authorised representatives.

For the following documents, when signed by representatives, tenderers must provide evidence for the delegation of the authorisation to sign:

- The Declaration on Honour of the tenderer (in case of a joint tender – the Declarations on Honour of all group members);

- (in the case of a joint tender) the Agreement/Power(s) of attorney drawn up using the model attached in ***Annex 3***).

The delegation of the authorisation to sign on behalf of the signatories (including, in the case of proxy(-ies), the chain of authorisations) must be evidenced by appropriate written evidence (copy of the notice of appointment of the persons authorised to represent the legal entity in signing contracts (together or alone), or a copy of the publication of such appointment if the legislation which applies to signatory requires such publication or a power of attorney). A document that the contracting authority can access on a national database free of charge does not need to be submitted if the contracting authority is provided with the exact internet link and, if applicable, the necessary identification data to retrieve the document.

[17] See here how to apply a QES on a document exchanged with a European institution, body or agency.

4.4. Confidentiality of tenders: what information and under what conditions can be disclosed?

Once the contracting authority has opened a tender, it becomes its property and shall be treated confidentially, subject to the following:

- For the purposes of evaluating the tender and, if applicable, implementing the contract, performing audits, benchmarking, etc., the contracting authority is entitled to make available (any part of) the tender to its staff and the staff of other Union institutions, bodies and agencies, as well to other persons and entities working for the contracting authority or cooperating with it, including contractors or subcontractors and their staff, provided that they are bound by an obligation of confidentiality.

- After the signature of the award decision, tenderers, whose tenders were received in accordance with the submission modalities, who are not subject to restrictive measures, have access to procurement, who are not found to be in an exclusion situation referred to in Article 138(1) of the FR, who are not rejected under Article 143 of the FR, whose tenders are not found to be incompliant with the procurement documents, and who make a request in writing, will be notified of the name of the successful tenderer to whom the contract is awarded for the lot(s) for which the tenderer applied, the characteristics and relative advantages of the successful tender and its total financial tender amount. The contracting authority may decide to withhold certain information that it assesses as being confidential, in particular where its release would prejudice the legitimate commercial interests of economic operators or might distort fair competition between them. Such information may include, without being limited to, confidential aspects of tenders such as unit prices included in the financial tender, technical or trade secrets[116].

- The contracting authority may disclose the submitted tender in the context of a request for public access to documents, or in other cases where the applicable law requires its disclosure. Unless there is an overriding public interest in disclosure[19], the contracting authority may refuse to provide full access to the submitted tender, redacting the parts (if any) that contain confidential information, the disclosure of which would undermine the protection of commercial interests of the tenderer, including intellectual property.

☝ The contracting authority will disregard general statements that the whole tender or substantial parts of it contain confidential information. Tenderers need to mark clearly the information they consider confidential and explain why it may not be disclosed. The contracting authority reserves the right to make its own assessment of the confidential nature of any information contained in the tender.

[18] For the definition of trade secrets please see Article 2 (1) of Directive (EU) 2016/943 on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use and disclosure.

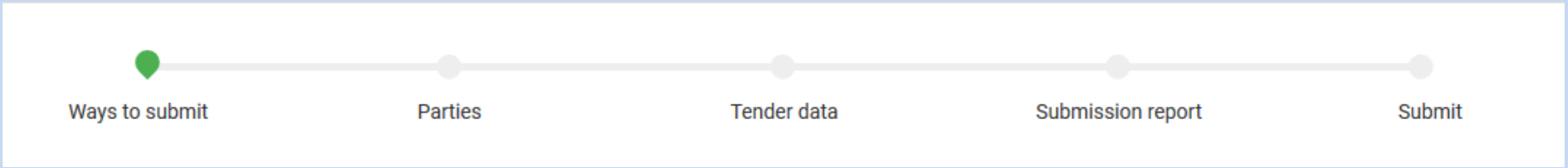
[19] See Article 4 (2) of the Regulation (EC) No 1049/2001 regarding public access to European Parliament, Council and Commission documents.

APPENDIX: LIST OF REFERENCES

Award criteria	See Section 3.4
Contracting authority	See Section 1.1
Entities on whose capacities the tenderer relies to fulfil the selection criteria	See Section 2.4.3
EU Validation services	See Section 2.3 EU Grants and Tenders Rules on Legal Entity Validation, LEAR appointment and Financial Capacity assessment
Exclusion criteria	See Section 3.1
Financial Regulation	Regulation (EU, Euratom) 2024/2509 of the European Parliament and of the Council of 23 September 2024 on the financial rules applicable to the general budget of the Union
Group leader	See Section 2.4.1
Group member	See Section 2.4.1
Identified subcontractors	See Section 2.4.2
Involved entities	See Section 2.4
Joint tender	See Section 2.4.1
Participant Register	See Section 2.3 https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/how-to-participate/participant-register
Selection criteria	See Section 3.2
Sole tenderer	See Section 2.4
Subcontracting/subcontractor	See Section 2.4.2
Treaties	The EU Treaties: https://europa.eu/european-union/law/treaties_en

