

Multi-stakeholder engagement and corrective measures to contribute to a transparent and inclusive online ecosystem FR (French)

[00:00:18] Bienvenue dans le module 5, le dernier de ce MOOC sur l'information et les élections à l'ère numérique. Vous savez que ce cours est organisé par le Centre Knight pour le journalisme dans les Amériques avec le soutien de l'UNESCO et du PNUD. Je m'appelle Albertina Piterbarg, je suis experte électorale à l'UNESCO et formatrice principale de ce cours. Je tiens à vous remercier tous.tes d'être venu.es ici ces dernières semaines. Nous sommes très heureux que vous trouviez ce cours intéressant et utile, c'est très gratifiant et nous allons continuer dans cette direction. Merci beaucoup pour vos commentaires, pour votre participation et encore une fois pour votre présence. Commençons notre dernier module. Nous allons parler de l'engagement multipartite et des mesures correctives visant à contribuer à un écosystème en ligne transparent et inclusif pendant les élections. Dans ce dernier module, nous aborderons d'autres pratiques qui peuvent aider à lutter contre les pratiques néfastes lors des élections, nous examinerons les mesures d'identification, de vérification et de veille sur des contenus problématiques, le suivi des médias liés aux élections et la vérification des faits, les mesures correctives telles que le signalement, le ciblage et le "blaklistage" (mise sur liste noire), les réponses techniques et algorithmiques, les mesures réglementaires et non réglementaires et la modération des contenus. Commençons par les mesures d'identification, de vérification des faits et de suivi. La vérification est une pratique traditionnelle du journalisme professionnel. Elle consiste à vérifier les sources d'information, à rechercher des preuves primaires auprès de témoins oculaires ou à vérifier des faits et des chiffres. C'est un moyen de contrôler la qualité de contenus avant qu'ils ne soient publiés ou diffusés. D'autre part, la vérification des faits a lieu après la publication ou la diffusion du contenu, et la vérification des faits vise à rendre les personnalités publiques, les institutions et les organes de presse responsables de ce qu'ils diffusent. Les "fact-checkeur.ses" recherchent des sources fiables qui peuvent confirmer ou infirmer les allégations faites au public. De plus en plus de vérification des faits impliquent également de démystifier de fausses informations. Cela implique de publier les preuves et démontrer les erreurs, souvent en expliquant le processus qui a amené à ces conclusions. Le suivi des médias liés aux élections implique un programme d'analyse de l'actualité et des contenus connexes diffusés sur les réseaux sociaux pendant une certaine période, généralement liée au calendrier électoral. Vous pouvez voir dans ce graphique que la vérification des faits, la démystification et la vérification sont trois moyens de corroborer la qualité des informations circulant sur les réseaux est très utile pour les processus électroniques. Voyons les trois étapes de vérification des faits. Comme nous l'avons dit, la vérification des faits consiste en l'analyse de contenu qui repose sur une question fondamentale, à savoir comment savons-nous que c'est vrai ou faux ? Et pour la vérification des faits, pour l'efficacité, la vérification des faits, nous devons suivre trois étapes. L'une consiste à déterminer quel contenu peut ou doit faire l'objet d'une vérification, c'est très important. Identifier les faits en examinant les preuves disponibles concernant l'affirmation en question et corriger l'information lors de la vérification à la lumière des nouvelles preuves, généralement selon une échelle de véracité. C'est ainsi que les professionnels procèdent habituellement, en suivant ces trois étapes. La vérification des sources et de l'origine des contenus visuels peut s'avérer très difficile. C'est encore plus difficile que de créer du contenu. Il existe différents types de contenus visuels faux ou trompeurs courants. Il y a les images très courantes où les lieux et l'époque sont erronées, c'est-à-dire des contenus qui peuvent être très facilement identifiés car ce ne sont que des photos et images qui ont été utilisées lors d'autres événements et qu'il est facile de retrouver dans les archives. Ensuite, il y a les contenus manipulés, c'est à dire des contenus numérisés et manipulés à l'aide d'un logiciel de retouche photo ou de montage vidéo. Et le dernier

