CEE Digital Democracy Watch

Overseeing the New Digital Era in Election Integrity

Warsaw, 14.05.2026

Position paper — Targeted Stakeholder Consultation
European Democracy Shield — Work Strand on AI in Electoral Processes
Submitted by: CEE Digital Democracy Watch

CEE Digital Democracy Watch welcomes the Commission's initiative to develop guidance on the responsible use of AI in electoral processes and to encourage voluntary commitments from political parties and other relevant actors. We support this initiative as a necessary complement to the existing legislative framework. Beyond the European Democracy Shield, the AI Act transparency guidelines, and multiple communications in the light of the Digital Services Act, on 9 April 2024 all major European political families signed a Commission-facilitated Code of Conduct committing to ethical AI use in campaigning, a precedent to build on.

The value of voluntary commitments lies in what they can do beyond the legal basis: build trust with citizens, establish norms ahead of enforcement, and signal that democratic actors take seriously their own role in shaping the information environment.

This position paper draws on our ongoing research into AI-driven influence operations, platform governance, and the conditions under which democratic participation remains meaningful in an age of generative AI. The research is conducted together with experts from leading institutions, with a particular focus on the peculiarities of the CEE region, a part of Europe that finds itself at the forefront of hybrid threats targeting democratic processes, from coordinated inauthentic behaviour and foreign-linked disinformation to AI-generated influence operations deployed at scale. In response, the region, despite limited institutional capabilities, developed civil society expertise that offers valuable insights for the design of EU-level guidance and voluntary commitment frameworks.

The comment is organised around the four stages of the electoral AI cycle (production, interaction, communication, and campaign) and concludes with recommendations on what credible voluntary commitments should contain and how they should be structured.

1. Production: Committing to transparency about message creation

AI-generated content is already a standard tool in European electoral campaigns. The question is no longer whether political parties use it, but what they use it for. This distinction matters as campaigns increasingly resort to AI slop: high-volume, low-quality synthetic content produced at scale to flood the information environment and drown out substantive political debate. Voluntary commitments at the production level should go beyond the minimum disclosure requirements of the AI Act and TTPA to establish a genuine culture of transparency with voters.

Concretely, this means committing to label all AI-generated or AI-assisted political communications (video, audio, and text) in ways that are citizen-legible, not merely technically compliant. The test of a label is whether a voter can meaningfully interpret and act on it. Commitments should also address the integrity of provenance: parties should commit not to strip, obscure, or circumvent technical markers such as C2PA watermarks on content they publish or amplify, and not to use synthetic reproductions of opponents' voices or likenesses for political communication purposes. While Article 50 of the AI Act introduces horizontal transparency obligations for AI-generated content, voluntary commitments should go further, as political content carries a particular sensitivity, given its direct capacity to shape electoral outcomes and distort voters' informed choices.

The 2026 Hungarian elections demonstrated concretely what the absence of such norms produces: a campaign environment saturated with unlabelled AI-generated videos, deepfake impersonations of candidates, and coordinated amplification through synthetic accounts with no political actor bound by any commitment to behave otherwise. Voluntary commitments, if designed seriously and adopted widely, can begin to establish the norm that AI transparency is a mark of democratic legitimacy. Platforms have demonstrated willingness to monitor and assess such content, as evidenced by their participation in the Rapid Response System under the Code of Practice on Disinformation. The CEE region offers particular evidence of this risk: weaker institutional oversight, smaller language markets with limited moderation capacity, and higher exposure to foreign-linked influence operations make the gap between legislative floor and actual practice especially consequential.

2. Interaction: Committing to accuracy and oversight

Political parties and campaigns increasingly use AI-powered tools (chatbots, personalised messaging systems, AI-integrated social media) to communicate with voters. These systems carry specific risks that voluntary commitments should address directly.

LLMs embedded in campaign communication voter-facing tools or deployed on platforms used for political outreach have been shown to generate inaccurate, outdated, or misleading information on electoral matters, particularly in real-time or rapidly developing contexts. Parties should commit to human oversight of AI-generated voter-facing communications, to accuracy-checking protocols for AI outputs on electoral facts, and to clear disclosure when voters are interacting with an AI system rather than a human representative. The private sector and leading AI providers are already developing relevant techniques that can be adapted in this context, including electoral safeguards such as redirecting sensitive political queries to authoritative sources, restricting AI-generated election-related content, and declining to produce materials that impersonate candidates or electoral officials.

The Commission guidance should make clear that voluntary commitments in this area are not merely about legal compliance but about the epistemic relationship between political actors and citizens. A party that knowingly deploys an AI system that misleads voters about polling dates, candidate positions, or electoral procedures has undermined democratic participation regardless of whether it has technically violated any law. Robust guidance already exists in this space: the DSA elections guidelines, which incorporate stress-testing requirements, provide a strong foundation. Companies should be encouraged to participate in that framework on equal terms, building on the structured relationships already developed with Very Large Online Platforms.

3. Communication: Committing to fair use of AI in reaching voters

AI-driven content recommendation and targeting systems give campaigns powerful tools to shape what voters see and when. Beyond algorithmic suggestions that fuel social media systems, LLMs offer tempting new ways to alter voters' information environments, from personalised messaging calibrated to individual psychological profiles to the manipulation of model outputs to systematically favour particular candidates or narratives. Voluntary commitments should address the responsible use of these tools by committing to transparency about how AI is used to segment and reach audiences, and to avoiding

targeting strategies that exploit psychological vulnerabilities or suppress participation among specific groups.

