CONSULTATION

Guidelines for Providers of Very Large Online Platforms and Very Large Online Search Engines on the Mitigation of Systemic Risks for Electoral Processes

DISCLAIMER: This provisional draft is for public consultation and contains specific questions where feedback from all relevant stakeholders is sought. It will be further developed, based on the input received during the consultation.

1. INTRODUCTION

1.1. Purpose and Legal Basis

(1) Online platforms and search engines have become important venues for civic discourse and for shaping public opinion and voter behaviour. Regulation (EU) 2022/2065 ("Digital Services Act", or "DSA") imposes obligation on providers of very large online platforms (VLOPs) and very large online search engines (VLOSEs), i.e. those with more than 45 million average monthly active recipients of their service in the European Union, to carry out specific risk assessments and put in place reasonable, proportionate and effective risk mitigation measures including for "*any actual or foreseeable negative effects on civic discourse and electoral processes*" ([1]).

(2) Pursuant to Article 35(3) of Regulation (EU) 2022/2065, the Commission may issue guidelines on the risk mitigation measures providers of VLOPs and VLOSEs are required to adopt in relation to specific risks. Such guidelines may, in particular, 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.

(3) The dissemination of illegal hate speech, foreign information manipulation and interference (FIMI) and disinformation, and content generated through new technologies such as generative AI) ([2]) on online platforms and online search engines gives rise to heightened risks to election integrity. In view of several elections planned in the Union in months to come, including the upcoming 2024 elections to the European Parliament, this document contains guidance aimed at supporting providers of VLOPs and VLOSEs to ensure that they comply with their obligation to mitigate specific risks linked to electoral processes. That guidance remains valid even after those elections have taken place.

(4) These guidelines build upon readiness dialogues on election integrity carried out by the Commission with several providers of VLOPs and VLOSEs after

[1] Article 34(1)(c) of the DSA.

[2] Artificial intelligence capable of generating text, images, or other media, using generative models.

Regulation (EU) 2022/2065 entered into application for the first 19 designated services at the end of August 2023 [3].

(5) To the extent relevant for compliance of VLOPs and VLOSE with Regulation (EU) 2022/2065, the guidelines also reflect several commitments and measures to reduce the spread of online disinformation contained in the Code of Practice on Disinformation [4], the first worldwide industry-led framework in the digital field and the source of industry best practices to address disinformation. They also take into account the work done by the Union's institutions and the Member States on foreign information manipulation and interference (FIMI), notably the comprehensive framework provided by the EU FIMI Toolbox and the recent European External Action Service (EEAS) Report on FIMI Threats [5] focusing on responses to FIMI also in the context of elections. In addition, the European Union and the United States have taken a number of actions to increase transatlantic cooperation to proactively address FIMI and disinformation, including by adopting a common standard for exchanging structured threat information on FIMI, under the EU-US Trade and Technology Council [6]. Finally, these guidelines take into account the forthcoming obligations imposed on providers of VLOPs and VLOSE by the Regulation on transparency of political advertising [7], and the forthcoming Regulation laying down harmonised rules on AI (AI Act) [8] both of which are in the process of adoption by the Union legislator, as well as the voluntary commitments undertaken by VLOPs and VLOSEs under the AI Pact to adhere to the obligations laid down in the AI Act prior to its entry into application [9].

1.2. Outline

(6) The structure of these guidelines is as follows:

a. Section 1 sets out the purpose and structure of the guidelines, as well as references to relevant initiatives;

b. Section 2 sets out the scope of these guidelines;

c. Section 3 sets out the main mitigation measures the Commission proposes providers of VLOPs and VLOSEs to adopt to address election-related systemic risks. Specific subsections cover: the identification of election-related systemic risks; the main mitigations measures to address those risks; specific mitigations

[3] The Commission organised ad-hoc meetings with providers of VLOPs and VLOSEs, both in bilateral settings, as well as in the presence of national authorities, where elections were taking place to gather information on existing practices and ad-hoc policies in place to address elections-related risks.

[4] https://disinfocode.eu/introduction-to-the-code/

[5] https://www.eeas.europa.eu/eeas/2nd-eeas-report-foreign-information-manipulation-and-interference-threats_en

[6] https://www.eeas.europa.eu/sites/default/files/documents/2023/Annex%203%20-%20FIMI_29%20May.docx.pdf

[7] https://ec.europa.eu/commission/presscorner/detail/en/ip_23_4843

[8] https://digital-strategy.ec.europa.eu/en/news/commission-welcomes-political-agreement-artificial-intelligence-act

[9] https://digital-strategy.ec.europa.eu/en/policies/ai-pact

measures linked to generative AI content; cooperation with authorities and other stakeholders; the process of putting into place risk mitigation measures before or after an electoral event; and specific guidance for elections to the European Parliament;

d. Section 4 sets out the general modalities for a dialogue with the Commission on systemic risks for electoral processes;

e. Section 5 sets out that these guidelines will be reviewed one year from adoption.

2. SCOPE OF THESE GUIDELINES

(7) The guidelines are addressed to providers of VLOPs and VLOSEs whose services create a risk of actual or foreseeable negative effects on electoral processes stemming from the design, functioning, and use of those services within the meaning of Article 34 of Regulation (EU) 2022/2065. Pursuant to Article 35(1) of that regulation, providers of VLOPs and VLOSEs shall put in place reasonable, proportionate, and effective mitigation measures, tailored to the specific systemic risks identified.

(8) Article 35 of Regulation (EU) 2022/2065 provides a non-exhaustive list of mitigation measures that providers of VLOPs and VLOSEs may adopt to address the systemic risks they identify in the risk assessment process to which their systems give rise. These guidelines further elaborate on that list and recommend best practices for mitigating risks related to electoral processes.

