



Researching Systemic Risks under the Digital Services Act

Dr. Oliver Marsh, 26-07-24

Summary

We conducted a series of activities to better understand opportunities and challenges facing external researchers - including Civil Society Organisations (CSOs) - specifically related to researching “systemic risks” and mitigations under Article 34 of the Digital Services Act. In order to encourage practical reflections, we grounded many of these discussions in real-world cases of potential systemic risk (including a development of a “risk repository”).

We found that many of the challenges are practical rather than conceptual. Despite the current vagueness over definitions of “systemic risks”, researchers are still able and willing to provide inputs on potential cases of systemic risks. In doing so they drew on a diverse range of frameworks and approaches. Although they did not always reach the same conclusions about whether cases were “systemic risks”, this is ultimately not a requirement for external researchers. **The diversity of approaches instead provided different, often complementary, perspectives and approaches to the cases. An overly prescriptive approach to systemic risks could limit this.**

The more pressing problems relate to ensuring researchers have the information and understanding needed to plan and resource impactful research activities and to help iteratively improve DSA implementation. We have often heard that much of the practical implementation of the DSA, in particular around systemic risks, will evolve over time and involve “learning together”. This approach has advantages, particularly given the complexity of the subject. However, it must be acknowledged that, compared to other parties, **external researchers have fewer opportunities and much less information than other parties from which we can prioritise our activities and help this learning.** It is currently unclear what plans and structures are envisaged for effectively learning together, and as such how we can support iterative improvement of DSA risk assessment and mitigation measures. This could make the difference between the Article 34 provisions “working well” in a few years, versus in *many* years (or never).

To address these problems, we make the following proposals to the EU Commission (EUCOM), Digital Services Coordinators (DSCs), and potentially even Very Large Online Platforms and Search Engines (VLOP/SEs):



- **More detailed and regular information on risk assessment work that has already been conducted, that is being prioritised, and that is not being addressed by EUCOM, DSCs, and VLOP/SEs, as well as how this assessment work is being conducted.** This information needs to be transparent and up-to-date enough to inform external research decisions about current gaps to focus on, methods which can be worked on, etc. Currently there are substantial concerns about the limits of Article 42(4) reporting, as well as RFIs and other communications, for informing this decisionmaking.
- **More structure around collaborations and information-sharing between regulators and the research community,** to ensure there can be productive, efficient, and transparent back-and-forth on points of detail. This applies both to engagement related to specific topics of research and enforcement, and also to ensure the desired “learning together” of DSA implementation is efficient and effective.
- **Clarity on how the concept of “systemic risk” may be iterated and clarified over time** – through guidelines, Article 40(4) decisions by DSCs, emerging enforcement cases, etc. – and how the process will involve stakeholders to ensure that the emerging **concept facilitates rather than excludes a diverse research ecosystem.** This includes questions of e.g. how researchers formulate data access requests, or if researchers are legally protected when we use methods such as scraping or share data. Some ideas of “systemic risks” (e.g. a highly quantitative definition) might even impede some research.
 - This process should also help to ensure other important aspects such as legal clarity and providing safeguards against overreach; however in this piece we focus on the implications for research.

In sum, there is room for more immediate and productive involvement of external researchers, both to develop specific research into “systemic risks” and to more broadly support the “learning together” of DSA implementation. This requires, at a minimum, that EUCOM and DSCs ensure more transparency and provide legal certainties to minimise existing barriers for researchers. At best, they can help facilitate the collaborations, best-practice sharing, and resourcing needed to maintain a diverse, active, and impactful systemic risk research ecosystem.

Our next steps in this work will include (i) extending our engagement with the research community, particularly to incorporate a wider range of academic researchers and (ii) sourcing more detailed proposals on methods and practices for researching systemic risks, which should be facilitated by the steps we have proposed in this document.

Contact: marsh@algorithmwatch.org



Contents

Summary	1
What We Did.....	4
Broad Themes.....	5
More Detail on Specific Themes.....	6
Text of the DSA.....	6
Risk vs. Impact.....	7
Metrics and Measurement.....	8
Different and Conflicting Rights/Risks	9
Scope and Specificity.....	11
Evidence Sources & Quality.....	12
Additional Points on Process and Clarity	13
Acknowledgements.....	17
Selected Bibliography	18
Annex A: Selected “Grey Area” Risk Cases	20
Annex B: Relevant Text from the DSA	23



What We Did

- We began by producing a “risk repository” of 30+ real cases which show *potential* systemic risks under the DSA. From these we selected 7 “grey area” cases (see Annex).
- We distributed the full repository, as well as the 7 “grey area” cases to external experts. We asked them to select cases and provide arguments for and against these cases being evidence of a “systemic risk” under the DSA, a conclusion on balance, and their thought process on how they made the assessments.
 - At present we have received 5 full written responses, so results from this part of the work should not, at present, be seen as widely representative across the research community. The opinions largely come from CSOs, so as a next step we want to further encourage input from academics.
- We also presented cases from the repository, including some of the themes raised by respondents, at events related to assessing systemic risks under the DSA at: the CDT CSO Roundtable (April 2024), the CPDP Conference (May 2024), and re:publica (May 2024), as well as smaller exchanges with other experts (CSOs and academics).
- We also used the cases to develop a *hypothetical* risk case for a collaborative workshop including representatives of VLOP/SEs at the Global Network Initiative European Rights & Risks Stakeholder Engagement Forum (June 2024).
- Finally, we have reviewed documentation related to DSA risk assessments from other CSOs, organisations with specialist knowledge in risk assessments, academics, and legal experts (see Bibliography).

This interim report summarises current findings and ideas emerging from all the above, **focussing on implications for external researchers (which includes CSOs, NGOs, academics, research institutes, and investigative journalists).**

There are additional questions, beyond the scope of this document, regarding the implications of different definitions of “systemic risk” for enforcers and VLOP/SEs. In particular these relate to (i) legal clarity and (ii) how definitions could risk bad practices (e.g. overly specific criteria meaning VLOP/SEs focus on meeting metrics not real harms, overextension of concepts such as “civic discourse” or “public security” as a political weapon, etc.). We also primarily focused on research into identifying systemic risks under Article 34, rather than mitigations under Article 35.



