

20 Février 2017

MSI-NET(2016)06rev

ETUDE SUR LES DIMENSIONS DES DROITS DE L'HOMME DANS L'APPLICATION DES ALGORITHMES

2^e PROJET (20 février 2017)

présenté lors de la 3^{ème} réunion MSI-NET (27-27 février 2017)

RAPPORTEUR : BEN WAGNER

TABLE DES MATIERES

1. INTRODUCTION	3
2. PORTÉE DU RAPPORT	5
A. AUTOMATISATION	6
B. ANALYSE DES DONNÉES.....	6
C. CONSTRUCTIONS SOCIALES AUTOUR DES ALGORITHMES	6
3. INCIDENCES DES ALGORITHMES SUR LES DROITS DE L'HOMME	8
A. DROIT À UN PROCÈS ÉQUITABLE - ARTICLE 6 DE LA CONVENTION EUROPÉENNE DES DROITS DE L'HOMME.....	9
B. DROIT AU RESPECT DE LA VIE PRIVÉE ET FAMILIALE - ARTICLE 8 DE LA CONVENTION EUROPÉENNE DES DROITS DE L'HOMME	10
C. LIBERTÉ D'EXPRESSION - ARTICLE 10 DE LA CONVENTION EUROPÉENNE DES DROITS DE L'HOMME	12
D. LIBERTÉ DE RÉUNION ET D'ASSOCIATION - ARTICLE 11 DE LA CONVENTION EUROPÉENNE DES DROITS DE L'HOMME	16
E. DROIT À UN RECOURS EFFECTIF - ARTICLE 13 DE LA CONVENTION EUROPÉENNE DES DROITS DE L'HOMME.....	17
F. INTERDICTION DE LA DISCRIMINATION - ARTICLE 14 DE LA CONVENTION EUROPÉENNE DES DROITS DE L'HOMME	19
G. DROITS SOCIAUX ET ACCÈS AUX SERVICES PUBLICS.....	20
H. DROIT À DES ÉLECTIONS LIBRES.....	22
4. MÉCANISMES DE GOUVERNANCE, RESPONSABILITÉ, TRANSPARENCE ET ÉTHIQUE	24
A. INTRODUCTION	24
B. EST-IL RAISONNABLE ET POSSIBLE DE RÉGLEMENTER LES ALGORITHMES ?	25
C. TRANSPARENCE.....	26
D. RESPONSABILISATION	27
E. CADRES ÉTHIQUES ET MEILLEURE ÉVALUATION DES RISQUES.....	28
5. CONCLUSIONS.....	29
RESUME ET CONCLUSIONS GENERALES	30
BIBLIOGRAPHIE	31

1. INTRODUCTION

Quelles informations voyez-vous sur votre compte Facebook ? Qui est un criminel ou un terroriste ? Avez-vous droit à une assurance-maladie ? Allons-nous vous donner un emploi ? Des algorithmes supplantent de plus en plus souvent les êtres humains pour répondre à ce genre de questions, généralement au moyen de processus décisionnels automatisés. Ces algorithmes peuvent ne pas prendre de décisions par eux-mêmes mais ils peuvent préparer et présenter des décisions à des décideurs humains. Leur mode de fonctionnement, toutefois, entraîne souvent une prise de décision quasi automatisée, estompant la frontière entre prise de décision humaine et prise de décision automatisée. Ces systèmes sont source d'enjeux considérables non seulement dans les divers domaines d'action où ils sont utilisés, mais également pour l'ensemble de la société qui doit trouver les moyens de protéger les droits fondamentaux et la dignité humaine face à des technologies en constante évolution. Ils ont des incidences sur le droit à des élections libres, les droits des travailleurs, le droit à la vie, à la liberté d'expression, au respect de la vie privée et même sur l'Etat de droit. Les défis posés par les « algorithmes » utilisés par le secteur public et le secteur privé, et en particulier par les intermédiaires internet, font actuellement partie des questions les plus vivement débattues en matière de droits de l'homme.

Les êtres humains ont l'impression de ne plus contrôler ni comprendre les systèmes techniques qui les entourent, d'où la perception croissante que les « logiciels mangent le monde » (Andreessen 2011). Si cela est déconcertant, ce n'est pas toujours négatif. C'est un produit dérivé de cette phase de la vie moderne où les évolutions économiques et technologiques au niveau mondial produisent un grand nombre d'articles techniques basés sur des logiciels et où les « objets codés » (Kitchin et Dodge 2011) intègrent d'importantes capacités décisionnelles pertinentes en matière de droits de l'homme. Quels choix instantanés doit faire un véhicule piloté par un logiciel à l'approche d'une collision ? Les algorithmes des entreprises internet en situation de quasi-monopole ont-ils le pouvoir d'influer sur les élections ? Quels sont les droits des travailleurs dont les relations avec l'employeur sont entièrement automatisées ? Qui aura droit à une assurance maladie et quelles informations sont communiquées dans les fils d'actualités Facebook ? Y a-t-il plus de risques de préjugés sexistes, ethniques ou raciaux dans un système automatisé et quel degré de préjugé doit-on considérer comme acceptable ?

Dans le passé, les entreprises privées prenaient leurs décisions en matière de développement de logiciels selon les cadres économique, juridique et éthique qu'elles jugeaient appropriés. Il n'existe pas de cadre normatif pour le développement des systèmes et des processus qui aboutissent à une prise de décision algorithmique, ni pour leur mise en œuvre. En réalité, étant donné que de nombreuses technologies fondées sur des algorithmes en sont toujours à leurs balbutiements, on ignore s'il est possible d'établir un cadre normatif relatif à l'utilisation des algorithmes ou une réglementation efficace des techniques de traitement automatisé des données. Les questions soulevées par l'utilisation des algorithmes dans le processus décisionnel sont multiples et complexes et suscitent des inquiétudes sur la qualité des données, sur le respect de la vie privée et sur d'éventuelles discriminations. En même temps, le débat sur les algorithmes et leurs conséquences éventuelles sur les personnes, les groupes et les sociétés ne fait que commencer. Cela ne doit toutefois pas empêcher de chercher à comprendre ce qu'ils font réellement, les conséquences qui en découlent pour la société et comment les éventuelles préoccupations liées aux droits de l'homme peuvent être prises en compte.

Ce rapport recense quelques-unes des inquiétudes suscitées par la domination croissante des algorithmes dans le domaine des droits de l'homme. Selon le type de fonction exécutée, leur impact sur l'exercice des droits de l'homme sera différent. Lorsque des algorithmes violent des droits de l'homme, qui est responsable ? La personne qui a programmé l'algorithme, son opérateur, ou l'être humain qui a mis en œuvre une décision fondée sur un algorithme ? Existe-t-il une différence entre une telle décision et une décision prise par un humain ? Quelles sont ses incidences sur l'accès aux droits de l'homme, la jouissance de ces droits et les garanties en la matière, telles que prévues par les normes établies, y compris les principes de l'État de droit et les processus judiciaires ?

Les défis liés aux conséquences des algorithmes et des techniques de traitement automatisé des données pour les droits de l'homme ne peuvent que croître, les systèmes associés étant de plus en plus complexes et interagissant entre eux d'une manière de plus en plus impénétrable pour l'esprit humain. Ce rapport n'a pas pour objectif d'examiner le sujet de manière exhaustive mais il cherche plutôt à recenser les principales préoccupations actuelles du point de vue du Conseil de l'Europe et à étudier les possibilités de réglementation qui s'offrent aux États membres en vue de minimiser les effets négatifs. Un nombre de thèmes liés nécessiteront une recherche plus détaillée afin d'évaluer de manière plus systématique les problèmes qu'ils soulèvent et les

possibilités qu'ils offrent du point de vue des droits de l'homme, y compris les questions concernant le traitement des mégadonnées, l'apprentissage automatique, l'intelligence artificielle ou l'internet des objets.

2. PORTÉE DU RAPPORT

Dans le cadre de l'évaluation des algorithmes et des techniques de traitement automatisé des données dans lesquelles ils interviennent, il est important de bien spécifier les types d'algorithmes étudiés ici. Cette étude s'appuiera sur les définitions existantes bien établies, notamment les travaux de Tarleton Gillespie (2014), Nicholas Diakopoulos (2015) et Frank Pasquale (2015). Il est également important de garder à l'esprit que le terme « algorithme » est largement utilisé et a plusieurs significations possibles selon qu'il est utilisé au sein de la communauté des sciences informatiques, par les mathématiciens et les techniciens de l'information, ou dans la sphère publique, y compris dans le discours politique. L'examen des dimensions des droits de l'homme dans l'application des algorithmes doit également tenir compte de la divergence entre les définitions formelles des algorithmes et l'utilisation populaire du terme. En réalité, la plupart des débats sont moins axés sur les algorithmes eux-mêmes que sur le rôle de la technologie au sein de la société (Bucher 2016).

La définition utilisée ici part de l'hypothèse de Tarleton Gillespie selon laquelle « les algorithmes ne sont pas nécessairement des logiciels : au sens le plus large, il s'agit de procédures codées qui permettent de transformer des données d'entrée en un produit souhaité, sur la base de calculs précis. Les procédures désignent à la fois un problème et la démarche suivie pour le résoudre » (Gillespie 2014:167). Par conséquent, on peut suggérer que les algorithmes sont « une série d'opérations réalisées afin de résoudre un problème particulier ou d'obtenir un résultat défini » (Diakopoulos 2015:400).