(9) In line with recital 103 the preamble to Regulation (EU) 2022/2065, these guidelines may also serve as a source of inspiration for providers of online platforms or search engines that have not been designated as a VLOP or VLOSE and whose services give rise to similar risks. It may also serve as a reference for the continuous research into, and analysis of, the effectiveness of risk mitigation measures in response to risks related to electoral processes.

(10) Where the mitigation measures and best practices recommended in these guidelines are of application to electoral processes in general, providers of VLOPs and VLOSEs should consider keeping these measures and practices in place on a continuous basis, and not only during electoral periods.

Q1: Are there any documents, reports, guidelines, academic studies or relevant independent research you recommend as further input for these guidelines?

Q2: How can the Commission further clarify the purpose and scope of these guidelines to better address systemic risks in electoral processes?

3. ELECTION SPECIFIC RISK MITIGATION MEASURES

3.1. Identification of systemic risks related to electoral processes

(11) For the identification and subsequent design of reasonable, proportionate, and effective mitigation measures, providers of VLOPs and VLOSEs should consider **reinforcing internal processes** in line with Article 35(1)(f) of Regulation (EU) 2022/2065. To tailor their mitigation measures to identified risks to electoral processes, those providers should consider including an **election-specific risk profile** in their internal processes for detecting systemic risks. As part of that risk profile, providers are encouraged to collect information on elements such as the presence and activity of political actors on the service, relevant discussions on and usage of the platform in the context of elections, the number of users in a Member State when a particular election is called in that Member State, and indications of previous instances of coordinated or intentional information manipulation. While conducting this risk profile providers of VLOPs and VLOSEs should ensure compliance with relevant data protection legislation ([10]). Providers of VLOPs or VLOSEs should also assess whether (part of) their service is being used to search, share, or access information on elections and electoral processes, political parties or candidates, party programmes, manifestos or other political material, or related information, to organise events such as demonstrations or rallies, conduct activism, fundraising or other related political activities, and assess any actual or foreseeable risks that follow from this.

(12) To adequately reinforce their internal processes to mitigate systemic risks effectively and to tailor their mitigation measures to those risks, providers of VLOPs and VLOSEs are encouraged to collect and analyse information on **local context-specific risks and Member State specific information** at the national, regional and/or local level. This will require having adequate content moderation resources with local language capacity and knowledge of the national and/or regional contexts and specificities. The Commission also recommends those providers to perform an analysis of the state of media freedom and pluralism, reference to media literacy initiatives and indicators, and information on the existence of an enabling space for civil society organisations to participate in policy-making and civic discourse. The capacity of all relevant mitigation measures to perform effectively in the local linguistic and electoral context should also be considered. Continuous engagement with local independent civil society organisations, researchers and fact-checkers are essential inputs for such an analysis. European Digital Media Observatory (EDMO) hubs covering the entire EU are an important resource in this regard, since they assemble fact-checkers, researchers and media literacy specialists with local expertise and on the ground experience.

3.2. Elections-specific risk mitigation measures

(13) To reinforce internal processes and resources in a particular electoral context, providers of VLOPs and VLOSEs should consider setting up **a clearly identifiable internal team** prior to each individual electoral period (see also

[10] This includes Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)

section 3.5. 'During an electoral period'). The resource allocation for that team should be proportionate to the risk profile identified for the election in question, including being staffed by persons with country-specific expertise, such as local contextual and language knowledge. The team should cover all relevant expertise including in areas such as content moderation, threat disruption, cybersecurity, FIMI and disinformation, fundamental rights and public participation and cooperate with relevant external experts, for example with EDMO hubs and independent fact-checking organisations ([11]).

(14) Considering the important role in judging the veracity of information that such organisations have, the Commission recommends that providers of VLOPs and VLOSEs collaborate with independent fact-checking organisations that adhere to high standards of methodology, ethics and transparency, for example by being a member of the European Fact-Checking Standards Network (EFCSN) and following its Code of Standards ([12]).

(15) Mitigation measures could draw, in particular, on industry standards established through the Code of Practice on Disinformation, other relevant EU industry codes, such as the Code of conduct on countering hate speech online, and from existing best practices such as those documented in the Content-Agnostic Election Integrity Framework for Online Platforms ([13]) and the Election Integrity Programme of the Integrity Institute ([14]) as well as recommendations from civil society, such as those from the Civil Liberties Union for Europe and European Partnership for Democracy ([15]).

(16) Specifically, mitigation measures aimed at addressing systemic risks to electoral processes could include measures in the following areas:

a. **Access to official information on the electoral process**: To improve voter turnout and prevent the spread of mis- and disinformation and FIMI on the electoral process itself, providers of VLOPs and VLOSEs could point their users to official information about the electoral process, including information on how and where to vote. Any information displayed should always stem from the competent electoral authorities of the Member States concerned. Such information could be provided for example by means of information panels, banners, pop-ups, search interventions, links to websites of the electoral authority, specific election information tabs or a dedicated part of the platform. When designing such mitigation measures, the Commission recommends that providers of VLOPs and VLOSEs take principles such as inclusiveness and accessibility into account.