Broad Themes

- The current lack of clear definition of “systemic risk” does not make it impossible for external researchers to support enforcement under the DSA. The current text does provide areas towards which external researchers can direct attention and produce potentially relevant outputs. Researchers brought in a wide array of potential methods, approaches, and frameworks which can be applied to the question of identifying systemic risks under the current DSA text.
 - Indeed overly strict or narrow definitions of “systemic risks” - e.g. requiring demonstration of causal links or rigorous quantitative probabilities - **could exclude many potentially relevant methods, parties, and insights.**
- However, researchers often reached very different conclusions about whether a given case was evidence of “systemic risk”, and how clearly this decision could be reached.
 - Main sources of disagreement were (i) balancing between conflicting risks, in particular between freedom of speech and other systemic risks, and (ii) whether levels of “real-world impact”, connected to probability and/or causation, needed to be demonstrated for a risk to be “systemic”.
 - As researchers are not expected or required to make enforcement decisions, **this lack of consensus need not be a fundamental problem. Such pluralism may even be good for a rich, multi-perspective ecosystem of research.**
 - However there are some potential practical issues if understandings of the concept are *too* disparate, particularly if misunderstandings from / between regulatory bodies emerge over time.
- The main issue for researchers is a broader lack of clarity, which goes beyond definitions of “systemic risk”, **and also includes questions about transparency, forms of evidence, data access, and regulatory outcomes.**
 - The definition of “systemic risks” contributes to this, in terms of e.g., not knowing the potential scope of data access requests, not knowing what sorts of evidence might be deemed relevant or useful for enforcement, how we can best find and fill gaps in existing assessment and enforcement actions, etc. But it is only part of the issue.



More Detail on Specific Themes

Text of the DSA

(See Annex B for text of referenced Articles and Recitals).

The subparagraphs under Article 34 were frequently used as a helpful guidance - e.g. asking the question of the extent to which a case really was evidence of a systemic risk under 34(1), and whether it could plausibly be related to a factor under 34(2). This was a relatively common and straightforward “first step” to approaching the question. This seems to provide useful clarity and shared reference points. **However it is worth considering what may be missed in this approach, given the lists are not intended to be exhaustive.** There can also be conflicts between items in the list, as discussed shortly.

It should also be noted that even these apparently simple points contain complications. For instance a case involving deep fakes of politicians can be interpreted differently depending on if it is seen as a risk to electoral processes, fundamental right to human dignity, fundamental right to freedom of expression and information, and/or gender-based violence – potentially with different conclusions.

The text of Article 35 was occasionally, though not as often, referred to as a guiding principle. Where it was often useful was as a “low hanging fruit” question: *“are there simple and relatively unproblematic mitigation steps, like more transparency or better resourced moderation, which the VLOP/SEs could take to minimise the risk?”* The answer to this question does not solve the *conceptual* problem of whether a potential-but-mitigated risk would be “systemic”, if it were to be unmitigated. **However in practice participants often saw failure to take simple mitigation steps as relatively clear evidence of failing to meet the expectations of the DSA** - clearer than whether the risk itself was “systemic enough”.

No-one made explicit reference to Recitals in providing views on the cases. Recitals 79-89 were occasionally referenced in events and discussions, though were often not seen as providing much clarity (indeed, some argued they further confused text from the Articles by gesturing towards specific aspects to consider, without clarity on how or why). By contrast Recital 90, and its references to who and what should be involved in the assessment *process*, was frequently referenced positively in discussions and events.

Finally, there were questions around how decisions related to Article 40 from DSCs (on whether data access requests were demonstrating that they were researching “systemic risks”) would interplay with decisions related to Article 34 and 35 from EUCOM (on whether platforms were exhibiting “systemic risks”). For example, might DSCs' decisions on what counts as research into “systemic risks” create a form of quasi-case-law upon which



VLOP/SEs, auditors, and even EUCOM might draw on regarding Article 34 risk assessments? A related point is that, to encourage a full “understanding of systemic risks”, data access requests arguably should also encompass broad research which aims to understand the systems and functioning of platforms in general (e.g. to provide baselines against which systemic risks can be assessed), not just research into specific items listed in Article 34 and only in the EU.¹

Risk vs. Impact

The question of whether research into “systemic risks” should demonstrate a real/probable off-platform impact, and whether this impact needed to be “sufficiently severe”, was raised repeatedly. For instance in relation to elections, researchers queried whether the *spread* of disinformation content should be considered a “systemic risk”, even if the probability of actually changing an election (or the ability to assess this) was low.² For some participants, the expectation that real or probable impact needed to be demonstrated was an unrealistically high bar - particularly given the assessment is explicitly about *risks*, not impacts. This was one of the disagreements which led to different conclusions in views on repository cases.

Some experts saw value in a risk framework based on severity, probability, number of affected persons (broken down by users and non-users), irreversibility, and possibility to remedy and restore. This has often been used in other risk assessments and impact assessments, including in the human rights field.³ It could also potentially be used to weigh conflicting risks in some cases, while acknowledging that some normative judgement may still remain.

However others argued that, in many cases related to systemic risks from VLOP/SEs, such categories may prove hard to use. This may particularly be the case when deploying novel functionalities, including mitigation measures. Nonetheless, where a known impact can be demonstrated and categorised in the above framework, this could still provide valuable evidence and weighting criteria for risk assessment.

¹ See Leerssen, Paddy, [Outside the Black Box: From Algorithmic Transparency to Platform Observability in the Digital Services Act](#), pp.23-25

² For an example of discussion based on empirical observations see [Who Targets Me, Disinformation ads, Systemic Risks and the Digital Services Act](#).

³ See e.g. [ECNL / Access Now: Towards meaningful fundamental rights impact assessments under the DSA](#)



Metrics and Measurement

Multiple participants raised an extremely wide range of on-platform metrics which can be used to understand user behaviour and impacts (dwell time, content moderation decisions, content volume, etc.). However these alone were often deemed not sufficient to show “risk” (see the Impact point above). There have also been numerous concerns, and some demonstrated cases, that on-platform metrics (including those provided under Article 40) are not always accurate.⁴ Methods to check these metrics, such as manual inspection or scraping, can be inefficient, or raise concerns for researchers around breaking platform terms and conditions. Off-platform approaches - e.g. user surveys - were also raised as important alternatives and complements, though also challenging and potentially expensive to apply at scale.⁵

Broader discussions raised **different opinions around trying to develop new metrics geared towards specifically understanding systemic risks under the DSA** (potentially even towards providing benchmarks for enforcement). Benefits of such metrics would include consistent cross-platform and change-over-time measurement, which could be particularly valuable in (i) developing “best practices” and “effective mitigations” and (ii) information pooling and sharing. Metrics also allow for practices like e.g. pre-registration and consistent comparison, which can help reveal if risk assessments are being designed or adapted to give more favourable results. A downside can be that these metrics become the focus of compliance and enforcement, rather than the actual risks they are supposed to represent; that some important issues are intrinsically challenging to capture effectively as metrics; and that metrics will still involve, but often conceal, normative assumptions and framings.