Ce rapport ne traitera pas des algorithmes qui automatisent les processus de fabrication ou exécutent d'autres tâches de routine. Il semble plutôt raisonnable de limiter la discussion aux algorithmes qui sont numériques et « d'intérêt public ». Ce rapport se concentrera sur la prise de décision algorithmique ayant des implications sur les droits de l'homme. Sans être exhaustives ni chercher à prévoir toutes les itérations possibles des algorithmes et leurs décisions futures, les caractéristiques suivantes des algorithmes intervenant dans le traitement automatisé de données et la prise de décision (semi-)automatisée sont considérées comme essentielles du point de vue des droits de

l'homme dans le cadre de ce rapport : automatisation, analyse des données et adaptabilité.

A. AUTOMATISATION

L'automatisation est l'un des principaux défis posés par la prise de décision algorithmique. La capacité des systèmes informatiques automatisés à remplacer les êtres humains dans un nombre croissant de situations est une caractéristique essentielle de la mise en œuvre pratique des algorithmes. Qu'il s'agisse de modèles simples qui aident les prestataires de services en ligne à effectuer des opérations pour le compte de leurs utilisateurs (Kim et al., 2014) ou d'algorithmes de profilage plus complexes (Hildebrandt, 2008) qui filtrent les systèmes pour proposer un contenu personnalisé, les algorithmes de décision automatisée sont utilisés dans une variété de domaines. La prise de décision algorithmique automatisée est généralement difficile à prévoir pour un être humain et sa logique sera difficile à expliquer après coup.

B. ANALYSE DES DONNÉES

Des algorithmes d'analyse des données sont appliqués à de vastes quantités de données afin de trouver des corrélations au sein de l'ensemble de données sans établir de lien de causalité (Grindrod, 2014). Le fait qu'ils utilisent l'exploration de données et de la reconnaissance de tendances sans « comprendre » les liens de causalité peut conduire à des erreurs et susciter des inquiétudes quant à la qualité des données. Ces algorithmes reproduisent les fonctions auparavant exécutées par les êtres humains mais font appel à une logique décisionnelle quantitativement différente qui s'applique à des masses beaucoup plus importantes d'entrées.

C. CONSTRUCTIONS SOCIALES AUTOUR DES ALGORITHMES

Si la prise de décision algorithmique démontre sa capacité croissante à imiter la prise de décision humaine, d'importants éléments (comme la discrétion) des processus décisionnels ne peuvent être automatisés et sont souvent perdus lorsque les processus décisionnels humains sont automatisés (Spiekermann 2015). Sans juger leur « qualité » respective, les processus décisionnels exécutés par les humains et par les algorithmes sont fondamentalement et catégoriquement différents, ont des conséquences différentes et font des erreurs différentes. Si la société et les gouvernements ont une vaste expérience et une compréhension approfondie de la prise de décision humaine et de ses

échecs, ils commencent à peine à comprendre les points faibles de la prise de décision algorithmique. Le principal problème semble être la perception fréquente selon laquelle les algorithmes sont capables d'élaborer des prévisions neutres et indépendantes des événements futurs¹. Néanmoins, cela tient moins aux algorithmes qu'à la perception et à l'interprétation humaines de leur mise en œuvre et de leurs résultats.

Traditionnellement, les développeurs programmaient les algorithmes à la main « afin de traiter et de transformer les entrées en un produit souhaité, sur la base de calculs précis » (Gillespie, 2014). Avec l'évolution technologique, toutefois, le résultat des algorithmes devient de plus en plus opaque, en particulier lorsqu'il repose sur des capacités d'apprentissage qui empêchent les êtres humains de percevoir non seulement le modèle de prise de décision mais aussi la logique qui le sous-tend. Même lorsqu'un être humain prend une décision de manière formelle, par exemple celle de supprimer un contenu donné d'une plateforme de réseau social (voir point C ci-après), il se contentera souvent d'approuver une décision préparée par un algorithme car il ne disposera ni du temps, ni du contexte, ni des compétences pour prendre une décision adéquate en l'espèce. Ainsi, s'il semble logique de faire la distinction entre prise de décision entièrement automatisée et prise de décision semi-automatisée, dans la pratique, les frontières ne sont pas nettes. Dans aucun de ces deux cas l'être humain ne sera en mesure d'avancer un argument raisonné expliquant la nécessité de telle ou telle décision dans le cas considéré. Cela a des répercussions sur le droit de l'individu concerné à disposer d'un recours effectif contre une violation des droits de l'homme (voir point E ci-après).

Il convient de noter que les algorithmes évoqués ici n'ont pas d'existence qui fasse sens sans interaction avec des êtres humains. Les concepts mathématiques ou informatiques peuvent ne pas avoir d'impacts négatifs sur les droits de l'homme, mais leur mise en œuvre et leur application à l'interaction humaine, elles, peuvent en avoir. Affirmer que les systèmes informatiques sont ou peuvent être neutres serait toutefois une erreur. Les technologies, dans leur application à l'interaction humaine, sont des constructions profondément sociales (Winner 1980, 1986) aux conséquences politiques importantes (Denardis 2012). Si un logiciel de décision peut être « partial mais ambivalent » (McCarthy 2011:90), il n'a aucun sens sans un système social autour de lui qui lui

¹ L'effervescence autour de Google Flu Trends en 2011, qui s'est ultérieurement avérée injustifiée car la capacité de prédiction de ce service était bien moindre que ce qui était affirmé, est un exemple du débat permanent autour des affirmations relatives à la précision des algorithmes prédictifs (Lazer et al. 2014; Lazer and Kennedy 2015).

confère un sens. Il est par conséquent trop simple de blâmer l’algorithme ou de suggérer de ne plus avoir recours aux ordinateurs ou à l’informatique. Au contraire, c’est la construction sociale et les normes et valeurs spécifiques intégrées aux algorithmes qui doivent être remises en question, critiquées et contestées. Ainsi, ce ne sont pas les algorithmes eux-mêmes mais les processus décisionnels afférents qu’il convient d’analyser pour déterminer leurs éventuelles conséquences pour les droits de l’homme.

Établir si la qualité de décisions au regard des droits de l’homme diffère selon qu’elles sont prises par un humain ou par un algorithme ou basées sur un calcul algorithmique ne peut être fait qu’en sachant comment fonctionne la prise de décision humaine. Il est démontré que celle-ci présente la particularité d’utiliser des connaissances et des normes tacites (Schulz et Dankert 2016). C’est ce qui, par exemple, permet aux humains de remarquer des situations exceptionnelles dans lesquelles il n’est pas approprié d’appliquer une règle même dans un cas qui en relèverait. Devant l’importance croissante des algorithmes dans la prise de décision, il apparaît indispensables de mieux comprendre la conception et les caractéristiques des procédures décisionnelles.

3. INCIDENCES DES ALGORITHMES SUR LES DROITS DE L’HOMME

Les principales réserves émises à l’égard des algorithmes et des techniques de traitement automatisé des données portent généralement sur leur opacité et leur imprévisibilité². Au-delà de ces préoccupations générales, des droits de l’homme spécifiques sont particulièrement concernés. Ils sont recensés ci-après avec des études de cas montrant comment et pourquoi l’utilisation des algorithmes peut entraîner des violations de droits.

A. DROIT À UN PROCÈS ÉQUITABLE - ARTICLE 6 DE LA CONVENTION EUROPÉENNE DES DROITS DE L’HOMME

²Voir *The great question of the 21st century: Whose black box do you trust?* sur le site https://www.linkedin.com/pulse/great-question-21st-century-whose-black-box-do-you-trust-tim-o-reilly?trk=eml-b2_content_ecosystem_digest-hero-22-null&midToken=AQGexvwxq0Q3iQ&fromEmail=fromEmail&ut=2SrYDZ8lkCS7o1.

Dans le domaine de la sécurité nationale et de la prévention de la criminalité, la tendance est de plus en plus à l'utilisation de processus décisionnels automatisés intégrés dans des algorithmes. Après la vague d'attentats aux États-Unis et en Europe, les responsables politiques ont demandé aux plateformes de réseaux sociaux en ligne d'utiliser leurs algorithmes afin de repérer les terroristes (Rifkind 2014; Toor 2016). Il semblerait que certaines de ces plateformes utilisent déjà des algorithmes pour détecter les comptes qui mettent en ligne des contenus extrémistes et les gouvernements demandent à ce que les résultats leur soient communiqués. Hormis son impact significatif sur la liberté d'expression (voir point C ci-après), cette application des algorithmes pose également des questions sur le respect des normes en matière de procès équitable définies à l'article 6 de la Convention européenne des droits de l'homme, notamment la présomption d'innocence, le droit d'être informé dans le plus court délai de la cause et de la nature d'une accusation et le droit de se défendre soi-même.

Dans le domaine de la prévention de la criminalité, les principaux débats politiques sur l'utilisation des algorithmes portent sur la police prédictive. Cette approche va au-delà de la capacité des êtres humains à tirer des conclusions des infractions passées afin d'anticiper les tendances en matière de criminalité. Elle comprend des systèmes automatisés développés qui prédisent quelles personnes sont susceptibles de commettre une infraction (Perry 2013) ou de récidiver et devraient, par conséquent, faire l'objet de peines plus sévères³.