[11] See section 3.4. on 'Cooperation with national authorities, independent experts and civil society organisations

[12] EFCSN | European Fact-Checking Standards Network Project – European Fact-Checking Standards Network Project

[13] Democracy-By-Design.pdf (accountabletech.org)

[14] Elections Program — Integrity Institute

[15] DSA: New Risk Assessments To Protect Civic Discourse and Electoral Processes | liberties.eu

b. **Media literacy initiatives**: Providers of VLOPs and VLOSEs are encouraged to collaborate implement, invest and engage in media literacy initiatives and campaigns focussing on elections to foster critical thinking and help to improve the skills required, e.g. to analyse complex realities and recognise the difference between opinion and fact as well as the risks related to generative AI. This could be achieved by:

   i. Collaborating with **local media literacy organisations**, financially supporting, sharing and integrating election related initiatives and campaigns on the platform, including by developing joint initiatives. Local media literacy organisations already have knowledge of local contexts and target audiences. The Commission recommends the use of the network of EDMO and its hubs and the Commission's Expert Group on Media Literacy to find the relevant organisations at Member State level.

   ii. Developing and applying **inoculation measures** that pre-emptively build psychological resistance to possible and expected disinformation narratives and manipulation techniques by informing users and preparing them to approach them critically. Inoculation measures can take different forms, including e.g., gamified interventions ([16]), video or other types of content ([17]).

   iii. Taking into account specific narratives as well as tactics, techniques and procedures (TTP's) that are likely to occur during the election in question in the Member State concerned when designing media literacy campaigns in line with the approach of adapting mitigation measures to the **relevant national context**.

c. **Measures to provide users with more contextual information** on the content and accounts they engage with. Examples include:

   i. **Fact-checking labels** on identified FIMI and disinformation content provided by independent fact-checkers. Fact-checking coverage should extend across the EU and its languages, inter alia through strengthening the cooperation with local fact-checkers during election periods, prioritising the integration of fact-checks related to elections as well as employing mechanisms to help increase the impact of the fact-checks on audiences.

   ii. **Prompts and nudges** urging users to read content and evaluate its accuracy and source before sharing it.

   iii. Clear and visible **indications of verified and official accounts,** as well as accounts providing authoritative information on the electoral process, such as the accounts of electoral authorities.

([16]) Traberg, C. S., Roozenbeek, J., & van der Linden, S. (2022). Psychological Inoculation against Misinformation: Current Evidence and Future Directions. The ANNALS of the American Academy of Political and Social Science, 700(1), 136-151. https://doi.org/10.1177/00027162221087936

([17]) Jon Roozenbeek *et al.*, Psychological inoculation improves resilience against misinformation on social media.*Sci. Adv.***8**,eabo6254(2022).DOI:10.1126/sciadv.abo6254

iv. Clear and visible **labelling of accounts** controlled by Member States, third countries and entities controlled or financed by entities controlled by third countries.

v. **Tools and information to help users assess the trustworthiness** of information sources, such as trust marks focused on the integrity of the source based on transparent methodologies and developed by independent third-parties.

vi. **Other tools to assess the provenance**, edit history, authenticity, or accuracy of digital content. These help users to check the authenticity or identify the provenance or source of content related to elections.

d. **Analysing and appropriately moderating virality** of content that threatens the integrity of the electoral process itself, for example by introducing friction or circuit-breakers, e.g., through providing contextual information, such as fact-checking or warning labels, and by disrupting the algorithmic amplification and spread of such content.

e. **Influencers** can have a significant impact on the electoral choices made by recipients of the service, as they get increasingly involved in political debates. To mitigate the risk that such activities may represent in relation to electoral processes, providers of VLOPs and VLOSEs should consider:

i. Providing a functionality to allow influencers to declare whether the content they provide is or contains political advertising, including the identity of the sponsor;

ii. Ensuring that other recipients of the service can identify in a clear and unambiguous manner and in real time, including through prominent markings, that the content provided is or contains political advertising, as described in the influencer's declaration.

f. **Political advertising**([18]): If a provider of a VLOP or VLOSE offers the possibility to place political advertisements on its service, the Commission recommends that these are **clearly labelled** as such in an efficient and visible way to allow users to understand that the content displayed contains political advertising. The labels applied should remain in place when shared by users on the same platform. Providers of VLOP or VLOSE are advised to prepare for the entry into application of the regulation on political advertising and to take particular care to consider the provision which will enter into application early following the publication of the enactment. The Commission also recommends that providers of VLOPS and VLOSEs:

i. Provide users with information about the political advertisements they see (such as sponsor identity, display period, how much was spent) and meaningful information about the targeting of the advertising. The Commission also recommends that relevant providers of VLOPs and

([18]) Section relevant pending formal approval by co-legislators and entry into force of the Regulation on transparency and targeting of political advertising. (press release).

VLOSEs appropriately identify and link political sponsors to the advertising.

ii. Maintain a publicly available, searchable **repository of political ads**, updated in as close as possible to real-time. This should include as a minimum the information required under the Digital Services Act ([19]) and could also include e.g., the amount spent on the ad, the number of impressions and the geographical areas in which the ad was presented.

iii. When they do not allow political advertising on their services, have efficient verification systems in place and take the necessary actions to ensure that the decision is appropriately enforced.

iv. Harmonise their definition of political advertising [with the one set out in the Regulation on transparency and targeting of political advertising]

g. **Demonetisation of disinformation content**: The Commission recommends that providers of VLOPs and VLOSEs have targeted policies in place to ensure that the placement of advertising does not provide financial incentives for the dissemination of FIMI and disinformation around the elections.

h. **Integrity of Service**s: Providers of VLOPs and VLOSEs should put in place appropriate procedures to ensure the timely detection and disruption of coordinated inauthentic manipulation of the service when this has been identified by them as a relevant systemic risk. For example, they may include in their terms and conditions specific rules against the creation of inauthentic accounts or botnets (which may include automated, partially automated, or non-automated accounts), or deceptive use of a service.

i. The Commission recommends that providers of VLOPs and VLOSEs develop and effectively enforce rules preventing deception by impersonation of candidates, the deployment of deceptive manipulated media, the use of fake engagements, non-transparent compensated messages, or non-transparent promotion by influencers, as well as inauthentic coordination of content creation or amplification.

ii. The Commission recommends cooperation between the relevant teams of different providers of VLOPs and VLOSEs to identify common threats and to counter cross-platform influence operations and migration of malicious actors (see section 3.4 on cooperation with national authorities, independent experts and civil society organisations).

