CDT Europe response to the public consultation for the draft Delegated Regulation on Data Access for Researchers in the Digital Services Act.

The Centre for Democracy and Technology Europe (CDT Europe) is a non-for-profit organisation advocating for the promotion and protection of democracy and human rights in European tech law and policy. We have previously responded to the call for evidence for the Delegated Act on Data Access and frequently publish research on approaches to enabling researcher access to data in ways that minimize human rights risks, for instance 1, 2, 3, 4.

Introduction

CDT Europe welcomes the European Commission's initiative to provide independent researchers with access to platforms' data, initially through Article 40 of the Digital Services Act and now with further specifications outlined in the Draft Delegated Act.

Several VLOPs have made access to their data more difficult in the last year, but transparency that allows for research remains the primary tool for understanding how online services contribute to systemic risks to society and the best avenues to mitigate them.

The draft delegated act clarifies many important details, and much of the feedback that was provided from stakeholders during the call-for-evidence (including CDT Europe's feedback) has been incorporated into the text. In what follows, we aim to build on this great effort by providing suggestions on what should be further detailed.

Overview of key recommendations:

- **Expand the independence requirements for applicant researchers to prevent government overreach**
- **Explicitly mention the role of CSOs, including CSOs outside the EU, as potential applicant researchers**
- **Further detail the requirements for the data inventory to ensure completeness**
- **Empower researchers to initiate the mediation process**
- **Empower independent experts to assess the quality of data inventories and be part of the mediation process**
- **Extend the timeline for DSCs responding to data access applications**

1. Who can access the data

1.1 Preclude Law Enforcement from becoming vetted researchers

Context:

Article 8 (2)(c) and recital (9) of the draft Delegated Act state that the Digital Services Coordinator of Establishment should verify the documentation demonstrating the applicant researchers' independence from commercial interests. This is important to ensure impartiality. However, we are concerned that – unless explicitly precluded in the Delegated Act – **there is a risk that law enforcement agencies may use DSA Article 40 for unjustified or abusive law enforcement surveillance.**

CDT has conducted research documenting that EU law enforcement and other governmental agencies have a demonstrated interest in gathering data from social media, and, in some instances, this data gathering may appear to have similarities to research. This raises serious human rights concerns as law enforcement personnel have been shown to use social media data in the past for illegitimate purposes such as monitoring protestors, dissidents, and members of religious or racial minorities.

With the current language of the Delegated Act, taking into account GDPR Article 89 and DSA Article 40, it will be at the discretion of each DSC receiving a data access application to decide whether a government or public body qualifies as a a 'vetted researcher'. This can be particularly problematic in Member States with track records of infringing on fundamental rights, where the independence of DSCs is not a guarantee. In such contexts, the DSA risks becoming a tool for surveillance and oppression, undermining public trust in the regulation, even if the European Commission may intervene under DSA Article 50.

Suggestion:

- Consider explicitly precluding law enforcement agencies from being eligible to qualify as vetted researchers by adding that, on top of the independence requirements listed in Article 2 of Directive (EU) 2019/790, vetted researchers may not be government bodies with law enforcement authority, employed or contracted by such government bodies, or otherwise affiliated with law enforcement.

1.2 Facilitate Civil Society Organisations becoming vetted researchers

Context:

Article 40(8) and recital 97 of the DSA specify the entities that are eligible to become 'vetted researchers' – this includes Civil Society Organisations (CSOs), but only if they are carrying out "scientific research", as defined in Article 2 of Directive (EU) 2019/790. The lack of clarity as to what constitutes 'scientific research', in combination with the data protection requirements of Article 5 of the Delegated Act and the fact that CSOs often employ contracts that may not constitute "a formal relationship" between the applicant researcher and the research organisation, as defined in Article 8 of the Delegated Act, means that – in practice – **civil society risks being largely excluded from Article 40 data access requests.**

In order to comply with the above requirements – and given the limited capacity of many organisations that may have in-house expertise to carry out useful research, but do not fulfill all the criteria listed above – CSOs may have to collaborate with each other, or with academic institutions, to apply for data access requests. This will make it easier to comply with the GDPR requirements, while also opening the door for concurrently pooling resources and expertise from different organisations.

Suggestion:

- Consider including language in the Delegated Act that will permit collaboration between CSOs, and between CSOs and universities, by extending 'formal relationships' to include joint applications for data access, partnership agreements, fellowships and other kinds of contractual or non-contractual arrangements to be decided upon on an ad hoc basis by each DSC. This will allow for the possibility of multiple researchers from different organisations having access to the data under a single request, but it will still be at the discretion of the DSC to decide whether the necessary privacy and security requirements are met, proportional to the sensitivity of the data requested.