ANNEXES

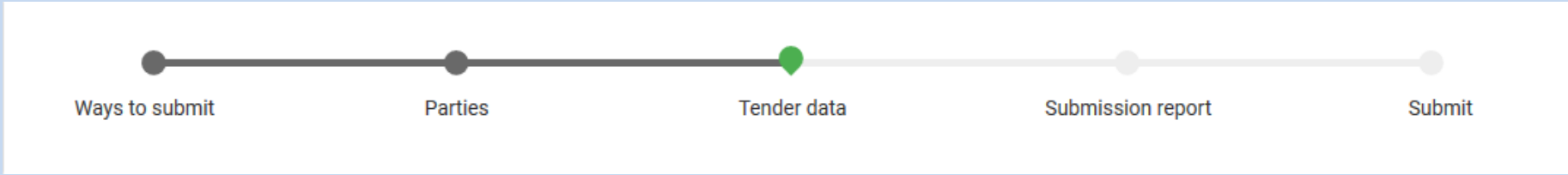
Annex 1. List of documents to be submitted with the tender or during the procedure

Description	Sole tenderer	Joint tender		Identified Subcontractor	Entity on whose capacity is being relied (that is not subcontractor)	When and where to submit the document?	Instructions for uploading in eSubmission (if applicable)	
		Group leader	Group member				How to name the file?	Where to upload?
1. Identification and information about the tenderer. *eSubmission view* 								
Declaration on Honour on Exclusion and Selection Criteria (see Section 3.1) *model in Annex 2*	☒	☒	☒	☒	☒	With the tender in eSubmission	'Declaration on Honour'	With the concerned entity under 'Parties' →'Identification of the participant' →'Attachments'→'Declaration on Honour'. For entities that are not subcontractors and on whose capacity the tenderer relies to fulfil the selection criteria, the document must be uploaded in the section of the sole tenderer or group leader: →'Identification of the participant' →'Attachments'→'Other documents'.
Evidence that the person	☒	☒	☒			With the tender	'Authorisation to	With the concerned entity

signing the documents is **an authorised representative** of the entity[217] (see Section 4.3)						in eSubmission	sign documents'	under 'Parties' →'Identification of the participant' →'Attachments'→'Other documents'.
Agreement/Power of attorney (see Section 2.4.1) *model in Annex 3*		☒	☒			With the tender in eSubmission	'Agreement Power of attorney'	In the group leader's section under 'Parties' →'Identification of the participant' →'Attachments'→'Other documents'.
List of identified subcontractors (see Section 2.4.2) *model in Annex 4*	☒	☒				With the tender in eSubmission	'List of identified subcontractors'	In the sole tenderer's or the group leader's section under 'Parties' →'Identification of the participant' →'Attachments'→'Other documents'.
Commitment letter (see Section 2.4.2 and 2.4.3)				☒ *(model in Annex 5.1)*	☒ *(model in Annex 5.2)*	With the tender in eSubmission	'Commitment letter'	With the concerned entity under 'Parties' →'Identification of the participant' →'Attachments'→'Other documents'.
Evidence of non-exclusion (see Section 3.1)	☒	☒	☒	☒	☒	Tenderers (sole tenderers/all group members in case of a joint tender) must provide the evidence when requested by the contracting authority and,	n.a.	n.a.

[20] A document that the contracting authority can access on a national database free of charge does not need to be submitted if the contracting authority is provided with the exact internet link and, if applicable, the necessary identification data to retrieve the document.

						in any event, if a tenderer is successful, before the award of the contract. Subcontractors and entities on whose capacity a tenderer relies to fulfil the selection criteria must provide the evidence only upon request by the contracting authority.		
Evidence of legal existence and status (see Section 2.3)	☒	☒	☒			Only upon request by *the EU Validation services* At any time during the procedure In the Participant Register	n.a.	n.a.
Evidence of legal capacity (see Section 3.2.1)	☒	☒	☒			With the tender in eSubmission	No specific requirements how to name the file(s).	With the concerned entity under 'Parties' →'Identification of the participant' →'Attachments'→'Legal and regulatory capacity'.
Evidence of economic and financial capacity F1 (see Section 3.2.2)	**The documents must be provided** **only by the involved entities** **which contribute to reaching the minimum capacity level** **for criterion F1**					Only upon request by the contracting authority At any time during the procedure	n.a.	n.a.
Evidence of technical and professional capacity T1-T2 (see Section 3.2.3)	**The documents must be provided** **only by the involved entities** **which contribute to reaching the minimum capacity level** **for criterion T1-T2**					With the tender in eSubmission	'Project reference No.1' 'Project reference No.2' ….	With the group leader or the sole tenderer under 'Parties' →'Identification of the participant' →'Attachments'→'Technical and professional capacity'.