(17) Considering the evolving nature of the understanding of online risks to electoral processes, mitigation measures should be tied to **rigorous and critical analysis, testing and review** of their intended and potentially unintended impact. As such, effective mitigation measures, should be based on the best available information and scientific insights. The Commission recommends that providers of VLOPs and VLOSEs pro-actively design, evaluate, and optimise conceptually valid performance metrics for the effectiveness of mitigation measures, for example via A/B testing of feature and design choices. These **performance metrics**

[19] Articles 26 and 39 DSA.

should be analysed as part of providers' risk management framework and set these to measure the success of relevant mitigation measures during a particular election. These metrics should be SMART (specific, measurable, achievable, relevant and time-bound) and they should be both qualitative and quantitative.

(18) **Third party scrutiny and research** into mitigation measures are important to help providers of VLOPs and VLOSEs ensure that the measures they put in place are effective and respect fundamental rights. In addition to their legal obligations under Article 40 of the DSA ([20]), the Commission recommends that providers of VLOPs and VLOSEs work closely with researchers and other relevant stakeholders and take into account their findings when designing and revising their risk mitigation measures. As underlined by the Integrity Institute ([21]), different measures for facilitating research by third parties could be considered. For instance, the provision of specific tools or features may include giving access to data points and keywords beyond those made available pursuant to Article 40(12) of Regulation (EU) 2022/2065, ad hoc cooperation projects with academia or civil society organisations, or expert consultations to gather insights on service-related risks with respect to election integrity.

(19) In the area of political advertising, the Commission recommends that relevant providers of VLOPs and VLOSEs ensure that the tools and application programming interfaces (APIs) enabling research on their political advertising repositories are fit-for-purpose and allow for meaningful research on FIMI and disinformation campaigns during elections, in accordance with the requirements of EU law, including on the protection of personal data. This includes a set of minimum functionalities and search criteria that enable users and researchers to perform customised searches for data in as close to real time as possible during the electoral period (e.g., searches per advertiser or candidate, election, geographic area or country, language).

(20) In addition to the reports referred to in Article 42(4) of Regulation (EU) 2022/2065, the Commission recommends that providers of VLOPs and VLOSEs are as **transparent** as possible to the public about the design, functioning, and execution of mitigation measures related to electoral processes to allow for public scrutiny which in turn may impact the design of effective mitigation measures. During electoral periods, it is of particular importance that providers of VLOPs and VLOSEs show that content moderation decisions do not affect the equality of candidates or disproportionately favour or promote voices representing certain (polarised) views.

(21) Risk mitigation measures, taken in line with Article 35 of Regulation (EU) 2022/2065, should have due regard for any actual or foreseeable negative effects on fundamental rights enshrined in the Charter of Fundamental Rights of the European Union, in particular the right to freedom of expression and of

([20]) Article 40.12 of the DSA already requires providers of designated VLOPs and VLOSEs to give access to eligible researchers to the information that are publicly available on their interface. Article 40(4) provides for a specific data access regime for vetted researchers which will be applicable as the dedicated delegated act will be adopted.

([21]) New Guide Provides Concrete Elections Integrity Recommendations for Online Platforms — Integrity Institute information, including media freedom and pluralism. In line with Recital 47 of that regulation, providers of VLOPs and VLOSEs should pay due regard to relevant international human rights standards such as the United Nations Guiding Principles on Business and Human Rights (UNGPs). Relevant independent reports [22] may also be considered when designing and enforcing the relevant mitigation measures.

(22) When mitigating systemic risks for electoral integrity, the Commission recommends that due regard is also given to the impact of measures to tackle illegal content such as incitement to violence and hatred to the extent that such illegal content may **inhibit or even silence voices representing certain groups in society**, in particular marginalised groups or minorities. In this respect, the Code of conduct on countering hate speech online can inspire action.

(23) In addition to the involvement of relevant actors during the risk assessment, the Commission recommends that providers of VLOPs and VLOSEs make available the **fundamental rights impact assessments**, as referred to in recital 90 of Regulation (EU) 2022/2065, performed as part of the risk assessments, to civil society organisations as soon as they are concluded, i.e. earlier than required under Article 42(4) of that regulation. This could provide a space for constructive open dialogue on possible good practices and potential improvements.

(24) Finally, there are also further mitigation measures conceivable for mitigating risks related to electoral process. In particular, journalists and media service providers perform a vital role in gathering, processing, and reporting fact-checked information to the public, a role that is rendered even more critical during election times, considering also that independent news media service providers and organisations with well-established internal editorial standards and procedures are widely regarded as trusted sources of information. In this context, for instance, the availability of trustworthy information from pluralistic sources is important for well-functioning electoral processes.

Q3: Do you agree with the recommended best practices in this section?

Q4: What additional factors should be taken into account by providers of VLOPs and VLOSEs when detecting systemic risks related to electoral processes??

Q5: Are there additional mitigation measures to be considered as best practices on the basis of their proven effectiveness mitigating risks to electoral processes?

Q6: How should providers of VLOPs and VLOSEs measure effectiveness of their risk mitigation measures in a reliable and conceptually valid way for electoral processes?