élément sont les contenus mis en scène, c'est à dire des contenus qui ont été créés et partagés dans le but de tromper. Il s'agit de mises en scène consistant à jouer, à préparer ou à fabriquer un décor, dans lesquelles les informations enregistrées ou photographiées sont totalement fausses. Il existe de nombreux outils pour démystifier la désinformation et la désinformation. Il existe différentes recherches d'images qui permettent de voir si telle ou telle image se trouve dans certaines bases de données ou si elle se trouve archivées dans sa version d'origine. Cela est très facile à faire, même avec Google, vous pouvez également utiliser la géolocalisation, qui permet de déterminer où la vidéo ou l'image a été capturée. La géolocalisation nécessite souvent de croiser les caractéristiques visuelles et les points de repère du contenu avec l'imagerie satellite, la vue sur la rue et le contenu visuel disponible. C'est comme croiser différents types d'images et vous pouvez avoir un aperçu de n'importe quel endroit, puis vous avez une analyse des images. Certains outils permettent de détecter des incohérences dans les métadonnées de l'image suggérant une manipulation sur une image. A titre d'exemple de réponse globale, nous pouvons citer l'International Fact-Checking Network (IFCN) qui est l'un des réseaux mondiaux de vérification des faits les plus importants et engagés à promouvoir l'excellence dans la vérification des faits dans le monde. IFCN a créé et élaboré un code de principes que vous pouvez voir à l'écran. Il existe cinq principes importants pour un contrôle professionnel et fiable. Le premier consiste à s'engager en faveur de l'impartialité et de l'équité, la transparence des sources, la transparence du financement et de l'organisation, la transparence de la méthodologie, et une politique de rectifications ouverte et intègre, même s'ils se trompent, ils doivent en avertir le public pour ne pas perdre leur crédibilité. Ce sont donc les cinq principes les plus importants pour la vérification des faits.

Regardons de plus près la vérification des faits et la veille des médias pendant les élections. Regardons de plus près cette question, la veille sur les médias en période électorale est une histoire de longue date qui se déroule généralement sur une courte période, par exemple, pendant la campagne électorale ou une autre phase du cycle électoral, comme la période pré-électorale, post-électorale ou pendant les élections, cela dépend, mais généralement, c'est pour une courte période. Elle peut également être mise en œuvre dans d'autres cas, bien entendu, par exemple, pendant le processus d'enregistrement. La méthodologie de veille traditionnelle sur les médias suit un ensemble prédéterminé et peut combiner des outils automatisés, mais il est également utile pour effectuer des exercices de suivi d'impliquer la composante humaine, c'est une combinaison de méthodes. Le suivi des médias pendant les élections se traduit généralement par un échantillon qui définit quels médias seront surveillés quand, comment et pourquoi, et généralement par des indicateurs qualitatifs et quantitatifs. La qualité peut être reflétée dans des mots clés et les indicateurs quantitatifs peuvent être reflétés sur des graphiques et des données qui sont traités pour arriver à des conclusions. Le suivi des médias nécessite un financement pour les infrastructures. L'activation du matériel et des logiciels coûte très cher. Bien entendu, il existe d'autres moyens de surveiller les médias beaucoup plus, disons, faciles et réalisables, mais en général, c'est un processus coûteux et parfois, il est fort probable que vous ayez besoin d'un tiers pour assurer un suivi automatique, de ressources humaines et des formations. Le suivi des médias est très utile et efficace lorsqu'il est associé à des mécanismes et à des plateformes d'alerte précoce et c'est ce que nous avons constaté dans de nombreux pays. Il y a des échantillons et des indicateurs, idéalement une série de mots clés sont identifiés tels que des éléments comme des indicateurs du niveau d'hostilité ou de discours de haine présents pendant les périodes électorales. Les rapports de suivi des médias sont extrêmement utiles, notamment lorsqu'ils sont associés à des mécanismes et à des plateformes d'alerte précoce et la vérification des faits peut faire partie des exercices de suivi des médias, en complément de l'ensemble des indicateurs. Il existe donc de nouvelles technologies et de multiples technologies qui peuvent être combinées pour