Il est en outre à craindre que la pratique de telles évaluations dans le contexte de la prévention de la criminalité ne crée des caisses de résonance qui ne feraient qu'amplifier les préjugés. Le parti pris ou les préjugés liés, par exemple, à l'origine raciale ou ethnique peuvent ne pas être reconnus comme tels par la police une fois intégrés dans un programme informatique automatisé jugé indépendant et neutre (voir également point F). Il y a dès lors un risque de normalisation du préjugé dont le caractère raciste a moins de chance d'être dénoncé que s'il reposait sur une décision humaine. Bien que l'on ne connaisse pas la prévalence des décisions par algorithmes dans l'ensemble du système de justice pénale, la simple possibilité de leur utilisation suscite de vives préoccupations au regard de l'article 6 de la Convention européenne des droits de

³ Voir également Article 19, *Algorithms and Automated Decision-Making in the Context of Crime Prevention: A Briefing paper*, 2016.

l'homme et du principe d'égalité des armes établi par la Cour européenne des droits de l'homme⁴.

B. DROIT AU RESPECT DE LA VIE PRIVÉE ET FAMILIALE - ARTICLE 8 DE LA CONVENTION EUROPÉENNE DES DROITS DE L'HOMME

Le débat le plus ancien et le plus vif sur la dimension des droits de l'homme dans l'application des algorithmes et l'utilisation des processus de traitement automatisé de données porte sur le droit au respect de la vie privée⁵. Les algorithmes facilitent la collecte, le traitement et la réutilisation de grandes quantités de données et d'images qui peuvent avoir de graves répercussions sur la jouissance du droit au respect de la vie privée et familiale, garanti par l'article 8 de la Convention européenne des droits de l'homme et par les normes européennes en matière de protection des données à caractère personnel.

Les algorithmes jouent un rôle dans le suivi et le profilage en ligne des personnes dont les habitudes de navigation sont enregistrées par des « cookies »⁶ et des technologies similaires telles que les empreintes numériques, et agrégées à des requêtes de recherche (moteurs de recherche) et à d'autres données (le suivi des réseaux sociaux et la collecte de données via les applications des dispositifs mobiles, par exemple) (Tene et Polonetsky 2012). L'une des principales applications du suivi et du profilage en ligne est la publicité ciblée basée sur le profil des intérêts présumés d'une personne.

Des efforts ont été entrepris pour moderniser la Convention du Conseil de l'Europe de 1981 pour la protection des personnes à l'égard du traitement automatisé des données à caractère personnel en phase avec l'évolution technologique, et mieux définir les droits des personnes concernées compte tenu des conséquences sur la vie privée des outils utilisés aujourd'hui pour la collecte, le traitement et la réutilisation de données et pour le profilage. L'article 8 du projet modernisé de Convention pour la protection des personnes

⁴ Voir, par exemple, l'affaire *Jespers c. Belgique* (requête n° 8404/78) du 15 octobre 1980.

⁵ Voir Sills 1970.

⁶ Un cookie est une petite quantité de données générées par un [site internet](#) et sauvegardées par un [navigateur web](#) afin d'enregistrer les informations concernant l'utilisateur, à l'instar d'un fichier de préférences créé par une [application](#) logicielle. Si les cookies peuvent remplir de nombreuses fonctions, leur principal but consiste à enregistrer les informations de [connexion](#) pour un site donné. Ils sont également utilisés pour enregistrer les préférences de l'utilisateur pour un site en particulier. Par exemple, un [moteur de recherche](#) peut enregistrer les paramètres de recherche dans un cookie.

à l'égard du traitement automatisé des données à caractère personnel établit expressément le droit de toute personne de ne pas être soumise à une décision l'affectant de manière significative qui serait prise uniquement sur le fondement d'un traitement automatisé de données, sans que son point de vue soit pris en compte, le droit d'obtenir connaissance du raisonnement qui sous-tend le traitement des données lorsque les résultats de ce traitement lui sont appliqués et le droit de s'opposer à tout moment, pour des motifs tenant à sa situation, au traitement de ses données à caractère personnel, à moins que le responsable du traitement ne démontre des motifs légitimes justifiant le traitement qui prévalent sur ses intérêts ou ses droits et libertés fondamentaux⁷.

Les cadres réglementaires de la protection des données à l'échelle de l'Union européenne, tels que le Cadre général de protection des données prévu par le règlement d'avril 2016 et qui doit entrer en vigueur en mai 2018, fixent également des normes pour l'utilisation des algorithmes dans la collecte de données, y compris éventuellement un « droit à l'explication » (Goodman et Flaxman 2016) et le droit pour la personne de « connaître la logique qui sous-tend le traitement automatisé des données la concernant » (Directive 95/46/CE du Parlement européen et du Conseil)⁸.

C'est surtout le recours à des courtiers en données, qui agrègent les informations contenues dans les profils personnels, qui suscite des préoccupations. Ces informations peuvent ensuite être exploitées au moyen d'algorithmes, ce qui crée un risque de surveillance à grande échelle (« dataveillance ») aussi bien par des entités privées que par des gouvernements (Rubinstein, Lee et Schwartz 2008). Le principal problème posé par l'utilisation des données extraites des profils à différentes fins au moyen d'algorithmes réside dans le fait que ces données perdent leur contexte originel. La réutilisation des données à de nouvelles fins est susceptible de nuire à l'autonomie informationnelle de la personne. Les moteurs de recherche peuvent avoir un effet similaire sur le droit au respect de la vie privée et la protection des données, dans la mesure où ils facilitent l'agrégation de données sur une personne en particulier et permettent de trouver des informations plus aisément en levant le voile sur les données anonymes.

⁷ Voir

<https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=09000016806a616c>

⁸ Voir <https://secure.edps.europa.eu/EDPSWEB/edps/EDPS/Ethics> pour plus de détails.

Un autre aspect important de l'utilisation des algorithmes pour le traitement automatisé des données est le stockage des données en nuage. Il s'agit de solutions qui permettent de stocker des fichiers et d'autres données non plus localement mais à distance sur des serveurs accessibles via internet. Cependant, du fait qu'elles ne sont plus stockées localement, les données des utilisateurs peuvent être traitées par des algorithmes selon des méthodes intrusives qui ne seraient pas habituellement mises en œuvre. Ce type de traitement automatisé des données peut avoir lieu à deux endroits : 1) lors du transfert vers l'emplacement de stockage à distance et 2) sur les serveurs à distance où les données sont stockées. Il peut s'avérer de plus en plus difficile pour les utilisateurs de savoir s'ils utilisent des services locaux ou à distance, les systèmes d'exploitation modernes et les services en nuage étant imbriqués de façon de plus en plus étroite. En ce qui concerne les données en transit, il peut donc être difficile de déterminer si elles sont suffisamment protégées par des technologies telles qu'un puissant cryptage de bout en bout et si elles ne sont pas manipulées d'une façon ou d'une autre⁹.

C. LIBERTÉ D'EXPRESSION - ARTICLE 10 DE LA CONVENTION EUROPÉENNE DES DROITS DE L'HOMME

L'utilisation des algorithmes a également une incidence sur le droit à la liberté d'expression. Si l'effet positif des algorithmes de recherche et des moteurs de recherche pour le droit fondamental à la liberté d'expression a été étudié à maintes reprises¹⁰, le risque qu'ils portent préjudice à la liberté d'information et d'expression des personnes, des groupes et de segments entiers de la société est de plus en plus souligné¹¹.

Les contenus qui ne sont ni indexés ni très bien classés par un moteur de recherche internet ont moins de chance d'atteindre un large public. Un algorithme de recherche

⁹ Par exemple, le service en nuage de Microsoft « SkyDrive » utilise un processus automatisé conçu pour supprimer certains contenus (les photos de nu par exemple). Voir Clay 2012.

¹⁰ Voir, par exemple, Conseil de l'Europe, *Recommandation CM/Rec(2012)3 du Comité des Ministres aux États membres sur la protection des droits de l'homme dans le contexte des moteurs de recherche*, adoptée par le Comité des Ministres le 4 avril 2012 lors de la 1139^e réunion des délégués des ministres, paragraphe 1, disponible sur le site https://search.coe.int/cm/Pages/result_details.aspx?ObjectId=09000016805caa93, observant que les *moteurs de recherche* « permettent au public du monde entier de rechercher, de recevoir et de communiquer des informations, des idées et d'autres contenus, en particulier, d'avoir accès au savoir, de prendre part à des débats et de participer aux processus démocratiques. »

¹¹ Voir, par exemple, le rapport présenté en 2016 par le Rapporteur spécial des Nations Unies sur la promotion et la protection du droit à la liberté d'opinion et d'expression, David Kaye, lors de la 32^e-deuxième session du Conseil des droits de l'homme (A/HRC/32/38), soulignant que « les algorithmes des moteurs de recherche déterminent ce que les utilisateurs voient et dans quel ordre, et peuvent être manipulés de manière à restreindre ou hiérarchiser les contenus ».

peut également privilégier certains types de contenus ou fournisseurs de contenu avec le risque de nuire à des valeurs comme le pluralisme et la diversité des médias¹². La question ici est de savoir comment les résultats fournis par les moteurs de recherche devraient répondre aux souhaits des utilisateurs et dans quelle mesure ces réponses devraient promouvoir le pluralisme des médias et la diversité.