The obvious conclusion is to ensure the research ecosystem allows for both metrics-driven and qualitative research, aiming to leverage the advantages of both. However it should be

⁴ See, for example, [forthcoming work by Philip Darius and colleagues](#).

⁵ An extensive list of possible metrics can be found in [CERRE: Systemic Risk in Digital Services: Benchmarks for evaluating the management of risks to electoral processes](#). For two examples of metrics being developed and used in practice, see Wagner et. al. [Mapping Interpretations of the Law in Online Content Moderation in Germany](#) in the context of content moderation, and [Who Targets Me, Disinformation ads, Systemic Risks and the Digital Services Act](#) in the context of elections.



noted that such combinations can, even unintentionally, disempower some research approaches.⁶

For example there may be categories of risks which are more amenable to measurement - whether through creating specific metrics or through drawing on precedents e.g. from fundamental rights. However this could unintentionally create a hierarchy where the “measurability” of a risk could determine how it is treated, including: how much attention it is given, in particular the potential for focusing on less impactful but easily measurable risks rather than high impact but harder-to-measure risks⁷; how a risk is treated in conflicts or trade-offs with less easily-measured risks; whether mitigating these risks can conceal ineffective mitigation of less easily measured risks; etc.

This is not an argument against any one form of research, but rather an argument for being attentive to how different forms of research are treated, used, and synthesised.

Developing metrics should also involve questions of the background against which measurements should be compared. Other VLOP/SEs? An “acceptable level” of risk? “Normal levels” of that risk in a given context? Also, where considering multi-VLOP/SE risk, the appropriateness of comparisons between different VLOP/SEs should be taken into account.

Some experts noted that similar clarity and consistency to metrics can be provided by other approaches, for instance drawing on legal precedent for Articles 34(1)(a) (illegal content) and (b) (fundamental rights), precedents could provide clarity and assessment methods. However such precedents may not exist for 34(1)(c) (civic discourse and electoral processes, and public security) and (d) (gender-based violence, the protection of public health and minors, physical and mental well-being). There are also similar questions as for metrics, regarding whether how “clearly” a risk can be defined will affect how it is treated by comparison with other risks.

Different and Conflicting Rights/Risks

⁶ For examples drawing on interdisciplinary research during COVID see Colman *et. al.* [Following the science? Views from scientists on government advisory boards during the COVID-19 pandemic: a qualitative interview study in five European countries](#)

⁷ One could imagine, for instance, measuring election disinformation with clearer but less consequential metrics such as “accuracy of detection of dis/misinformation content”, rather than harder but more consequential impacts such as “impact on voting patterns”.



Extending on the above differences between 34(a) and (b) versus 34(c) and (d), there are questions around whether (c) and (d) could, in theory, correspond to particular fundamental rights (including some not listed in 34(1)(b) but are in the European Charter of Fundamental Rights), for instance aligning protection of elections with rights to freedom of information and the dignity of political candidates. Other rights frameworks can also be drawn on, including the Universal Declaration of Human Rights and the International Covenant on Civil and Political Rights. However many experts also noted that concepts like “illegal content” and “fundamental rights” can be less clear than they first seem, particularly when considered in the full context of national and international frameworks. As noted previously, drawing on rights frameworks can provide useful precedents, but may still leave open questions of how to assess potentially novel risks at a “systemic” scale. These questions are beyond the scope of this current document.⁸

However, one of the substantial complicating factors in the risk repository exercise was related to the question of how to weigh up potentially conflicting rights at a systemic scale. **In particular, the right to freedom of speech was often seen as conflicting with other parts of Article 34.** Many/all of the other listed risks tend towards suggesting VLOP/SEs should, as mitigation measures, restrict certain kinds of content or behaviour (hate speech, disinformation, etc.); this is not a *necessary* conclusion, but in practice it often is the case. By contrast minimising risks to freedom of speech often pushes in the other direction; it can even be argued that many mitigation measures may *create* systemic risks to freedom of speech.

These issues are well known in debates over platform governance, but the DSA does not seem to provide guidance in how to balance these. This was another factor underlying differing views on the cases. At a normative level, some participants argued that DSA enforcement should be directed at ensuring that VLOP/SEs outline their own balance, and consistently and transparently stick to it. By contrast others argued the nature of potential harms, with potentially severe negative externalities for individuals and society, means that VLOP/SEs should not be the arbiters of this balance.

As noted above, it is not essential for the functioning of the DSA that *researchers* reach consensus on how the DSA should function when different requirements pull in different directions. **Nonetheless, researchers are not and should not be seen as simply “neutral” conduits of information; our views on what we regard as risks, and behaviour by VLOP/SEs which we believe fails to appropriately balance conflicting issues, will inform our research.** We should also have an understanding of how our work could feature in actions which

⁸ But see, for example, discussion on the legality of using the DSA to counter disinformation in Husovec, Martin, [The Digital Service Act’s Red Line: What the Commission Can and Cannot Do About Disinformation.](#)



shape the online ecosystem, and the balancing tests that will be applied.⁹ Therefore more clarity and transparency on how DSA enforcement foresees addressing such conflicts, and the process by which this would be judged, should still be provided.

Scope and Specificity

When selecting “grey area” cases from the repository, we deliberately provided a range of scopes/specificities of cases (see Annex A for full list). Some were extremely broad, such as “academic research suggests that Twitter (X) use predicts substantial changes in well-being, polarization, sense of belonging, and outrage”. Others were much more specific, for instance deep fakes on Meta platforms in the Slovakian election. Our intention was to understand whether participants saw potential for very specific cases to be “windows” into broader systemic risks (and if so, how); or whether they felt such a question *by necessity* required a broader starting point.

In general, respondents suggested that even specific cases *could* provide useful windows into potentially risk-inducing activities on VLOP/SEs. Also broad cases could actually, on inspection, be broad-but-shallow; for instance the specific academic study on Twitter/X presented only had a small sample size despite asking a broad question (see some of the trade-offs described in the Evidence section).

There have been questions raised around whether “systemic risk”, by comparison with finance, would require risks to affect multiple “nodes” in a system. One interpretation of this, that a systemic risk *must* affect multiple VLOP/SEs, was generally considered to be an unhelpfully restrictive requirement; in the words of one of our correspondents, “the fundamental first step is each individual VLOP or VLOSE’s own assessment.”¹⁰ Nonetheless risks can arise from interlinkages between VLOP/SEs and researchers, as well as regulators and companies themselves, should be attentive to these.

