ASSEMBLÉE NATIONALE

11 février 2022

IVERSES DISPOSITIONS D'ADAPTATION AU DROIT DE L'UNION EUROPÉENNE EN
MATIÈRE DE PRÉVENTION DE LA DIFFUSION DE CONTENUS À CARACTÈRE
TERRORISTE EN LIGNE (4883 RECTIFIÉ) - (N° 5024)

Commission	
Gouvernement	

Rejeté

AMENDEMENT

N° 3

présenté par

M. Bernalicis, Mme Autain, M. Coquerel, M. Corbière, Mme Fiat, M. Lachaud, M. Larive,
M. Mélenchon, Mme Obono, Mme Panot, M. Prud'homme, M. Quatennens, M. Ratenon,
Mme Ressiguier, Mme Rubin, M. Ruffin et Mme Taurine

ARTICLE ADDITIONNEL**APRÈS L'ARTICLE UNIQUE, insérer l'article suivant:**

L'article L. 111-7 du code de la consommation est complété par un III ainsi rédigé :

« III. – Tout opérateur de plateforme est tenu de rendre public l'ensemble des documents permettant de prouver que les algorithmes utilisés n'ont pas pour conséquence, intentionnelle ou non, directement ou indirectement, un traitement des données discriminant au regard de l'origine, du sexe, de la situation de famille, de la grossesse, de l'apparence physique, de la particulière vulnérabilité résultant de la situation économique, apparente ou connue d'un groupe historiquement discriminé, du patronyme, du lieu de résidence, de l'état de santé, de la perte d'autonomie, du handicap, de caractéristiques génétiques, de l'orientation sexuelle, de l'identité de genre, de l'âge, des opinions politiques, de l'activité syndicale, de la capacité à s'exprimer dans une langue autre que le français, de l'appartenance ou de la non-appartenance, vraie ou supposée, à une ethnie, une Nation, une prétendue race ou une religion déterminée ; ainsi que les mesures mises en place pour garantir l'absence de traitement défavorable au fondement des motifs précités. La publicité se fait une fois par an et à tout moment à la demande d'une personne utilisatrice ou d'une association agissant dans le domaine de la lutte contre les discriminations ou des droits numériques ayant un motif raisonnable de penser que l'algorithme mis en cause a un effet discriminatoire au regard des motifs précités.

« Cette publicité doit être faite de manière à garantir son intelligibilité, sa compréhensibilité et sa facilité d'accès. »

EXPOSÉ SOMMAIRE

Cet amendement a pour objectif que les algorithmes utilisés par les plateformes n'aient pas pour conséquence de perpétuer, de produire ou de renforcer des préjugés justifiant et instituant des discriminations entre les personnes en raison de ce qu'elles sont.

Plusieurs études l'ont montré, les grosses plateformes, dont nous avons souvent un usage quotidien, ne sont pas exemptes des structures de discriminations existant historiquement et socialement dans notre société. C'est ce qu'a analysé l'ancienne analyste de Wall Street, Cathy O'Neil, dans son ouvrage « Weapons of Math destruction », préfacé dans sa version française par le député et mathématicien Cédric Villani, qui a souligné la justesse de ses analyses. Elle y démontre notamment que derrière l'objectivité et la neutralité proclamée des algorithmes du fait de leur fonctionnement mathématique se cache en réalité les opinions et intérêts subjectifs des personnes les créant ou les utilisant.

De même, dans « Algorithms of Oppression : How Search Engines Reinforce Racism », Safiya Umoja Noble démontre en quoi les algorithmes des moteurs de recherches renforcent les stéréotypes. Par exemple : cherchez « jeune noire », « jeune asiatique » « jeune fille arabe » sur Google, Qwant ou sur Yahoo les premières propositions sont des sites pornographiques, cherchez « jeune blanche » est le résultat sont les pages blanches en premier lieu sur Qwant, puis une adaptation cinématographique de « Blanche-neige » et un site d'agence matrimoniale liant des femmes ukrainiennes avec des hommes sur Yahoo. Sur google, de nouveau les recherches proposent en revanche des sites pornographiques. De même lorsque l'on cherche « lesbienne », l'intégralité des résultats des deux premières pages, à l'exception de la proposition wikipédia, sont des sites pornographiques.

