Submission on behalf of:
Center for Data Innovation

Center for Data Innovation Submission to the Multi-stakeholder Consultation FUTURE-PROOF AI ACT: TRUSTWORTHY GENERAL-PURPOSE AI

On behalf of the Center for Data Innovation, we are pleased to submit this response to the EU AI Office's ("the AIO") call for submissions to its first consultation on the AI Act's ("the AIA") Code of Practice ("the Code").

The Center for Data Innovation studies the intersection of data, technology, and public policy. Its mission is to formulate and promote pragmatic public policies designed to maximize the benefits of data-driven innovation in the public and private sectors. It educates policymakers and the public about the opportunities and challenges associated with data, as well as technology trends such as open data, artificial intelligence, and the Internet of Things. The Center is part of the Information Technology and Innovation Foundation (ITIF), a nonprofit, nonpartisan think tank.

EXECUTIVE SUMMARY

The AIA came into law on 12th July 2024, triggering several actions by the AIO to ensure streamlined compliance with the new rules. These rules place varying degrees of responsibility on stakeholders across the AI value chain depending on the risk categorisation of the AI system. One of the first actions the AIO needs to take is to develop the Code, which details separate AIA rules for general-purpose AI (GPAI) providers with systemic risks and will apply twelve months after the AIA enters into force. This consultation on the Code evaluates four key sections: transparency and copyright-related provisions; risk identification and assessment measures for systemic risks; risk mitigation measures for systemic risks, and internal risk management and governance for GPAI model providers. We put forward six key recommendations to support AI innovation and adoption within the new framework.

Transparency and copyright-related provisions:

1. The AIO should establish different levels of disclosure for different audiences across the AI value chain to effectively operationalise the principle of transparency; and
2. The AIO should create the Code with consideration of existing sector-specific legislation that may already fulfil the requirements of the AI Act.

Risk identification and assessment measures for systemic risks:

3. The AIO should use the work of the international AI community to establish risk domains and assessment measures; and
4. The AIO should segregate actual risk from speculative risk when establishing GPAI provider obligations to curtail systemic risk.

General Considerations:

5. The AIO should explicitly state that the Code's core focus is to promote AI innovation and adoption; and
6. The AIO should develop the Code iteratively and root it in technical feasibility.

Working Group 1 - Transparency and copyright-related rules

The first Working Group focuses on detailing out documentation to downstream providers and the AI Office on the basis of Annexes XI and XII to the AI Act, policies to be put in place to comply with Union law on copyright and related rights, and making publicly available a summary about the training content.

Specific Objective	Measure	Key Performance Indicator (KPI)	Rationale / Justification
1. The AIO should establish different levels of disclosure for different audiences across the AI value chain to effectively operationalise the principle of transparency	See free format text below.	See free format text below.	See free format text below.
2. The AIO should create the Code with consideration of existing sector-specific legislation that may already fulfil the requirements of the AI Act	See free format text below.	See free format text below.	See free format text below.

If you prefer to provide **input in a free-form format**, please enter your comments below. This section is ideal for providing insights that may not be captured by the structured input above or provide any contextual information that may be relevant to this Working Group.

THE AIO SHOULD ESTABLISH DIFFERENT LEVELS OF DISCLOSURE FOR DIFFERENT AUDIENCES ACROSS THE AI VALUE CHAIN TO EFFECTIVELY OPERATIONALISE THE PRINCIPLE OF TRANSPARENCY

Transparency of different aspects of the model development and deployment process fosters accountability, facilitates knowledge sharing that informs decision-making, and builds trust among stakeholders. However, the AIO should recognise that transparency is not a one-size-fits-all approach. Different audiences will require varying levels of disclosure based on their roles and needs within the AI value chain.

The AIO should account for this spectrum of transparency in the Code. The Code should clearly outline the knowledge-sharing requirements for GPAI model providers, AI system providers, the AIO, national competent authorities, and downstream users who interact with both the broader AI system and specific models. By acknowledging the diverse needs of these stakeholders, the AIO can create a more effective framework that promotes clarity without overwhelming recipients with unnecessary information.