Parties should further commit to refraining from the deliberate manipulation of large language model outputs through training data poisoning or prompt injection techniques. These practices are now offered commercially by marketing agencies as a form of AI-era SEO positioning, designed to make models surface favourable narratives about a party or candidate and suppress unfavourable ones. Unlike traditional advertising or even targeted content campaigns, LLM poisoning is structurally hidden from the citizen encountering its effects: there is no label, no sponsored tag, no disclosure mechanism, but a distorted information environment whose manipulation is invisible by design.

Parties should also commit to supporting, rather than circumventing, platform moderation systems. The Hungarian case is instructive here: AI-generated political videos were systematically routed through nominally independent pages to bypass VLOPs political advertising restrictions and their terms and conditions, exploiting the gap between platform policy and enforcement reality. This behaviour was directly inconsistent with the spirit of the 2024 Code of Conduct, which committed signatories to ethical online campaigning standards, yet the Code offered no mechanism to address it. A voluntary commitment not to use AI to circumvent content governance systems, however imperfect those systems are, must this time be paired with a credible accountability structure.

Smaller languages and minority communities face structural disadvantages in AI-mediated political communication: lower moderation quality, weaker content provenance infrastructure, and greater vulnerability to AI-generated disinformation. Voluntary commitments should therefore explicitly address linguistic equity, including a commitment not to exploit moderation blind spots in smaller EU languages for campaign advantage.

4. Campaign: Committing to special electoral-period safeguards

The electoral period is the moment at which AI risks are most acute and the stakes for democratic legitimacy are highest. Voluntary commitments should include specific electoral-period provisions that go beyond business-as-usual conduct standards.

These should include: a commitment to disclose all AI use in political communications during the formal campaign period, in a format accessible to electoral authorities and civil society monitors; a commitment to refrain from publishing or amplifying synthetic

impersonations of candidates or electoral officials; and a commitment to cooperate with election incident protocols, including national and EU-level mechanisms for public communication about AI-driven threats to electoral integrity.

Canada's public protocol on foreign interference offers a useful model for the latter. The planned Stakeholder Platform at the Centre for Democratic Resilience, alongside the established Code of Practice on Disinformation model operating under the DSA, represent a valuable template that should be actively developed and extended to AI providers, bringing them into the same structured accountability frameworks that have already proven effective for Very Large Online Platforms. Parties and campaigns that commit to cooperate with such protocols by sharing information about suspected AI-driven attacks on their own communications would make a material contribution to collective electoral resilience.

Further cooperation

CEE Digital Democracy Watch remains available to contribute to the development of the Commission guidance, including:

- Evidence on AI use in recent EU electoral campaigns and the governance gaps that voluntary commitments should address, with particular depth on CEE electoral cycles
- Input on commitment design and monitoring frameworks, drawing on comparative analysis of existing voluntary instruments
- Practical recommendations for operationalising commitments in smaller EU languages and higher-risk electoral environments

Selected references

Road to the Ballot: AI Governance and Election Integrity, CEE Digital Democracy Watch, 2026 (forthcoming)

Election safeguards update, Anthropic, April 2026 (https://www.anthropic.com/news/election-safeguards-update)

Code of Conduct for the 2024 European Parliament Elections, European Commission, April 2024 (https://commission.europa.eu/document/download/bebd9b72-fbb9-42f3-bcea-dbace7e0650f_en)

How Europe's Far Right Used Unlabelled AI to Win Votes, EUobserver, December 2025 (https://euobserver.com/4811/how-europes-far-right-used-unlabelled-ai-to-win-votes-and-now-writes-the-rules/)

Synthetic Influence — Deepfakes and Artificial Intelligence in the Hungarian Election Campaign, Political Capital / Hungarian Digital Media Observatory, April 2026 (https://politicalcapital.hu/pc-admin/source/documents/HDMO2_PC_SyntheticInfluence_260410.pdf)

Circumventing Meta's Ban with AI-Generated Campaign Videos, EDMO / Political Capital, March 2026 (https://edmo.eu/publications/circumventing-metas-ban-with-ai-generated-campaign-videos-and-super-forecasting-of-fideszs-defeat/)

Adding Fuel to the Fire: AI Information Threats and Crisis Events, Alan Turing Institute / CETAS, February 2026 (https://cetas.turing.ac.uk/sites/default/files/2026-02/cetas_research_report_-_adding_fuel_to_the_fire_-_ai_information_threats_and_crisis_events.pdf)

Tech Accord to Combat Deceptive AI in Elections, Munich Security Conference, 2024 (https://securityconference.org/en/aielectionsaccord/)

Canada's Protocol on Foreign Interference in Federal Elections, Government of Canada (https://www.canada.ca/en/democratic-institutions/services/protecting-democracy/public-protocol.html)

CEE Digital Democracy Watch
Mokotowska 43/104, 00-551 Warsaw, Poland
www.ceeddw.org
KRS 0001090110 | EU Transparency Register 114284692189-05