

Deepfakes en campagne électorale : quelle réponse juridique ?

La désinformation est devenue une technique couramment utilisée en période de campagne électorale. Conscient de ce danger, le législateur a introduit un dispositif de référé spécial pour lutter contre la diffusion massive de fausses informations en campagne électorale. Codifié à l'[article L. 163-2 du code électoral](#), ce référé d'urgence – parfois surnommé « *référé fake news* » – permet de faire cesser rapidement la propagation de fausses nouvelles en ligne durant les campagnes. Dans ce le cadre du scrutin municipal à venir, il pourrait être particulièrement utile face à la prolifération des contenus audiovisuels manipulés par l'intelligence artificielle (deepfakes).

Quand s'applique-t-il ?

L'article L. 163-2 du Code électoral ne s'applique que « *pendant les trois mois précédant le premier jour du mois d'élections générales et jusqu'à la date du tour de scrutin où celles-ci sont acquises* ». Autrement dit, il couvre la période de campagne officielle des principales élections (présidentielle, législatives, régionales, départementales, municipales etc.).

D'autre part, il est matériellement ciblé : les contenus visés doivent consister en des « *allégations ou imputations inexactes ou trompeuses d'un fait de nature à altérer la sincérité du scrutin à venir* ». Cette définition particulièrement large pourrait donc englober les deepfakes véhiculant des faits mensongers susceptibles d'égarer les électeurs (par exemple, une vidéo falsifiée d'un candidat tenant des propos qu'il n'a jamais tenus). Le texte exige en outre que la diffusion de ces fausses informations soit effectuée « *de manière délibérée, artificielle ou automatisée et massive* » via un service de communication au public en ligne.

Cette condition vise les campagnes de désinformation organisées et virales : par exemple, l'utilisation de bots ou de moyens artificiels pour amplifier la portée d'un deepfake. Un simple contenu isolé ou une rumeur artisanale, même malveillante, échappe donc, *a priori*, à ce dispositif – ce qui traduit la volonté du législateur de ne cibler que les ingérences numériques les plus graves et coordonnées.

La procédure

La procédure de référé prévue par l'article L. 163-2 du Code électoral mobilise le juge judiciaire dans des délais extrêmement brefs, afin d'apporter une réponse rapide à la désinformation virale. La loi prévoit que le juge des référés – uniquement le Tribunal judiciaire de Paris – peut être saisi par « *le ministère public, tout candidat, tout parti ou groupement politique (...) ou toute personne ayant intérêt à agir* ».

L'intérêt à agir est présumé pour les candidats et partis, directement concernés par la sincérité du scrutin. D'autres personnes, comme une association de défense de la démocratie ou un électeur, doivent justifier en quoi la fausse information est susceptible de porter atteinte à un intérêt légitime les concernant. Le défendeur visé par le référé sera généralement un intermédiaire internet : la loi habilite le juge à prescrire « *toutes mesures proportionnées et nécessaires* » aux acteurs de l'internet tels que les hébergeurs de contenu ou, à défaut, aux fournisseurs d'accès. En pratique, cela peut consister à ordonner le retrait d'un contenu mensonger, le déréférencement d'un site ou le blocage d'un compte diffusant massivement le *deepfake* contesté.

Le caractère d'urgence de la procédure est affirmé par des délais stricts : le tribunal saisi doit statuer dans les 48 heures. En cas d'appel, la cour d'appel désignée doit également rendre sa décision sous 48 heures. Ce calendrier resserré est censé garantir une réactivité compatible avec la vitesse de propagation des infox en ligne. En théorie, un *deepfake* viral publié en pleine campagne pourrait ainsi faire l'objet d'une ordonnance de retrait en quelques jours maximum, limitant le risque qu'il altère la sincérité du scrutin et entraîne la confusion des électeurs.

Un dispositif rarement mobilisé

Fait notable, depuis son entrée en vigueur il y a plus de quatre ans, ce référé anti-inbox a été très peu utilisé. La doctrine relève qu'il donne l'impression d'être « *très peu appliqué* », notamment parce que le texte est « *tellement restrictif* » que seules des situations exceptionnelles peuvent y satisfaire.

Dans cette première application du référé, deux élus avaient demandé le retrait d'un message diffusé sur Twitter par le ministre de l'Intérieur de l'époque, qu'ils estimaient mensonger et de nature à fausser la campagne des élections européennes de 2019. Le juge des référés a rejeté la demande, considérant qu'« *aucun des critères nécessaires pour retirer un contenu n'était respecté* ». En effet, après examen, le tribunal a jugé que le tweet litigieux – quoique possiblement « *exagéré* » – n'était pas « *manifestement inexact ou trompeur* » dès lors qu'il se rattachait à un fait réel (une intrusion dans un hôpital).

Par ailleurs, la diffusion du tweet n'avait pas été *artificiellement amplifiée* par des moyens automatisés (pas de bots ni de campagne coordonnée rémunérée) et son impact sur la sincérité du scrutin n'était pas avéré : le message avait été aussitôt contesté publiquement, permettant aux électeurs de se forger leur opinion sans risque manifeste de manipulation. Le juge des référés n'entend pas devenir l'arbitre des polémiques politiques ordinaires, et n'interviendra qu'en présence d'une manipulation avérée, grave et caractérisée. En pratique, cela signifie que seuls des deepfakes manifestement falsifiés et massivement propagés de façon coordonnée pourraient justifier une mesure de référé sur ce fondement. Les simples *fake news* isolées ou les intox mineures de campagne, aussi regrettables soient-elles, échappent au périmètre de l'article L. 163-2 du Code électoral, ce qui limite d'autant son champ d'action.

Réactivité et obstacles procéduraux

Même lorsque le référé *fake news* est envisageable, son déploiement effectif peut se heurter à des contraintes pratiques. D'abord, l'exemple de 2019 a montré que le délai théorique de 48 heures n'est pas toujours tenu. Dans l'affaire précitée, malgré l'urgence affichée, l'audience n'a eu lieu que *six jours* après la saisine, et la décision a été rendue sept jours après la saisine. Certes, l'encombrement juridictionnel et la nécessité d'appeler en cause les plateformes concernées (ici Twitter) peuvent expliquer ce léger dépassement. Il n'en demeure pas moins que, face à la viralité de certaines fausses vidéos (susceptibles d'atteindre des millions de vues en quelques heures), si la décision est rendue près d'une semaine après la saisine, la diffusion se poursuit et le risque d'altérer le scrutin s'intensifie. Ensuite, la question de l'identification du bon défendeur peut ralentir la procédure. La loi désigne prioritairement les hébergeurs ou réseaux sociaux, mais encore faut-il assigner la bonne entité juridique. Ainsi, Twitter France s'est avérée incompétente, le vrai opérateur étant Twitter International basé à l'étranger. Ce type de complication, tout comme l'éventuelle nécessité de faire exécuter l'ordonnance à l'étranger ou via les fournisseurs d'accès internet français (en dernier recours), peut nuire à la rapidité et à l'effectivité de la mesure.

Il n'en demeure pas moins que l'intégration massive de l'intelligence artificielle dans la communication politique laisse à craindre que, à l'occasion du prochain scrutin des élections municipales, le référé *fake news* prenne de l'ampleur.

David Conerardy, Avocat à la Cour

[Cabinet Seban et Associés](#)