For example, model cards can be particularly useful for downstream AI system providers so they can better understand the model's development process and limitations. Model cards typically include training parameters, training procedures, and evaluation metrics, and stated risks and limitations can guide the limits AI system providers place on their own products for general use. The AIO however should not make this level of disclosure necessary for the general public, as doing so would not serve the purpose of fostering accountability, informed decision-making or trust. It would also open the door to concerns around trade secrets, dual-use capabilities, and competition, particularly if a disclosure requirement would include publishing training data. Such data is of little use to the majority of the general public, who do not have the level of subject matter expertise to appropriately digest or understand the significance of the data for the model's outputs.

The AIO should work with broader stakeholders, including civil society and specific interest groups to understand what information different stakeholders require, including the general public, in order to foster trust and encourage AI adoption. The AIO should also evaluate the minimum amount of information necessary to achieve this goal. For instance, the information exchanged between GPAI model providers and AI system providers should be adequate to comply with the Code and promote AI innovation and deployment. Additionally, the relevance of the information shared should be tailored to the specific needs of the intended audience. Stakeholders in the AI value chain should have the autonomy to determine what information is useful, while the AIO can guide this tailoring process to ensure that the shared information aligns with the overarching goals of the Code. This approach allows for a more nuanced understanding of transparency, ensuring that stakeholders receive the right amount of information at the right time. Finally, the AIO should work with industry to strike the right balance between transparency and the protection of sensitive, competitive information.

Therefore, the AIO should adjust the level of disclosure required between different stakeholders to ensure the right level of information is accessed at the right time and serves the purpose that the principle of transparency seeks to achieve.

THE AIO SHOULD CREATE THE CODE WITH CONSIDERATION OF EXISTING SECTOR-SPECIFIC LEGISLATION THAT MAY ALREADY FULFIL THE REQUIREMENTS OF THE AI ACT

The AIA does not exist in a vacuum. It operates within a broader regulatory framework that encompasses various sector-specific legislation. The AIO should recognise this interconnectedness when determining compliance requirements for stakeholders. The AIO should carefully navigate the existing regulatory landscape to ensure that the Code complements, rather than conflicts with, pre-existing regulations. This awareness will help prevent redundant compliance obligations that could overwhelm stakeholders.

The AIO should avoid contradictions, duplications, or excessive burdens related to compliance already covered by existing legislation. The complexities of the EU's digital regulatory landscape mean that businesses will find themselves facing a myriad of compliance requirements. If these requirements become overly burdensome, they will disincentivise businesses from entering, remaining, or scaling within the EU market, ultimately harming the digital economy. Therefore, it is crucial for the AIO to streamline compliance processes to maintain a competitive environment.
To achieve this, the AIO's approach to developing the Code should be cognisant of the existing regulatory landscape to effectively support stakeholders within the AI value chain. The AIO should conduct thorough analyses of existing sector regulations to identify overlaps and gaps in compliance requirements. Understanding these pre-existing frameworks will allow the AIO to adapt the Code in a manner that either fits within the existing regulatory environment or streamlines its provisions. The goal should be to not subject stakeholders across the AI value chain to more than the necessary regulatory obligations. This [what?] will leave room for businesses to voluntarily commit to exceeding the requirements, which may empower providers to take proactive steps in enhancing their compliance and ethical standards, fostering a culture of responsibility and trust within the AI community.

Union copyright law is one example where there is a risk of contradictory, duplicative, or excessive regulation and the AIO should use its position to avoid this risk. The AIO should take a proactive role in clarifying how stakeholders across the AI value chain can comply with existing Union copyright law, particularly Article 4 of the Directive on Copyright in the Digital Single Market (DSM), which relates to the Text and Data Mining (TDM) exception. The AIO should communicate that the Code will not impose additional burdens on providers; rather, the Code should focus on how providers can adhere to these copyright requirements effectively. Rather than mandating specific technologies or processes, the Code should emphasise the importance of utilising current best practices to ensure compliance with the TDM exception. This approach will not only provide flexibility for both the AIO and other stakeholders to adapt as new best practices emerge, but also allow for adjustments in response to any amendments or evolving policy discussions surrounding copyright and AI model development. By fostering clarity in this area, the AIO can support innovation while ensuring that stakeholders respect existing intellectual property rights.