There are questions of how to design and support research to establish risks which may be severe when aggregated across VLOP/SEs, but appear insignificant on any one individual VLOP/SE. Cross-platform mitigation efforts may also be impeded by organisational and legal limits to information-sharing between teams inside VLOP/SEs with those in other VLOP/SEs, though this is beyond the scope of this paper. Also, cross-platform analysis

⁹ See [previous work by AlgorithmWatch outlining a “5E’s Framework”](#) for ensuring legitimate use of work by external stakeholders

¹⁰ Correspondence with Sally Broughton Micova, in addition to [CERRE: Elements for Effective Systemic Risk Assessment under the DSA](#)



needs to account for the very wide range of variation across VLOP/SEs (including user base, functionalities, primary uses, etc.). Nonetheless it is important to consider the potential of cross-platform risks for multiple reasons, including (i) their role in malicious online activity and (ii) the increased levels of risk and complexity of mitigation when phenomena occur on multiple platforms.

Another important question is the time frame within which risks are considered. In some cases, e.g. for elections or before releasing a new functionality, this may be relatively clear. In other cases it may not. Similarly to the question of cross-platform risks, whether a risk appears to be one-off or systemic, and the evidence available to support such decisions, may be completely different depending on the time frame used.¹¹

Evidence Sources & Quality

Participants drew on a wide range of existing and potential evidence. Even though the repository provided some quite specific cases, without much supporting detail in each case beyond e.g. one or two articles, researchers were often able to find a range of additional relevant evidence for their arguments. **However the question of what count as legitimate sources of evidence to draw on, whether evidence was “good enough”, and what to do in the face of conflicting evidence, was a challenge.**

As noted in “the Metrics” section above, evidence collected directly from platforms (e.g. through APIs) versus evidence collected through other means have advantages and disadvantages. Some questions are conceptual, such as how accurately user surveys reflect actual behaviour and experiences; others are practical, such as whether VLOP/SE-provided metrics are accurate. There are also trade-offs between slower, multi-data-source, more systematic evidence vs. faster, more focused, and/or “indicative” evidence-gathering.

While the importance of highly rigorous evidence was widely acknowledged, going slowly – as many academics felt institutionally compelled to do – could allow risks to persist or worsen (and could still lead to unclear and conflicting conclusions). But equally, over-responding to insufficient evidence could also create risks (see “conflicting rights/risks”

¹¹ See, for instance, the high-profile [Facebook and Instagram Elections Study](#) in which evidence was collected, in collaboration with platforms, through experimental interventions. Supporters of this research point to the potential of experimental approaches to demonstrate causal links, however critics argue that the relatively short time period of the interventions minimises their relevance for understanding the impacts of interest (opinion formation).



section). The question of how to ensure expert and efficient scrutiny of evidence is therefore an important one.

Metaphors used on multiple occasions were the historic cases of building evidence bases against smoking or climate change; though it took time, aggregated studies ultimately produced a clear picture of risks and harms. However this does raise questions of (i) whether such consensus around the more diverse, dynamic, and less clear/direct potential harms of online VLOP/SEs would ever clearly emerge (even very wide-scale systematic studies at present show complex effects)¹² (ii) whether such an evidence base could have been (or can be) assembled more efficiently, potentially minimising harms earlier and (iii) whether “systemic risks” online raise harder normative questions.

Additional Points on Process and Clarity

Many of the complications in the discussions were about practical, rather than conceptual, questions. E.g. if we identify relevant information we need, how quickly and reliably can we obtain it (and iterate on emerging findings if required)? How can we ensure our research is aligned with expectations and needs of end-users (in particular EUCOM and DSCs), while also ensuring we our own independent and expert perspectives? What level of direct collaboration can and should researchers expect from VLOP/SEs? If internal understandings of what is treated as a “systemic risk” within EUCOM, DSCs, and/or VLOP/SEs is going to gradually develop over time, how will this be shared in an ongoing and useful format with external researchers?

These issues were raised more often throughout our discussions than issues with the definition of “systemic risks”. General questions around barriers to transparency in reportage of systemic risks were also raised repeatedly. A key example was the lack of clarity over what exactly will be published under the Transparency Provisions in 42(4) – see Annex B – and whether these publications will actually help guide further research (e.g. provide good information on where more research would be beneficial). As another example, the unpredictable outcomes and timelines of Article 40(4) requests (including follow-up requests given the iterative nature of much research) could cause issues for many researchers, including in e.g. applying for funding and project management questions.

¹² See [Lorenz-Spreen et. al. A systematic review of worldwide causal and correlational evidence on digital media and democracy.](#), as well as debates around the [impacts of social media on teenagers' mental health](#) prompted by e.g. the [US Surgeon General's Report](#) or responses to [the work of Jonathan Haidt](#).



It is widely predicted that early DSA risk assessments and audit reports will probably be of low quality,¹³ given the newness and complexity of the task, but that there will be learning and improvement over time. However it is not clear what the structure or process for this learning will look like; and **it is possible that learning and improvement will (i) largely be informed by very limited input into very infrequent and potentially low-information reporting and therefore (ii) be extremely slow to reach intended goals or react to new developments.** In very specific terms: On our current understanding, in late 2025, external researchers will only be seeing the second round of Article 42(4) reporting, which will have been written in 2024, with very little input from the 2023 assessments – i.e. even by late 2025, much of our input on risk assessments will be based on potentially low-quality and out-of-date materials. By contrast more opportunities for substantial external input on approaches to VLOP/SEs risk assessments outside of these annual cycles could create a virtuous circle, whereby each round of reporting improves much faster in quality and thereby allows for more detailed feedback.

Requests for Information (RFIs) can give a sense of immediate interests of the EUCOM, and can make researchers aware that relevant evidence or ongoing projects they already have may be of relevance. However they are unpredictable and hard to plan for, so it is unclear how they can support planning and prioritisation over the long term. There are also risks that they direct too much research into specific (and potentially high-profile) areas while directing attention and resource away from other areas. This need not be the case if other information is available to guide decisions, but in the absence of other information their impact could be outsized.

There are also occasional opportunities to discuss priorities with EUCOM and DSCs directly, such as through large roundtable events. There have already been proposals on how such engagements should be structured.¹⁴ To reiterate key points from these proposals relevant to this piece: for the wider research ecosystem – many of whom do not have easy access to engaging with EUCOM – such engagements need to be structured to ensure transparency and legitimate representation of their views, and also that these engagements are genuinely collaborative. This means substantive discussion, based on sharing of relevant and detailed information, such that all parties can leave knowing what they can do to best support the wider ecosystem of DSA enforcement.

¹³ Though the unpredictability of what information we can expect to receive is an additional issue.

¹⁴ See [Suzanne Vergnolle, Putting collective intelligence to the enforcement of the Digital Services Act](#), AlgorithmWatch, [Ensuring Legitimacy in Stakeholder Engagement: The '5 Es' Framework](#), Julian Jaurisch [Here is why Digital Services Coordinators should establish strong research and data units](#).



