



HAL
open science

La Commission européenne publie ses lignes directrices sur l'atténuation des risques systémiques pour les processus électoraux

Philippe Mouron

► **To cite this version:**

Philippe Mouron. La Commission européenne publie ses lignes directrices sur l'atténuation des risques systémiques pour les processus électoraux. *Revue européenne des médias et du numérique*, 2024, printemps-été 2024 (69/70), pp.15-19. hal-04629621

HAL Id: hal-04629621

<https://amu.hal.science/hal-04629621v1>

Submitted on 10 Sep 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0
International License



LA COMMISSION EUROPÉENNE PUBLIE SES LIGNES DIRECTRICES SUR L'ATTÉNUATION DES RISQUES SYSTÉMIQUES POUR LES PROCESSUS ÉLECTORAUX

-

Revue européenne des médias et du numérique, n° 69-70, printemps 2024, pp. 15-19

Philippe MOURON

Professeur de droit privé

LID2MS – Aix-Marseille Université

En application du règlement sur les services numériques (DSA), il s'agit d'atténuer les risques systémiques liés aux grandes plateformes et aux grands moteurs de recherche, afin de garantir la sincérité du scrutin en protégeant les électeurs de sources d'information erronées ou trompeuses.

Les périodes électorales sont, comme on le sait, propices à la diffusion de fausses informations et autres tentatives d'ingérences dans le débat public. C'est ce phénomène qui avait justifié en son temps l'adoption en France de la loi du 22 décembre 2018 sur la lutte contre la manipulation de l'information (voir *La rem*, n°45, p.66-73), qui comporte un important volet destiné à prévenir la propagation de tels contenus pendant les campagnes électorales, que ce soit dans les services de médias audiovisuels comme au sein des plateformes de partage de contenus en ligne. À l'approche des élections européennes, la vigilance s'impose d'autant plus que de nouveaux moyens de manipulation ont fait leur apparition, et sont susceptibles de perturber encore davantage le discours politique. Tel est le cas, en particulier, des systèmes d'intelligence artificielle (SIA), qui rendent possible la création de *deepfakes* visuels ou sonores. Les avancées techniques prodigieuses qu'ont connues ces systèmes permettent désormais de **fabriquer de toutes pièces un message faussement attribué à un candidat, avec un réalisme singulièrement troublant**. L'usage de tels procédés en période électorale peut se révéler extrêmement risqué, comme cela s'est produit récemment en Slovaquie. Dans ce contexte, la Commission européenne a publié le 26 avril dernier des lignes directrices éclairant l'application de l'article 35 du règlement sur les services numériques (DSA)¹.

Le précédent slovaque : un *deepfake* diffusé 48 heures avant le scrutin

La Slovaquie a fait l'expérience de ce qui semble être une tentative de manipulation de l'information ayant eu recours à un SIA lors de ses dernières élections législatives, en septembre 2023.

Un enregistrement audio d'une conversation entre l'un des candidats favoris des sondages, connu pour ses positions pro-Union européenne et Otan, et une journaliste a ainsi été diffusé quarante-huit heures avant le scrutin, alléguant l'existence d'une fraude électorale en préparation. Bien que rapidement identifié comme un *deepfake*, le fichier est devenu viral sur les plateformes, pendant la période d'embargo sur les sondages préélectoraux que doivent respecter les médias de presse écrite et de médias audiovisuels². C'est finalement une autre liste qui est arrivée en tête de l'élection, sous la houlette d'un candidat ne cachant pas ses sympathies pour Viktor Orbán et pour Vladimir Poutine. S'il n'est pas possible de savoir dans quelle mesure ce contenu a pu influencer une partie des électeurs, tant le contexte que la nature du procédé employé laissent présumer une ingérence russe dans le scrutin³. De là, Reporters sans frontières a appelé à **la création d'un délit spécifique visant à sanctionner les *deepfakes* portant atteinte aux journalistes et au droit du public à l'information**⁴.

