

# Comments on the Committee of Ministers' draft recommendation on online safety and empowerment of content creators and users

Submitted by

Luke Thorburn  
King's College London  
[lukethorburn.com](http://lukethorburn.com)

Responding to

[Draft Recommendation CM/Rec\(20XX\)XX of the Committee of Ministers to member States on online safety and empowerment of content](#)

## Comments

Overall, I think the draft recommendations, if implemented, would build constructively on current policy settings and are well targeted to promote empowerment and reduce harms without impinging on the right to freedom of expression. I've included some granular feedback on a few of the clauses below — feel free to disregard if not at the intended level of abstraction.

- In (o) and (51), it may be helpful to acknowledge that there are better and worse forms of “consultation”, and that in particular, representative deliberative processes are recognized by the OECD as being particularly well-equipped for addressing “value-driven policy dilemmas, complex problems that require trade-offs, and long-term issues that go beyond short-term incentives.”
  - [Innovative Citizen Participation and New Democratic Institutions: Catching the Deliberative Wave](#)  
OECD (2020)
- In (49), the terms “amplify” and “amplification” are used without acknowledging the significant lack of consensus on how those terms should be understood. Collaborators and I have previously argued that they are too fuzzy to be used in law, and also described some properties that a good measure of amplification would have, and why they are difficult to achieve (references below) which might be a helpful resource. I would suggest avoiding the terms in favour of alternative language that is more precise.
  - [What Will “Amplification” Mean in Court?](#)  
Luke Thorburn, Jonathan Stray & Priyanjana Bengani  
Tech Policy Press (2022)
  - [Making Amplification Measurable](#)  
Luke Thorburn, Jonathan Stray & Priyanjana Bengani  
Tech Policy Press (2023)
- In (50), the text states: “If such assessments conclude that proposed interventions pose risks, they should also include concrete measures to prevent or mitigate such risks, which platforms should be bound to implement before introducing design changes.” I would suggest that not all risks that might be posed by proposed design changes can be effectively mitigated, and so it would be good to acknowledge that, in some cases, the appropriate response to a risk assessment is to *abandon* the proposed design changes.

- In (55b), the text mentions the right to inspect the provenance of advertising, but I would suggest adding also a right to understand the *social provenance* of organic content — the communities from which it originates, and the degree to which it is resonating with different communities. Such social provenance and context is critically missing on almost all major social media platforms, which likely contributes to so-called ‘perception gaps’ and other failures of common knowledge that degrade the social fabric. Making such provenance transparent would significantly empower everyone (users, creators, researchers, ...) to better understand society. See the article with Glen Weyl, Audrey Tang and others for more on this phenomenon.
  - [Prosocial Media](#)  
E. Glen Weyl, Luke Thorburn, Emillie de Keulenaar, Jacob Mchangama, Divya Siddarth, Audrey Tang  
SSRN (2025)
- In (67), I suggest explicitly stating that platforms should make it possible for users to move their entire *social graph*, including the connections to those they subscribe to or follow, and the connections to those that subscribe or follow *them*, to other complementary or competing platforms. With the emergence of widely adopted protocols like [ActivityPub](#) and the [At Protocol](#) (among others), this is technically possible, and would ensure that anyone who builds a business or following on a single platform is not vulnerable to the ongoing goodwill or reliability of that platform.
- In (70), when mentioning “accuracy”, it would be helpful to specify that platforms should share the full [confusion matrix](#), recall, precision, and where possible, the [ROC curve](#) for classifiers used in content moderation, so that the public conversation can have a sophisticated conversation about the *trade-offs* that platforms face related to respecting freedom of expression while addressing harms through content moderation.
- I strongly support (74) — while there are many practical details that would need to be figured out to operationalise such a requirement, in many cases it is currently not possible as an academic researcher to conduct ecologically valid public interest research about the effects of platforms on society.