Les plateformes de réseaux sociaux prévoient également les préférences des utilisateurs grâce à des algorithmes et, par conséquent, orientent les publicités qui leur sont présentées, l'organisation de leurs fils sur les réseaux sociaux et l'ordre d'apparition des résultats de leurs recherches, compromettant ainsi largement la liberté d'expression et le droit à l'information des utilisateurs. Cela n'est pas anodin, si l'on considère la taille des plateformes telles que Google ou Facebook, leur rôle central dans la pratique d'internet en tant que sphère quasi publique (York 2010) et leur capacité à amplifier massivement certaines voix (Bucher 2012).

Selon l'article 10 de la Convention européenne des droits de l'homme, toute mesure de filtrage ou de suppression bloquant l'accès à un contenu doit être prévue par la loi, poursuivre l'un des buts légitimes visés à l'article 10.2 et constituer une mesure nécessaire dans une société démocratique. Comme l'indique la jurisprudence de la Cour européenne des droits de l'homme, toute restriction à la liberté d'expression doit correspondre à un « besoin social pressant » et être proportionnée au(x) but(s) légitime(s) poursuivi(s).

Les algorithmes sont largement utilisés pour les processus de filtrage et de suppression de contenus (Urban, Karaganis et Schofield 2016), y compris sur les plateformes de réseaux sociaux, ce qui a un impact direct sur la liberté d'expression et pose problème au regard de l'État de droit (questions de légalité, de légitimité et de proportionnalité). La suppression de contenus sur les plateformes de réseaux sociaux est souvent effectuée à l'aide de processus semi-automatisés ou automatisés. Bien que les grandes plateformes de réseaux sociaux comme Google ou Facebook affirment fréquemment que toute suppression de contenu est effectuée par des êtres humains (Buni et Chemaly 2016), le processus est en grande partie automatisé (Wagner 2016) et basé sur des opérations semi-automatisées. Selon un rapport de la commission parlementaire britannique du

¹² Proposition d'Aleksandra Kuczerawy, Brendan van Alsenoy et Jef Ausloos.

renseignement et de la sécurité¹³, il existe plusieurs techniques automatisées pour détecter les contenus présumés contraires aux conditions de service du fournisseur concerné – contenus extrémistes ou liés à l’exploitation d’enfants ou à des actes illégaux tels que l’incitation à la violence. Ces techniques peuvent également être utilisées pour désactiver ou suspendre automatiquement des comptes d’utilisateurs (Rifkind 2014).

Aux États-Unis, l’administration Obama a plaidé en faveur de l’utilisation de la détection et de la suppression automatisées des vidéos et images extrémistes¹⁴. En outre, certains ont proposé de modifier les algorithmes de recherche afin de « cacher » les sites web qui encouragent et soutiennent l’extrémisme. Des mécanismes de filtrage automatisé de vidéos extrémistes ont été adoptés par Facebook et YouTube. Cependant, aucune information n’a été communiquée sur le processus ou les critères retenus afin d’en déterminer le caractère « extrémiste »¹⁵.

Des initiatives similaires ont été mises en place en Europe. Un an après sa création en juillet 2015, l’unité de signalement des contenus sur internet au sein d’Europol a évalué et traité 11 000 messages contenant des documents au contenu extrémiste violent sur 31 plateformes en ligne dans huit langues, ce qui a entraîné la suppression de 91,4 % du contenu total de ces plateformes¹⁶. Le système aurait été automatisé avec la mise en place de la plateforme commune annoncée en avril 2016¹⁷.

Ces pratiques suscitent des préoccupations considérables sous l’angle des droits de l’homme quant au caractère prévisible et à la légalité des ingérences dans la liberté d’expression. En particulier, les données traitées par Europol recouvrent non seulement des contenus illégaux dans les États membres du Conseil de l’Europe, mais aussi les matériels qui enfreignent les conditions de service des intermédiaires internet. Par ailleurs, il est bien souvent difficile d’identifier un contenu extrémiste ou un matériel incitant au terrorisme car il faut réussir à démêler des facteurs comme le contexte

¹³ Voir <http://isc.independent.gov.uk/committee-reports/special-reports>.

¹⁴ Voir <https://www.article19.org/resources.php/resource/38579/en/algorithms-and-automated-decision-making-in-the-context-of-crime-prevention>

¹⁵ Voir <http://www.reuters.com/article/us-internet-extremism-video-exclusive-idUSKCN0ZB00M>

¹⁶ Voir <https://www.europol.europa.eu/newsroom/news/europol-internet-referral-unit-one-year>.

¹⁷ Voir Communication de la Commission au Parlement européen, au Conseil européen et au Conseil sur le Programme européen en matière de sécurité pour lutter contre le terrorisme et ouvrir la voie à une union de la sécurité réelle et effective https://ec.europa.eu/home-affairs/sites/homeaffairs/files/what-we-do/policies/european-agenda-security/legislative-documents/docs/20160420/communication_eas_progress_since_april_2015_en.pdf. Voir également Article 19, *Algorithms and Automated Decision-Making in the Context of Crime Prevention: A Briefing paper*, 2016.

culturel et l'humour. Selon la Cour européenne des droits de l'homme, l'article 10 protège également les contenus qui heurtent, choquent ou inquiètent. Le blocage, le filtrage ou la suppression de contenu par algorithme risque d'avoir une forte incidence négative sur les contenus légitimes. Le problème déjà très répandu de la suppression de grandes quantités de contenus licites en raison des conditions de service des intermédiaires internet est encore accentué par la pression qu'ils subissent afin d'opérer un filtrage actif selon des notions vagues telles que l'« extrémisme ».

Depuis les élections américaines en 2016, les publics européens et américains s'inquiètent de plus en plus de la création et de la diffusion de fausses informations, y compris par des techniques automatisées et sur les plateformes de réseaux sociaux, pouvant avoir une forte influence sur les processus décisionnels démocratiques (voir également point H ci-après). Des appels ont été une nouvelle fois lancés pour que soient appliquées aux plateformes de réseaux sociaux les normes de responsabilité des médias traditionnels. Certains universitaires ont comparé Facebook à un « rédacteur [à qui] incombe la responsabilité éditoriale de ses sujets tendances » (Helberger et Trilling 2016). La question se pose donc de savoir si les plateformes de réseaux sociaux, par le biais de leurs algorithmes qui classent et organisent les contributions de tiers, exercent une forme de contrôle éditorial traditionnellement assuré par les professionnels des médias et, par conséquent, créent des responsabilités spécifiques incombant aux médias¹⁸.

D. LIBERTÉ DE RÉUNION ET D'ASSOCIATION - ARTICLE 11 DE LA CONVENTION EUROPÉENNE DES DROITS DE L'HOMME

Internet, et en particulier les services de réseaux sociaux, sont des outils indispensables à l'exercice et à la jouissance du droit à la liberté de réunion et d'association et augmentent les possibilités de participation des individus à la vie politique, sociale et culturelle¹⁹. La liberté d'utiliser des plateformes internet telles que les réseaux sociaux, pour nouer des relations et créer des associations et s'organiser afin de se réunir

¹⁸ Voir aussi <http://reutersinstitute.politics.ox.ac.uk/news/editors-vs-algorithms-who-do-you-want-choosing-your-news>

¹⁹ Voir Recommandation CM/Rec(2012)4 du Comité des Ministres aux États membres sur la protection des droits de l'homme dans le cadre des services de réseaux sociaux.

pacifiquement, y compris pour manifester, conformément à l'article 11 de la Convention européenne des droits de l'homme, a également été soulignée²⁰.

En vertu de l'article 11, toute restriction du droit à la liberté de réunion pacifique et à la liberté d'association doit être prévue par la loi, poursuivre un but légitime et constituer une mesure nécessaire dans une société démocratique. L'utilisation d'algorithmes sur les plateformes de réseaux sociaux pouvant entraîner l'exclusion automatique de certaines personnes ou de certains groupes d'appels à des rassemblements par exemple, peut avoir un impact significatif sur la liberté de réunion, étant donné que les utilisateurs qui se servent de ces plateformes pour leurs contacts peuvent, sans le savoir, ne pas recevoir certaines communications. L'utilisation de filtres algorithmiques par les pouvoirs publics peut également empêcher des manifestations pacifiques.

²⁰ Voir Recommandation CM/Rec(2016)5 du Comité des Ministres aux États membres sur la liberté d'internet et Recommandation CM/Rec(2014)6 du Comité des Ministres aux États membres sur un Guide des droits de l'homme pour les utilisateurs d'internet.