Si le projet de loi SREN (Sécuriser et réguler l'espace numérique, voir *infra*) entend étendre le délit d'atteinte à la représentation de la personne de l'article 226-8 du code pénal aux contenus générés par intelligence artificielle, force est de constater qu'il ne permettra pas de prévenir tout risque d'atteinte au droit du public à l'information ou à la sincérité du scrutin. L'infraction est, en effet, exclue lorsque le *deepfake* est identifié comme tel ou bien lorsqu'il porte sur une personne fictive, ce qui relève de la liberté de création. Or, de tels contenus pourraient quand même avoir une certaine influence en contexte électoral, ne serait-ce qu'en brouillant le discours officiel des candidats en ajoutant une « couche » de contenus trompeurs. Le cas récent des *deepfakes* présentant les fausses nièces de Marine Le Pen en est un bon exemple⁵.

Des lignes directrices précisant la portée de l'article 35 du DSA

Faute de pouvoir recourir systématiquement à une action judiciaire, les lignes directrices publiées par la Commission européenne tendent à prévenir de tels risques en période électorale.

Ces recommandations ont été prises sur la base de l'article 35, paragraphe 3 du DSA, relatif aux mesures d'atténuation des risques systémiques que doivent mettre en œuvre les fournisseurs de très grandes plateformes en ligne et les très grands moteurs de recherche. Ces risques, qui sont visés à l'article 34, incluent entre autres **la diffusion de contenus illicites, l'atteinte à certains droits et libertés protégés par la Charte des droits fondamentaux de l'Union européenne, mais aussi, et surtout, « tout effet négatif réel ou prévisible sur le discours civique, les processus électoraux et la sécurité publique »**. De façon non exhaustive, les mesures peuvent porter sur la conception et le fonctionnement des interfaces en ligne, des systèmes de modération, des algorithmes et autres systèmes de recommandation, sur la présentation et le référencement des publicités, sur l'identification de contenus trompeurs générés par intelligence artificielle, ou encore prévoir une coopération avec d'autres fournisseurs de services ou avec les signaleurs de confiance.

Si les lignes directrices de la Commission portent spécifiquement sur ces mesures, elles sont, bien sûr, à corréler à d'autres dispositions du DSA, tel que l'article 25 qui interdit les interfaces trompeuses, ainsi que d'autres règlements récemment promulgués ou en passe de l'être, tels que le règlement du 13 mars 2024 sur la publicité à caractère politique⁶ (voir *La rem*, n°64, p.5) et le règlement sur l'intelligence artificielle⁷ (voir *infra*). De plus, elles peuvent également inspirer les pratiques des fournisseurs de services qui n'ont pas été désignés comme « très grandes plateformes », au vu de l'importance qu'ils peuvent jouer dans la circulation en ligne de l'information. Enfin, si elles ont vocation à s'appliquer dans la mesure du possible dès la campagne des prochaines élections européennes, elles valent plus généralement pour toutes les élections à venir, y compris celles des États membres. Elles sont donc vouées à devenir le document de référence quant à l'application du DSA pendant les périodes électorales à tous les niveaux, que ce soit régional ou local.

Leur contenu comporte plusieurs séries de mesures, portant principalement sur la création de processus internes et sur le renforcement de la canalisation des contenus. Une exigence d'identification permanente des contenus générés par intelligence artificielle et une coopération soutenue avec d'autres acteurs de la société civile sont aussi à l'ordre du jour.

La création de processus internes de suivi des risques en période électorale

Tout d'abord, il appartient aux très grandes plateformes et très grands moteurs de recherche de tenir à jour une base de documentation complète sur les élections (identité des candidats, positionnement et programme des partis, manifestations et rencontres, enjeux nationaux ou locaux...), afin de bien comprendre les circonstances dans lesquels des risques systémiques pourraient survenir.

On imagine que des contenus qui peuvent sembler anodins, ou sans risques à première vue, peuvent, en fait, participer d'une campagne de manipulation si on les replace dans le contexte électoral. Ces informations pourront utilement être employées par les modérateurs, lesquels devront attester de compétences linguistiques adéquates et d'une bonne connaissance des enjeux des élections. Dans l'idéal, la Commission recommande **la constitution d'équipes spécialisées chargées d'une veille sur le déroulement des campagnes électorales**. Le but est de pouvoir mettre en œuvre des processus internes propres à chaque élection, et d'en faire état dans les conditions d'utilisation et de modération. Les informations ainsi collectées doivent ensuite servir de bases aux mesures concrètes d'atténuation des risques, qui sont elles-mêmes variées.