[22] Examples include the Access Now and the European Center for Not-for-Profit Law policy paper “Towards meaningful fundamental rights impact assessments under the DSA”, Danish Institute for Human Rights, Guidance on Human Rights Impact Assessment of Digital Activities, Julian Jaursch, Josefine Bahro, Asha Allen, Claire Pershan and Katarzyna Szymielewicz, DSA risk mitigation: Current Practices, ideas and open questions

3.3. Mitigation measures linked to generative AI

(25) Recent technological developments in generative AI have enabled the creation and widespread use of artificial intelligence capable of generating text, images, videos, or other synthetic content. While such developments may bring many new opportunities, they may lead to specific risks in the context of elections. generative AI can notably be used to mislead voters or to manipulate electoral processes by creating and disseminating inauthentic, misleading synthetic content regarding political actors, false depiction of events, election polls, contexts or narratives. Generative AI systems can also produce incorrect, incoherent, or fabricated information, so called "hallucinations", that misrepresent the reality, and which can potentially mislead voters.

(26) Pursuant to Article 35(1) of the DSA, providers of VLOPs and VLOSEs should also put in place reasonable, proportionate, and effective mitigation measures tailored to risks related to both the **creation** [23] and potential large-scale **dissemination** of generative AI content, depending on the nature of their service and the conducted risk assessment. Best practices which may inform the relevant risk mitigation measures may be drawn from the AI Act and the AI Pact, [24], which aims to anticipate the AI Act's early voluntary application before the date of applicable. Particularly relevant in this context are the obligations envisaged in the AI Act for providers of general-purpose AI models, including generative AI, requirements for labelling of 'deep fakes' and for providers of generative AI systems to use technical state-of-the-art solutions to ensure that content created by generative AI is marked as such, which will enable its detection by providers of VLOPs and VLOSEs.

(27) Following from the specific actual or foreseeable risks for electoral processes identified, the Commission recommends that providers of VLOPs and VLOSEs whose services can be used for the **creation** of deceptive, false or misleading generative AI content have the following risk mitigation measures in place:

a) Ensure that generative AI content, other types of synthetic and manipulated media, is clearly distinguishable for users – notably by using **watermarking,** including by relying on added relevant metadata. This is particularly important for any generative AI content involving candidates, politicians, or political parties. Watermarks may also apply to content that is based on real footage (such as videos, images or audio) that has been altered through the use of generative AI.

b) Make reasonable efforts to ensure that generative AI **provided information relies to the extent possible on reliable sources** in the electoral context, such as official information on the electoral process from relevant electoral authorities, and that any quotes or references made by the system to external sources are accurate and do not misrepresent the cited content, thus limiting the effects of 'hallucinations'.

c) Warn users about potential errors in content created by generative AI suggesting them to consult authoritative sources to check the veracity of such

[23] For example, refer to this academic study by Stanford university or this analysis by AlgorithmWatch and AI Forensics.

[24] AI Pact | Shaping Europe's digital future (europa.eu)

information as well as put safeguards in place to prevent the creation of false content that may have a strong potential to influence user behaviour.

d) Conduct and document **red-teaming** exercises with a particular focus on electoral processes, with both internal teams and external experts, before releasing generative AI systems to the public and follow a staggered release approach when doing so to better control unintended consequences.

e) Set appropriate performance metrics, including for safety and factual accuracy of answers given to questions on electoral content, and continually **monitor the performance of generative AI systems**, and take appropriate actions when needed.

f) Integrate into generative AI systems safeguards that increase their **safety**, such as prompt classifiers, content moderation and other filters, to detect and prevent prompts that go against terms of service of the provider of a VLOP or VLOSEs concerning electoral processes; take other appropriate measures that seek to prevent the misuse of the generative AI system for illegal, manipulative and disinformation purposes in the context of electoral processes.

g) For text content: indicate, where possible, in the outputs generated the concrete sources of the information used as input data to enable users to verify the reliability and further contextualise the information.

(28) Following from the specific actual or foreseeable risks for electoral processes identified, the Commission recommends that providers of VLOPs and VLOSEs **whose services can be used to disseminate** deceptive, false or misleading generative AI content consider the following risk mitigation measures:

a) Adapt their terms and conditions and ensure their efficient enforcement, to significantly decrease the reach and impact of generative AI content that falsely depicts disinformation on the electoral process, such as election irregularities.

   i. The Commission recommends that providers of VLOPs and VLOSEs provide clear information on which internal processes and mitigation measures, such as labelling, marking, demoting or removing, are in place to enforce these policies.

   ii. The Commission recommends that providers of VLOPs and VLOSEs cooperate and share information about such deceptive content with fact checkers to ensure that the risk of amplification in other platforms is minimised.

b) **Clearly label,** or otherwise make distinguishable through prominent marking, synthetic or manipulated images, audio or videos that appreciably resemble existing persons, objects, places, entities, events, or depict events as real that did not happen or misrepresent them (**deepfakes**).

   i. The Commission recommends that providers of VLOPs and VLOSEs provide users with standard and easy to use interfaces and tools to add labels to AI generated content.

   ii. When labelling generative AI content, the Commission recommends that providers of VLOPs and VLOSEs apply efficient labels, easily recognised by users, taking into account aspects such as graphics, position and timing, drawing on scientific research on the effectiveness of labels ([25]).

iii. The Commission recommends that providers of VLOPs and VLOSEs make sure the labelled generative AI content retains its label once it is shared by other users on the platform.

c) The Commission recommends that providers of VLOPs and VLOSEs adapt their advertising systems, for example by providing advertisers with options to clearly label content created with **generative AI in advertisements** or promoted posts and require in their advertising policy that this label is used when the advertisement includes generative AI content.

d) To enforce these policies, providers of VLOPs and VLOSEs should adapt their content moderation processes and algorithmic systems in such a way as to detect watermarks and other content provenance indicators.

i. In this context, providers of VLOPs and VLOSEs should cooperate with providers of generative AI systems and follow leading state of the art measures to ensure that such watermarks and indicators are detected in a reliable and effective manner; they are also recommended to support new technology innovations to improve the effectiveness and interoperability of such tools.

e) **Media literacy measures** mentioned in section 2 should also focus on generative AI, for instance to explain how the technology works and the possibilities for its misuse.