E. DROIT À UN RECOURS EFFECTIF - ARTICLE 13 DE LA CONVENTION EUROPÉENNE DES DROITS DE L'HOMME

L'article 13 de la Convention européenne des droits de l'homme dispose que toute personne dont les droits ont été violés a droit d'un recours effectif devant une instance nationale. Par conséquent, les États doivent s'assurer que les personnes ont accès à des procédures judiciaires ou administratives à même de statuer en toute impartialité sur leurs allégations de violations de droits de l'homme en ligne, y compris des mécanismes non judiciaires effectifs, des moyens administratifs ou d'autres voies de recours, comme les institutions nationales des droits de l'homme. Etant responsables au premier chef de tous les droits prévus par la Convention européenne des droits de l'homme, les États doivent prendre les mesures appropriées pour assurer une protection contre les violations des droits de l'homme, y compris celles commises par des acteurs du secteur privé, et veiller à ce que les personnes concernées aient accès à un recours effectif. Ils doivent donc encourager tous les acteurs du secteur privé à respecter les droits de l'homme dans l'ensemble de leurs opérations, en particulier en mettant en place des mécanismes de réclamation effectifs, afin de traiter rapidement et de redresser directement les griefs des personnes.

De plus en plus d'entreprises, notamment les grandes, utilisent des algorithmes et des techniques de traitement automatisé des données pour le fonctionnement de leurs procédures de traitement des réclamations. Cela peut avoir un effet important sur le délai dans lequel une personne reçoit une réponse satisfaisante. Dans le cadre des processus automatisés de suppression de contenu sur les plateformes de réseaux sociaux (voir point C ci-avant), l'utilisation d'algorithmes se ressent particulièrement dans les délais de réponse en fonction des types de contenus et de leur classement par ordre de priorité, processus évidemment automatisé. Il en va de même pour le seuil de réclamations d'internautes requis pour qu'un contenu soit révisé. Tout laisse à penser que les réponses complètes des intermédiaires internet, tels que Facebook, Google ou Microsoft, aux demandes des utilisateurs sont automatisées pour de nombreux types de demandes et de réclamations (Wagner 2016; Zhang, Stalla-Bourdillon, and Gilbert 2016). Bien souvent, il faut attendre qu'un grand nombre d'internautes se soient plaints d'un type de contenu particulier pour qu'un algorithme automatisé l'identifie comme pertinent pour être soumis au contrôle d'un opérateur humain. Ces opérateurs sont réputés travailler souvent sous la pression et avec des consignes minimales quant à ce qu'ils doivent précisément supprimer conformément aux « règles de suppression »

internes²¹. Le droit à un recours effectif implique explicitement le droit à une décision motivée et individuelle. Jusqu'à présent, toutes ces décisions ont été prises par des êtres humains qui, dans l'exercice de leurs fonctions et après avoir suivi une formation approfondie, bénéficient d'une grande liberté d'appréciation. En principe, il appartient à un juge ou à un responsable administratif de décider comment l'équilibre des droits individuels, tels que la liberté d'expression et la protection contre la violence, doit être mis en pratique sur la base d'une analyse minutieuse au cas par cas du contexte, des conditions et de la nature de la situation considérée. Toutefois, en raison de l'utilisation accrue des techniques algorithmiques de traitement des données dans les procédures de réclamation, les algorithmes remplacent progressivement les humains.

En outre, la question de savoir si les processus de traitement automatisé des réclamations constituent un recours effectif est extrêmement préoccupante. Si la fameuse vidéo YouTube d'un débat du Parlement européen sur la torture, qui avait été supprimée, a été remise en ligne seulement quelques heures après la réclamation d'une députée européenne qui a même reçu des excuses publiques de la part de Google, il est largement permis de douter que toutes les réclamations soient traitées avec autant d'attention²². Au contraire, les algorithmes empêchent souvent d'accéder à une explication motivée sur les mesures qui ont été prises dans une affaire donnée.

Les décisions des pouvoirs publics visant à restreindre l'accès à un site web ou à un contenu particulier reposent souvent sur des termes vagues tels que « discours de haine » ou « extrémiste », bien souvent sans évaluation sous l'angle du respect des droits de l'homme (Husovec 2014). Ce faisant, il se peut que les pouvoirs publics délèguent le choix des outils et des mesures à un opérateur privé qui, seulement à ce moment-là, peut mettre en œuvre des solutions (comme la restriction d'accès ou la suppression de contenu) que la loi ne les autoriseraient pas à ordonner eux-mêmes. Les partenariats public/privé peuvent ainsi permettre aux acteurs publics « d'imposer des réglementations en matière de liberté d'expression qui pourraient outrepasser les règles constitutionnelles » (Mueller 2010:213), en violation des normes de l'État de droit. En outre, ce type de demandes faites par les instances publiques à des acteurs privés conduisent à un filtrage automatisé et par trop indiscriminé des contenus, qui offre le

²¹ Voir <http://international.sueddeutsche.de/post/154513473995/inside-facebook>.

²² Voir <https://www.marietjeschaake.eu/en/when-youtube-took-down-my-video>.

meilleur rapport coût-efficacité pour répondre à une demande publique de « suppression de tout discours de haine ».

En ce qui concerne le droit au respect de la vie privée, les techniques automatisées et les algorithmes facilitent des formes de surveillance secrète et de « dataveillance » dont la personne concernée ne peut avoir connaissance. La Cour européenne des droits de l'homme a souligné que l'absence de notification à quelque moment que ce soit compromet le caractère effectif des recours contre ces mesures²³.

F. INTERDICTION DE LA DISCRIMINATION - ARTICLE 14 DE LA CONVENTION EUROPÉENNE DES DROITS DE L'HOMME

Une autre liberté fondamentale qui est fréquemment citée en relation avec l'application des algorithmes est le droit à la protection contre toute discrimination.

Par définition, les algorithmes de recherche et les moteurs de recherche ne traitent pas toutes les informations de la même manière. Si les processus de sélection et d'indexation des informations peuvent être utilisés de manière systématique, les résultats de recherche seront généralement classés en fonction de leur pertinence supposée. Ainsi, différents éléments d'information bénéficieront de degrés de visibilité différents en fonction des facteurs pris en compte par l'algorithme de classement²⁴. En raison de l'agrégation de données et du profilage, il est possible que les algorithmes et les moteurs de recherche classent les publicités des petites entreprises immatriculées dans des régions moins favorisées en dessous de celles des grandes entités, ce qui peut les désavantager. Les moteurs et les algorithmes de recherche peuvent aussi ne pas traiter tous les utilisateurs de la même manière. Différents utilisateurs peuvent obtenir des résultats différents, sur la base des profils de comportement ou autres, y compris les profils de risque individuels qui peuvent être établis à des fins d'assurance ou de crédit ou, plus généralement, pour pratiquer une tarification différentielle (à savoir l'offre de prix différents pour les mêmes biens ou services à différents consommateurs en fonction de leur profil)²⁵.

²³ Voir *Roman Zakharov c. Russie* (requête n° 47143/06) du 4 décembre 2015.

²⁴ L'algorithme peut aussi, délibérément ou non, être influencé par divers facteurs extérieurs qui peuvent avoir trait aux modèles économiques, aux contraintes juridiques (les droits d'auteur par exemple) ou à d'autres facteurs contextuels.

²⁵ Proposition d'Aleksandra Kuczerawy, Brendan van Alsenoy et Jef Ausloos.

Un algorithme biaisé au sein d'un grand moteur de recherche en situation de quasi-monopole qui discrimine systématiquement un groupe de la société, par exemple en fonction de l'âge, de l'orientation sexuelle, de la race, du genre ou de la situation socio-économique, peut susciter de graves inquiétudes non seulement en ce qui concerne l'accès aux droits des consommateurs ou des utilisateurs finaux concernés par ces décisions, mais aussi pour la société dans son ensemble²⁶. On peut dès lors faire valoir que les personnes devraient avoir le droit de voir une version « impartiale » et non individuellement ciblée de leurs résultats de recherche. Cela pourrait permettre à un internaute de sortir de sa « bulle de filtres » et de voir une version non ciblée du contenu de sa recherche, de son activité sur les réseaux sociaux ou de tout autre service ou produit internet qu'il utilise. De fait, les algorithmes peuvent être des outils utiles pour réduire les discriminations dans des domaines où elles sont fréquentes, comme dans les processus de recrutement.

G. DROITS SOCIAUX ET ACCÈS AUX SERVICES PUBLICS

Le travail est un autre domaine important où la prise de décision automatisée est devenue de plus en plus courante ces dernières années. Les algorithmes peuvent être utilisés dans les décisions concernant le recrutement et le licenciement de personnel, l'organisation et la gestion du personnel, ainsi que l'évaluation individuelle des employés. Des boucles de rétroaction automatisées, parfois liées aux commentaires des clients, peuvent décider de l'évaluation des performances des employés (Kocher and Hensel 2016). Les processus décisionnels sont loin d'être parfaits lorsqu'ils sont exécutés par des humains. Des discriminations liées à la race (Bertrand and Mullainathan 2004), à la catégorie sociale et au sexe (Altonji et Blank 1999; Goldin et Rouse 1997) ont été démontrées à maintes reprises dans les pratiques et les processus de gestion des ressources humaines. Cependant, le fait que de plus en plus d'entreprises optent pour des méthodes de recrutement algorithmiques (Rosenblat, Kneese et al. 2014) soulève de nouvelles préoccupations quant à l'absence de transparence dans les décisions prises, au cours du processus de recrutement et au-delà. De plus, nombre de ces processus décisionnels automatisés reposent sur des données reçues via des intermédiaires internet. Permettre à la « sagesse des foules » de prendre des décisions sur l'emploi d'individus n'est pas seulement hautement discutable d'un point de vue éthique ; cela limite en outre la possibilité pour les travailleurs de contester ces décisions car elles

²⁶ Proposition de Sophie Stalla-Bourdillon, Steffen Staab et Laura Carmichael.

semblent constituer des mesures « objectives » de leur performance (Tufekci et al. 2015). Cela peut susciter des inquiétudes quant au respect des droits énoncés dans la Charte sociale européenne révisée.