obtenir de meilleurs résultats, pour identifier les résultats et être en mesure de résoudre les problèmes liés aux nouvelles technologies et aux données numériques. Jusqu'à présent, tout va bien, nous allons faire une pause de 5 minutes. Ensuite, nous poursuivrons avec les mesures correctives. On se retrouve dans 5 minutes. Bon retour. Nous allons maintenant continuer à jeter un œil sur les mesures correctives. Dans le cadre de leurs efforts visant à accroître la transparence et à sensibiliser aux menaces croissantes qui pèsent sur la liberté d'expression en ligne, les principales sociétés internet ont commencé à mettre en place une série de mesures correctives. Elles peuvent être le résultat de l'initiative unilatérale de ces entreprises, ou elles peuvent également être le résultat du cadre juridique en place dans certains pays ou une région spécifique. Mais ces mesures correctives sont très pertinentes pour les processus électoraux. Les plateformes sociales, y compris les plus connues comme Facebook, Twitter, Instagram, TikTok, Snapchat, etc., ont mis en place des mécanismes de plainte et d'appel qui fonctionnent sur la base de l'accord signé des utilisateur.rices, lorsqu'ils décident d'utiliser ces services. Ces mécanismes déclenchent théoriquement des mesures correctives internes de la part de ces entreprises. Par exemple, les plateformes en ligne permettent généralement aux utilisateur.rices de signaler des hôtes qui postent des contenus violents, offensants, etc. Le ciblage permet aux utilisateur.rices d'identifier les publicités ambiguës, particulièrement importantes lorsqu'elles sont liées à de la propagande politique et la mise sur liste noire peut même s'appliquer à la suppression de contenu ou d'un utilisateur.rice en particulier d'une plateforme de médias sociaux. Néanmoins, toutes ces pratiques ne devraient pas affecter le droit à la liberté d'expression et le droit d'accès à l'information, ce qui est un équilibre très délicat et très difficile à mesurer et à maintenir. C'est le principal enjeu de toutes ces mesures. Vous voyez à l'écran un exemple de signalement sur les réseaux sociaux, Twitter, qui est très populaire dans les débats politiques lors des élections. Vous pouvez voir à l'écran comment le problème est signalé. Les réponses techniques et algorithmiques pour la détection semi-automatique de la désinformation en ligne et leur utilité pratique tout au long du cycle de vie des campagnes de désinformation, y compris les réponses au contenu, l'analyse de crédibilité, la diffusion du réseau, la mesure de l'impact sur les croyances et les actions des citoyens méthodes, ces réponses techniques peuvent être mises en œuvre par les plateformes sociales et les moteurs de recherche eux-mêmes, mais peuvent également être des outils tiers, par exemple des plug-ins de navigateur ou des méthodes expérimentales issues de recherches universitaires. Quoi qu'il en soit, la réglementation des entreprises privées relative à la régulation est souvent automatisée avec un nombre limité d'humains, n'est pas toujours très transparente ou facile à comprendre par tous. La plupart des outils automatiques de détection de la désinformation visent actuellement à contribuer à la prise de décision humaine. Au niveau des contenus, ils fournissent des informations permettant l'analyse et la manipulation humaines. L'hypothèse qui sous-tend les approches techniques et algorithmiques est qu'elles peuvent freiner et réduire le partage de désinformation ou de mésinformation ou de contenus nuisibles. Les entreprises internet déploient une gamme de modèles de détection automatisés pour les types de contenus et les services. Il s'agit notamment d'outils permettant de suivi de la diffusion artificielle d'information, ainsi que d'identifier le contenu qui répond aux critères de classement, de ciblage ou de suppression, mais cela est très difficile, comme je l'ai déjà mentionné, car les contenus circulent dans des milliers de langues et le contexte n'est pas toujours facile à analyser de façon automatique. Passons maintenant aux mesures réglementaires et non réglementaires, car la régulation de contenus en ligne pendant les périodes électorales est également une question très complexe car elle peut menacer les droits à la liberté d'expression et à l'accès à l'information. Et le processus d'élaboration des lois et des réglementations doit également suivre une approche ouverte, transparente et participative. C'est pour cette raison que de plus en plus de pays essaient de faire appliquer certaines réglementations pendant les

