



**SYSTEMIC RISK IN DIGITAL
SERVICES: BENCHMARKS FOR
EVALUATING THE
MANAGEMENT OF RISKS
TO ELECTORAL PROCESSES**

REPORT

May 2024

Sally Broughton Micova
Daniel Schnurr





As provided for in CERRE's bylaws and procedural rules from its “Transparency & Independence Policy”, all CERRE research projects and reports are completed in accordance with the strictest academic independence.

The project, within the framework of which this report has been prepared, received the support and/or input of the following CERRE member organisations: Arcom, Booking, Google, Microsoft, Ofcom, and Tencent. However, they bear no responsibility for the contents of this report. The views expressed in this CERRE report are attributable only to the authors in a personal capacity and not to any institution with which they are associated. In addition, they do not necessarily correspond either to those of CERRE, or of any sponsor or of members of CERRE.

© Copyright 2024, Centre on Regulation in Europe (CERRE)

info@cerre.eu – www.cerre.eu



TABLE OF CONTENTS

ABOUT THE AUTHORS.....	4
ACKNOWLEDGEMENTS.....	4
EXECUTIVE SUMMARY.....	5
1. INTRODUCTION.....	7
2. BENCHMARKS IN THE ELECTORAL INFORMATION SPACES.....	10
2.1 Benchmarks for Process Information Space.....	11
2.2 Benchmarks for Choice Information Space.....	15
3. RISKS OF NEGATIVE EFFECTS IN THE PROCESS INFORMATION SPACE THROUGHOUT THE ELECTORAL CYCLE.....	20
3.1 Risks of Negative Effects in the Pre-election period.....	21
3.2 Risks of Negative Effects in the Electoral Period.....	23
3.3 Risks of Negative Effects in the Post-election Period.....	25
4. RISK OF NEGATIVE EFFECTS IN THE CHOICE INFORMATION SPACE.....	26
4.1 Candidates and Campaigns.....	26
4.2 Citizens' Information and Participation.....	28
5. METRICS AND DATA FOR EVALUATING EFFECTS AND LEARNING.....	31
6. CONCLUSIONS AND RECOMMENDATIONS.....	38



ABOUT CERRE

Providing top quality studies and dissemination activities, the Centre on Regulation in Europe (CERRE) promotes robust and consistent regulation in Europe's network and digital industries. CERRE's members are regulatory authorities and operators in those industries as well as universities.

CERRE's added value is based on:

- its original, multidisciplinary and cross-sector approach;
- the widely acknowledged academic credentials and policy experience of its team and associated staff members;
- its scientific independence and impartiality;
- the direct relevance and timeliness of its contributions to the policy and regulatory development process applicable to network industries and the markets for their services.

CERRE's activities include contributions to the development of norms, standards and policy recommendations related to the regulation of service providers, to the specification of market rules and to improvements in the management of infrastructure in a changing political, economic, technological and social environment. CERRE's work also aims at clarifying the respective roles of market operators, governments and regulatory authorities, as well as at strengthening the expertise of the latter, since in many Member States, regulators are part of a relatively recent profession.



ABOUT THE AUTHORS



Sally Broughton Micova is a CERRE Academic Co-Director and an Associate Professor in Communications Policy and Politics at the University of East Anglia (UEA). She is also a member of UEA's Centre for Competition Policy.

Her research focuses on media and communications policy in Europe.

She completed her PhD in the Department of Media and Communications at the London School of Economics and Political Science (LSE), after which she was an LSE Teaching and Research Fellow in Media Governance and Policy and Deputy Director of the LSE Media Policy Project.



Daniel Schnurr is a CERRE Research Fellow and a Professor of Information Systems at the University of Regensburg, where he holds the Chair of Machine Learning and Uncertainty Quantification.

Previously, he was head of the research group Data Policies at the University of Passau. He received his Ph.D. in Information Systems from the Karlsruhe Institute of Technology, where he previously studied Information Engineering and Management. Daniel Schnurr has published in leading journals in Information Systems and Economics.

His current research focuses on the role of artificial intelligence for competition, privacy and data sharing in digital markets, as well as regulation of AI and the data economy.

ACKNOWLEDGEMENTS

The authors are extremely grateful for the substantive input from fellow project team members from the UEA Centre for Competition Policy (CCP), CCP Member Andrea Calef and CCP Researcher Bryn Eystone. They also thank Peter Courridge and Juliette Hardman of CCP for their assistance the executing the empirical investigation and all those who gave their time to participate as interviewees, as well as the steering committee members for their comments.



EXECUTIVE SUMMARY

Among its many other requirements, the Digital Services Act (DSA) requires providers of designated very large online platforms (VLOPs) and search engines (VLOSEs) to assess and mitigate the systemic risk of “negative effects on civic discourse and electoral processes” from their services. With elections for the European Parliament and in several member states taking place in mid-2024, **we are at a crucial point for assessing the approach of the DSA and the effects of the measure taken by service providers** as part of a process of continual improvement in the mitigation of risk to electoral process.

In March 2024 the European Commission issued its Guidelines for VLOPs and VLOSEs on the mitigation of systemic risks for electoral processes, which set out a list of mitigation measures that services are recommended to undertake to prevent negative effects on electoral processes. They are not exhaustive, however, and do not contain benchmarks against which to assess the success or failure of the suggested measures. This report contributes to filling this gap.

We propose **benchmarks and a framework for understanding the negative effects to electoral processes that VLOP and VLOSE providers are expected to mitigate**. These benchmarks provide a vision, generated from fundamental rights and established standards, of what ‘good’ electoral integrity looks like and set out how we could identify the positive roles that VLOPs and VLOSEs can play and if/when they contribute to systemic failure.

VLOPs and VLOSEs contribute, in differing ways and to varying extents, to the *electoral information space*, which should provide citizens accurate information about the electoral process and enable them to access sufficient information about their options and opportunities for discourse to make well-informed choices. Among the many other actors that contribute to the information space of any given electoral process, these very large services can be sources of information, public spheres for debate, and vehicles for expression. Having this in mind, our benchmarks are divided into those related to the *electoral process information space* and the *electoral choice information space*.

From these benchmarks, **we define a set of 13 potential negative effects to electoral processes that we argue should serve as the basis for conducting coordinated individual and cross-platform assessment after the upcoming European elections**.

Implementation of the DSA’s provisions on these very large services should be about learning how to effectively prevent “negative impacts of systemic risks on society and democracy”. The effective implementation of the DSA cannot be a box-ticking compliance exercise, but should be an iterative process in order to achieve an overall reduction in harm. Therefore, **we suggest metrics and types of data that will help us to understand whether proposed mitigation measures contribute to preventing the identified negative effects and that will allow for learning on a systemic level**.

The recommendations of this paper are derived from the data gathered through focus groups and interviews with a total of 27 individuals including election monitors and practitioners, election integrity experts, and representatives from civil society groups and regulatory authorities, as well as



from an extensive review of academic literature, civil society reports, and publicly available information from industry.

Based on the identified benchmarks, the derived negative effects, and the suggested metrics and data for determining the effectiveness of risk mitigation measures, we make the following five main recommendations:

1. The Digital Services Board and the European Cooperation Network on Elections should lead cooperation on the development of an *evaluation and learning strategy* to steer the use of data access requests under Article 40 of the DSA and on the application of the additional election-specific transparency, disclosure, and data sharing recommended in the EC Guidelines.
2. An effects-based framework as developed in this paper with specific metrics and types of data, some of which are already generated by services that are signatories to the Code of Practice on Disinformation or by election-related institutions, can contribute to such a strategy.
3. Standard practices and formats should be developed based on the experiences of information disclosure and sharing between VLOP and VLOSE providers and election-specific civil society groups, monitoring organisations, and Election Management Boards during the electoral period.
4. VLOP and VLOSE providers should invest, with input from those investigating campaign irregularities and foreign or other manipulative interference, in tools that help them hold political actors to account.
5. Drawing on commitments made by many VLOP and VLOSE providers in the AI Elections Accord to engage in collective evaluative action and with civil society and researchers, common adaptive standards on what can be considered appropriate levels and types of automation in election campaigns and discourse and what constitutes manipulative intervention should be developed through an inclusive multi-stakeholder process.



1. INTRODUCTION

In March 2024 the European Commission issued its *Guidelines for very large online platforms and very large search engines on the mitigation of systemic risks for electoral processes* (Guidelines).¹ The Digital Services Act (DSA) requires providers of designated very large online platforms (VLOPs) and search engines (VLOSEs) to assess and mitigate the systemic risk of “negative effects on civic discourse and electoral processes” from their services.² Discourse and electoral process were among many in the broad risk areas for which the service providers had to submit the first of their annual risk assessments to the Commission by the end of August 2023. With elections for the European Parliament and in several member states taking place in mid-2024, it is not surprising that the Commission prioritised improving risk mitigation in this area.

The EC Guidelines set out a list of mitigation measures that services are recommended to undertake to prevent negative effects on electoral processes. They are not exhaustive, however, and do not contain benchmarks against which to assess the success or failure of the suggested measures. This report contributes to filling this gap. It proposes benchmarks and a framework for understanding the negative effects to electoral processes that VLOP and VLOSE providers are expected to mitigate. These benchmarks provide a vision, generated from fundamental rights and established standards, of what ‘good’ electoral integrity looks like and set out the positive roles that VLOPs and VLOSEs can play as well as how to recognize systemic failure. From these benchmarks, we define a set of 13 potential negative effects on electoral processes.

