Centre for Information Policy Leadership
HUNTON ANDREWS KURTH

2 October 2024

Allocating Controllership across the Generative AI Supply Chain

Consultation Response

Centre for Information Policy Leadership (CIPL)

Centre for Information Policy Leadership
HUNTON ANDREWS KURTH

2 October 2024

Response by the Centre for Information Policy Leadership to the Information Commissioner's Office's Fifth Consultation on Allocating Controllership across the Generative AI Supply Chain

Submitted October 2, 2024

The Centre for Information Policy Leadership (CIPL) welcomes the opportunity to respond to the Information Commissioner's Office's (ICO) fifth consultation on allocating controllership across the AI value chain.

For more than 20 years, CIPL has been a thought leader on organisational accountability and a risk-based approach as key building blocks of smart regulation, responsible governance, and use of data, including the accountable development and deployment of artificial intelligence (AI). CIPL's *Ten Recommendations for Global Regulation*[1] proposes a layered, three-tiered approach to AI regulation that would protect fundamental human rights and minimise the potential risks of harm to both individuals and society while enabling the responsible development and deployment of AI. Our recent report, *Building Accountable AI Programs: Mapping Emerging Best Practices to the CIPL Accountability Framework*[2] documents, through the lens of CIPL's Accountability Framework, best practices and case studies that reflect how 20 leading organisations are responsibly developing and deploying AI. Based on CIPL's independent research and observations, we provide input to the ICO public consultation below.

CIPL agrees with the ICO's assertion that the genAI supply chain is growing in complexity. CIPL therefore supports the ICO developing guidance that provides clarity to developers, deployers, and other actors that will support innovation and privacy protection in relation to this transformative technology. Such guidance should maintain a fact-based approach to the allocation of controllership, to ensure that it reflects the realities of genAI technologies. Joint controllership should not be the starting assumption.

CIPL welcomes the ICO's consideration of the differing qualities of closed-access and open-access models which could affect the allocation of controllership. However, the ICO should exercise caution in making any default assumptions. As the ICO rightfully states, "[w]hether an organisation is a controller, joint controller or processor is not necessarily determined by a contract". In equal measure, controllership is not necessarily determined by whether the model is closed- or open-source.

Regardless of whether a model is open- or closed-source, there will be an initial developer who creates the genAI model. The model will then be deployed either by the developer directly or by a separate deployer or deployers. CIPL agrees that the developer of the AI model will likely be a controller for the processing activities which occur during the training and development of the AI model. Subsequent

---

[1] CIPL, "Ten Recommendations for Global AI Regulation", October 2023, https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_ten_recommendations_global_ai_regulation_oct2023.pdf.

[2] CIPL, "Building Accountable AI Programs: Mapping Emerging Best Practices to the CIPL Accountability Framework", February 2024, https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_building_accountable_ai_programs_23_feb_2024.pdf.

controllership during the deployment phase should then be determined by a factual assessment of the level of control that entities involved downstream exert over the purposes and means of the processing. Thus, in many instances, the developer of an AI model might act as a controller regarding the processing of training data, and potentially as a processor for input and output data provided by the deployer or user – similar to other algorithms, which are included in software or services provided "as a service". To that end, CIPL welcomes the ICO's acknowledgement that in some circumstances a model developer, similar to traditional software developers, may also have no role as a controller or processor in relation to a given deployment, and appreciates the different examples provided by the ICO for possible controller-processor relationships along the AI value chain. Additional examples that reference specific processing activities would be useful.

The ICO proposes that the nature of open-access genAI may be more suggestive of separate controllership, but CIPL would caution against pre-supposing separate controllership for open-access models or joint-controllership for closed-access ones as a starting assumption.

The ICO suggests that closed-access models may for instance provide deployers with less information than open-access models (and the ICO rightly notes that "open" and "closed" should be thought of as concepts on a spectrum rather than binary). The ICO further highlights that such (closed access) generative AI models may exhibit behaviour that has been pre-determined in the development stage and for which a third-party deployer does not have adequate influence or control to identify and mitigate risks. The ICO concludes that it is more likely in these situations that the developer and the third-party deployer will be joint controllers.