processus électoraux, car dans les situations complexes, il est important de s'assurer que la liberté et l'équité des élections ne sont pas compromises. En même temps, ces mesures doivent être compatibles avec l'article 19 relatif à la liberté d'expression et le Pacte international relatif aux droits et droits civils. Cependant, la plupart des lois et règles régissant les élections et les outils d'autorégulation des médias connexes ne s'appliquent pas toujours aux plateformes numériques, aux réseaux sociaux, c'est donc un autre défi. Par exemple, la période de silence que de nombreux pays appliquée juste avant le jour du scrutin, ou même après, n'est pas toujours respectée par les réseaux sociaux même si un gouvernement interdit la publication des résultats du scrutin avant une élection ou le jour du scrutin, les utilisateur.rices des réseaux sociaux peuvent rendre ces résultats publics plus tôt sans respecter les règles électorales. De même, la publicité politique est souvent diffusée en ligne et par le biais de messages sociaux et même pendant les périodes de silence. Et ce qui pose problème ici, c'est que ces contenus, comme nous l'avons vu dans les modules précédents, peut être formé à partir de données, de sondages et de résultats inventés et peuvent affecter négativement et discréditer l'ensemble du processus électoral. De plus, comme nous l'avons vu dans les modules précédents, ces actions peuvent également s'inscrire dans le cadre d'une campagne de désinformation plus vaste orchestrée pour déstabiliser et saper la crédibilité des institutions démocratiques et d'un gouvernement particulier dans un pays donné. La législation contre la désinformation s'est développée ces dernières années et de nombreux pays ont adopté des lois contre la désinformation, telles que la réglementation du contenu par le biais du blocage ou de la suppression, la criminalisation pour diffamation ou discours de haine, des mandats réglementaires ou juridiques pour surveiller les réseaux sociaux et des lois sur la protection de la vie privée, ainsi que les réglementations relatives au financement politique qui concernent la réglementation des médias sociaux d'une entreprise technologique ou d'une plateforme de médias sociaux spécifique pendant des périodes spécifiques. Néanmoins, les initiatives réglementaires ont été très critiquées dans certains cas car elles tendent parfois à conférer un pouvoir de censure disproportionné à l'État ou à un gouvernement. Cela implique parfois de donner ce pouvoir à un parti politique particulier au détriment des autres. Toute tentative visant à réglementer le contenu en ligne doit donc trouver un équilibre entre les droits à la liberté d'expression et à l'accès à l'information et la protection d'autres droits civils et politiques, tels que la participation politique, la vie privée et le droit de ne pas être victime de discrimination. Les restrictions à la liberté d'expression en ligne doivent répondre aux trois parties décrites à l'article 19 du Pacte international relatif aux droits civils et politiques, à savoir: les lois et règlements doivent être prescrits par la loi; toute attention à un objectif légitime doit être nécessaire et proportionnée; l'autorégulation est un mécanisme totalement différent et constitue un secteur de conformité volontaire, au niveau de l'industrie, où la législation ne joue pas nécessairement un rôle dans l'application des normes. Comme nous l'avons vu, certaines entreprises mettent en place des mesures par elles-mêmes. En d'autres termes, il s'agit d'initiatives privées de modération de contenus. Mais qu'est-ce que la modération de contenu exactement ? La modération est le processus de suivi pour s'assurer que les contenus soumis à un site web sont conformes à ses règles et directives, qu'ils sont appropriés pour être publiés sur le site, et s'ils sont adaptés au cadre juridique en place dans le pays ou la région. Cela implique de définir des règles et des directives que l'ensemble des contenus doivent respecter et de filtrer tout ce qui est considéré comme dangereux, sensible ou inapproprié. Essentiellement, la modération de contenus permet de s'assurer que les contenus publiés sur le site ne sont pas illégaux, inappropriés, ne harcèlent pas, ou ne nuisent pas à autrui. La modération est couramment utilisée dans les applications qui s'appuient fortement sur les contenus générés par les utilisateur.rices, telles que les forums, les plateformes de médias sociaux, les sites de rencontres et autres sites de vente en ligne. Les réseaux et médias sociaux et d'autres intermédiaires en ligne