The designated services differ greatly in design, function, and purpose, among other characteristics, and therefore the potential for negative effects and appropriate mitigations will also differ greatly. The DSA went into effect relatively recently, making the 2024 European elections a crucial point for testing the approach of the DSA and beginning a process of continual improvement in the mitigation of risk. We suggest ways to evaluate the success of efforts by VLOP and VLOSE providers to avoid contributing to these negative effects in the upcoming elections, because the implementation of the DSA’s provisions on these very large services should be about learning how to effectively prevent “negative impacts of systemic risks on society and democracy”.³ These proposals are derived from the data gathered through four focus groups and ten interviews with a total of 27 individuals including election monitors and practitioners, election integrity experts, and representatives from civil society groups and regulatory authorities, as well as from an extensive review of academic literature, civil society reports, and publicly available information from industry.

The management of systemic risk from the use, design, and function of VLOPs and VLOSEs will unavoidably be a learning process that involves improving risk assessment and mitigation with each

¹ Communication from the Commission on Guidelines for providers of Very Large Online Platforms and Very Large Online Search Engines on the mitigation of systemic risks for electoral processes pursuant to the Digital Services Act (Regulation (EU) 2022/2065)

² Article 34c, REGULATION (EU) 2022/2065 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act)

³ Recital 104, REGULATION (EU) 2022/2065 DSA



cycle. To ensure better prevention of negative effects over time, we need to thoroughly and accurately identify the negative effects and monitor using indicators for their successful prevention. With some risk areas listed in the DSA, such as from the dissemination of illegal content or negative effects on fundamental rights, this can be a continual process, however, the cyclical nature of electoral processes means that there are a crucial times in which to act, evaluate, and learn.

As argued in a previous CERRE Report,⁴ the concept of systemic risk in the DSA covers risks to individual users and wider society, and VLOPs and VLOSEs are part of integrated platform ecosystems that are characterised by various types of relationships with other services and actors. Service providers should be considering how the risks that stem from their services contribute to the overall risks to electoral processes. This means acknowledging their role while recognising that many sources of risk in the context of elections are only partly or not at all under their control. **The effective implementation of the DSA cannot be a box ticking compliance exercise, but should be an iterative process in order to achieve an overall reduction in harm.** Therefore, we propose metrics that will help us to understand if those mitigation measures have helped prevent negative effects and that will allow for learning on a systemic level. As one Digital Services Coordinator (DSC) respondent pointed out in the interview conducted for this investigation, “We can only check whether the reporting is in line with the requirements. We definitely need to go further, and also see whether their measures and mitigation of risk have a real effect.”

This paper is organised in a progression from setting benchmarks, to defining negatives effects and then suggesting metrics and data for determining effectiveness. Based on this analysis, we recommend:

1. The Digital Services Board and the European Cooperation Network on Elections should lead cooperation on the development of an *evaluation and learning strategy* to steer the use of data access requests under Article 40 of the DSA and on the application of the additional election-specific transparency, disclosure, and data sharing recommended in the EC Guidelines.
2. An effects-based framework as developed in this paper with specific metrics and types of data, some of which are already generated by services that are signatories to the Code of Practice on Disinformation or by election-related institutions, can contribute to such a strategy.
3. Standard practices and formats should be developed based on the experiences of information disclosure and sharing between VLOP and VLOSE providers and election-specific civil society groups, monitoring organisations, and Election Management Boards during the electoral period.
4. VLOP and VLOSE providers should invest, with input from those investigating campaign irregularities and foreign or other manipulative interference, in tools that help them hold political actors to account.

⁴ Broughton Micova and Calef, ‘Elements for Effective Systemic Risk Assessment under the DSA’.



5. Drawing on commitments made by many VLOP and VLOSE providers in the AI Elections Accord to engage in collective evaluative action and with civil society and researchers, common adaptive standards on what can be considered appropriate levels and types of automatised in election campaigns and discourse and what constitutes manipulative intervention should be developed through an inclusive multi-stakeholder process.



2. BENCHMARKS IN THE ELECTORAL INFORMATION SPACES

In an earlier CERRE report, Broughton Micova and Calef, pointed out that unlike for systemic risk in other sectors, there were not yet clear benchmarks for determining system failure in the risk areas covered by the DSA.⁵ That report argued that benchmarks should be set based on relevant international standards and norms and on input from stakeholders and experts. This section presents the results of such a process for the risk to civic discourse and electoral processes and suggests some benchmarks for success or failure in the role that VLOPs and VLOSEs play. VLOPs and VLOSEs are designated as such because of the extent to which they are used by EU consumers and the role they serve as public spaces for discourse and the dissemination of information. Their role in electoral processes is as contributors to an *information space* for the exercise of epistemic rights and conduct of civic discourse.⁶

The benchmarks presented in this section can be used to assess whether the ‘negative effects’ on electoral processes that the DSA aims to prevent are being prevented. They can aid in assessment of the effectiveness of risk mitigation but should not be used for determining compliance with the DSA. To verify whether the goals of the DSA have been achieved, an important step is, of course, to assess whether certain mitigation measures have been taken by service providers. However, it should also involve assessment of whether mitigations are effective, which requires dialogue with service providers, experts, and civil society about how to improve measures if needed.

“One potential failure is that the VLOPs are engaging with meaningless reporting, and we are not able to really supervise whether their measures are adequate to mitigate risks.”

- DSC representative

The *electoral information space* can be defined as all the information available for citizens to access and to which they are exposed throughout an election cycle. A healthy information space enables citizens to exercise their democratic rights. Citizens enjoy specific rights during an election period that include the right to be informed in order to make informed decisions and the right to participate in the process. Garnett and James, who lead the Electoral Integrity Project, which sets benchmarks for and measures perceptions of electoral integrity worldwide, argue that, in addition to robust electoral management, the principles of equality of participation in the election process and of equality and fairness in deliberation are crucial to electoral integrity.⁷ Our interviewees also consistently demonstrated an understanding that the information space serves both to enable voters to participate in the election discourse and to provide citizens with the information they require to make an informed election choice. For electoral information spaces, we therefore distinguish

⁵ Broughton Micova and Calef.

⁶ As defined by Hannu Nieminen, epistemic rights refers to the requirement that “society should guarantee that all its citizens will be given truthful information and knowledge and the competence to use these for their own benefit and that of society as a whole.” Nieminen, ‘Why We Need Epistemic Rights’, 11; Aslama Horowitz et al., *Epistemic Rights in the Era of Digital Disruption*.

⁷ Garnett and James, ‘Cyber Elections in the Digital Age: Threats and Opportunities of Technology for Electoral Integrity’. See also the Electoral Integrity Project <https://www.electoralintegrityproject.com/>



between the *process-related information space* and the *choice-related information space* and define each further in the next two sections.

2.1 Benchmarks for Process Information Space

The process information space, to which VLOPs and VLOSEs contribute, needs to provide citizens with accurate information on how to exercise their right to vote. This can include information on registration, districts, the logistics of election day, and how counting and reporting of results will take place. Participation in the electoral process is also conditioned by citizens' trust in the electoral process. The process information space shapes the level of trust that citizens have that the vote will be secret, counted fairly, and ultimately citizens' willingness to accept the results as valid.

Multiple actors, including election management boards (EMBs), governments, mainstream media, political parties, and others, all play a role in shaping the process information space in any given state. VLOPs and VLOSEs have been playing an ever-greater role in this as citizens increasingly rely on them as sources of information and devote a lot of their daily attention to these platforms and services. Although overall we should hope citizens are fully and equally informed about the process and have complete faith in it, the expectations for VLOPs and VLOSEs in this space relate only to the role they can play in enabling or undermining that. Their role is linked to that of the national authorities, EMBs, media services, political parties, civil society, and other actors whose contributions also determine whether citizens are well-informed.

As several of the election experts with whom we spoke pointed out, not all elections are run perfectly, and some information that might undermine trust in the process is not dis/mis-information. Generally, within the EU, elections are run freely and fairly, however problems with the management of elections affecting citizens' ability to participate have been documented, especially for those 'mobile' citizens living outside their home state. Observers, civil society groups, candidates, parties, and citizens have to be able to call out any legitimate irregularities, including by sharing or finding information on VLOPs and VLOSEs.

Table 1 suggests high level benchmarks for success or systemic failure with a focus on the role that VLOPs and VLOSEs might play in the process information space. These are the basis for the understandings of negative effects elaborated in Section 3 and the metrics for mitigation suggested in Section 4. The points of departure for developing these benchmarks are those in the first column for electoral integrity drawn from international human rights law, and the standards used by international monitoring organisations. Article 25 of the International Covenant of Civil and Political Rights (ICCPR) is the most relevant part of international human rights law. The guidance of the Council of Europe (CoE) and those of the election monitoring services of the EU's Election Observation and Democracy Support (EODS) programme and of the OSCE's Office of Democratic Institutions and Human Rights (ODHIR) provided more operational standards, as did input from our interviewees.

As hosting and intermediary services respectively, VLOPs and VLOSEs play an important role in the dissemination of information related to electoral processes. However, they are not the only sources of information on these processes. Media, education systems, various direct communications from



authorities and parties, and even friends and family are also part of electoral information space. The benchmarks set here reflect that. They focus on the potential role of VLOPs and VLOSEs in contributing to healthy electoral process information space or to what might be a wider systemic failure in electoral integrity. It is important to note that they are to be read as accounts of what could be – potentially successful and potentially contributing to systemic failure – and not as descriptions of what is currently the situation of any given service.