Ces associations nourrissent l'hypersexualisation des jeunes femmes et particulièrement de celles subissant également le racisme. Elles perpétuent donc stéréotypes à la croisée du racisme et de sexisme. Pour désigner ce phénomène, on parle de « biais algorithmique ». Comme le souligne la CNIL dans son rapport « Tout algorithme est [...] en un sens, biaisé, dans la mesure où il est toujours le reflet – à travers son paramétrage et ses critères de fonctionnement, ou à travers les données d'apprentissage qui lui ont été fournies – d'un système de valeurs et de choix de société. Le débat autour des biais et des discriminations qu'ils peuvent générer n'est donc qu'un miroir grossissant mettant en valeur cette caractéristique essentielle dans ce qu'elle a de plus problématique ».

Ces biais sont produits de deux manières. D'un côté, il y a les biais par apprentissage (opération d'entraînement des algorithmes) : les algorithmes traitent les informations par association et reproduisent les associations racistes, sexistes, homophobes, validisme, xénophobie qui leur sont soumises. D'autre part, il existe les « biais dans la formule » : les algorithmes sont, dans leur configuration, soit non pensés afin de corriger les biais dans la société en créant des paramètres correctifs, soit sont pensés par des personnes qui elles-mêmes défendent des positions discriminantes ou ont un intérêt économique à la discrimination (intégration plus ou moins consciente).

Ces deux aspects sont cumulatifs et peuvent se renforcer.

Cela a des conséquences concrètes sur le quotidien. Dans la recherche d'emploi : il a été révélé par exemple qu'un système de tri de CV au profit d'entreprises proposé par AMAZON pénalisait les

CVS contenant le mot « femme » soit en les mettant de côté, soit en proposant des salaires plus bas pour elles. Il s'appuyait sur les données des entreprises en terme de recrutement sur 10 ans. De même pour LinkedIn, Indeed ou Monster qui, selon Anja Lambrecht et Catherine Tucker, deux chercheuses du MIT et de la London Business School, proposent plus d'annonces (notamment dans les métiers de l'ingénierie, des mathématiques etc.) aux hommes qu'aux femmes. A cela s'ajoute les associations négatives liées au lieu de domiciliation qui peuvent croisées elles-aussi d'autres discriminations liées au racisme, au validisme etc.

Ces biais algorithmiques peuvent également créer des difficultés lorsqu'une personne souhaite contracter un prêt, louer ou acheter un logement et que les entreprises utilisent un système de traitement automatisé de données afin de classer et hiérarchiser les demandes. De même, pour l'estime de soi : il n'est pas tolérable que des personnes faisant une recherche anodine sur internet se voient associer en premier lieu ou comme unique résultat un site pornographique. Ainsi les discriminations sur internet et la haine qui les produisent ou en découlent sont également produits par les algorithmes de grandes plateformes de mises en réseaux ou de recherches et impactent la vie quotidienne réelle des gens qui en sont victimes. En cela, les biais algorithmiques renforcent et perpétuent les discriminations systémiques de la société, et sont d'autant plus problématiques qu'ils ont un impact massif : google, linkedin, amazon font partie de quotidien de la majorité des gens. Or ils sont difficilement identifiables pour les personnes victimes, celles-ci ne voyant qu'un résultat final sans connaître les opérations y ayant menées. Elles sont également difficiles à identifier pour les personnes produisant les algorithmes si une réflexion sur ceux-ci n'est pas systématiquement intégrée au processus de création et de correction des algorithmes : nous ignorons en effet souvent la plupart des préjugés qui nous habitent.

Cet amendement vise ainsi à créer une obligation pour les plateformes de rendre publics les moyens mis en œuvre pour que leurs algorithmes ne discriminent pas. Il est un outil pour, à la fois forcer les plateformes à se confronter à ce phénomène massif, et pour les personnes utilisatrices de contrôler que les outils internet qu'elles utilisent ne sont pas des instruments perpétuant des idées déshumanisantes contre elles. La haine sur internet ne se concrétise pas uniquement dans les attaques violentes et facilement identifiables mais également dans des structures permettant, sans que cela soit visible et facilement identifiable et donc dénonçable, de discriminer des personnes en raison de ce qu'elles sont ou plutôt, de ce qu'on croit savoir qu'elles sont.