CIPL is cautious of the ICO's reasoning and instead urges the ICO to consider a typical relationship between third-party deployers and developers of AI models. In practice, deployers generally have discretion to determine the means and purposes of processing when deploying a generative AI model, where that processing takes place within a range of uses indicated as appropriate by the developer. Deployers choose what to use a model for, assess whether it is appropriate for their use case, and decide whether personal data will be processed in any given deployment. Even in situations where a model is not fully "open-access," (as described in the ICO's analysis) the deployer will still often be the only party that determines whether that model is suitable for their use case, whether they will undertake additional fine-tuning, and/or how personal data will be used in the deployment of that model. Developers, by contrast, often have no oversight of or insight into the purposes for which a deployer processes personal data when using a model. Deployers or third parties will often be able to process personal data without further involvement from the developer or link to the developer's processing purpose. In such circumstances, they do not have to "jointly determine" (i.e. agree) the means and purposes of any further processing of the deployment with the developer.

For example, where "Developer A" trains an LLM with publicly available data and licenses the trained LLM to "Deployer B" (without access to the training data), while "Deployer B" of the (trained) LLM has not participated in this training phase, "B" will decide to **further** use the trained model for its **own purposes**, distinct from the LLM model creation as such. "Deployer B" may decide to fine-tune the licensed original LLM or to use it to power "Deployer B"'s own products, which may involve injecting other personal data. In this instance, "Deployer B" is an independent data controller, as the Developer has no further control over the decisions of the deployer. The lack of control on how the model was built is not determinative in this case; rather, it is essential to consider which party has control over

processing at each point of the AI lifecycle. It is likely to be an example of sequential, separate controllership rather than joint controllership. The point of being a joint controller is precisely that both can share specific purposes and means of the processing of the same personal data, in a manner that justifies why they would become joint and several liable for this processing, which would not be the case in this example.

Furthermore, the level of information provided to the deployer should be a consideration for transparency and supporting wider GDPR compliance, rather than controllership. And to the extent the ICO assumes that closed model deployers may lack expertise, or resources for example, the same can be true for deployers in the context of open access models. In any event, even a less sophisticated deployer can utilize tooling to assess how a model performs for their use case and be accountable for the risks and mitigations appropriate for their use case. An organization's failure to take basic responsible AI best practices or otherwise perform their own compliance obligations cannot shift controller obligations to a model developer.

In the context of open-source models, the ICO suggests erasure or anonymising in advance of publication to mitigate potential risk inherent in releasing the model or training data sets. Apart from the question of whether the model itself can be considered to contain personal data at all (and its release subsequently to constitute processing), CIPL outlined the challenges regarding erasure and machine-unlearning (which are not specific to open access models) in its response to the ICO's Fourth consultation[3], including our support for the continued development, adoption, and implementation of privacy-enhancing and privacy-preserving technologies (PETs/PPTs) in the context of the entire AI lifecycle.

In closing, CIPL encourages the ICO to take a judicious approach to the concept of joint controllership, recognizing that in many instances, developers and deployers will exercise control over the purpose and means of processing at different stages of the AI life cycle. As the ICO states, "[d]ifferent processing activities should not be lumped together when they serve different objectives or have distinct data protection risks." CIPL also notes that joint controllership may be challenging to implement in practice as it requires careful coordination and agreement between developers and deployers on the purpose and means of data processing in situations where alignment on these the purpose and means may be neither practical nor realistic. Controllership must remain a fact-based assessment, and regardless of the type of model employed, the facts and context of each deployment scenario should be determinative of the assignment of controllership.

[3] CIPL, "CIPL Response to the ICO's 4th Consultation on Engineering Individual Rights into Generative AI Models", June 2024, https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipls_response_to_4thico_gen_ai_consultation_on_engineering_individual_rights_into_generative_ai_models.pdf.