Table 1 Benchmarks for the electoral process information space

ELECTORAL INTEGRITY BENCHMARK	SOURCES OF THE BENCHMARK	ROLES OF VLOPS/VLOSES	A SUCCESSFUL ROLE OF VLOP/VLOSE (VARYING DEPENDING ON TYPE) WOULD BE IF:	FAILURE BY A VLOP/VLOSE (VARYING DEPENDING ON TYPE) IN RELATION TO SYSTEMIC RISKS TO ELECTORAL PROCESSES WOULD BE IF:
<p>Elections conducted lawfully:</p> <ul style="list-style-type: none"> • secrecy of vote; • security of election management personnel; • integrity of the count and results; • The public trusts the process and accepts result. 	<p>ICCPR Art. 25; Article 3 of the Protocol to the European Convention for the Protection of Human Rights and Fundamental Freedoms; OSCE/ODHIR Handbook for the Observation of Information and Communication Technologies (ICT) in Elections; EODS Handbook for EU Election Observation; Electoral Integrity Project (EIP).</p>	<p>They can serve as vehicles for dissemination of and access to:</p> <ul style="list-style-type: none"> • Information about how the electoral process should work; • information about election conduct; • information about or by election management personnel; • information about results. 	<p>They serve as a timely and broad source of accurate information on the conduct of the election, contributing to equality of information access.</p> <p>Election management personnel can safely maintain a presence on VLOPs.</p> <p>They provide a space for legitimate evidencing of irregularities and for voicing any concerns about the conduct of elections.</p>	<p>They contribute to widespread exposure to inaccurate information about the conduct of the election.</p> <p>They consistently enable the harassment or threatening of election management personnel or contribute to their inability to do their job.</p> <p>They cannot be used for voicing and receiving legitimate concerns or sharing evidence of election irregularities.</p> <p>They contribute to widespread, unfounded mistrust of the electoral process.</p>



<p>Voting access is universal and equal:</p> <ul style="list-style-type: none"> • voter registration unhindered; • voters freely able to vote on election day. 	<p>ICCPR Art. 25; OSCE/ODHIR Handbook for the Observation of Information and Communication Technologies (ICT) in Elections; EODS Handbook for EU Election Observation; EIP.</p>	<p>They can serve as vehicles for dissemination of and access to information about the registration process and requirements, and about election day processes and requirements</p>	<p>They serve as sources of accurate information on registering to vote and participating on election day.</p> <p>They provide a vehicle for election officials, monitors and parties to share critical information on election day.</p>	<p>Their use contributes to the deliberate suppression of voter registration, confusion about voting obligations where mandatory and/or voter participation in an election.</p> <p>They contribute to widespread dissemination of inaccurate information on an election day.</p>
<p>Citizens have equal ability to exercise their right to stand for office:</p> <ul style="list-style-type: none"> • standing for office is not subject to undue interference; • procedures are transparent. 	<p>ICCPR Art. 25; EIP; OSCE/ODHIR 2002 International Standards and Commitments on the Right to Democratic Elections: A Practical Reference Guide to Democratic Elections Best Practice; EODS Handbook for EU Election Observation; EIP.</p>	<p>They can serve as vehicles for dissemination of and access to information about or by individuals or parties intending to stand for election, and about the process of standing.</p>	<p>They serve as a source of accurate information on standing for office.</p> <p>They provide a vehicle for candidates to communicate effectively and safely with the public.</p>	<p>They contribute to widespread dissemination of inaccurate information on the processes of standing for office.</p> <p>They consistently enable the harassment or threatening of candidates or contribute to the undue silencing of contestants in the election.</p>



Ultimately, systemic failure of the process information space would be marked by a significant number of voters failing to exercise their right to vote due to being wrongly informed about how to register, how to vote, or convinced not to vote by information aimed at voter suppression. Several interviewees also described systemic failure as when there is so little trust in the process among citizens that there is widespread unacceptance of the results, and some pointed out that this can lead to violence. Failure can also occur when the management of the elections is hampered by the process information space being polluted by online attacks on or harassment of electoral officials, or individuals or parties are prevented from standing in the election due to being ill-informed, threatened, or silenced. As the last column in Table 1 shows, VLOPs and VLOSEs can contribute to these systemic failures if the risks are not effectively mitigated. The Table also shows that VLOPs and VLOSEs could play very constructive roles, providing important channels for communicating to the public and contributing to equality of information access.

2.2 Benchmarks for Choice Information Space

The choice information space is where deliberation and contestation take place, and where citizens should get the information they need to make a choice. This is often conceived of as a public sphere in which rational discourse takes place.⁸ The public sphere, however, is also recognised as being an agonistic and often fragmented space where emotion does have a place in the discourse⁹ and where participation parity for various minority groups is important.¹⁰ Civic discourse is always underway. As was pointed out by several of our interviewees, the electoral choice information space is coloured by longer-running narratives that build up from discourse and interactions over time, and that can often be polarising and even radicalising. During an election cycle there is an acute need to ensure there is a healthy information space in relation to the specific choices to be made in the election, even though it can be hard to separate from the wider issues in civic discourse. The benchmarks suggested here therefore do not address all the long term issues around polarization, civility, and extremism in online civic discourse. They do cover fragmentation of debate about the electoral choices and pluralism of views, as well as disinformation on relevant policy issues, which are also longer-term issues.

⁸ Habermas, *The Structural Transformation of the Public Sphere: An Inquiry into a Category of Bourgeois Society*.

⁹ Fraser, 'Rethinking the Public Sphere: A Contribution to the Critique of Actually Existing Democracy'.

¹⁰ Broughton Micova, 'The Collective Speech Rights of Minorities'.



The rights and interests of both the potential voters and those standing in the election are at stake in the choice information space. The contestants in an electoral process should have fair access to the means for communicating to citizens and the ability to be found by those seeking information. Citizens should have access to sufficient and balanced information from those contestants, and about the policies and positions they represent. Citizens should also be able to participate in debate about them, including those from marginalised or minority groups. These standards stem from the fundamental right to freedom of expression enshrined in international human rights law and EU law. They are also crucial to the exercise of the right to participate in the conduct of public affairs in Article 25 of the ICCPR.¹¹ Again, relevant recommendations from the CoE and the standards of the EODS and OSCE/ODHIR also provide a basis for these benchmarks along with the input of our interviewees.

“At one level there are the discussions about the parties, party leaders and candidates. What do citizens think about them? Are they ‘likeable’ or ‘competent’? At second level there are also the substantive discussions of policy and manifestoes. What is the actual nature of the policy problem? And are the proposed solutions ‘credible.’”

- Prof. Toby James, Electoral integrity expert

¹¹ ICCPR, Article 25 (a)



Table 2 Benchmarks for the Electoral Choice Information Space

ELECTORAL INTEGRITY BENCHMARK	SOURCES OF BENCHMARK	ROLES OF VLOPS/VLOSES	A SUCCESSFUL ROLE OF VLOP/VLOSE (VARYING DEPENDING ON TYPE) WOULD BE IF:	FAILURE BY A VLOP/VLOSE (VARYING DEPENDING ON TYPE) IN RELATION TO SYSTEMIC RISKS TO ELECTORAL PROCESSES WOULD BE IF:
<p>Candidates and contesting parties have equal opportunity to inform citizens of their positions and views.</p>	<p>Venice Commission Guidelines on Media Analysis During Election Observation Missions; CoE Recommendation CM/Rec(2022)12; EODS Handbook for EU Election Observation; EIP.</p>	<p>They serve as a vehicle for the dissemination of and access to content by or about candidates and parties, which could include political advertising.</p>	<p>They enable contesting parties to communicate to the public in a fair and transparent manner, without undue interference.</p>	<p>They contribute to widespread dissemination of disinformation about candidates, their policies, and positions.</p> <p>They fail to treat contesting parties fairly in terms of access to the extent that any are notably disadvantaged in their opportunities to communicate.</p>
<p>Electoral campaigns are conducted freely and fairly:</p> <ul style="list-style-type: none"> • fair access to public resources; • campaign finance transparency; • free from undue interference. 	<p>Venice Commission Guidelines on Media Analysis During Election Observation Missions; CoE Recommendation 2208 (2021) and Recommendation CM/Rec(2022)12; EODS Handbook for EU Election Observation; EIP.</p>	<p>They serve as a vehicle for the dissemination of content by or about candidates and parties, which could include political advertising.</p> <p>They serve as an evidence base for party and campaign behaviour online.</p>	<p>They enable EMBs and monitoring groups to assess whether candidates have equal access to opportunities to inform the public and are conducting their campaigns in a fair manner.</p>	<p>They facilitate or enable widespread irregularities in campaigning by contesting parties or manipulative interference of any kind.</p> <p>They facilitate or enable parties to engage in widespread circumvention of campaign financing rules or fail to enable adequate oversight of spending on their services in relation to the electoral contest.</p>



Citizens receive sufficient information to inform their choice and form their opinion free from manipulation.	ICCPR Art. 19 & 25 (and General Comment 25); Venice Commission Guidelines on Media Analysis During Election Observation Missions; EODS Handbook for EU Election Observation; EIP	They serve as a source of information about candidates, parties, societal issues, policy options, and the opinions of others.	They serve as a source of accurate, thorough and diverse information about candidates, parties, societal issues, policy options, and the opinions of others.	They contribute to the widespread dissemination of inaccurate or malign information about candidates, parties, societal issues, policy options, and the opinions of others. They facilitate or enable manipulative interference of any kind.
Citizens are able to participate in political deliberation, voice opinions, and engage with candidates and contesting parties	ICCPR Art. 19; EODS Handbook for EU Election Observation	They serve as a place for exchange of information and deliberation among citizens and with candidates and contesting parties, and as a vehicle for citizen expression.	They provide space for constructive deliberation and means through which citizens can participate in this deliberation and engage with candidates and contesting parties without undue interference.	They contribute to widespread fragmentation of the deliberation space and erosion of pluralism. They overly constrain citizens' ability to express or receive information. They facilitate or enable manipulative interference in civic discourse.