Concluding Remarks

The DSA could, in principle, bring benefits for researchers. It directs attention towards areas (systemic risks) with clearly indicated potential for positive impact via regulation, and provides opportunities through data access requests and transparency reporting. However, from a researcher perspective, some of the realities are still strange and sub-optimal: for example having to pre-guess what data a VLOP/SE might have when formulating a data access request, hoping to receive something useful; or fitting research programmes into rather specific boxes like “functionalities which may increase risks to fundamental rights within the European Union”, potentially ignoring a wider landscape of things which may better improve our understanding of how VLOP/SEs function and impact on people and society; or trying to plan external risk assessment related research without knowing what risks platforms have already assessed, and how.

Such problems may require longer-term and/or more creative solutions; but they may also stem from immediate power imbalances and legal risks facing researchers into the online world, which the DSA is ultimately intended to address. In the meantime, based on the above work, we make the following recommendations:

- **More information on risk assessment work that has already been conducted and is currently being prioritised by EUCOM, DSCs, and VLOP/SEs.** This needs to be detailed and up-to-date enough to inform research decisions about current gaps to focus on, methods which need to be improved, etc. Current methods for information sharing, in particular Article 42(4) reporting, are not sufficient.
- **More clarity and structure around how collaborations between regulators and the research community will take place** – both around specific topics of research and enforcement, and also to ensure the desired “learning together” of DSA implementation is efficient and effective.
- **Clarity on how the concept of “systemic risk” may be iterated and clarified over time** – e.g. through guidelines, Article 40 decisions by DSCs, etc. Also, process to ensure that these iterations:
 - will have external scrutiny such that – in addition to questions around legal clarity and safeguards against misuse – the **concept facilitates rather than excludes a diverse research ecosystem.**
 - Account for the **practical and process issues** outlined in this document – not focus solely on conceptual issues of what a “systemic risk” is.



The ultimate aim of the above should be providing clarity and transparency, facilitate resourcing and information-sharing, and give a clearer plan for effective learning-by-doing; while still leaving space for a pluralistic research ecosystem to support identifying risks, developing mitigations, and (if necessary) enforcement actions.



Acknowledgements

This document was composed from discussions and works from numerous experts, and we thank them for their time and input.

From within AlgorithmWatch input on the underlying ideas, research into the cases, and producing the text, has been particularly provided by Dr. Michele Loi, research scientist and collaborator on the “Auditing Algorithms for Systemic Risks” project from which this work emerges. Support was also provided by Clara Helming and Kirsten Morehouse.

The individuals and organisations named below supported through either (i) providing views on our risk repository cases, (ii) playing a substantial role in events and other discussions related to this work, and/or (iii) providing views on a draft of this document:

Alex Hohlfeld

Ann-Kathrin Watolla, HIIG

Anna-Katharina Meßmer

Eliška Pírková, Access Now

Sofia Calabrese, European Partnership for Democracy

Gemma Galdon Clavell & Eticas

The Global Network Initiative

The Institute for Strategic Dialogue

Julian Jaurisch, Interface

Karolina Iwańska, ECNL

Lena-Maria Böswald, Das Netz

Marie-Therese Sekwenz, TU Delft

Martin Degeling

Orsolya Reich, Liberties

Philipp Darius, Hertie School of Digital Governance

Sally Broughton Micova, University of East Anglia

There are many others who provided support but are not named, potentially for privacy reasons, and we are nonetheless grateful to them. Although we have tried to accurately summarise views expressed to us, this document was composed solely by AlgorithmWatch and should not be seen as a collaborative document, or that all views would be endorsed by all those involved in this work. Also as many of the relevant events above were held under Chatham House and an expectation of caution around information-sharing, we have generally presented amalgamated rather than individually identified views.



Selected Bibliography

[AlgorithmWatch: Ensuring Legitimacy in Stakeholder Engagement: The '5 Es' Framework](#)

[AlgorithmWatch: How to define platforms' systemic risks to democracy](#)

[BSR: A Human Rights-Based Approach to Content Governance](#)

[CERRE: Elements for Effective Systemic Risk Assessment under the DSA](#)

[CERRE: Systemic Risk in Digital Services: Benchmarks for evaluating the management of risks to electoral processes](#)

[ECNL / Access Now: Towards meaningful fundamental rights impact assessments under the DSA](#)

[ECNL / Mozilla: Navigating the Digital Services Act: exploring key elements and scenarios](#)

[Global Network Initiative: Assessment Toolkit](#)

[Global Network Initiative: Implementing Risk Assessments Under the DSA](#)

[Husovec, Martin, The Digital Service Act's Red Line: What the Commission Can and Cannot Do About Disinformation.](#)

[Integrity Institute: On Risk Assessment and Mitigation for Algorithmic Systems](#)

[Jaurisch, Julian Here is why Digital Services Coordinators should establish strong research and data units](#)

[Leerssen, Paddy: Digital Services Act: Summary report on the call for evidence on the Delegated Regulation on data access](#)

[Leerssen, Paddy, Outside the Black Box: From Algorithmic Transparency to Platform Observability in the Digital Services Act,](#)

[Liberties / EPD: Identifying, analysing, assessing and mitigating potential negative effects on civic discourse and electoral processes](#)

[Shiffman *et. al.* Burden of Proof: Lessons Learned for Regulators from the Oversight Board's Implementation Work](#)

[Panoptikon / Irish Council for Civil Liberties / People vs Big Tech: Fixing Recommender Systems: From identification of risk factors to meaningful transparency and mitigation](#)



Ruscheimer, Hannah in: Spindler/Schuster/Kaesling, Recht der elektronischen Medien, 5th Edition, Art. 33 DSA, forthcoming

[Stiftung Neue Verantwortung / Interface: Auditing Recommender Systems: Putting the DSA into practice with a risk-scenario-based approach](#)

[UNGP: Guide to Human Rights Impact Assessment and Management](#)

[Vergnolle, Suzanne, Putting collective intelligence to the enforcement of the Digital Services Act](#)

[Wagner, Ben and Kettmann, Matthias C. and Tiedeke, Anna Sophia and Rachinger, Felicitas and Sekwenz, Marie-Therese, Mapping Interpretations of the Law in Online Content Moderation in Germany.](#)

[Weizenbaum Institut: What the Scientific Community Needs from Data Access under Art. 40 DSA](#)

[Who Targets Me, Disinformation ads, Systemic Risks and the Digital Services Act](#)



Annex A: Selected “Grey Area” Risk Cases

Cases of “potential systemic risks” were real observations related to one or more VLOP/SEs which could impact people within the European Union and were potentially related to one of the risk areas outlined in Article 34(1). “Grey area” cases had features which, in previous work and discussions on systemic risks in the DSA, have been raised as potential indicators of a risk being “systemic” (e.g. happening at scale, repeatability, closely related to the design of a system), but also features which raised questions around evidence quality, trade-offs between different risks and responsibilities, and the point at which a risk might become “systemic”.