Une canalisation renforcée des contenus de nature politique

C'est là qu'on retrouve en premier lieu l'objectif de canalisation des contenus, qui figurait déjà dans le Code de bonnes pratiques de lutte contre la désinformation (voir *La rem*, n°63, p.10). Il consiste à hiérarchiser les contenus en fonction de leur source et de la qualité de celle-ci. Sa portée se trouve affinée pour les périodes électorales de plusieurs manières.

Les fournisseurs de services sont censés mettre en avant les informations officielles relatives aux élections (périodes de campagne, lieux de vote, communications des candidats...) par des moyens souples et facilement accessibles (pop-ups, bannières, onglets, redirection *via* des liens hypertextes...). De même, ils doivent mettre en place tout un ensemble de dispositifs pratiques pour différencier les sources d'information : marquage des comptes officiels des candidats et partis, des comptes institutionnels ; marquage des contenus erronés ou trompeurs ayant fait l'objet d'un *fact checking* ; démonétisation des contenus de désinformation... À l'image des journaux classés dans un kiosque, **l'utilisateur doit pouvoir aisément distinguer les contenus en fonction de leur provenance**, et bénéficier d'incitations à consulter les informations

électorales officielles. Les systèmes de recommandation devront également être calibrés en fonction de ces paramètres, de manière à mieux (dé)référer les comptes et les contenus.

On doit garder à l'esprit que ces mesures interviennent en complément d'autres dispositions et recommandations qui restent applicables en période électorale. C'est ainsi que la mise en avant des contenus issus de sources journalistiques et médiatiques, qui est une recommandation prévue par le Code de bonnes pratiques contre la désinformation, devra aussi être respectée. La position des services de médias est d'autant plus essentielle au sein des très grandes plateformes qu'ils bénéficient désormais d'un « privilège » opposable en termes de modération au sens de l'article 18 du règlement sur la liberté des médias du 11 avril 2024⁸ (EMFA, voir *infra*). Hors des cas intéressant la modération de contenus illicites, les fournisseurs de très grandes plateformes devront avertir les services de médias de leur décision de procéder au retrait ou à la restriction de visibilité de l'un de leurs contenus vingt-quatre heures à l'avance, en leur laissant la possibilité d'apporter une réponse adéquate aux griefs invoqués.

Le règlement invite aussi à distinguer les services de médias indépendants de ceux qui ne le sont pas par une obligation de déclaration rendue publique, qui peut être contestée devant l'autorité de contrôle nationale.

Une coopération entre fournisseurs de services, l'identification des influenceurs politiques et diverses mesures d'éducation aux médias

Des mesures complémentaires sont prévues par les lignes directrices pour contribuer à cette organisation. Les fournisseurs de services sont invités à **échanger leurs informations quant aux risques systémiques qu'ils auraient identifiés, afin de pouvoir mener une action coordonnée**. De même sont-ils invités à anticiper l'application du règlement sur la publicité à caractère politique, là encore par le biais d'un marquage et d'obligations de transparence (identité du parraineur, montants et avantages reçus, etc.), ainsi que par la tenue d'un répertoire à jour des annonces publicitaires à caractère politique et par la prévention des utilisations abusives.

Les influenceurs politiques sont également ciblés par les lignes directrices : les fournisseurs de services devront ainsi proposer une fonctionnalité leur permettant de déclarer la nature de leurs contenus, et de publier les informations afférentes à leur parrainage, afin d'être clairement identifiés comme des comptes de nature politique. L'existence de faux influenceurs générés

par intelligence artificielle intervenant dans le débat électoral, à l'image des prétendues nièces de Marine Le Pen, atteste du besoin de transparence sur ces pratiques.