The roles played by VLOPs and VLOSEs in the choice information space are as locations of civic discourse and vehicles through which participants in that discourse access and express information and opinions. However, they are not neutral conduits. They select and rank information and therefore influence the prominence and visibility of content to users. Some, such as social media platforms, do this organising on an individual, often personalized level relying on algorithms to handle the vast quantities of content or products that they host and present to the user. Others, especially search engines, rank based on their criteria for the relevance, veracity, and perhaps authoritativeness of the content or source. The effects of this curation are conditioned by the design of the services and their functionalities, the commercial interests, and the context in which they are used, and thus there is a risk of amplification of malign behaviour and content, or the enabling of manipulation.¹² At the same time, unlike the rather obvious editorial policies and commercial incentives of broadcast media, the conditions, criteria, or instructions shaping algorithmically curated feeds, recommender systems, and search results are difficult to monitor. Table 2 therefore imagines a successful or constructive role for VLOPs and VLOSEs as well as what failure would like should systemic risks not be effectively mitigated.

The biggest failure would be that the result of the elections would be one that did not represent what, in fact, people would like to be... because they were convinced that something was happening that was not.

- Anonymous interviewee

VLOPs and VLOSEs can be important vehicles for candidates and parties to disseminate information about themselves, their positions and policies. Some also provide spaces for others to challenge their positions and for citizens to engage with them and others in deliberation about issues and policies. They can be vital sources of diverse and accurate information about all the options for voters making choices about how to place their vote or even whether to vote at all. However, without effective mitigation, they can also enable

widespread manipulative interference in contestants' campaigns and voters' choices by malign actors, which may be foreign but can include the candidates and parties involved in the election. They can contribute to the pollution of civic discourse and fragmentation or the siloing of that discourse to the extent that citizens no longer have access to or are presented with sufficiently accurate and diverse information and opinion to freely make an informed choice.

Ultimately systemic failure could happen if these very large services that are so widely used contribute to the results of an election not actually representing the will of the citizens. Though this sentiment was expressed by many of our interviewees, it was also acknowledged that demonstrating that a result does not actually reflect preferences or interests would not likely be possible in any situation. Assessing the exact contribution of VLOPs and VLOSEs to general failure of the choice information space to serve its purpose and enable democracy will also be hard. Therefore, for the purpose of advancing the implementation of the DSA we now take these high-level benchmarks and use them to elaborate the negative effects for which metrics and assessment tools can be developed.

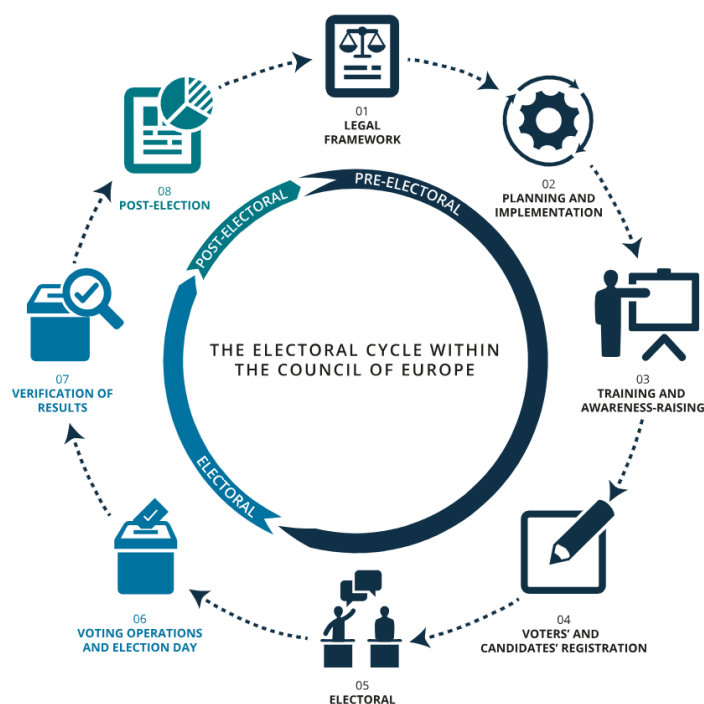
¹² Saurwein and Spencer-Smith, 'Automated Trouble: The Role of Algorithmic Selection in Harms on Social Media Platforms'; Broughton Micova, *What's the Harm in Size? Very Large Online Platforms in the Digital Services Act*.



3. RISKS OF NEGATIVE EFFECTS IN THE PROCESS INFORMATION SPACE THROUGHOUT THE ELECTORAL CYCLE

Ensuring the integrity of electoral processes requires attention to the whole electoral cycle. In the Guidelines the Commission points out that the requirement to mitigate risk applies to the whole cycle, referring to the CoE's depiction of the pre-electoral, electoral, and post-electoral phases.¹³ In this section we use the CoE depiction (see Figure 1) as a basis and draw on other accounts of the cycle in academic and grey literature,¹⁴ evidence from our interviews and focus groups, and other sources to set out specific risks of negative effects. As was pointed out in a previous CERRE report on systemic risk in the DSA, sources of risks of negative effects can be exogenous, coming from sources outside the VLOP and VLOSE providers' ecosystems, or they can be endogenous, arising from the relationships, functionalities, and systems of those ecosystems.¹⁵

Figure 1: The Council of Europe Electoral Cycle



Source: Council of Europe, *The Electoral Cycle* has been designed by [ACE Project - The Electoral Knowledge Network](#)

¹³ <https://www.coe.int/en/web/elections/electoral-cycle>

¹⁴ A very useful elaboration and illustration of the electoral cycle is also in Norris, P. 2014 *Why Electoral Integrity Matters*, page 34. Another depiction is in Bicu, I. 2023 *The Information Environment Around Elections*, a report for International IDEA, <https://www.idea.int/theme/information-communication-and-technology-electoral-processes/information-environment-around-elections>

¹⁵ Broughton Micova and Calef, 'Elements for Effective Systemic Risk Assessment under the DSA'.



3.1 Risks of Negative Effects in the Pre-election Period

The pre-election period can vary in length depending on the type of electoral process in question and the state in which it is taking place. The EC Guidelines advise service providers to consider at least 6 months prior to the election date to be the pre-election phase of the cycle. As illustrated in Figure 1 above, this phase includes several important steps, namely the setting of election rules including any changes to district boundaries, voting processes, or voting requirements,¹⁶ the selection or declaration of candidates, the registration of voters if not automatic, the recruitment and training of election management personnel, and campaigns to inform voters.

Negative Effect 1: Erosion of public trust in the fairness of the organisation and administration of the election

As one of our interviewees put it, the erosion of public trust in electoral processes is a snowballing effect that begins with doubts sown about how the elections are organised. The erosion of trust also builds up over successive electoral cycles. Another interviewee who is involved in election monitoring worldwide stated that there was evidence of increasing distrust among citizens in the election management bodies themselves and in the electoral processes in nearly every election monitoring report they had read in the last several years.

Certain steps in the pre-election phase are specific targets for disinformation campaigns by domestic political actors, foreign information manipulation and interference (FIMI), or other forms of malign use aimed at eroding public confidence that the electoral processes are free and fair. The EC Guidelines suggest that VLOP and VLOSE providers should consider ‘relevant discussions’ that are taking place on their services when making decisions about how to reinforce their internal processes ahead of an election.¹⁷ Discussions around each of the following topics in the pre-election phase merit particular attention by VLOPs and VLOSEs in their internal processes and later in the assessment of success in preventing this negative effect:

1. The topic of defining constituency districts can engender volatile debate and disinformation. At the same time, it can be a complex issue that average citizens are not very aware of and there is a need to ensure that any attempts at gerrymandering can be called out. As VLOPs and VLOSEs attempt to mitigate disinformation, care will need to be taken to mitigate disinformation about constituency defining while ensuring that legitimate concerns about gerrymandering are not stifled. This will likely require close collaboration with fact-checkers, civil society groups, and even political parties, as service providers cannot, nor should be judging the quality of district-defining processes.
2. Voter registration has increasingly moved online in EU member states, which for some citizens makes it more suspect than before. Where online systems are the main way for voters to register, VLOPs and VLOSEs should be aware that discussions about these may be particularly vulnerable to disinformation narratives aimed at undermining public confidence. They may

¹⁶ Examples of these would be the introduction of new voting technologies such as machines or internet voting systems, or new requirements for voter identification.

¹⁷ EC Guidelines para 18.



need to boost content moderation attention or additionally prioritise authoritative sources depending on the type of service. Information from EMBs and other local authorities will likely be important to service providers' mitigation measures.

3. There is significant evidence that the independence and competence of EMBs are being challenged by both the dissemination of disinformation and by online harassment of EMB members. Research by International IDEA found that the planning and implementing of electoral processes by EMBs, which included budgeting, logistics, and the recruitment of electoral officials, were popular targets for disinformation.¹⁸ This same research also found EMBs to be overwhelmingly the targets of disinformation aimed at organisations involved in the implementation of elections (95% of all targeted organisations) and that often disinformation targeted specific individuals. These results were corroborated in our interviews with election practitioners. Both the topics related to the EMB and its work as well as the individual members of the EMB will therefore be 'relevant discussions' to be attended to in designing mitigation measures. Again, information and input from EMBs would be useful to service providers.