se sont principalement concentrés sur l'élaboration de règles d'autorégulation, généralement sous la forme de standards ou de conditions d'utilisation ou sur des questions telles que la suppression de contenus dans le cadre de traitement des données. Un exemple très intéressant de l'un de ces mécanismes est le Meta Oversight Board qui est un organisme indépendant de modération de contenus. En 2020, il a créé un conseil de surveillance pour Facebook et Instagram, composé de personnalités renommées des droits humains et de la liberté d'expression. L'idée principale était de disposer d'un organisme indépendant chargé de prendre les décisions de modération les plus complexes. Dans un premier temps, dans le cadre délicat de retrait de contenus par l'entreprise. Le Conseil accepte les dossiers d'utilisateur.rices qui ont épuisé la procédure d'appel sur Facebook ou Instagram. Il examine chaque cas de manière indépendante par rapport à la direction de Meta et prend des décisions, par exemple en ce qui concerne le retrait de contenus. Ils fournissent également des recommandations stratégiques qui poussent Meta à faire des règles et à agir de manière équitable, transparente et traiter correctement tous les utilisateur.rices. C'est donc un exemple intéressant. Enfin pour finaliser notre MOOC, après ces cinq semaines, nous allons parler de l'approche coréglementaire hybride pour les processus électoraux. Il s'agit d'une approche très spécifique qui ne peut pas être appliquée partout, mais qui donne de nombreux résultats positifs et cela a été minutieusement analysé. Dans ce modèle, le cadre juridique existant n'impose pas d'exigences restrictives supplémentaires aux plateformes de réseaux sociaux, mais plutôt expriment des attentes de résultats. Par exemple, un code de bonnes pratiques est une pratique pour les entreprises de médias sociaux en consultation avec ces entreprises ainsi qu'avec d'autres partis politiques et politiques, ou encore des besoins plus larges peuvent être développés, examinés, rendus publics de manière collaborative, créant ainsi un écosystème beaucoup plus solide dans lequel chaque acteur.rice participe aux décisions concernant les mesures prises contre les contenus nocifs. Nous avons donc un rôle à jouer à cet égard en matière d'élection de tous.tes les acteur.rices électoraux.ales, en particulier en ce qui concerne la surveillance électronique par ces autorités médiatiques, les organisations de la société civile, le gouvernement et de nombreuses fois également par les partis politiques. Cela implique également que tous ces engagements soient, bien entendu, volontaires et dépendent du cadre juridique existant. Nous ne pouvons pas généraliser ce modèle car, dans de nombreux contextes, il ne serait pas possible à réaliser. Un exemple d'initiative réglementaire est l'adoption d'un code de conduite et d'engagements convenus, par exemple, par les institutions de l'État, les partis politiques et les principales sociétés de médias sociaux, les annonceurs, les médias et d'autres acteur.rices. Ces codes de conduite visent généralement à garantir la transparence de la publicité politique et à renforcer les mesures visant à fermer les faux comptes ou à mettre fin aux activités préjudiciables, signaler les messages diffusés par les robots et collaborer avec des "fact-checkeurs" dans les universités pour détecter des campagnes de désinformation et accroître la visibilité et la portée de contenus vérifiés. Faisons un bref résumé de ce dernier module sur l'information et les élections à l'ère numérique. Au cours de ce module, nous avons examiné l'engagement multipartite et les mesures correctives pour contribuer à un écosystème transparent et inclusif; la plupart des différentes mesures qui permettent aujourd'hui de lutter contre la désinformation, la mésinformation et d'autres pratiques nuisibles dans le cadre des élections; l'identification, la vérification des fait, le suivi des médias, les mesures correctives; et les mesures réglementaires et non réglementaires, y compris les approches hybrides destinées à mettre en œuvre ces mesures. Merci beaucoup pour votre présence. Je suis très fière d'avoir pu être votre formatrice principale pendant ces cinq semaines. Encore une fois, j'apprécie tous vos commentaires, toutes vos suggestions et vos contributions. Tout a été très positif. Merci beaucoup. J'espère que ce MOOC a été intéressant et utile pour vous. A bientôt.