(29) Pursuant to Article 35(1) of the DSA, when providers of VLOPs and VLOSEs address legal but harmful forms of generative AI content that can influence voters’ behaviour, they should consider the impact their policies and measures may have on **fundamental rights**, particularly considering the impact it may have on political expression, parody and satire. Such a fundamental rights assessment is in particular required when developing policies on what type of deceptive generative AI content a provider of a VLOP or VLOSE does not allow on their service and will remove from it.

(30) As AI generated content bears specific risks, it should be specifically scrutinised, also through the development of ad hoc tools to perform research aimed at identifying and understanding specific risks related to electoral processes. Providers of online platforms and search engines are encouraged to consider setting up dedicated tools for researchers to get access to and specifically identify and analyse AI generated content that is known as such, in line with the obligation under Article 40.12 for providers of VLOPs and VLOSEs in the DSA.

[25] See for example Tom Dobber, Sanne Kruikemeier, Fabio Votta, Natali Helberger & Ellen P. Goodman (2023) The effect of traffic light veracity labels on perceptions of political advertising source and message credibility on social media, Journal of Information Technology & Politics

Q7: Do you agree with the recommended best practices in this section?

Q8: Which risks of generative AI for electoral processes should additionally be considered in this section?

Q9: What additional evidence-based best practices on risk mitigation for electoral processes related to the creation of generative AI content should be considered?

Q10: What additional evidence-based best practices on risk mitigation for electoral processes related to the dissemination of generative AI content should be considered?

Q11: What are best practices for providers of VLOPs and VLOSEs to ensure that their risk mitigation measures keep up with technological developments and progress?

3.4. Cooperation with national authorities, independent experts and civil society organisations

(31) Contributing to protecting the integrity of a specific election cannot be done without knowledge of the specific national, legal, societal, and political context as well as timely reactions to real-time developments affecting the risk profile on a VLOP's or VLOSE's service. To this end, in the design and implementation of risk mitigation measures related to electoral processes, the Commission recommends that providers of VLOPs and VLOSEs regularly exchange information with and have contact points for responsible national authorities and other local actors to facilitate the escalation of problems and deliberation of solutions. The Digital Service Coordinators designated under the DSA in each Member State may serve as contact point for providers of VLOPs or VLOSEs should it not be clear which is the national authority responsible for elections.

(32) Procedures and organisational structures for elections differ from country to country and even from one election to another. Providers of VLOPs and VLOSEs should know the applicable national **election governance structure** for the elections at hand and the role of various actors. By gaining a good understanding of specific national procedures such as the delimitation of the electoral campaign periods, timing of the official designation of election candidates, and election silence periods providers of VLOPs and VLOSEs may design risk mitigation measures taking into account the specific regional or linguistic aspects of the relevant Member State.

(33) Prior to electoral campaigns, providers of VLOPs and VLOSEs should **establish contacts with relevant national authorities** to foster an efficient exchange of information before, during and after the election. Following from Commission Recommendation (EU) 2023/2829, The Commission recommended to Member States to strengthen their **national election networks** ([26]) and to facilitate their

([26]) https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32018H0234 cooperation with relevant stakeholders ([27]). Relevant providers of VLOPs and VLOSEs could make sure they are connected to these networks and establish two-way communication channels. These contacts, for example with electoral authorities, could also lead to integrating authoritative information about the voting process into the services interface before and during the elections, or can inform the design of other risk mitigating measures.

(34) Especially during an election campaign, the Commission recommends that providers of VLOPs and VLOSEs establish **efficient and timely communication** with the authorities with swift, efficient and appropriate follow-up mechanisms to issues flagged. The information provided by the authorities may be used by the provider of the VLOP or VLOSE in order to assess the mitigation measures and to determine whether additional measures are required. For the sake of efficacy, the Commission recommends that the communication be streamlined via pre-established points of contact (and/or a limited number of points of contact) on both sides ([28]). In order to improve the effectiveness of the mitigation measures taken, providers of VLOPs and VLOSEs should maintain records of their interactions with authorities, including any requests made and actions taken by the companies in response ([29]).

(35) Prior to the elections, providers of VLOPS and VLOSEs may also organise meetings as well as establish channels of regular communication with non-state actors active in electoral processes such as **academics, independent experts, civil society organisations and representatives of various communities**, and invite them to share their independent expertise, insights and observations that can help identify risks that may require mitigation measures and contribute to the development of such mitigation measures.

(36) Establishing channels for communication during the election campaign with non-state actors, including campaign organisations and election observers will help the providers of VLOPs and VLOSEs to better understand the context of the elections to react promptly in emergency situations and understand better how their mitigation measures work in the local context. The Working Group on Elections of the Code of Practice – and its rapid response system - is a good example of such an existing and active multistakeholder forum, including NGOs and fact-checkers with important election specific experience. The EDMO Task-force on Elections – composed of independent fact-checkers, academics, and media literacy specialists – as well as the EDMO hubs across the EU – can also provide important input in this respect.

[27] Commission Recommendation (EU) 2023/2829 of 12 December 2023on inclusive and resilient electoral processes in the Union and enhancing the European nature and efficient conduct of the elections to the European Parliament

[28] Articles 11 and 12 DSA require providers of intermediary services to have a point of contact for the Commission, the Board, national Digital Service Coordinators and recipients of the service.

[29] Article 10 DSA orders to act against illegal content can only be issued by the relevant national judicial or administrative authorities, on the basis of the applicable Union law or national law in compliance with Union law, following separate procedures for this

Q12: Do you agree with the recommended best practices in this section?