Title / Theme	Platform(s)	Summary	Source(s)
Slovakian Election Deep Fakes	Facebook, Instagram	<p>Deep faked audio clips circulating on Facebook and Instagram during Slovakian election.</p> <p>Clips purported to be of Michal Šimečka, leader of Progressive Slovakia party, and Monika Tódová from the daily newspaper Denník N. discussing how to rig the election, partly by buying votes from the country's marginalized Roma minority. Clips remained up but were labelled by fact checkers after verification by local partners (Demagog), a process which took some hours. There have been claims (see Wired) that Meta's policy had loopholes as it mentioned video but not audio</p>	<p>Wired Demagog Euractiv Criticism in Euractive</p>
Amazon Covid misinformation	Amazon	<p>In November 2023, 71.7% of Amazon France's search results to COVID related search queries contained books by authors known for spreading misinformation. This fraction increased to 91.1% when ranked by decreasing average user ratings. Similarly, in Late September to mid-October 2023, Amazon Belgium's search results for "vaccine" and for "COVID" consistently showcased books by authors questioning the pandemic's reality or severity or delve into conspiracy theories within the top 10 results. On "COVID", 80% of books in the top 10 are either questioning the existence of the pandemic, minimising its health effects or framing it as a conspiracy. The proportion is even more alarming on “vaccine” as 90% of results feature anti-vaccination narratives.</p>	<p>AI Forensics & CheckFirst: The Amazing Library: An Analysis of Amazon's Bookstore Algorithms within the DSA Framework (especially pp.18-21)</p>
Bing Chatbot and election misinformation	Bing	<p>Study of Bing Chat / Microsoft Copilot responses to German and Swiss elections finds that one third of Bing answers to election-related questions contained factu</p>	<p>AlgorithmWatch: Summary and link to full article</p>



		include wrong election dates, outdated candidates, or scandals concerning candidates. .	
Meta Oversight Board ruling on Facebook post targeting transgender people	Facebook	The Oversight Board overturned Meta’s original decision to leave up a Facebook post in which a user targeted transgender people with violent speech advocating for members of this group to commit suicide. The Board finds the post violated both the Hate Speech and Suicide and Self-Injury Community Standards. However, the fundamental issue in this case is not with the policies, but their enforcement. Meta’s repeated failure to take the correct enforcement action, despite multiple signals about the post’s harmful content, leads the Board to conclude the company is not living up to the ideals it has articulated on LGBTQIA+ safety. The Board urges Meta to close enforcement gaps, including by improving internal guidance to reviewers.	Oversight Board Decision
Automated censoring by TikTok	TikTok	<p>TikTok application of their community guidelines is to “ensure any content that may be promoted by our recommendation system is appropriate for a broad audience.” This has led to more removal of content and searches than other platforms, which has been the subject of much criticism and accusations of bias</p> <p>Optional extra: Comment on other platforms' responses to situation in Gaza as described by Atlantic Council or other sources</p> <p>Optional extra: Tageschau reporting found that TikTok systematically uses word filters in Germany. At least 20 words prevent comments from appearing publicly, new research shows. The users don't find out about it.</p>	<p>TikTok section here: https://view.atlanticcouncil.org/social-media-gaza/p/1</p> <p>TikTok response: https://newsroom.tiktok.com/en-us/the-truth-about-tiktok-hashtags-and-content-during-the-israel-amas-war</p> <p>TikTok reporting in Tageschau (DE) https://www.tageschau.de/investigativ/ndr/tik-tok-begriffe-101.html</p>
Academic research suggests that Twitter (X) use predicts substantial changes in well-being, polarization, sense of	X/Twitter	In public debate, Twitter (now X) is often said to cause detrimental effects on users and society. Here we address this research question by querying 252 participants from a representative sample of U.S. Twitter users 5 times per day over 7 days (6,218 observations). Results revealed that Twitter use is related to decreases in well-being, and increases in political polarization, outrage, and sense of belonging over the course of the following 30 minutes. Effect sizes were comparable to the	<p>Paper in Nature</p> <p>Could also contextualise the above in Lorenz-Spreen et. al. meta-study (also in Nature)</p>



belonging, and outrage		effect of social interactions on well-being. These effects remained consistent even when accounting for demographic and personality traits. Different inferred uses of Twitter were linked to different outcomes: passive usage was associated with lower well-being, social usage with a higher sense of belonging, and information-seeking usage with increased outrage and most effects were driven by within-person changes.	
X/Twitter reintroduction of political ads	X/Twitter	X/Twitter has lifted previous bans on political advertising. This is despite criticisms regarding effectiveness of their reporting and moderation, concerns around permissiveness towards borderline content, and potential privacy violations around gathering of political views.	Criticisms of reporting/moderation of ads: Politico Accusations of privacy violations: NOYB



Annex B: Relevant Text from the DSA

Article 34: Risk assessment

1. Providers of very large online platforms and of very large online search engines shall diligently identify, analyse and assess any systemic risks in the Union stemming from the design or functioning of their service and its related systems, including algorithmic systems, or from the use made of their services.

They shall carry out the risk assessments by the date of application referred to in Article 33(6), second subparagraph, and at least once every year thereafter, and in any event prior to deploying functionalities that are likely to have a critical impact on the risks identified pursuant to this Article. This risk assessment shall be specific to their services and proportionate to the systemic risks, taking into consideration their severity and probability, and shall include the following systemic risks:

- a. the dissemination of illegal content through their services;
 - b. any actual or foreseeable negative effects for the exercise of fundamental rights, in particular the fundamental rights to human dignity enshrined in Article 1 of the Charter, to respect for private and family life enshrined in Article 7 of the Charter, to the protection of personal data enshrined in Article 8 of the Charter, to freedom of expression and information, including the freedom and pluralism of the media, enshrined in Article 11 of the Charter, to non-discrimination enshrined in Article 21 of the Charter, to respect for the rights of the child enshrined in Article 24 of the Charter and to a high-level of consumer protection enshrined in Article 38 of the Charter;
 - c. any actual or foreseeable negative effects on civic discourse and electoral processes, and public security;
 - d. any actual or foreseeable negative effects in relation to gender-based violence, the protection of public health and minors and serious negative consequences to the person's physical and mental well-being.
2. When conducting risk assessments, providers of very large online platforms and of very large online search engines shall take into account, in particular, whether and how the following factors influence any of the systemic risks referred to in paragraph 1:
 - a. the design of their recommender systems and any other relevant algorithmic system;
 - b. their content moderation systems;
 - c. the applicable terms and conditions and their enforcement;
 - d. systems for selecting and presenting advertisements;



- e. data related practices of the provider.