Evidence of technical and professional capacity T3 (see Section 3.2.3)	**The documents must be provided** **only by the involved entities** **which contribute to reaching the minimum capacity level for criterion T3**						'CV No.1' 'CV No.2' ….	With the group leader or the sole tenderer under 'Parties' →'Identification of the participant' →'Attachments'→'Technical and professional capacity'.
Annex 7. Administrative identification form	☒	☒	☒			With the tender in eSubmission	'Administrative Identification form'	With the concerned entity under 'Parties' →'Identification of the participant' →'Attachments'→'Other'.
2. Tender data. *eSubmission view*  ***Failure to upload the following documents in eSubmission will lead to rejection of the tender.***								
Technical tender (see Section 4.2)	☒	☒				With the tender in eSubmission	'Technical tender'	Under section 'Tender Data' →'Technical tender'
Financial tender (see Section 4.2) *model in Annex 6*	☒	☒				With the tender in eSubmission	'Financial tender'	Under 'Tender Data' →'Financial tender'

Annex 2. Declaration on Honour on exclusion and selection criteria

Annex 2 is published as a separate document

Annex 3. Agreement/Power of attorney

Call for tenders **EC-CNECT/2025/OP/0032** – **ARTIFICIAL INTELLIGENCE ACT: TECHNICAL ASSISTANCE FOR AI SAFETY – Lot [*lot number*]**

AGREEMENT/POWER OF ATTORNEY

The undersigned:

[- Signatory 1 (Name, Function, Legal entity name, Registered address, VAT Number)]

- Signatory 2 (Name, Function, Legal entity name, Registered address, VAT Number)

- …

- Signatory N (Name, Function, Legal entity name, Registered address, VAT Number)]

having the legal capacity required to act on behalf of the entities they represent,

HEREBY AGREE TO THE FOLLOWING:

1) To submit a joint tender (the tender) as members of a group of tenderers (the group), constituted by ***[Insert names of Legal entity 1, Legal entity 2, ... Legal entity N – the name of the group leader must be included here!]*** (the group members), and led by ***[Insert name of Legal entity 1]*** (the group leader), in accordance with the conditions of the procurement documents and the terms of the tender to which this Agreement/Power of attorney is attached.
2) If the contracting authority awards a contract resulting from this call for tenders (the contract) to the group on the basis of the tender to which this Agreement/Power of attorney is attached, all group members (including the group leader) shall be considered parties to the contract in accordance with the following conditions:
   (a) All group members (including the group leader) shall be jointly and severally liable towards the contracting authority for the performance of the contract.
   (b) All group members (including the group leader) shall comply with the terms and conditions of the contract and ensure the proper delivery of their respective share of the services and/or supplies subject to the contract.
3) Payments by the contracting authority related to the services and/or supplies subject to the contract shall be made through the bank account of the group leader indicated in the contract.
4) The group members grant to the group leader all the necessary powers to act on their behalf in the submission of the tender and the conclusion of the contract, including:
   (a) The group leader shall submit the tender on its own behalf and on behalf of the other group members and indicate in the "Contact Person" section in eSubmission the name and e-mail address of an individual as a single point of contact authorised to communicate officially with the contracting authority in connection with the submitted tender on behalf of all group members, including in connection with all relevant questions, clarification requests, notifications, etc., that may be received during the evaluation, award and until the contract signature.
   (b) The group leader shall sign any contractual documents — including the contract and amendments thereto — and shall warrant the submission of any invoices related to the

performance of the contract on behalf of all group members.

(c) The group leader shall act as a single contact point with the contracting authority in the delivery of the services and/or supplies subject to the contract. It shall coordinate the delivery of the services and/or supplies by the group to the contracting authority, and shall see to a proper administration of the contract.

This Agreement/Power of attorney may be executed in counterparts, each of which shall be deemed to be an original, but all of which, taken together, shall constitute one and the same document.

Any modification to the present Agreement/Power of attorney shall be subject to the contracting authority's express approval. This Agreement/Power of attorney shall expire when all the contractual obligations of the group have ceased to exist. The parties cannot terminate it before that date without the contracting authority's consent.