Q13: What other mechanisms should be considered to foster more effective collaboration with relevant stakeholders, such as national authorities and civil society organisations?

Q14: Are there any additional resources that could help providers of VLOPS and VLOSEs identify relevant organisations/experts at the national level?

3.5. During an electoral period

(37) Certain risk mitigation measures such as additional internal procedures or dedicated teams may only be needed during a specific electoral period, depending on the risk profile of a given provider and the specificities of a given Member State. Some Member States have a set period of time for election campaigning, while others do not. Therefore, the Commission recommends that providers of VLOPs and VLOSEs **define the period** during which measures and resources would be in place. In line with insights on expected threat progression during elections from the 2nd EEAS Report on Foreign Information Manipulation and Interference (FIMI) ([30]), the Commission recommends that measures are already in place and functioning one to six months before an electoral period, and continue at least one month after the elections, depending on the risk assessment for the particular election.

(38) During the electoral period the Commission recommends that providers of VLOPs and VLOSEs pay specific attention to risk mitigation measures that reduce the impact of incidents that can have a **significant impact on the election outcome or turnout**. Measures to prevent voter suppression, include providing users with access to reliable, timely and intelligible information from official sources on how to vote as well as on the voting process – to pre-empt claims undermining the trust in the electoral system - or measures like those mentioned in section 3.3 to reduce the potential harm of high impact issues such as manipulated images, voice recordings or deepfakes, for example of political actors contending in elections.

(39) Elections are high impact events where incidents occurring on- or off-platform during an electoral period can have rapid consequences for the integrity of elections or public security. The Commission recommends, as a result, that providers of VLOPs and VLOSEs put in place **an incident response mechanism**, involving also the senior leadership, as well as a mapping of the stakeholders involved in responding to the incident. This procedure needs to be set-up, agreed-upon and tested, including through red teaming exercises, beforehand so it can be applied quickly.

(40) The rapid response system to be established by the signatories of the Code of Practice on Disinformation also provides a good example and useful forum for cooperation during elections, feeding into the platforms incident response mechanisms. The signatories should set out the procedural framework for

[30] EEAS-2nd-Report on FIMI Threats-January-2024_0.pdf (europa.eu)

cooperation and coordination between them during elections, including a rapid feedback mechanism with the need of swift, efficient, and appropriate follow-up by platforms. Another example on how to organise the work on responses to FIMI and disinformation can be found in the second EEAS Report on FIMI Threats [31] which puts forward a "Response Framework" linking effectively analysis to evidence-based responses and highlighting the importance of cooperation between various stakeholders.

(41) A **timely response** to incidents is often key. The Commission recommends that providers of VLOPs and VLOSEs consider a 'follow the sun' model in which offices around the world would be able to cover all time zones.

(42) To react in a timely matter, providers of VLOPs and VLOSEs should integrate their possible collaboration with electoral authorities in incident response mechanisms. These recommendations complement Article 84 of the DSA on crisis protocols.

Q15: Do you agree with the recommended best practices in this section?

Q16: Are there any additional measures that providers of VLOPs and VLOSEs should take specifically during an electoral period?

Q17: How can rapid response mechanisms be improved for handling election-related incidents on VLOPs or VLOSEs?

Q18: What other mechanisms should be considered to foster more effective collaboration with national authorities and civil society organizations?

Q19: Are there any additional resources that help providers of VLOPS and VLOSEs identify relevant organisations/experts at the national level?

3.6. After an electoral period

(43) After an electoral period, the Commission recommends that providers of VLOPs and VLOSEs conduct a post-election review including an assessment of the effectiveness of the risk mitigation measures employed in that context with a view to adapting the measures, if necessary. This internal report is recommended to include an assessment of whether the internal performance metrics and any other assessment criteria were met, lessons learnt and possible areas for improvement.

(44) The Commission recommends providers of VLOPs and VLOSEs to take into account specific contributions from independent researchers on the impact of VLOPs and VLOSEs mitigation measures in the election review exercise. In addition, providers of VLOPs and VLOSEs may engage with credible and independent election observer groups who may be able to provide information on the use and impact of their services in that context.

[31] EEAS-2nd-Report on FIMI Threats-January-2024_0.pdf (europa.eu)

(45) In particular, the report should include information on the average response time for terms and conditions violations, the number of violations of certain policies pertaining to elections, instances of information manipulation and the reach of certain measures such as media literacy initiatives and authoritative initiatives. Such reports can be shared in a confidential manner with the Commission ([32]).

(46) The Commission recommends providers of VLOPs and VLOSEs publish a public version of such post-election review documents. This can include information on actions taken by the provider of a VLOP or VLOSE and any incidents that might have occurred aiming at gathering public feedback on how to improve the risk mitigation measures in place or share successful measures with other providers. As a further example for such post-election reporting, signatories of the Code of Practice on Disinformation have developed a reporting template through that they will report ahead and after elections on their measures taken and relevant metrics regarding their impact.

Q20: Do you agree with the recommended best practices in this section?

Q21: What elements should be included in voluntary post-election review by providers of VLOPs or VLOSEs to assess the effectiveness of their risk mitigation strategies?

3.7. Specific guidance for the elections to the European Parliament

(47) As stated in the Communication on the Defence of Democracy package ([33]) the upcoming elections to the European Parliament will be a crucial test case for the resilience of our democratic processes. In that context, and due to their unique cross-border nature, providers of VLOPs and VLOSEs are expected to put in place robust preparations for the elections to the European Parliament specifically, for example those taking place from 6 to 9 June 2024.

(48) This means that providers of VLOPs and VLOSEs have to make sure **sufficient resources** and risk mitigation measures are available and these are **distributed** in a way that is proportionate to the current risk assessments and include access to relevant local expertise across all EU Member States.