Enfin, les lignes directrices prévoient **le recours à des mesures d'éducation aux médias destinées aux utilisateurs**. Outre la collaboration avec des organisations dédiées, les fournisseurs de services sont invités à proposer des outils ludiques de compréhension, tels que des jeux vidéo « *sur la génération de désinformation, favorisant une réflexion critique sur les tactiques utilisées pour influencer* ».

Une exigence d'identification permanente des contenus générés par intelligence artificielle

Les lignes directrices de la Commission alertent aussi sur les risques que présentent les SIA, qui peuvent être la source de nombreuses informations trompeuses et autres « hallucinations » de nature à induire les électeurs en erreur, telles celles que nous avons mentionnées plus haut. Sur ce point, les fournisseurs de SIA sont invités à anticiper l'entrée en vigueur du règlement sur l'intelligence artificielle, **en intégrant des fonctionnalités de marquage durable et d'identification des contenus procédant d'un hypertrucage**, afin qu'ils puissent être plus facilement décelés par les fournisseurs de très grandes plateformes et de très grands moteurs de recherche. Pour les mêmes raisons, ces derniers devront déployer tous les procédés techniques utiles et efficaces permettant d'identifier des contenus générés à l'aide de SIA, en particulier ceux qui intéressent les hommes et femmes politiques candidats ou candidates et les partis politiques. Les *deepfakes* sont bien sûr en ligne de mire, en particulier ceux qui sont générés en utilisant l'image et la voix de personnalités identifiées ou identifiables. La Commission n'entend pas interdire la création et la diffusion de tels contenus, ne serait-ce qu'au titre de la liberté d'expression, qui inclut la parodie et la satire. Ils devront simplement être mieux canalisés à l'instar des autres éléments d'information erronés ou trompeurs.

Le cas des fausses nièces de Marine Le Pen est à ce niveau éclairant. En effet, les auteurs de ces vidéos hypertrucées n'ont nullement menti sur leur origine, le recours à des SIA étant clairement indiqué dans la description de leurs comptes. **Cette mention n'est cependant plus visible une fois les vidéos partagées**, celles-ci se fondant ensuite dans la masse des informations nourrissant le fil des utilisateurs. Ces derniers ne sont donc plus à même de distinguer clairement les contenus truqués de ceux qui ne le sont pas, comme semblent en

attester certains commentaires postés sous les *deepfakes* précités. Pour cette raison, les lignes directrices insistent sur la nécessité de recourir à un étiquetage permanent de ces contenus, y compris chaque fois qu'ils sont partagés par des utilisateurs.

Un contrôle exercé par des tiers à des fins de recherche

Dans la continuité de l'article 40 du DSA, la Commission recommande aux fournisseurs de services d'ouvrir l'accès à leurs données à des tiers, en particulier à des chercheurs mais pas seulement, spécialement dans le cadre des périodes électorales.

Ainsi, il est prévu de recourir à des outils simples d'utilisation, tels que des tableaux de bord ou des interfaces aidant à suivre le plus rapidement possible les éventuelles campagnes de manipulation de l'information et autres tendances générales de désinformation, à des fins d'étude, de recherche et de publication. Si l'article précité est dédié principalement à l'accès des chercheurs, **les lignes directrices précisent que cet accès peut être ouvert à un plus grand nombre de personnes en fonction des circonstances.** On pense naturellement aux journalistes et autres services de *fact-checking* des services de médias, qui disposent d'une expertise certaine sur ce type de problématiques. Cette ouverture serait d'autant plus judicieuse que ces professionnels sont à même d'intervenir sur des temps plus courts que ceux des chercheurs. Les périodes électorales ne durent au mieux que quelques semaines, un tel élargissement devrait logiquement être encouragé. À ce titre, on ne peut que déplorer la décision récemment prise par Meta de se dispenser de l'outil CrowdTangle, lequel permettait aux journalistes d'identifier et pister les contenus de désinformation avec une grande précision (voir *infra*).

Au-delà, la Commission appelle les fournisseurs de très grandes plateformes et de très grands moteurs de recherche à **collaborer activement avec les autorités nationales chargées du contrôle des élections, mais aussi avec des organes non étatiques, tels que les universitaires ou bien les organisations de la société civile**, parmi lesquelles peuvent figurer des signaleurs de confiance au sens du DSA, mais aussi des fournisseurs de services de médias qui sont censés disposer « *de normes et de procédures éditoriales internes bien établies* » et « *sont généralement considérés comme des sources d'information fiables* ».