- We also suggest that you consider including an explicit reference to the fact that civil society organisations and universities not based in the EU can apply for access to data, something that is currently only implied by reference to international transfers in Article 9 (3), but could benefit from further legal clarity.

2. The data provided

2.1 Ensure that data inventories are useful

Context:

Article 6(4) mandates that platforms publish an "overview of the data inventory of their services," including examples of datasets and access modalities. Recital 26 recognises the importance of data providers making available to vetted researchers "the relevant metadata and documentation describing the data made available, such as codebooks, changelogs and architectural documentation" in order to put the data "in the proper context". However, recital 6, which describes the kinds of data that should be made available in the data inventory, is less descriptive, creating the **risk of making it harder for researchers to put the data "in the proper context" at the inception of their research project**. This could limit their imagination i.e. the ability of researchers to understand what data can be useful and in what way, but could also make it **more difficult to foresee what kind of privacy and security measures will be necessary to include in the data access application.**

Furthermore, the definition of "data" under Article 40, although open to interpretation, remains narrow. Currently, the examples of data listed revolve around user-related data. **This fails to include data on the organisational structure, internal policies, procedures for enforcement of such policies, and records of enforcement actions, which can be crucial in understanding how companies' operations may lead to systemic risks.**

Suggestion:

- Consider including explicit references in recital 26 to types of data that are not user-related, but company-related, such as organisational structures and internal policies and processes for their adoption, implementation, and enforcement.
- Consider expanding the language in recital 6 to include the examples given in recital 26.
- Include a requirement for frequent update of the data inventories by data providers in Article 6(4).

3. The request process

3.1 Ensure an equitable and effective dispute settlement procedure

Context:

Article 13 of the Delegated Act stipulates that only the data providers may initiate a dispute settlement procedure. This, presumably, seeks to reflect the fact that it is also the data provider that is responsible for covering the costs of the mediation process. The principal researcher may then only participate in the mediation process upon invitation by the DSC, while it is not explicitly specified what happens in the case of a failed mediation process.

Though we appreciate that this is intended to shield researchers from potentially burdensome and time-consuming procedures, this set-up of allowing the data provider to initiate the mediation process and to choose the mediating body (that the data provider is also responsible for paying), while not specifying what happens in the case of failure, **may provide perverse incentives for data providers to abuse the dispute settlement mechanism.**

Suggestion:

- We suggest that researchers be empowered to seek mediation, for instance upon receiving data that they believe fails to meet accuracy, veracity or completeness standards, meaning the data provider would not be complying with the reasoned request. Another situation where researchers could seek mediation is if they do not agree with the decision of the DSC in response to an amendment request. To avoid undue financial burdens on the applicant researchers that would discourage them from seeking such mediation, the data provider should still cover the costs.

- We also suggest that the principal researcher be empowered to choose whether they want to be part of the mediation process, even if the process is initiated by the data provider. This would not require them to be invited by the DSC, leaving it at the discretion of the principal researcher to determine whether they are in a capacity to carry that burden.

- Finally, we suggest that the text specifies that, in case of a failed mediation initiated by the data provider, the data provider is obligated to provide the data pursuant to the decision of the DSC following the amendment request.

3.2 Further empower the Intermediary Bodies of the Independent Advisory Mechanism

Context:

We welcome the creation of an independent advisory mechanism in Article 14, as a way to include external experts in the process. In the current text, such experts are to be consulted by the DSCs only 'before formulating a reasoned request, or taking a decision on an amendment request'. Their involvement is not foreseen in the mediation process, nor is there mention of the role that such external experts could play in assessing the quality of the data inventories. Building the necessary capacity and expertise within all DSCs will likely take time, **at an initial stage the VLOPs and VLOSEs have the potential to exert considerable influence regarding all technical aspects of the data request and access process, including when sending amended requests and in particular during mediation processes**. Additional possibilities for the involvement of external experts in more parts of the process could help add balance to this dynamic and ensure a more impartial process.

Suggestion:

- We suggest that the external experts be able but not obliged, upon invitation by the DSC, to weigh-in during a dispute settlement procedure.

- We also suggest that you consider adding language that would empower independent experts to advise the DSCs and the European Commission, upon request, on possible improvements to the structure of data inventories and the quality of the data provided in the data inventories.

3.3. Avoiding bottlenecks – extending the timeline

Context:

We appreciate the European Commission prescribing an ambitious timeline for DSCs responding to applications, as described in Article 7 of the Delegated Act. However, we are concerned that – at least at an initial stage – DSCs may not have sufficient capacity and experience to confirm that a given application has all the information and supporting documentation required within only five working days. This may lead to undue rejection of legitimate applications, or may compromise the quality of the applications that are accepted, leading to amendment requests that can create unnecessary bottlenecks.

Suggestion:

- The deadline for responding to a data access application should be extended to +/- 10 working days.