À mesure que les plateformes d'emploi « transforment les gens en intelligence collective » (« *human computation* ») (Irani 2015:227), des questions se posent au sujet des droits des travailleurs, de l'autodétermination des employés et de la façon dont les sociétés dans leur ensemble pensent que les êtres humains devraient être traités au travail. En particulier, l'automatisation accrue du travail pose des défis considérables en ce qui concerne les droits au respect de la vie privée des employés et leur protection dans le cadre professionnel (Hendrickx et van Bever 2013). Alors que de plus en plus de systèmes sont automatisés et que de plus en plus de données sont collectées dans le cadre du travail, les droits des employés en vertu de l'article 8 sont évidemment en danger, même s'ils ne sont pas directement ciblés par ces mesures de collecte des données (voir point B ci-avant). Enfin, il existe d'autres problèmes liés à l'utilisation des algorithmes par les organisations tant du secteur public que du secteur privé pour contrôler les communications du personnel. Ces pratiques sont généralement utilisées afin de s'assurer que les employés représentent bien l'entreprise ou l'administration, et elles ont des conséquences manifestes pour la liberté d'expression des employés (Voorhoof et Humblet 2013) et leurs droits de l'homme en vertu de l'article 10 de la Convention (voir point C ci-avant).

Les organismes et les services publics automatisent de plus en plus leur prise de décision au moyen d'algorithmes (van Haastert 2016). La question de savoir si ces systèmes peuvent ou non accroître l'efficacité fait encore l'objet de vifs débats, mais il est évident que leur utilisation soulève d'importantes questions quant à la transparence et la responsabilité du processus décisionnel public, qui est tenu de satisfaire à des normes plus strictes que celles du secteur privé ou non lucratif. Aujourd'hui en Europe, le secteur public a recours à la prise de décision automatisée dans des domaines aussi divers que la sécurité sociale, la fiscalité, les soins de santé et le système judiciaire (van Haastert 2016; Tufekci et al. 2015). Il existe un fort risque de tri social dans les données médicales, étant donné que les algorithmes peuvent sélectionner des groupes de citoyens ou des profils humains spécifiques, et ainsi les empêcher d'accéder aux services sociaux. Un autre exemple est celui de la pratique de profilage des demandeurs d'emploi en Pologne, qui a été analysée par des chercheurs dans le but d'évaluer les conséquences sociales et politiques d'une prise de décision algorithmique associée à des prestations sociales (Jędrzej Niklas, Karolina Sztandar-Sztanderska et Katarzyna

Szymielewicz 2015). Cette analyse a mis en évidence plusieurs problèmes qui valent également pour l'utilisation d'algorithmes dans d'autres secteurs de la prestation de services publics, comme l'application de règles non transparentes et algorithmiques dans la répartition des services publics et des défaillances informatiques entraînant des décisions arbitraires, concernant par exemple eu égard l'octroi de prestations sociales.

H. DROIT À DES ÉLECTIONS LIBRES

L'utilisation d'algorithmes et de systèmes de recommandation automatisés capables de créer des « bulles de filtres », des chambres d'écho entièrement automatisées dans lesquelles les personnes voient uniquement les informations qui confirment leurs opinions ou correspondent à leur profil (Bozdag 2013; Pariser 2011; Zuckerman 2013), peut avoir des effets très importants sur les processus démocratiques de la société. Ces chambres d'écho entièrement automatisées présentent le risque de créer des « bulles idéologiques » (O'Callaghan et al. 2015) dans lesquelles il peut être relativement facile d'entrer mais dont il est difficile de sortir (Salamatian 2014), et d'avoir une conséquence déterminante, en particulier dans le contexte d'élections.

Si, depuis l'avènement d'internet, il a été allégué que les campagnes en ligne et les réseaux sociaux étaient susceptibles de changer la façon de mener les politiques et les élections, ce n'est que récemment que la recherche universitaire a révélé dans quelle mesure l'organisation et la manipulation du contenu en ligne sur les plateformes de réseaux sociaux pouvaient « faire basculer » des élections. Selon certaines informations, au cours des élections américaines, des chercheurs auraient manipulé la plateforme Facebook afin d'influencer le comportement électoral des internautes à leur insu, en leur faisant savoir pour qui leurs amis avaient dit avoir voté, et ils auraient réussi à convaincre une part statistiquement importante de la population de voter lors des élections de mi-mandat au Congrès le 2 novembre 2010 (Bond et al. 2012)²⁷. Il existe de bonnes raisons de penser que, depuis, Facebook vend des services de publicité politique aux partis politiques du monde entier, un comportement similaire ayant été observé lors

²⁷ Dans le cadre d'une expérience, les chercheurs ont présenté à certains utilisateurs de Facebook, dans leur flux d'actualité, un graphique indiquant combien de leurs amis avaient voté ce jour-là et proposant un bouton sur lequel cliquer pour confirmer qu'ils avaient eux aussi voté. Il s'est avéré que le fait d'être informés du vote de leurs amis avait accru de 0,39 % la probabilité que les utilisateurs aillent voter, et que leur décision avait eu à son tour des répercussions sur le comportement électoral de leurs amis. Les chercheurs ont conclu que leur simple message sur Facebook, communiqué de manière stratégique, avait augmenté directement la participation de 60 000 électeurs et, grâce à l'effet de ricochet, finalement permis l'expression de 340 000 voix supplémentaires (sur un total de 82 millions) ce jour-là. Voir Jonathan Zittrain, *Engineering an election*, Harvard Law Review Forum Vol. 127, 335 – 339 (2014).

des élections locales au Royaume-Uni, en 2016 (Griffin 2016). Le fait de savoir si Facebook, et d'autres plateformes en ligne en situation de quasi-monopole, utilisent leur pouvoir pour influencer le vote humain, de manière bien intentionnée ou non, est moins important que celui de savoir que ces plateformes sont, en principe, capables d'influencer massivement des élections.

Le droit à des élections libres, prévu par l'article 3 du Protocole n° 1, a été reconnu par la Cour européenne des droits de l'homme en tant que principe fondamental d'un régime politique véritablement démocratique. Et surtout, comme le note le projet d'étude de faisabilité sur l'utilisation d'internet dans les élections préparé par le Comité d'experts sur le pluralisme des médias et la transparence de leur propriété (MSI-MED) du Conseil de l'Europe, les difficultés réglementaires liées aux élections ne sont pas dues à l'augmentation du nombre d'intermédiaires mais plutôt à l'absence d'une réglementation adéquate. Comme l'étude le souligne, « [l]'effet le plus fondamental, le plus pernicieux et en même temps le plus difficilement détectable du recours accru aux médias sociaux n'est pas le pouvoir grandissant des intermédiaires, mais l'incapacité de la réglementation à garantir l'équité des règles de la lutte politique et à limiter le rôle de l'argent dans les élections. » (Renvoi au rapport de Damian Tambini du Comité MSI-MED du Conseil de l'Europe). Utilisation et effets des bots, des fausses informations, effets sur le pluralisme et la cohésion sociale.

4. MÉCANISMES DE GOUVERNANCE, RESPONSABILITÉ, TRANSPARENCE ET ÉTHIQUE

A. INTRODUCTION

Jusqu'à présent, les problèmes soulevés par le traitement automatisé des données étaient réglés par la législation relative à la protection des données. Aujourd'hui, des approches pertinentes et innovantes, telles que le « droit à l'explication » (Goodman and Flaxman 2016), sont également le fruit de cette législation. Cependant, il existe une différence importante entre la réglementation du droit au respect de la vie privée et de la protection des données qui, au final, reste un mécanisme de gouvernance conçu pour protéger la vie privée, et la réglementation d'autres droits. S'il est clair que les problèmes concernant la discrimination des contenus ou la manipulation des élections dépassent les questions du respect de la vie privée et de la protection des données et posent des questions relatives au droit de la concurrence ou peuvent relever des commissions électorales et des parlements, il est néanmoins possible de s'inspirer de l'expertise du secteur de la protection des données pour essayer d'identifier des réponses réglementaires adaptées à la gouvernance des algorithmes.

Enfin, des questions éthiques et juridiques tout à fait fondamentales se posent autour de la personnalité juridique des systèmes automatisés tels que les algorithmes, questions qui ne peuvent aisément être résolues dans le cadre de ce rapport. Sans vouloir disculper les personnes engagées dans le développement, la programmation et la mise en œuvre de systèmes autonomes, il convient de reconnaître que l'automatisation, l'analyse et l'adaptabilité de vastes ensembles de données et l'auto-apprentissage sont sources de défis considérables sous l'angle de la responsabilité des décisions algorithmiques. Cela a conduit certains auteurs à suggérer que de nombreuses formes de transparence, de responsabilisation et de réglementation des algorithmes sont impossibles car les programmeurs eux-mêmes sont incapables de prédire ou de comprendre entièrement la façon dont l'algorithme prend ses décisions (Kroll 2016).