Considerations for this Working Group

You can use this space to elaborate on your viewpoint on state-of-the-art practices for general-purpose AI models in the context of transparency and copyright-related rules. Take the opportunity to provide further context, clarify your assumptions, outline the principles that underlie your approach and the key considerations that inform your position.

Working Group 2 - Risk identification and assessment measures for systemic risks

The Code of Practice should help to establish a risk taxonomy of the type and nature of the systemic risks at Union level, including their sources. The second Working Group will focus on detailing the risk taxonomy based on a proposal by the AI Office and identifying and detailing relevant technical risk assessment measures, including model evaluation and adversarial testing.

Specific Objective	Measure	Key Performance Indicator (KPI)	Rationale / Justification
1. The AIO should use the work of the international AI community to establish risk domains and assessment measures	See free format text below.	See free format text below.	See free format text below.
2. The AIO should segregate actual risk from speculative risk when establishing GPAI provider obligations to curtail systemic risk	See free format text below.	See free format text below.	See free format text below.

If you prefer to provide **input in a free-form format**, please enter your comments below. This section is ideal for providing insights that may not be captured by the structured input above or provide any contextual information that may be relevant to this Working Group.

THE AIO SHOULD USE THE WORK OF THE INTERNATIONAL AI COMMUNITY TO ESTABLISH RISK DOMAINS AND ASSESSMENT MEASURES

To effectively enhance AI safety and governance, the Code should align with internationally agreed-upon risks and safety testing protocols, leveraging the deepening network of the AI Safety Institutes, standards-setting bodies, and international organisations. In pursuit of international commitments, the AIO should ensure that its classification of system risks and associated testing measures corresponds with the efforts of the global community.

Promoting standardisation in risk classification and testing is another key benefit of aligning the Code with international frameworks. By adopting consistent practices, the AIO can facilitate a smoother compliance process for stakeholders, reducing confusion and enhancing clarity in regulatory expectations. This standardisation will streamline compliance for existing EU entities and make it easier for non-EU firms to operate within the European market. Fostering a competitive environment in Europe is crucial for attracting global talent and businesses. By aligning with international safety and governance standards, the AIO can position Europe as an attractive destination for AI innovation. This approach simplifies the regulatory landscape, encouraging international firms to engage with the EU market without facing significant barriers.

Moreover, aligning the Code with international efforts serves as a practical solution to address the challenges of limited resources typically experienced at government level. The ability of the AIO to leverage established international standards can mitigate the difficulties associated with recruiting top researchers and experts in AI safety. By collaborating with existing institutions and frameworks, the AIO can enhance its effectiveness while maximising the impact of its regulatory efforts.

THE AIO SHOULD SEGREGATE ACTUAL RISK FROM SPECULATIVE RISK WHEN ESTABLISHING GPAI PROVIDER OBLIGATIONS TO CURTAIL SYSTEMIC RISK

The AIO should prioritise the segregation of actual risk from speculative risk when establishing obligations for GPAI providers to effectively mitigate systemic risk. While it is essential to conduct research into potential future risks associated with AI technologies, the obligations imposed on GPAI providers should focus primarily on addressing known and observed risks. This targeted approach ensures that resources and efforts are directed towards tangible threats that have already been identified and documented.

To achieve this, the Code should explicitly differentiate between speculative and observed risks, placing a greater emphasis on the latter. This distinction will help stakeholders prioritise their safety measures and compliance efforts based on the most pressing challenges facing the AI landscape.

Focussing on observed risks will also strengthen the Code's role as a practical mechanism for enhancing AI safety, rather than merely serving as a hypothetical framework for forecasting future risks. This practical orientation will improve the effectiveness of compliance measures and foster greater trust among stakeholders, who will see that the regulations are grounded in reality and geared towards immediate safety concerns.

Considerations for this Working Group

You can use this space to elaborate on your viewpoint on state-of-the-art practices for general-purpose AI models in the context of transparency and copyright-related rules. Take the opportunity to provide further context, clarify your assumptions, outline the principles that underlie your approach and the key considerations that inform your position.