L'objectif des lignes directrices est, certes, louable, **mais on peut légitimement s'interroger sur les moyens qui seront alloués à la création de ces processus ainsi que sur la formation**

qui sera dispensée en la matière aux modérateurs. Il sera, par ailleurs, difficile de mettre en œuvre l'ensemble des recommandations à court terme, et en particulier pour les élections européennes, qui sont très proches. Sur ce point, la Commission demande aux fournisseurs de faire leurs meilleurs efforts pour adapter les processus existants, en tenant compte du contexte transfrontière dans lequel la campagne et le scrutin vont se dérouler.

S'agissant de la France, l'Arcom a, elle aussi, adressé aux fournisseurs de services une liste de préconisations à respecter durant la campagne des élections européennes⁹. Celles-ci reprennent les lignes directrices de la Commission sous une forme plus synthétique, mais identique sur le fond. Elles s'ajoutent à la recommandation concernant les services de médias audiovisuels linéaires¹⁰, dont la portée reste classique, si ce n'est pour ce qui concerne le relevé des temps d'intervention des soutiens aux listes de candidats, sujet sur lequel l'Autorité se montre déjà plus exigeante¹¹, suite à une décision du Conseil d'État¹². La régulation de la liberté d'expression en matière électorale se doit bien d'évoluer pour prendre en compte la multiplicité et le caractère protéiforme des sources d'information des électeurs.

PM

¹ Lignes directrices de la Commission à l'intention des fournisseurs de très grandes plateformes en ligne et de très grands moteurs de recherche en ligne sur l'atténuation des risques systémiques pour les processus électoraux, présentées en vertu de l'article 35, paragraphe 3, du règlement (UE) 2022/2065, communication de la Commission n° C/2024/3014, 26 avril 2024.

² Meaker Morgan, « Slovakia's Election Deepfakes Show AI Is a Danger to Democracy », *wired.com*, October 3, 2023.

³ Devine Curt, O'Sullivan Donie, Lyngaas Sean, « A fake recording of a candidate saying he'd rigged the election went viral. Experts say it's only the beginning », *cnn.com*, February 1, 2024.

⁴ Berthier Vincent, « Réouverture de l'enquête sur le deepfake de la journaliste slovaque Monika Tódová : RSF appelle à la création d'un délit pour ce type d'attaque », *Reporters sans frontières*, 6 mars 2024.

⁵ Voir notre éclairage sur le sujet : Mouron Philippe, « Quelle régulation pour les deepfakes de type "Amandine Le Pen" ? », *Les Surligneurs*, 18 avril 2024.

⁶ Règlement n° 2024/900 du Parlement européen et du Conseil du 13 mars 2024 relatif à la transparence et au ciblage de la publicité à caractère politique.

⁷ La dernière version du règlement, en date du 16 avril 2024, est consultable sur le site : <https://artificialintelligenceact.eu/fr>

⁸ Règlement n° 2024/1083 du Parlement européen et du Conseil du 11 avril 2024 établissant un cadre commun pour les services de médias dans le marché intérieur et modifiant la directive 2010/13/UE (règlement européen sur la liberté des médias).

⁹ Préconisations relatives à la lutte contre la manipulation de l'information sur les plateformes en ligne en vue des élections au Parlement européen du 6 au 9 juin 2024, 6 mars 2024.

¹⁰ Recommandation de l'Arcom du 6 mars 2024 aux services de radio et de télévision en vue de l'élection des représentants au Parlement européen les 8 et 9 juin 2024.

¹¹ Voir notre éclairage : Mouron Philippe, « L'Arcom est-elle orientée lorsqu'elle comptabilise le temps d'intervention audiovisuelle de Philippe de Villiers ? », *Les Surligneurs*, 13 mars 2024.

¹² CE, 5^e et 6^e Ch. réunies, 13 février 2024, n° 463162, *RLDI*, n° 212, mars 2024, p. 8-12, note E. Derieux.