Negative Effect 2: Recruitment or representation problems EMBs

Harassment and other forms of online violence against EMB members were mentioned by several of the election practitioners and experts we interviewed, with social media platforms particularly implicated as locations for this kind of malign use. According to research by International IDEA, intimidation, harassment, and other forms of online violence disproportionately affect women on EMBs, leading to resignations and deterring women from taking on this public service role.¹⁹ As of 2022 even in Europe only 25% of EMBs were chaired by women.²⁰ How quickly and thoroughly VLOPs used for disseminating user-generated content address and prevent such behaviour, which is almost universally against their terms and guidelines, and the experiences of EMB members will be important measures of success in mitigating this negative effect.

Negative Effect 3: Voter suppression or disenfranchisement through misdirection about registration

Where voter registration is not automatic, the pre-election period is when anyone not registered to vote or who has changed their residence needs to register. Increasingly this can be done online in EU member states and there are deadlines by which it needs to be done ahead of election day. Malign actors including contestants in the election can use VLOPs and to some extent VLOSEs to disseminate targeted disinformation about how to register aimed at misdirecting or dissuading specific constituencies. Several of the mitigation measures called for in the EC Guidelines address this, namely compliance with the Code of Practice on Disinformation, providing contextual information in the various ways set out in paragraph 27(c), taking the actions expected on political advertising described

¹⁸ Bicu, 'The Information Environment Around Elections'.

¹⁹ <https://www.idea.int/news/breaking-barriers-woman-electoral-management>

²⁰ Bicu, 'Few Women at the Top of Electoral Management Bodies Worldwide'.



in paragraph 27(e), and working with election authorities, fact-checkers, media, and other civil society groups.

Negative Effect 4: Candidate suppression due to online harassment and violence

Content and behaviour on VLOPs used for disseminating user-generated content, and sometimes content accessible on VLOSEs, disproportionately suppresses the candidature of women and minorities. From multiple established democracies there is evidence of women choosing not to stand or stand again for election because of harassment and other forms of violence against them on social media VLOPs and on smaller platforms, or because they expect such violence having observed the treatment of other women in politics.²¹ This is a negative effect that can be observed in the pre-election period when candidates are selected or declared, but the sources of the risk are much more long-term. The mitigation of this effect requires measures both during and between electoral cycles in cooperation with multiple other actors.

3.2 Risks of Negative Effects in the Electoral Period

As illustrated in Figure 3, the electoral period includes the day or days of voting and the counting and verification of those results. It ends with the announcement of the results. Key concerns in this period are voters' ability to exercise their right to vote freely and public trust in the processes that take place in this period. This is a period in which time matters significantly as the windows for mitigating risks of negative effects may be quite short.

Negative Effect 5: Erosion of public trust in the fairness of the electoral process

In this period the snowball of public mistrust discussed under Negative Effect 1 can gain speed and grow rapidly. The risk of VLOPs being used to undermine the credibility of EMBs and others involved in the administration of the election covered in point three under Negative Effect 1 remains very relevant and could be exacerbated. In addition, VLOP and VLOSE providers should pay particular attention to 'relevant discussions' on their services and the need to prioritise authoritative information about the following issues which may be points of vulnerability:

1. Where technology is used for voting, our investigation found evidence that this is a consistent target of disinformation and conspiracy theories that can undermine confidence in the electoral process. Two election monitoring specialists we interviewed relayed the seemingly paradoxical situation that while these technologies have usually been implemented to protect the integrity of the voting process, in their experience they are often met with distrust, which gets fostered and amplified by disinformation or just misinformation circulating online. This particular risk was also noted in the OSCE/ODHIR election monitor report on the most recent elections in Estonia,²² which is the only member state to have implemented online voting as standard.²³ If voting machines or information technology systems are used in an election, the

²¹ Collignon and Rüdiger, 'Increasing the Cost of Female Representation? The Gendered Effects of Harassment, Abuse and Intimidation towards Parliamentary Candidates in the UK'; Bigio and Vogelstein, 'Women under Attack'.

²² OSCE/ODHIR, 'ESTONIA PARLIAMENTARY ELECTIONS ODIHR Election Expert Team Final Report' 3 March 2019, <https://www.osce.org/files/f/documents/8/e/424229.pdf>

²³ France uses online voting in limited ways for voting from abroad.



nature of this use and any evidence on public perceptions of it should be part of service providers' analysis of the context-specific risks,²⁴ and given attention in content moderation and counter-disinformation measures.

2. Tied to the issue of the credibility of electoral administration mentioned above is that of conduct at polling stations. Debunking malicious and unfounded reports of fraud, voter intimidation, violations of secrecy, or mismanagement at polling stations circulating on VLOPs or appearing in results on VLOSEs will require close and timely coordination with EMBs, monitors, media, and other local actors. Some reports may be legitimate, so the incident response mechanisms and cross-platform coordination with other actors called for in the EC Guidelines²⁵ will be particularly important for mitigating risk of negative effects from disinformation and other malign use related to this issue. False reports of irregularities during the vote may be paired with incitement to further interfere in the process or to violence, making the tracking of the evolution and propagation of narratives an important element of mitigation that will likely rely on local and context-specific knowledge.
3. The counting of votes in most countries is a relatively opaque process to the average citizen and is observed only by registered monitors and possibly party representatives. This makes the conduct of the count another 'relevant discussion' topic for VLOP and VLOSE providers to address with mitigation measures within the electoral period. Disinformation and conspiracy theories about the conduct of the count may circulate before it officially begins and as part of sowing doubts about the results in the post-electoral period as well. Mitigation measures should pay attention both to discussion of incidents at particular polling stations and to the overall counting process, as well as to content covering the verification of the count. There also can be legitimate concerns about these processes so mitigation measures must be carefully applied, in close collaboration with local actors.

Negative Effect 6: Voter suppression or other manipulative interference through disinformation and misdirection about how to vote

This potential for voters to be disenfranchised by intentional efforts to misdirect them or from being misinformed due to lack of access to accurate information seems to be a key concern in the EC's Guidelines and was brought up by many of those we interviewed. Many of the mitigation measures for this negative effect cited by those we interviewed already appear in the EC Guidelines or the Code of Practice on Disinformation. Those are pro-active provision of accurate information through nudges and prompts, the use of labels to identify credible and official sources, as well as fact-check labels.²⁶ Insight from our interviews also highlighted some specific sources of risk that indicate a need for additional tools to identify and combat sophisticated manipulation techniques. VLOP and VLOSE providers should be aware of the risks from scams mimicking phishing attempts that try to lure users with links to fake online voting websites. This can trick citizens who intend to vote into thinking they have done so, thus interfering with voters' ability to exercise their right to participate.

²⁴ Initial context analysis is referred to in EC Guidelines paragraph 20.

²⁵ EC Guidelines paragraphs 53 & 54.

²⁶ EC Guidelines paragraph 27



3.3 Risks of Negative Effects in the Post-election Period

The post-election period begins with the announcement of the results. In national level elections in many member states this is followed by periods of negotiations among parties over forming a government. Overwhelmingly, those we interviewed viewed acceptance of the result as the ultimate marker for a successful electoral process and a lack of acceptance as a sign of serious failure. The core concern in this period is public trust in the process and perception of fairness. It is when VLOP and VLOSE providers are expected to conduct a ‘post-election review’ with input from independent researchers, civil society organisations, fact-checkers, and independent election monitors.²⁷

Negative Effect 7: Widespread non-acceptance of the result

Significant erosion in public confidence in the previous two phases of the electoral cycle can lead to many citizens not accepting the results of the process. The recent US electoral cycles are often mentioned as examples of this problem, because of the well-publicised violence that occurred on 6 January 2021, the role of then President Trump, and his continued contestation of the results of the 2020 election. However, it is not the only example of this negative effect leading to violence. Protesters stormed the parliament in North Macedonia in 2017 and that same year in Kenya protests against the validity of the results ended in violence.

Though no such violent examples occurred recently in EU member states, it is still a concern raised by some interviewees in relation to the upcoming European elections. VLOP providers can help mitigate this risk of violent response by identifying where ‘relevant discussion’ about the validity of the results co-occurs with incitement content. Mitigation on their services can be paired with communication with member state authorities. Harassment, threats and other online violence against EMB members and others in election administration can be part of this, which can have significant personal negative effects as well as the broader effect of undermining faith in the validity of the result. Non-acceptance of the result does not necessarily result in violence. Where incitement to violence has not followed, it can still lead to citizen disengagement from electoral processes and contribute to polarisation and radicalisation.

²⁷ EC Guidelines paragraphs 59-61.



4. RISK OF NEGATIVE EFFECTS IN THE CHOICE INFORMATION SPACE

The benchmarks set for VLOPs and VLOSEs in the choice information space stem from the important role they play in enabling expression, deliberation, and access to information for millions of EU citizens. The ‘choice’ category of information refers to the information that voters need in order to make an informed decision about how to utilise their vote. This includes information about candidates and parties as well as information about policy issues that are central to the platforms and civic discourse in the lead up to the election. Arguably mitigating risk from disinformation is more challenging in the choice information space because, whereas there are usually clear facts about the election process and relatively easily identifiable sources of accurate information, the choice information space involves exchange of views and opinion and often contested information about complex policy issues.

Mitigating the negative effects in the choice information space can be as much about achieving balance, pluralism, and access to a diversity of views, as it about countering disinformation and prioritising authoritative sources. It can also be about ensuring transparency and access to data for those institutions with regulatory or watchdog roles over the conduct of campaigns. In addition, there are different timeframes for the risks of negative effects in this space. Polarization and the silencing of voice can be gradual, therefore requiring careful tracking and perhaps subtle responses, while manipulative interventions can have an acute effect where timely reaction is crucial. The recent Slovakian election provides a clear example of how disinformation, an audio deepfake, about a candidate was disseminated through social media VLOPs during the pre-election silence period that hindered debunking by others.²⁸

As set out in Table 2, the benchmarks for this space cover the rights of candidates and whether their campaigns have been free and fair, and they address both the rights of citizens to participate in civil discourse and the need for them to make an informed choice. Therefore, we divide the negative effects in this space along those lines.