Working Group 3 - Risk mitigation measures for systemic risks

The Code of Practice should be focused on specific risk assessment and mitigation measures. The third Working Group will focus on Identifying and detailing relevant technical risk mitigation measures, including cybersecurity protection for the general-purpose AI model and the physical infrastructure of the model.

Specific Objective	Measure	Key Performance Indicator (KPI)	Rationale / Justification
1. N/A	N/A	N/A	N/A

If you prefer to provide **input in a free-form format**, please enter your comments below. This section is ideal for providing insights that may not be captured by the structured input above or provide any contextual information that may be relevant to this Working Group.

Considerations for this Working Group

You can use this space to elaborate on your viewpoint on state-of-the-art practices for general-purpose AI models in the context of transparency and copyright-related rules. Take the opportunity to provide further context, clarify your assumptions, outline the principles that underlie your approach and the key considerations that inform your position.

Working Group 4 - Internal risk management and governance for GPAI providers:

The Code of Practice should also be focused on specific risk assessment and mitigation measures. The fourth Working Group will focus on identifying and detailing policies and procedures to operationalise risk management in internal governance of general-purpose AI model providers, including keeping track of, documenting, and reporting serious incidents and possible corrective measures.

Specific Objective	Measure	Key Performance Indicator (KPI)	Rationale / Justification
2. N/A	N/A	N/A	N/A

If you prefer to provide **input in a free-form format**, please enter your comments below. This section is ideal for providing insights that may not be captured by the structured input above or provide any contextual information that may be relevant to this Working Group.

Considerations for this Working Group

You can use this space to elaborate on your viewpoint on state-of-the-art practices for general-purpose AI models in the context of transparency and copyright-related rules. Take the opportunity to provide further context, clarify your assumptions, outline the principles that underlie your approach and the key considerations that inform your position.

General Considerations for the drawing-up of the Code of Practice

Please provide any general considerations that should be taken into account during the drawing-up of the first Code of Practice for providers of general-purpose AI models.

THE AIO SHOULD EXPLICITLY STATE THAT THE CODE'S CORE FOCUS IS TO PROMOTE AI INNOVATION AND ADOPTION

The core focus of the Code should be to promote AI innovation and adoption across the EU, and the AIO should reflect this goal in the Code and the AIO's broader decision-making. The biggest challenge for the AIO is that both the technical and policy landscapes are quickly evolving, which means there are many unanswered questions for both regulators and industry. To manage this uncertainty, the AIO should maintain a flexible, outcome-driven approach to AI governance, rooted in a single overarching purpose that can withstand an evolving technical and policy landscape.

THE AIO SHOULD DEVELOP THE CODE ITERATIVELY AND ROOTED IN TECHNICAL FEASIBILITY

The AIO should adopt an iterative development process when creating the Code, grounded in technical feasibility. This approach, commonly used in the technology industry, focuses on establishing a minimum viable product (MVP) that can be progressively refined through successive cycles. Developers typically begin by identifying the essential core functionalities of a product, and with each subsequent release, they broaden and enhance the MVP based on user feedback and technological advancements. The AIO should similarly apply this methodology to the Code, ensuring that it evolves in a practical and achievable manner.

The AIO should first identify the core obligations that are currently attainable by industry stakeholders. By establishing this baseline of technically feasible requirements, the AIO can set realistic expectations aligned with current best practices and avoid the pitfalls of setting unrealistic expectations that could lead to frustration and non-compliance among stakeholders. This approach will help reduce the compliance burdens on industry players, providing them with greater certainty while allowing the Code to incrementally adjust and become more fit-for-purpose as research, technology, and policy evolve.

Moreover, the AIO should be careful not to include unrealistic expectations in the Code by trying to satisfy the interests of all stakeholders in this consultation. To effectively address varying perspectives, it is essential for the AIO to first assess what is currently possible in the field and compare this with the practices already being implemented by industry. This careful examination will help in crafting regulations that are not only aspirational but also grounded in the realities of technological capabilities.

By taking an iterative, technically feasible approach, the AIO can foster a regulatory environment that encourages innovation while ensuring that safety and compliance are not compromised. This strategy will ultimately lead to a more adaptable and responsive Code that reflects the evolving landscape of AI technologies and their applications.