B. EST-IL RAISONNABLE ET POSSIBLE DE RÉGLEMENTER LES ALGORITHMES ?

L'utilisation des algorithmes en tant que telle suscite des inquiétudes croissantes en Europe au niveau public et politique, car elle soulève des problèmes considérables au regard des droits de l'homme et devrait par conséquent être réglementée²⁸. Bien qu'il n'y ait pas de consensus sur des mécanismes appropriés pour la réglementer, dans de nombreux cas, des gouvernements ou des contrôleurs indépendants réglementent déjà les algorithmes avant qu'ils ne soient appliqués.

Les logiciels et les algorithmes utilisés dans les « machines à sous » en Australie et en Nouvelle Zélande doivent, en vertu d'une réglementation publique, être « justes, sûrs et vérifiables » (Woolley et al. 2013). Ainsi les développeurs de ces machines sont tenus de soumettre leurs algorithmes aux autorités de régulation avant de pouvoir les mettre à disposition des consommateurs. La norme nationale australienne/néo-zélandaise relative aux machines de jeu, dans sa version 10.3 la plus récente, explique de manière remarquablement technique et détaillée comment ces machines devraient fonctionner. Par exemple, l'« écart-type nominal d'un jeu doit être égal ou inférieur à 15 » et « l'algorithme de hachage pour la vérification des logiciels, des matériels et des PSD des machines de jeu est l'algorithme HMAC-SHA1 »²⁹. Au Royaume-Uni, les machines de jeu sont également contrôlées par un régime d'autorisation spécifique. De plus, un débat est en cours dans le secteur financier sur une réglementation des algorithmes de trading à grande vitesse qui pourraient avoir un effet fortement déstabilisant sur l'ensemble du système financier. En 2012, un haut responsable politique a suggéré que les « algorithmes de trading soient soumis à un test de résistance afin de vérifier leur stabilité » (Steinbrück 2012). Une réglementation similaire est envisagée dans le domaine des services d'assistance en ligne et de la régulation des contenus en ligne. Le Centre de la police britannique pour la protection en ligne et la lutte contre l'exploitation des enfants a exigé que son « bouton Facebook » soit proposé par défaut à tous les internautes (Wagner 2016). Si cette tentative de faire pression sur Facebook pour que la plateforme change son code par défaut sur son site web britannique s'est soldée par un

²⁸ Voir, par exemple, le projet de loi pour une République numérique adopté le 26 janvier 2016 par l'Assemblée nationale française. Ce projet contient des dispositions relatives à la transparence des algorithmes et à l'obligation de « loyauté » ou équité, des plateformes en ligne et des décisions algorithmiques » (Rosnay 2016).

²⁹ La norme nationale australienne/néo-zélandaise relative aux machines de jeu peut être consultée ici : <https://publications.qld.gov.au/dataset/a-nz-gaming-machine-national-standards>

échec, elle laisse entrevoir le type de réponses réglementaires auxquelles on pourrait s'attendre si les États entreprennent de définir le contenu des algorithmes sur les grandes plateformes internet.

C. TRANSPARENCE

Pour de nombreux consommateurs et régulateurs, les algorithmes ressemblent à des boîtes noires (Pasquale 2015). Comme l'ont noté Tufekci et al., « une préoccupation éthique couramment évoquée au sujet de la prise de décision algorithmique est celle du caractère opaque de nombreux algorithmes. Lorsque les algorithmes sont utilisés pour prendre des décisions directes, comme dans le cas d'un diagnostic médical ou dans l'aviation, l'absence de transparence soulève d'importantes questions de responsabilité » (Tufekci et al. 2015:11). La transparence des algorithmes fait-elle ainsi l'objet d'un débat récurrent et, de plus en plus souvent, les gouvernements demandent aux entreprises de faire contrôler leurs algorithmes par des auditeurs indépendants, les autorités de réglementation ou l'opinion publique (Diakopoulos 2015; Rosnay 2016) avant de les mettre en œuvre.

Il est peu probable que des algorithmes soient divulgués au public dans leur intégralité, les entreprises privées les considérant comme leur principal secret commercial³⁰. Cependant, il y a aussi débat autour de la possibilité de fournir au public des sous-ensembles importants d'informations sur les algorithmes, par exemple les variables utilisées, les valeurs moyennes et les écarts-types des résultats produits, ou la quantité et le type de données traitées.

Toutes ces mesures visent à renforcer la transparence des systèmes automatisés qu'il est d'autant plus difficile d'assurer que les algorithmes utilisés subissent de fréquents changements. Google, par exemple, modifie son algorithme des centaines de fois par an (Tufekci et al. 2015). Il existe aussi bien souvent un risque de manipulation des algorithmes s'ils sont rendus publics. Par ailleurs, les techniques d'apprentissage automatique compliquent encore la situation, de sorte que même la divulgation de tous les codes sources d'un algorithme pourrait ne pas suffire à garantir la transparence et qu'il serait préférable d'avoir une explication précise de la façon dont les résultats d'un

³⁰ Dans une décision du 28 janvier 2014, la Cour suprême fédérale allemande (Bundesgerichtshof) a rejeté une demande d'information concernant l'algorithme d'un organisme de crédit au motif qu'il s'agissait d'un secret commercial protégé. Elle a, toutefois, autorisé une demande d'information concernant les données utilisées pour calculer la solvabilité au moyen de l'algorithme. (SOURCE ?)

algorithmes ont été produits. Les premières étapes vers un droit à une transparence *effective* peuvent s'inspirer du Règlement général européen sur la protection des données (GDPR), y compris un éventuel droit à l'explication (Goodman and Flaxman 2016).

Étant donné que l'utilisation des algorithmes dans la prise de décision risque de nuire aux droits des personnes, un mécanisme de contrôle pourrait permettre de garantir que l'algorithme est appliqué de manière juste et durable. Par exemple, l'article 28 b de la loi fédérale allemande sur la protection des données prévoit qu'il doit exister un processus mathématique statistique fondé sur des preuves scientifiques pour calculer la probabilité d'un comportement spécifique d'un individu avant qu'un tel algorithme puisse être utilisé pour prendre une décision sur un contrat.

D. RESPONSABILISATION

Quelle est la responsabilité des personnes ou des entreprises concernant les algorithmes qu'elles mettent en œuvre ? Cela dépend principalement de la nature des algorithmes et de leurs résultats. Dans nombre de cas, si les résultats sont diffamatoires, lèsent des droits d'auteur ou posent d'autres problèmes sur le plan juridique, les mécanismes de gouvernance existants s'assurent qu'ils soient réduits (Staab, Stalla-Bourdillon et Carmichael 2016). L'affaire Max Mosley contre Google n'est qu'un exemple parmi d'autres (Stanley 2011). Cependant, ces mécanismes ne concernent généralement que des règles de second niveau, à savoir des modifications apportées aux résultats des algorithmes. En revanche, il manque généralement de cadres réglementaires permettant d'influer sur les règles de premier ordre et de garantir que les algorithmes produisent d'emblée des résultats qui respectent et protègent les valeurs fondamentales ou les principes éthiques et sociétaux de base.

Quoi qu'il en soit, il a été suggéré que « les techniciens ont une conception de la confiance et de l'assurance à l'égard des systèmes informatiques un peu différente de celle des décideurs politiques ; ils cherchent de solides garanties formelles ou des preuves numériques fiables démontrant qu'un système fonctionne comme prévu ou respecte une règle ou un objectif politique plutôt que la simple assurance qu'un logiciel fonctionne d'une certaine manière » (Kroll et al. 2016).

Cela vient alimenter le débat plus large sur un contrôle des algorithmes selon lequel ceux-ci pourraient vraisemblablement générer des « preuves à divulgation nulle de

connaissance » pour démontrer qu'ils sont conformes à certaines propriétés, sans que la personne utilisant la preuve n'ait une vision du véritable algorithme (Kroll 2016).

E. CADRES ÉTHIQUES ET MEILLEURE ÉVALUATION DES RISQUES

Outre les mécanismes réglementaires directs destinés à influencer le code des algorithmes, des mécanismes indirects pourraient également être envisagés. Ils concernent le processus de production ou les producteurs d'algorithmes et cherchent à garantir qu'ils sont conscients des problèmes juridiques, des dilemmes éthiques et des préoccupations en matière de droits de l'homme suscités par la prise de décision automatisée. Cet objectif pourrait être atteint au moyen d'une éthique professionnelle normalisée ou d'un système d'autorisation pour les ingénieurs de données et les concepteurs d'algorithmes, analogue à ceux qui existent pour les médecins, les juristes ou les architectes³¹. Une autre proposition souvent avancée consisterait à améliorer les mécanismes existants pour les processus de gestion et de développement de logiciels (Spiekermann 2015). Cela pourrait concerner plus particulièrement les techniques de développement de logiciels agiles dans lesquelles la modularité, la temporalité et la capture posent des problèmes considérables sur le plan du droit au respect de la vie privée (Gürses and Hoboken 2017) et d'autres droits de l'homme (Mannaro 2008).