4.1 Candidates and Campaigns

Negative Effect 8: Candidates and campaigns do not have fair access to or ability to be found by VLOP and VLOSE users.

There are two main ways that contestants in elections use VLOPs and VLOSEs to communicate to potential voters: advertising and their own organic content through profiles and websites. Candidates and parties use social media accounts and have websites that are often discovered and reached through search engines. Three of our interviewees expressed concerns about risks to fair access to the electoral campaigns that could stem from intentional or unintentional bias in the content moderation algorithms on VLOPs or the criteria behind search results. There may be difficult calls to make by

²⁸ Meaker, M. (2023). Slovakia’s Election Deepfakes Show AI Is a Danger to Democracy. Wired. <https://www.wired.com/story/slovakias-election-deepfakes-show-ai-is-a-danger-to-democracy/>



service providers when candidates or parties violate terms or community guidelines, therefore transparency about such consequential decisions is an important risk mitigation. Furthermore, several mitigations measures listed in the EC Guidelines address opportunities and tools for advertising, which can represent a second important channel for candidates and campaigns to reach users.

Negative Effect 9: Widespread manipulative interference in campaigning through malign use of VLOPs

Manipulative interference can come from foreign sources, but according to the accounts of our interviewees it seems to be more often from actors within states. This includes other candidates running but often the actors remain unidentifiable. Candidates lie and campaigns are known to misrepresent their opponents, however, the scale of VLOPs and the distinct ways of information dissemination on these services, as well as the use of AI, enable them to do this at a scale and manner that could be considered manipulative interference.²⁹ The DISARM framework, which maps tactics and techniques of disinformation, identifies several tactics found to be used in manipulation campaigns, such as developing and testing false or misleading narratives and disseminating them in a targeted strategic manner and through flooding and cross-posting. It also lists a variety of techniques of inauthentic use such as creating fake accounts and posts, using bots, creating fake experts, or manipulating search engine optimization.³⁰ Manipulative interference increasingly relies on generative AI to create deepfakes and also can include threats or harassment of candidates. This can have negative consequences for contestants in an election to conduct their campaigns and their ability to disseminate information to citizens without interference.

Negative Effect 10: Campaigns cannot be effectively monitored by EMBs and civil society monitors.

In any member state, a combination of institutions are in place to ensure that contestants follow the rules for the conduct of campaigns, which usually include ones on spending, transparency, and behaviour. The scale and functionality of some VLOPs can make it challenging for those monitoring campaigns to have sufficient oversight of campaign conduct or spending, especially when they lack access to the necessary data. Interviewees from election monitoring organisations reported that ad libraries and CrowdTangle, a free Meta tool that allows election monitors and journalists to follow and analyse public content on social media have been important assets in this kind of work. Mitigation of this negative effect requires the sharing of information called for in the EC Guidelines,³¹ but also continuous access to data and tools for EMBs, other member state authorities, and independent monitoring organisations.

²⁹ Saurwein and Spencer-Smith, 'Automated Trouble: The Role of Algorithmic Selection in Harms on Social Media Platforms'.

³⁰ The DSARM list with explanations

https://github.com/DISARMFoundation/DISARMframeworks/blob/main/generated_pages/techniques_index.md

³¹ EC Guidelines paragraph 48



4.2 Citizens' Information and Participation

Negative Effect 11: Citizens do not receive sufficient, diverse and accurate information to make an informed choice.

This negative effect was raised by most of those interviewed and mentioned by ones from various backgrounds. They spoke of echo chambers, silos, polarization, and radicalisation into extremist factions, as well as disinformation. Interviewees feared that citizens may not be sufficiently informed because they are only exposed to a limited range of reinforcing information, they are persistently exposed to disinformation, or both. Where business models depend on maintaining levels of user engagement, interviewees saw particularly high risks of users being served content that feeds existing fears and prejudices, shocks or surprises, and reinforces positions.

The electoral choice information space is often affected by longer term issues mentioned earlier in the wider civic discourse in which many VLOPs and VLOSEs play an important role because of their reach and their function as public spaces and sources of information. Election monitors were able to relate frequently observed narratives that often appear over longer periods of time contributing to this negative effect. Such narratives included claims of foreign allegiances by candidates or parties, anti-EU or anti-democratic positions, climate change denial, and hateful or inciting narratives about particular groups often minorities. Where feeds of content and underlying recommender systems or search results fail to offer alternative viewpoints and access to information from a variety of sources, including fact-checked credible ones, citizens may enter an election cycle already conditioned by positions based on inaccurate information and extreme narratives. Mitigation of this negative effect is not limited to the election cycle and will require longer term efforts to identify and better understand the dynamics and mechanisms at work as well as carefully constructed interventions to encourage diversity of exposure to general interest content from a variety of sources.³²

Negative Effect 12: Citizens cannot engage in constructive deliberation and express themselves in civic discourse.

The fragmentation of the civic discourse mentioned above not only can lead to polarization and the reinforcing of false or extremist narratives, but also limits citizens' ability to engage in deliberation themselves.³³ Citizens' own ability to express themselves and engage in debate may be restricted if there is a lack of opportunities to be exposed to and confronted with alternative views. This affects both their right to free expression and their right to assembly. It can be an issue of how feed algorithms are constructed, or search criteria are designed. Another source of risk of this negative effect is the silencing of voices due to online harassment and violence. Individuals cease to engage in the important public spaces created by some VLOPs because of this type of behaviour from other

³² Options for implementing this and metrics for assessing are presented in Chapter A of European Commission et al., *Study on Media Plurality and Diversity Online : Final Report*. For options for news information in particular see Helberger, 'On the Democratic Role of News Recommenders'.

³³ Lorenz-Spreen et al., 'A Systematic Review of Worldwide Causal and Correlational Evidence on Digital Media and Democracy'.



users, which can include the use of threatening or sexualised deepfakes.³⁴ This particularly affects women and minority groups.

Due to variations in terms of use, functionality, design, and userbase, some services may be more prone to polarization or threatening speech. The lack of accessibility tools on some arguably safer services leaves citizens with disabilities with limited options for engagement, as was pointed out by one of our interviewees. Mitigations for this specific problem have not been explicitly included in the EC Guidelines, though a wide interpretation of the suggestions that providers consider accessibility and inclusivity in relation to fundamental rights impact and measures to ensure access to official information could cover it.³⁵

Negative Effect 13: Voters' choices are manipulated by deliberate interference

The negative effect of voters' choices being a result of manipulation rather than informed free will is closely linked to Negative Effect 9. As mentioned above, lies by politicians are assumed to be part of civic discourse and are still protected speech unless they cross lines into hate speech, incitement, war propaganda, or other unprotected categories. However, deliberate manipulative interference, usually characterised by inauthentic use that capitalises on the scale, speed and functionalities of many very large services and by careful and purposeful timing in the electoral cycle can undermine the integrity of an electoral process. The DSA's limitations on the use of personal data for targeting, the Regulation on Political Advertising, the advertising related commitments in the Code of Practice on disinformation, and the EC Guidelines' provisions on political advertising³⁶ might curtail the kind of microtargeted disinformation that was widespread in the UK's referendum on EU membership and the 2016 elections in the US, Brazil and other countries.³⁷ However as these instruments are all quite new and have not been tested and assessed thoroughly in electoral cycles. Malign use of search optimization strategies,³⁸ deliberate virality campaigns, interlinkages between VLOPs, and private messaging services are also used to manipulate voters and are not covered by rules on political advertising.

In examples given by our interviewees from observations during electoral periods, there is a pattern of using social media and video-sharing VLOPs to lure users into groups on messaging services. Another pattern was identified when manipulative content gains traction in private messaging groups and then crosses over to VLOPs. These tactics are also described in the DISARM framework.³⁹ This is therefore an effect exacerbated by interlinkages with other services that may not be designated as very large services under the DSA and may not be governed in some way by relationships with VLOP providers.⁴⁰ The use of deepfakes, bots, and other forms of inauthentic use as well as concerted efforts

³⁴ Pawelec, 'Deepfakes and Democracy (Theory): How Synthetic Audio-Visual Media for Disinformation and Hate Speech Threaten Core Democratic Functions'; Saurwein and Spencer-Smith, 'Automated Trouble: The Role of Algorithmic Selection in Harms on Social Media Platforms'.

³⁵ EC Guidelines paragraph 4 & 27a

³⁶ DSA article 26.3, Regulation 2024/900 of 13 March 2024, and EC Guidelines 27(e).

³⁷ Loos and Nijenhuis, 'Consuming Fake News'.

³⁸ Samantha Bradshaw, 'Disinformation Optimised'.

³⁹ Supra note 30.

⁴⁰ Broughton Micova and Calef, 'Elements for Effective Systemic Risk Assessment under the DSA'.



of content dissemination through these channels are markers of risk of this negative effect. FIMI is a particular form of deliberate manipulative intervention that has justifiably gained a lot of attention recently.⁴¹

However, according to our interviewees, much of what they observe comes from domestic actors, mainly the candidates and parties involved in the elections. Our interviewees and others provide evidence that both FIMI and domestic attempts at manipulation make use of the whole information ecosystem, not just VLOPs and VLOSEs,⁴² and increasingly make use of generative AI for deception.⁴³ The use of generative AI poses specific challenges to measures aimed at mitigating negative effects from disinformation.⁴⁴ These may not be easily addressed by the measures to detect and identify AI generated content or counter it with warnings and nudges to authoritative sources contained in the EC Guidelines⁴⁵ when it is being deliberately used to interfere with the choice information space in an electoral cycle.