Name ***Function*** ***Name of the legal entity***	***Name*** ***Function*** ***Name of the legal entity***
signature[s]: _____	***signature[s]:*** _____
Done at ..... on ***.....***	***Done at ..... on*** ***.....***
Name ***Function*** ***Name of the legal entity***	***Name*** ***Function*** ***Name of the legal entity***
signature[s]: _____	***signature[s]:*** _____
Done at ..... on ***.....***	***Done at ..... on*** ***.....***

Annex 4. List of identified subcontractors and proportion of subcontracting

EC-CNECT/2025/OP/0032 – ARTIFICIAL INTELLIGENCE ACT: TECHNICAL ASSISTANCE FOR AI SAFETY – Lot [*lot number*]

Identification details	Roles/tasks during contract execution	Proportion of subcontracting (% of contract volume)
[Full official name of the identified subcontractor, registered address, statutory registration number, VAT registration number]		
[Full official name of the identified subcontractor, registered address, statutory registration number, VAT registration number]		
[REPEAT AS MANY TIMES AS THE NUMBER OF IDENTIFIED SUBCONTRACTORS]		
Other subcontractors that do not need to be identified under Section 2.4.2[21]		
	TOTAL % of subcontracting	0,00%

[21] For this category of subcontractors, please provide in a general manner their intended roles/tasks during contract execution, as well as the aggregated % of contract volume for all non-identified subcontractors.

Annex 5.1. Commitment letter by an identified subcontractor

[Letterhead, if any]

EUROPEAN COMMISSION

Call for tenders Ref. EC-CNECT/2025/OP/0032

Attn:

[Insert date]

Commitment letter by identified subcontractor

I, the undersigned,

Name:

Function:

Legal entity:

Registered address:

VAT Number:

having the legal capacity required to act on behalf of ***[insert name of the entity]***, hereby confirm that the latter agrees to participate as subcontractor in the tender of *[insert name of the tenderer]* for the call for tenders **EC-CNECT/2025/OP/0032 – ARTIFICIAL INTELLIGENCE ACT: TECHNICAL ASSISTANCE FOR AI SAFETY - Lot [*lot number*]**.

In the event that the tender of the aforementioned tenderer is successful, ***[insert name of the subcontractor]*** commits itself to make available the resources necessary for performance of the contract as a subcontractor and to carry out the services that will be subcontracted to it in compliance with the terms of the contract. It further declares that it is not subject to conflicting interests, which may negatively affect the contract performance, and that it accepts the terms of the procurement documents for the above call for tenders, in particular the contractual provisions related to checks and audits.

Done at:

Name:

Position:

Signature:

Annex 5.2. Commitment letter by an entity on whose capacities is being relied

[Letterhead, if any]

EUROPEAN COMMISSION

Call for tenders Ref. EC-CNECT/2025/OP/0032

Attn:

[Insert date]

Commitment letter by an entity on whose capacity is being relied

I, the undersigned,

Name:

Function:

Legal entity:

Registered address:

VAT Number:

having the legal capacity required to act on behalf of ***[insert name of the entity]***, hereby confirm that the latter **authorises the *[insert name of the tenderer]* to rely on its [financial and economic capacity] [technical and professional capacity] in order to meet the minimum levels** required for the call for tenders **EC-CNECT/2025/OP/0032 – ARTIFICIAL INTELLIGENCE ACT: TECHNICAL ASSISTANCE FOR AI SAFETY - Lot [*lot number*]**.

In the event that the tender of the aforementioned tenderer is successful, ***[insert name of the entity]*** commits itself to make available the resources necessary for performance of the contract. It further declares that it is not subject to conflicting interests which may negatively affect the contract performance, and that it accepts the terms of the procurement documents for the above call for tenders, in particular the contractual provisions related to checks and audits.

Done at:

Name:

Position:

Signature:

Annex 6. Financial tender form

Annex 6 is published as a separate document

Annex 7. Administrative identification form
EC-CNECT/2025/OP/0032

TENDERER'S ID	
Name	
Legal form	
Date of registration	
Country of registration	
Registration number	
VAT number	
Address of registered office	
Contact address (if different)	
URL	
Yes / No	The tenderer is Small or Medium Size Enterprise in accordance with Commission Recommendation 2003/361/EC
Bank account (**lead partner only**) Name of bank: Full address of branch: Exact denomination of account holder: IBAN code:	
AUTHORISED REPRESENTATIVE(S)[1] [name and position]	
CONTACT PERSON	
Name	
Forename	
Position	
Telephone	
Fax	
Email	
DECLARATION BY THE AUTHORISED REPRESENTATIVE(S): I, the undersigned, certify that the information given in this tender is correct and that the tender is valid.	

Place and date:
Name (in capital letters) and signature:

1 Please include the names of the legal representative(s) whose contract signature is required in accordance with the statutes of the organisation and the official document to be provided as required in Part 2 under Section 2.3