(49) For elections to the European Parliament no predetermined campaigning period exists. This means that in Member States the campaigns for these elections can start at different points in time. Providers of VLOPs and VLOSEs are encouraged to take this into account when planning their risk mitigation measures for election to the European Parliament.

([32]) DSA Article 84

([33]) COM(2023) 630 final. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions on Defence of Democracy

(50) Providers of VLOPs and VLOSEs should also take into account the **unique cross-border and European dimension** of these elections, when assigning appropriate risk mitigating resources. Not only can individual Member States be targeted by malicious actors, but also the EU institutions. In addition to establishing contact with the relevant national authorities, the Commission recommends that providers of VLOPs and VLOSE establish contact with Union-level authorities before elections to the European Parliament. EU-wide networks of national experts in the areas of FIMI and disinformation, elections and cybersecurity such as the EU Rapid Alert System, European Cooperation Network on Elections and the NIS Cooperation Group can be useful networks for providers of VLOPs and VLOSEs in case of cross-border incidents during the electoral period that require a rapid response and deployment of risk mitigating measures. Furthermore, the Commission recommends establishing contact with the European Parliament’s administration and European political parties before elections to the European Parliament, similarly to what was proposed above for national elections.

(51) Providers of VLOPs and VLOSEs who are signatories of the Code of Practice of Disinformation should engage fully in the work related to the elections to the European Parliament, including through effective participation in the rapid response system and feedback mechanism with appropriate and timely follow-up actions. They should also provide – ahead of and after the elections - targeted reporting on the measures put in place to reduce the spread of disinformation and information manipulation in relation to the elections to the European Parliament, including relevant metrics on their impact (based on commitment 37.2 and 42). The Commission also recommends that – taking into account these reports, the work under the rapid response system, the inputs of fact-checker and civil society signatories, as well as other relevant input – signatories take stock of the lessons learned after the elections.

(52) To tailor their risk mitigation measures for elections to the European Parliament, the Commission recommends that providers of VLOPs and VLOSEs establish contact and cooperate with the EDMO Task Force on the elections to the European Parliament. For the 2024 elections to the European Parliament, this Task Force will produce reports and regular updates about the main disinformation trends, challenges, and phenomena. This should inform providers of VLOPs and VLOSEs actions and mitigating measures.

Q22: What are your views on the best practices proposed in this section?

Q23: What additional mitigation measures should be considered for the elections for the European Parliament present for online platforms?

4. READINESS DIALOGUE WITH THE COMMISSION

(53) The Commission is committed to enforcing Regulation (EU) 2022/2065, including in the area of elections. These guidelines give guidance to providers of VLOPs and VLOSEs on how to assess and mitigate systemic risks for electoral processes that stem from their service or use thereof. The Commission strongly encourages providers of VLOPs and VLOSEs to swiftly follow these guidelines and welcomes

assessments from researchers and civil society organisations on the effectiveness of the risk mitigation measures taken by the providers of these VLOPs and VLOSEs.

(54) At the same time, in particular given the early stage of implementation of Regulation (EU) 2022/2065, and the specific nature of systemic risks to electoral processes, the Commission stands ready to provide support to providers of VLOPs and VLOSEs to ensure they can adjust the design and functioning of their services and related systems early enough in case problems are detected, so as to prevent breaches of that regulation and, more importantly, harm to election processes in the EU.

(55) In this context, the Commission is available to facilitate a periodic review of the risk mitigation measures adopted by providers of VLOPs and VLOSEs on a voluntary basis. This could take the form of ex ante and ex post reviewing after specific elections. The feedback provided by the Commission in that context will be based on the information provided by the providers of VLOPs and VLOSEs and would not constitute a fully-fledged assessment of the compliance measures they had adopted. As such, it is without prejudice to the Commission's investigatory and enforcement powers.

5. CONCLUSION

(56) The risk-mitigation measures identified in these guidelines have been considered based on the readiness dialogues on election integrity with providers of VLOPs and VLOSEs, as well as the experience gained with the Code of Practice on Disinformation and the EU FIMI Toolbox. As such, the measures outlined in the guidelines can be considered as best practices at this moment in time.

(57) Nonetheless, it is early stage in the implementation of the DSA: while the objective of the guidance is to support VLOPs and VLOSEs in ensuring compliance with their obligations under Article 35 of Regulation (EU) 2022/2065 in a such unprecedented content, the Commission's understanding of the issues at stake in the interpretation and implementation of Article 35 of that regulation may evolve with further experience.

(58) In addition, the fast-evolving landscape in which providers or VLOPs and VLOSEs operate, and the tactics of malicious actors are constantly evolving, thereby requiring constant updates and adjustments to respond to the ever-emerging challenges. Moreover, following the assessment of the Commission and the European Board for Digital Services, the Code of Practice on Disinformation is expected to be converted into a Code of Conduct tying it to the legal framework of Regulation (EU) 2022/2065: in this context, the Commission expects signatories to continue implementing their commitments to fight disinformation under the Code of Practice on Disinformation. At the same time, additional pieces of Union legislation are set to come into force in the months to come and complement this regulation with specific rules also relevant for the subject matter of these guidelines: notably the Political Advertising Regulation and the AI Act (whose relevant obligations have been taken into account in these guidelines).

(59) These guidelines are without prejudice to future technological, societal and regulatory developments. To ensure such developments can be fully taken into consideration by the Commission in the assessment of DSA compliance, **the**

present guidelines will be subject to **review after one year from their adoption**. At that time the Commission may decide to withdraw or amend the present Communication. The European Commission encourages the providers of VLOPs and VLOSEs, research community and civil society organisations to contribute to this process.

Q24: What additional feedback or suggestions do you have regarding these guidelines?