⁴¹ EEAS 2nd EEAS Report on Foreign Information Manipulation and Interference Threats: A Framework for Networked Defence, January 2024. https://www.eeas.europa.eu/eeas/2nd-eeas-report-foreign-information-manipulation-and-interference-threats_en

⁴² Lukito, 'Coordinating a Multi-Platform Disinformation Campaign: Internet Research Agency Activity on Three US Social Media Platforms, 2015 to 2017'; Bennett and Livingston, 'The Disinformation Order: Disruptive Communication and the Decline of Democratic Institutions'.

⁴³ Wirtschafter, 'The Impact of Generative AI in a Global Election Year'.

⁴⁴ Bontcheva and et. al., 'Generative AI and Disinformation: Recent Advances, Challenges, and Opportunities'.

⁴⁵ EC Guidelines paragraphs 37-40



5. METRICS AND DATA FOR EVALUATING EFFECTS AND LEARNING

The EC Guidelines recommend that each provider of a VLOP or VLOSE should individually conduct a post-election review and sets out a list of what should be included in the internal review report. As shown in Table 3, the metrics and measures suggested are mix of ones that can assess the execution of the VLOP or VLOSE’s mitigation measures, such as response times to flagging or violation of terms and conditions, and a couple that could contribute to insight into the effectiveness of risk mitigation measures, such as information from third parties on the impact of measures and the reach and engagement on actioned content.

Table 3 Suggested elements for post-election review by VLOP and VLOSE providers as stated in the EC Guidelines

SUGGESTED POST ELECTION REVIEW REPORT ELEMENTS	EC GUIDELINES PARAGRAPH
Whether internal performance metrics and other assessment criteria for before, during, and after the election were met	59
Lessons learnt	59
Areas for improvement identified	59
Information on the impact of measures from independent researchers, CSOs, and fact-checkers	60
Information on the use and impact of their services from established independent election observer groups	60
Their average response time for terms and conditions violations	61
The distribution of their response time for terms and conditions violations	61
Their average response time content flagged by users and non-state actors	61
The distribution of their response time content flagged by users and non-state actors	61
The reach and engagement of content acted upon	61
The number of violations of certain polices pertaining to elections	61
Number of instances of information manipulation	61
The reach of any media literacy or authoritative initiatives	61



The EC Guidelines also state that VLOP and VLOSE providers should make a version of their post-election review reports public and elicit feedback on them aimed at improving their own measures for future election cycles. This will provide information on what mitigation measures were taken by the services and how they were assessed by the service providers. While this specific feedback loop is very important, the post-election review will be a significant step in the effective implementation of the DSA and should also involve much wider analysis.

Monitoring trends and tracking metrics on specific types of content, such as authoritative source or fact-checked content, and on user engagement with specific nudges, tools, or labels is crucial throughout an electoral cycle. The Integrity Institute's advice on monitoring and trend detection can help with real-time optimization of measures and responses to harmful content or manipulative behaviour on specific platforms.⁴⁶ The very large services that are signatories to the Code of Practice on Disinformation have already committed to sharing trends, information about influence operations, and foreign interference identified on their services among the signatories to the Code.⁴⁷ However, the accounts from our interviewees made clear that the detection of harmful trends from content or user behaviour such as serious disinformation based narratives, harassment, or manipulative interference may require further insight into what is happening on other services, such as messaging apps, on niche online media, and even on mainstream media.

Real-time mitigation of risk and optimisation of measures would benefit greatly from maximum use of the "cooperation and swift and efficient exchange of information cross-platform and with relevant non-state actors" called for by the EC Guidelines from the very start of the cycle.⁴⁸ To be able to contribute effectively to this cooperation, some non-state actors will need access to some of the monitoring insight gathered by the VLOP and VLOSE providers. Election monitors, fact-checkers, and even EMBs will likely be monitoring messaging groups and off-line discussions, as well as speeches by candidates and conversations at rallies. Some of these organisations may be able put their monitoring together with information from VLOP and VLOSE providers to help identify specific sources of risk and malign actors to provide evidence and insight on the wider impact of positive measures such as nudges, bridging algorithms,⁴⁹ or the prioritisation of authoritative sources.

To evaluate the extent to which VLOPs and VLOSEs have played a successful or constructive role in the electoral information spaces as set out in Tables 1 and 2, and ultimately whether the DSA provisions are effective in preventing negative effects on electoral processes, there is a need for more data than is needed for short-term optimisation of specific risk mitigation measures and the post-election review of individual services. This will require looking at metrics and other data across services, comparing

⁴⁶ Integrity Institute, 'Elections integrity best practices: Defining and achieving success Part 2' 4 October 2023. <https://drive.google.com/file/d/1Hl-T9WlQK0Gm6O4gkd8lYduwzKUUpd5b/view?usp=sharing>

⁴⁷ Code of Practice on Disinformation, commitment 16

⁴⁸ EC Guidelines paragraph 54

⁴⁹ The Integrity Institute explains that "In contrast to how engagement-based ranking tends to promote polarizing content, bridging algorithms overweight content that is broadly socially acceptable, in some ways replicating the agenda-setting of traditional media." In Elections integrity best practices: Defining and achieving success Part 2' 4 October 2023. page 32



across member states, and situating data from and about VLOPs and VLOSEs in the wider information ecosystem.

In Table 4, we propose a number of metrics and types of data for evaluating success in preventing each of the negative effects described in Section 3. These include quantitative measures and qualitative data. Many of these metrics or types of data are already available or at least should be due to the commitments made by providers who are signatories to the Code of Practice on Disinformation or the transparency obligations in the DSA. For example, signatories have committed to sharing metrics on the impressions and penetration achieved by inauthentic accounts and violating content, as well as data on the political advertisements refused for not complying with standards and the on the reach and engagement with fact-checked content. Other types of data are already tracked and may be published by EMBs or election observation missions through post-election reports. Some of the metrics and types of data listed in Table 4 can also be used to gain insight into user engagement with or their response to specific mitigation measures, but the aim of this table is to relate them to achievement in relation to the negative effects. Several of the suggested metrics and types of data listed in the second column of Table 4 apply to multiple negative effects, which are noted in the third column. It was beyond the resources of this investigation to track whether and where all the metrics and data might be available, but the fourth column indicated our assumptions as to whether the data would be held by service providers, by other actors, or a combination.



Table 4 Proposals for metrics and data for evaluating prevention of negative effects and learning

	METRIC OR TYPE OF DATA	POTENTIAL NEGATIVE EFFECT	DATA SOURCE
1	Cross platform and cross electoral process analysis of amount and type of resources invested in election-specific content moderation, fact-checking, nudges, inauthentic use detection, rapid response mechanisms, and other measures	All	Platform providers and other institutions/organisations
2	Cross platform analysis of level, reach, type and sources of inauthentic use, including a list of top sources	1, 3, 5, 6, 7, 8, 9, 11, 12, 13	Platform providers
3	Individual and cross platform data on the number of complaints or appeals to content moderation and the speed and outcome of reactions	4, 8, 12	Platform providers
4	Individual and cross platform data on reports of harassment or threats and response metrics	2, 4, 8, 12	Platform providers
5	Election-specific information on resources dedicated to handling complaints from EMBs, election administrators, candidates, and parties	2	Platform providers
6	Cross platform analysis of fact checking rates and exposure to official, labelled, and fact-checked information	1, 3, 5, 6, 7, 9, 11, 12, 13	Platform providers
7	Cross platform analysis of data on responses to nudges (e.g. change of user behaviour or consumed content)	1, 3, 5, 6, 7, 9, 11, 12, 13	Platform providers
8	Individual and cross platform analysis of outcomes of rapid response mechanisms and time taken to respond to EMB, civil society, candidate or political party requests	3, 5, 7, 9, 13	Platform providers



9	Qualitative data from reports of election monitors	1, 3, 5, 6, 7, 10, 13	Other institutions/organisations
10	Compilation of accounts of the quality of cooperation around data access from EMBs and civil society monitors	1, 2, 3, 6, 7, 10	Other institutions/organisations
11	Results of polling or qualitative data on user perceptions of labels and fact-checking	1, 11	Other institutions/organisations
12	Results of polling of citizens through Eurobarometer, member state level polls, and VLOP/VLOSE user polls on trust in the electoral process and acceptance of the result	1, 5, 7	Other institutions/organisations
13	Results of polling of citizens through Eurobarometer, member state level polls, and VLOP/VLOSE user polls on policy preferences, vote choice, result satisfaction, and other preference indicators.	13	Other institutions/organisations
14	Individual and cross platform discourse 'toxicity' measures	2, 4, 12	Platform providers
15	Cross member state analysis of diversity of candidate lists including change between elections, and evidence from parties and civil society	4	Other institutions/organisations
16	Comparison of voter participation metrics over time and across electoral cycles	6	Other institutions/organisations
17	Individual and cross platform analysis of political advertising expenditures and types of targeting used by campaigns and by other actors	8	Other institutions/organisations
18	Individual and cross platform data on the extent of exposure of users to targeted political advertising and the diversity per user	8	Platform providers
19	Individual and cross platform analysis of the numbers, types and sources of any violating advertising or targeting attempted and mitigated by VLOP/VLOSEs	8, 9,13	Platform providers



20	Exposure diversity metrics for recommender systems for individual platforms by user categories and cross platform for each electoral cycle	8, 11, 12	Platform providers
21	Individual and cross platform data on the design and reach of any election specific media literacy initiatives by VLOPs/VLOSEs	3, 5, 6, 7, 9, 11	Platform providers
22	Ranking signals analysis and bias detection outcomes across services	8, 11	Platform providers
23	Individual and cross platform analysis of EMB, candidate, and party satisfaction with VLOPs/VLOSEs measures and responses to requests	8, 9, 10	Other institutions/organisations
24	VLOP/VLOSE providers' accounts of the tools and data access provided to EMBs and civil society monitors	10	Platform providers
25	Metrics on exposure to disinformation or other deceptive use	11	Platform providers
26	Cross platforms metrics on implementation and response to prominence measures taken, use of bridging, or other adaptations to recommender systems	11	Platform providers



As can be seen in the Table, we suggest a great deal of analysis of cross-platform analysis of data. This raises the obvious questions of who should do this and how they can access the data, much of which will be held by VLOP and VLOSE providers and not necessarily be granularly available in post-election review reports.