The assessments shall also analyse whether and how the risks pursuant to paragraph 1 are influenced by intentional manipulation of their service, including by inauthentic use or automated exploitation of the service, as well as the amplification and potentially rapid and wide dissemination of illegal content and of information that is incompatible with their terms and conditions.

The assessment shall take into account specific regional or linguistic aspects, including when specific to a Member State.

3. Providers of very large online platforms and of very large online search engines shall preserve the supporting documents of the risk assessments for at least three years after the performance of risk assessments, and shall, upon request, communicate them to the Commission and to the Digital Services Coordinator of establishment.

Article 35: Mitigation of risks

1. Providers of very large online platforms and of very large online search engines shall put in place reasonable, proportionate and effective mitigation measures, tailored to the specific systemic risks identified pursuant to Article 34, with particular consideration to the impacts of such measures on fundamental rights. Such measures may include, where applicable:
 - a. adapting the design, features or functioning of their services, including their online interfaces;
 - b. adapting their terms and conditions and their enforcement;
 - c. adapting content moderation processes, including the speed and quality of processing notices related to specific types of illegal content and, where appropriate, the expeditious removal of, or the disabling of access to, the content notified, in particular in respect of illegal hate speech or cyber violence, as well as adapting any relevant decision-making processes and dedicated resources for content moderation;
 - d. testing and adapting their algorithmic systems, including their recommender systems;
 - e. adapting their advertising systems and adopting targeted measures aimed at limiting or adjusting the presentation of advertisements in association with the service they provide;
 - f. reinforcing the internal processes, resources, testing, documentation, or supervision of any of their activities in particular as regards detection of systemic risk;



- g. initiating or adjusting cooperation with trusted flaggers in accordance with Article 22 and the implementation of the decisions of out-of-court dispute settlement bodies pursuant to Article 21;
 - h. initiating or adjusting cooperation with other providers of online platforms or of online search engines through the codes of conduct and the crisis protocols referred to in Articles 45 and 48 respectively;
 - i. taking awareness-raising measures and adapting their online interface in order to give recipients of the service more information;
 - j. taking targeted measures to protect the rights of the child, including age verification and parental control tools, tools aimed at helping minors signal abuse or obtain support, as appropriate;
 - k. ensuring that an item of information, whether it constitutes a generated or manipulated image, audio or video that appreciably resembles existing persons, objects, places or other entities or events and falsely appears to a person to be authentic or truthful is distinguishable through prominent markings when presented on their online interfaces, and, in addition, providing an easy to use functionality which enables recipients of the service to indicate such information.
2. The Board, in cooperation with the Commission, shall publish comprehensive reports, once a year. The reports shall include the following:
 - a. identification and assessment of the most prominent and recurrent systemic risks reported by providers of very large online platforms and of very large online search engines or identified through other information sources, in particular those provided in compliance with Articles 39, 40 and 42;
 - b. best practices for providers of very large online platforms and of very large online search engines to mitigate the systemic risks identified.

Those reports shall present systemic risks broken down by the Member States in which they occurred and in the Union as a whole, as applicable.

3. The Commission, in cooperation with the Digital Services Coordinators, may issue guidelines on the application of paragraph 1 in relation to specific risks, in particular to present best practices and recommend possible measures, having due regard to the possible consequences of the measures on fundamental rights enshrined in the Charter of all parties involved. When preparing those guidelines the Commission shall organise public consultations.



Article 42(4)

Providers of very large online platforms or of very large online search engines shall transmit to the Digital Services Coordinator of establishment and the Commission, without undue delay upon completion, and make publicly available at the latest three months after the receipt of each audit report pursuant to Article 37(4):

- (a) a report setting out the results of the risk assessment pursuant to Article 34;
- (b) the specific mitigation measures put in place pursuant to Article 35(1);
- (c) the audit report provided for in Article 37(4);
- (d) the audit implementation report provided for in Article 37(6);
- (e) where applicable, information about the consultations conducted by the provider in support of the risk assessments and design of the risk mitigation measures.

Recitals 79-90

(79)

Very large online platforms and very large online search engines can be used in a way that strongly influences safety online, the shaping of public opinion and discourse, as well as online trade. The way they design their services is generally optimised to benefit their often advertising-driven business models and can cause societal concerns. Effective regulation and enforcement is necessary in order to effectively identify and mitigate the risks and the societal and economic harm that may arise. Under this Regulation, providers of very large online platforms and of very large online search engines should therefore assess the systemic risks stemming from the design, functioning and use of their services, as well as from potential misuses by the recipients of the service, and should take appropriate mitigating measures in observance of fundamental rights. In determining the significance of potential negative effects and impacts, providers should consider the severity of the potential impact and the probability of all such systemic risks. For example, they could assess whether the potential negative impact can affect a large number of persons, its potential irreversibility, or how difficult it is to remedy and restore the situation prevailing prior to the potential impact.



(80)

Four categories of systemic risks should be assessed in-depth by the providers of very large online platforms and of very large online search engines. A first category concerns the risks associated with the dissemination of illegal content, such as the dissemination of child sexual abuse material or illegal hate speech or other types of misuse of their services for criminal offences, and the conduct of illegal activities, such as the sale of products or services prohibited by Union or national law, including dangerous or counterfeit products, or illegally-traded animals. For example, such dissemination or activities may constitute a significant systemic risk where access to illegal content may spread rapidly and widely through accounts with a particularly wide reach or other means of amplification. Providers of very large online platforms and of very large online search engines should assess the risk of dissemination of illegal content irrespective of whether or not the information is also incompatible with their terms and conditions. This assessment is without prejudice to the personal responsibility of the recipient of the service of very large online platforms or of the owners of websites indexed by very large online search engines for possible illegality of their activity under the applicable law.

(81)

A second category concerns the actual or foreseeable impact of the service on the exercise of fundamental rights, as protected by the Charter, including but not limited to human dignity, freedom of expression and of information, including media freedom and pluralism, the right to private life, data protection, the right to non-discrimination, the rights of the child and consumer protection. Such risks may arise, for example, in relation to the design of the algorithmic systems used by the very large online platform or by the very large online search engine or the misuse of their service through the submission of abusive notices or other methods for silencing speech or hampering competition. When assessing risks to the rights of the child, providers of very large online platforms and of very large online search engines should consider for example how easy it is for minors to understand the design and functioning of the service, as well as how minors can be exposed through their service to content that may impair minors' health, physical, mental and moral development. Such risks may arise, for example, in relation to the design of online interfaces which intentionally or unintentionally exploit the weaknesses and inexperience of minors or which may cause addictive behaviour.