We recommend collaborative efforts that draw on the expertise and capacities of various actors. Some of the metrics we suggest, such as ones on bias/fairness of recommender systems,⁵⁰ toxicity measures,⁵¹ and exposure diversity measures require large data sets and sophisticated methodologies appropriately handled by groups of independent researchers. Arguably, they are also best used as quantitative indicators to guide qualitative investigation, rather than metrics that alone can generate clear answers to questions about the prevention of these effects. Others will require the insight of civil society groups with election-specific expertise or EMBs or can only be generated by service providers.

Relying on self-initiated Article 40 requests from independent researchers or spontaneous cross-national collaboration among civil society organisations will not be sufficient. It would overburden the DSCs and spread too thin the already taxed resources of academia and civil society. However, the Digital Services Board and the European Cooperation Network on Elections, which brings together all the member state EMBs, could be very useful in providing direction, coordinating cross-platform analysis, and identifying opportunities for collaboration. Based on the advisory role given to it in the DSA's Article 61 that includes coordinating analysis and assisting the Commission in the supervision of VLOPs and VLOSEs, the Board is well placed to lead coordination and strategising for this purpose.

⁵⁰ Amigó et al., 'A Unifying and General Account of Fairness Measurement in Recommender Systems'.

⁵¹ For an example using Google's Perspectives API see Jiménez Durán, Rafael and Müller, Karsten and Schwarz, Carlo, (February 23, 2024). The Effect of Content Moderation on Online and Offline Hate: Evidence from Germany's NetzDG Available at SSRN: <https://ssrn.com/abstract=4230296> or <http://dx.doi.org/10.2139/ssrn.4230296>



6. CONCLUSIONS AND RECOMMENDATIONS

The implementation of the DSA's provisions on VLOPs and VLOSEs should not be reduced to ticking the boxes of compliance for individual services. The aim is to mitigate the risk of negative effects on institutions that are core to how we organise ourselves as societies – fundamental rights, civil discourse, electoral process, among others. It is important to keep in mind the benchmarks for what 'good' versions of these look like, also to be able to recognise systemic failure. This paper has presented the results of an investigation aimed at doing that in an inclusive manner. More specifically we need to hold up the vision for the constructive roles that VLOPs and VLOSEs could be playing as important public spaces, as enablers of diverse voices and connections, and as sources of information. Thinking about what good and failure look like in the narrow risk area of electoral processes and bringing a variety of stakeholders into discussions about them enabled us to define the list of negative effects presented here. It is not an exhaustive list but could be a starting point for determining whether the mitigation of risk to electoral processes under the DSA is working. It can be a framework through which to keep learning with each electoral cycle as services and conditions, as well as the tactics of malign actors, change.

Determining whether the assessment and mitigation of the risk of negative effects on electoral process are working essentially requires finding out if services whose business model is based on garnering attention and engagement can avoid facilitating and exacerbating the kinds of political discourse and electoral behaviour that undermine democracy, free will, and fundamental rights.

VLOPs or VLOSEs are not the instigators of disinformation, polarising or extremist narratives, or threats and harassment on their services. These exist offline as well and come from various political actors, including too often the main candidates and parties participating in an electoral process. What the DSA requires VLOP and VLOSE providers to do it to take a good close look at the role they might be playing in facilitating, amplifying, and exacerbating the effects of these kinds of content and behaviour. An essential step is looking within the ecosystem and users of any single service through internal performance metrics and other means of assessing service-specific measures, but this is not enough. Next to providers' own self-reflective evaluations, looking across services and across electoral processes will be necessary for the Commission, the DSCs, the service providers, and others to understand how the various VLOPs and VLOSEs are contributing – for good or for ill. It will require independent external eyes and election-specific expertise, and, we argue, coordination of efforts.

The EC Guidelines' call for VLOP and VLOSE providers to cooperate with non-state actors, EMBs, smaller digital services, and each other in the acute electoral period are a good start. The cooperation already underway among signatories to the Code of Practice on Disinformation, that includes a joint rapid response mechanism for the EP elections, and the further cooperation through the AI Elections Accord to which most VLOP and VLOSE providers have also committed to are also important venues for cross-service sharing of information as part of the mitigation process. However, there is a need for cross cutting analysis driven more by cooperation with more independent actors, especially ones with election expertise, aimed at longer term learning and assessment of approaches.



1. We therefore recommend an evaluation and learning strategy be developed, with the Digital Services Board and the European Cooperation Network on Elections playing leading roles, that can facilitate the development of effective cross-platform mitigation strategies for systemic risks and, to this end, especially steer the use of data access requests under Article 40 and the additional election-specific transparency, disclosure, and data sharing recommended in the EC Guidelines.
2. We suggest that the framework, specific metrics and types of data set out in Section 5, some of which are already generated by services that are signatories to the Code of Practice on Disinformation or by election-related institutions, could contribute to such a strategy.

VLOPs and VLOSEs could play positive, constructive roles as sources of information and places for deliberation. Because some of them are such important locations of discourse and spaces for the exercise of public participation, the mitigation of risks of negative effects could be interpreted to include some public interest expectations, not too unlike those we have for mainstream media. They cannot, nor should they, be expected to be watchdogs of the behaviour of political actors or electoral processes themselves, but they could be expected to help or at least make it possible for civil society actors, election monitors, and EMBs to hold political actors to account for manipulative interference or campaign irregularities taking place on their services. As could be seen in the discussion of negative effects in Section 4, our interviews with representatives from such stakeholders and others identified several things that could be useful during the pre-electoral and electoral phases of the cycle and to enable them to contribute to evaluation and learning after.

3. We suggest VLOP and VLOSE providers work with election-specific civil society groups, monitoring organisations, and EMBs to build on the experiences of information sharing during the electoral period to develop standard practices and formats. Coordination among providers would be useful to avoid overwhelming these often under-resourced counterparts. This should allow these groups to develop the capacity to make use of shared information over successive electoral cycles. It can also establish channels for discussion and innovation so information sharing can adapt to changes in the design and functionalities of the services or in the electoral contexts.
4. We recommend building on the positive experience those investigating campaign irregularities, FIMI, and other manipulative interference have with ad repositories, CrowdTangle, and other tools. Additional tools should be designed and put in place to enable third parties to effectively study the wider risks to electoral processes. The various third-party actors whose work contributes to protecting electoral integrity have limited capacity to engage directly with large quantities of data and may be unaware of what data exists that can help them. Particularly in efforts to combat disinformation narratives, polarisation, and to identify inauthentic malign use, these third parties need tools that enable them not only to access, but also to easily make use of data held by VLOPs and VLOSEs. At the same time, such tools and efforts need to be protected against targeted attacks of malign actors.



As was evident in the accounts and views of practitioners presented here, the inauthentic use of VLOPs and VLOSEs is a significant source of risk to electoral processes. Polarisation and extremism in civic discourse and society were often viewed as long-term issues that threaten electoral processes and effective democracy, and as trends that are exacerbated by manipulation campaigns feeding and amplifying existing divisions. For many of those interviewed, it was inauthentic use that made the difference between the regrettably standard exaggerations, incendiary remarks, and accusations of political speech and something that requires mitigation.

Generative AI exacerbates the threats posed by disinformation and harassment by elevating the sophistication of deception. Social bots achieve a speed and reach in dissemination that can put a post on par with regulated broadcast media, especially in smaller countries. Furthermore, the pervasiveness of inauthentic content and the difficulty to identify it as such can threaten the general trust of users in any type of content, thus making it harder to establish a common ground of trusted facts that enable constructive discourse and debate. Cross-platform efforts such as the AI Election Accord and the AI-related commitments in the Code of Practice on Disinformation justifiably focus on the identification, labelling, and prevention of inauthentic use that can contribute to the negative effects identified in this report. Especially during and shortly before an election period, quick and coordinated responses of service providers to the spread of inauthentic content by malign actors are important.

On a more fundamental level, this may also call for control measures at the platform level that can limit the automatised content creation during an electoral cycle. However, automatised content has a range of positive uses for candidates and parties as well, such as facilitating multi-homing and promoting the diffusion of accurate and timely messages. Some use of AI created content may be benign and within the bounds of acceptable political speech.

5. Therefore, we recommend that common standards on what constitutes appropriate levels of automatised content may be developed together with a range of election stakeholders including political party representatives and civil society observer and watchdog groups. These could draw on the understanding of impermissible manipulative behaviours agreed by the signatories to the Code of Practice on Disinformation, but should be election-specific and integrate an understanding of permissible use of AI and automated tools. To protect the integrity of elections such a standard may be adaptive and allow for different levels of automatised content depending on the period of the election cycle.

cerre Centre on Regulation in Europe



Avenue Louise 475 (box 10)
1050 Brussels, Belgium
+32 2 230 83 60
info@cerre.eu
www.cerre.eu

 Centre on Regulation in Europe (CERRE)
 CERRE Think Tank