(82)

A third category of risks concerns the actual or foreseeable negative effects on democratic processes, civic discourse and electoral processes, as well as public security.

(83)

A fourth category of risks stems from similar concerns relating to the design, functioning or use, including through manipulation, of very large online platforms and of very large online search engines with an actual or foreseeable negative effect on the protection of public health, minors and serious negative consequences to a person's physical and mental well-being, or on gender-based violence. Such risks may also stem from coordinated disinformation campaigns related to public health, or from online interface design that may stimulate behavioural addictions of recipients of the service.

(84)

When assessing such systemic risks, providers of very large online platforms and of very large online search engines should focus on the systems or other elements that may contribute to the risks, including all the algorithmic systems that may be relevant, in particular their recommender systems and advertising systems, paying attention to the related data collection and use practices. They should also assess whether their terms and conditions and the enforcement thereof are appropriate, as well as their content moderation processes, technical tools and allocated resources. When assessing the systemic risks identified in this Regulation, those providers should also focus on the information which is not illegal, but contributes to the systemic risks identified in this Regulation. Such providers should therefore pay particular attention on how their services are used to disseminate or amplify misleading or deceptive content, including disinformation. Where the algorithmic amplification of information contributes to the systemic risks, those providers should duly reflect this in their risk assessments. Where risks are localised or there are linguistic differences, those providers should also account for this in their risk assessments. Providers of very large online platforms and of very large online search engines should, in particular, assess how the design and functioning of their service, as well as the intentional and, oftentimes, coordinated manipulation and use of their services, or the systemic infringement of their terms of service, contribute to such risks. Such risks may arise, for example, through the inauthentic use of the service, such as the creation of fake accounts, the use of bots or deceptive use of a service, and other automated or partially automated behaviours, which may lead to the rapid and widespread



dissemination to the public of information that is illegal content or incompatible with an online platform's or online search engine's terms and conditions and that contributes to disinformation campaigns.

(85)

In order to make it possible that subsequent risk assessments build on each other and show the evolution of the risks identified, as well as to facilitate investigations and enforcement actions, providers of very large online platforms and of very large online search engines should preserve all supporting documents relating to the risk assessments that they carried out, such as information regarding the preparation thereof, underlying data and data on the testing of their algorithmic systems.

(86)

Providers of very large online platforms and of very large online search engines should deploy the necessary means to diligently mitigate the systemic risks identified in the risk assessments, in observance of fundamental rights. Any measures adopted should respect the due diligence requirements of this Regulation and be reasonable and effective in mitigating the specific systemic risks identified. They should be proportionate in light of the economic capacity of the provider of the very large online platform or of the very large online search engine and the need to avoid unnecessary restrictions on the use of their service, taking due account of potential negative effects on those fundamental rights. Those providers should give particular consideration to the impact on freedom of expression.

(87)

Providers of very large online platforms and of very large online search engines should consider under such mitigating measures, for example, adapting any necessary design, feature or functioning of their service, such as the online interface design. They should adapt and apply their terms and conditions, as necessary, and in accordance with the rules of this Regulation on terms and conditions. Other appropriate measures could include adapting their content moderation systems and internal processes or adapting their decision-making processes and resources, including the content moderation personnel, their training and local expertise. This concerns in particular the speed and quality of



processing of notices. In this regard, for example, the Code of conduct on countering illegal hate speech online of 2016 sets a benchmark to process valid notifications for removal of illegal hate speech in less than 24 hours. Providers of very large online platforms, in particular those primarily used for the dissemination to the public of pornographic content, should diligently meet all their obligations under this Regulation in respect of illegal content constituting cyber violence, including illegal pornographic content, especially with regard to ensuring that victims can effectively exercise their rights in relation to content representing non-consensual sharing of intimate or manipulated material through the rapid processing of notices and removal of such content without undue delay. Other types of illegal content may require longer or shorter timelines for processing of notices, which will depend on the facts, circumstances and types of illegal content at hand. Those providers may also initiate or increase cooperation with trusted flaggers and organise training sessions and exchanges with trusted flagger organisations.

(88)

Providers of very large online platforms and of very large online search engines should also be diligent in the measures they take to test and, where necessary, adapt their algorithmic systems, not least their recommender systems. They may need to mitigate the negative effects of personalised recommendations and correct the criteria used in their recommendations. The advertising systems used by providers of very large online platforms and of very large online search engines can also be a catalyser for the systemic risks. Those providers should consider corrective measures, such as discontinuing advertising revenue for specific information, or other actions, such as improving the visibility of authoritative information sources, or more structurally adapting their advertising systems. Providers of very large online platforms and of very large online search engines may need to reinforce their internal processes or supervision of any of their activities, in particular as regards the detection of systemic risks, and conduct more frequent or targeted risk assessments related to new functionalities. In particular, where risks are shared across different online platforms or online search engines, they should cooperate with other service providers, including by initiating or joining existing codes of conduct or other self-regulatory measures. They should also consider awareness-raising actions, in particular where risks relate to disinformation campaigns.

(89)

Providers of very large online platforms and of very large online search engines should take into account the best interests of minors in taking measures such as adapting the design of



their service and their online interface, especially when their services are aimed at minors or predominantly used by them. They should ensure that their services are organised in a way that allows minors to access easily mechanisms provided for in this Regulation, where applicable, including notice and action and complaint mechanisms. They should also take measures to protect minors from content that may impair their physical, mental or moral development and provide tools that enable conditional access to such information. In selecting the appropriate mitigation measures, providers can consider, where appropriate, industry best practices, including as established through self-regulatory cooperation, such as codes of conduct, and should take into account the guidelines from the Commission.

(90)

Providers of very large online platforms and of very large online search engines should ensure that their approach to risk assessment and mitigation is based on the best available information and scientific insights and that they test their assumptions with the groups most impacted by the risks and the measures they take. To this end, they should, where appropriate, conduct their risk assessments and design their risk mitigation measures with the involvement of representatives of the recipients of the service, representatives of groups potentially impacted by their services, independent experts and civil society organisations. They should seek to embed such consultations into their methodologies for assessing the risks and designing mitigation measures, including, as appropriate, surveys, focus groups, round tables, and other consultation and design methods. In the assessment on whether a measure is reasonable, proportionate and effective, special consideration should be given to the right to freedom of expression.