Il est important de noter que ces problèmes se posent non pas uniquement aux professionnels qui développent des algorithmes, mais aussi aux analystes qui utilisent les données. Il a souvent été affirmé que l'application des algorithmes dans l'apprentissage automatique se fait en grande partie sans que les relations de cause à effet soient « comprises » (corrélation au lieu de causalité), ce qui peut provoquer des biais et des erreurs et susciter des préoccupations quant à la qualité des données (O'Neil 2016). Le problème toutefois, concerne moins les algorithmes eux-mêmes que la façon dont les êtres humains perçoivent et interprètent leurs résultats. L'idée selon laquelle les algorithmes informatiques produisent des résultats impartiaux et neutres (Chun 2006) exempt de toute forme de politique (Denardis 2008) est au cœur du problème. C'est pourquoi il serait plus utile d'assurer une participation plus critique aux débats publics sur les algorithmes que d'essayer de les modifier.

³¹ Proposition de Markus Oermann, Université de Hambourg.

Une réglementation directe des codes des algorithmes et des logiciels devrait être adordée avec la plus grande prudence. C'est l'approche réglementaire qui présente le plus d'écueils et qui risque le plus d'exacerber les problèmes. En particulier, la réglementation directe pose des problèmes considérables concernant la liberté d'opinion et d'expression et le droit au respect de la vie privée. En outre, dans la mesure où les régulateurs possèdent rarement une connaissance approfondie des algorithmes, il serait bien plus approprié de renforcer les mesures visant à en garantir la transparence.

5. CONCLUSIONS

Le fonctionnement des systèmes décisionnels automatisés est très difficile à comprendre et soulève de nombreuses questions au regard des droits de l'homme. S'agissant d'un domaine relativement nouveau, nombre de ces problèmes ne peuvent être aisément appréciés et il reste difficile de trouver des solutions efficaces. Dans un premier temps, les décideurs politiques devraient chercher à en apprendre davantage sur la mise en œuvre des systèmes décisionnels automatisés dans leur pays respectif. Dans un deuxième temps, ils devraient tenter de s'assurer que la législation et les cadres juridiques en vigueur sont effectivement appliqués pour faire face aux difficultés soulevées par la prise de décision automatisée dans les différents domaines où elle est appliquée. Les conclusions de la présente étude sont ici similaires à celles du projet d'étude de faisabilité du comité MSI-MED sur l'utilisation d'internet dans les élections, ce qui laisse à penser que les principales difficultés rencontrées ne sont pas liées à l'importance croissante du rôle des intermédiaires mais bien à des défaillances réglementaires en matière de gouvernance.

Les conclusions de ce rapport ne doivent pas être entendues comme un appel à réglementer le développement des algorithmes ou d'autres codes logiciels. Une ingérence dans le droit des personnes à rechercher, développer et tester pourrait constituer en soi une violation de leur liberté d'opinion, d'expression, de pensée et de recherche. En plus des incidences non négligeables sur les droits de l'homme, une réglementation de la recherche et du développement d'algorithmes empêcherait de mieux comprendre comment les algorithmes fonctionnent et quels sont leurs effets.

Néanmoins, les discussions politiques relatives aux algorithmes et aux techniques de traitement automatisé des données devraient être guidées par des considérations juridiques, sociales et éthiques étroitement liées et interdépendantes et, plus largement,

poser la question de la nécessité d'assurer une transparence et une responsabilité effectives et de permettre la poursuite de la recherche-développement.

RESUME ET CONCLUSIONS GENERALES

Un résumé sera rajouté à la version finale de l'étude.

BIBLIOGRAPHIE

- Altonji, JG and RM Blank. 1999. 'Race and Gender in the Labor Market'. Pp. 3143–3259 in Handbook of labor economics. Elsevier B.V. Retrieved (<http://www.sciencedirect.com/science/article/pii/S1573446399300390>).
- Andreessen, Marc. 2011. 'Why Software Is Eating The World'. Wall Street Journal, August 20. Retrieved 1 September 2016 (<http://www.wsj.com/articles/SB10001424053111903480904576512250915629460>).
- Bertrand, Marianne and Sendhil Mullainathan. 2004. 'Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination'. The American Economic Review 94(4):991–1013.
- Bond, Robert M. et al. 2012. 'A 61-Million-Person Experiment in Social Influence and Political Mobilization'. Nature 489(7415):295–298.
- Bozdag, Engin. 2013. 'Bias in Algorithmic Filtering and Personalization'. Ethics and Information Technology 15(3):209–227.
- Bucher, Taina. 2012. 'Want to Be on the Top? Algorithmic Power and the Threat of Invisibility on Facebook'. New Media & Society 1461444812440159.
- Bucher, Taina. 2016. 'The Algorithmic Imaginary: Exploring the Ordinary Affects of Facebook Algorithms'. Information, Communication & Society 1–15.
- Buni, Catherine and Soraya Chemaly. 2016. 'The Secret Rules of the Internet'. The Verge. Retrieved 9 September 2016 (<http://www.theverge.com/2016/4/13/11387934/internet-moderator-history-youtube-facebook-reddit-censorship-free-speech>).
- Chun, Wendy Hui Kyong. 2006. *Control and Freedom: Power and Paranoia in the Age of Fiber Optics*. Cambridge Mass.: MIT Press.
- Denardis, Laura. 2008. 'Architecting Civil Liberties'. in Global Internet Governance Academic Network Annual Meeting. Hyderabad (Andhra Pradesh), India: GIGANET. Retrieved (<http://worldcat.org/oclc/619234880/viewonline>).
- Denardis, Laura. 2012. 'Hidden Levers of Internet Control'. Information, Communication & Society (September):37–41.
- Diakopoulos, Nicholas. 2015. 'Algorithmic Accountability'. Digital Journalism 3(3):398–415.
- Gillespie, Tarleton. 2014. 'The Relevance of Algorithms'. Pp. 167–94 in Media technologies: Essays on communication, materiality, and society, edited by T. Gillespie, P. J. Boczkowski, and K. A. Foot. Cambridge Mass.: MIT Press.
- Goldin, Claudia and Cecilia Rouse. 1997. *Orchestrating Impartiality: The Impact Of 'blind' auditions on Female Musicians*. National bureau of economic research. Retrieved 9 September 2016 (<http://www.nber.org/papers/w5903>).
- Goodman, Bryce and Seth Flaxman. 2016. 'European Union Regulations on Algorithmic Decision-Making and a Right to Explanation'. in 2016 ICML Workshop on Human Interpretability in Machine Learning. New York, NY: ArXiv e-prints.
- Griffin, Andrew. 2016. 'How Facebook Is Manipulating You to Vote'. The Independent. Retrieved 31 August 2016 (<http://www.independent.co.uk/life-style/gadgets-and-tech/news/uk-elections-2016-how-facebook-is-manipulating-you-to-vote-a7015196.html>).
- Gürses, Seda and Joris Hoboken. 2017. 'Privacy After the Agile Turn'. in The Cambridge Handbook of Consumer Privacy, edited by Selinger. Retrieved (<https://osf.io/ufdvb/>).
- van Haastert, Hugo. 2016. 'Government as a Platform: Public Values in the Age of Big Data'. Oxford Internet Institute.
- Helberger, Natali and Damian Trilling. 2016. 'Facebook Is a News Editor: The Real Issues to Be Concerned about'. Media Policy Project. Retrieved 9 September 2016 (<http://blogs.lse.ac.uk/mediapolicyproject/2016/05/26/facebook-is-a-news-editor-the-real-issues-to-be-concerned-about/>).
- Hendrickx, Frank and Aline van Bever. 2013. 'Article 8 ECHR: Judicial Patterns of Employment Privacy Protection'. Pp. 183–208 in The European Convention on Human Rights and the Employment Relation, edited by F. Dorssmont, K. Lörcher, and I. Schömann. Oxford: Hart Publishing.
- Husovec, Martin. 2014. 'CJEU Allowed Website-Blocking Injunctions with Some Reservations'. Journal of Intellectual Property Law & Practice jpu101.

- Irani, L. 2015. 'Difference and Dependence among Digital Workers: The Case of Amazon Mechanical Turk'. *South Atlantic Quarterly* 114(1):225–234.
- Jędrzej Niklas, Karolina Sztandar-Sztanderska, and Katarzyna Szymielewicz. 2015. Warsaw, Poland: Panoptykon Foundation. Retrieved (<https://en.panoptykon.org/articles/profiling-unemployed-poland-%E2%80%93-report>).
- Kitchin, R. and M. Dodge. 2011. *Code/Space Software and Everyday Life*.
- Kocher, Eva and Isabell Hensel. 2016. 'Herausforderungen Des Arbeitsrechts Durch Digitale Plattformen – Ein Neuer Koordinationsmodus von Erwerbsarbeit'. *Neue Zeitschrift Für Arbeitsrecht* (16/2016):984–89.
- Kroll, Joshua A. et al. 2016. 'Accountable Algorithms'. Retrieved 1 September 2016 (<http://balkin.blogspot.co.at/2016/03/accountable-algorithms.html>).
- Kroll, Joshua A. 2016. 'Accountable Algorithms (A Provocation)'. Media Policy Project. Retrieved 9 September 2016 (<http://blogs.lse.ac.uk/mediapolicyproject/2016/02/10/accountable-algorithms-a-provocation/>).
- Lazer, David and Ryan Kennedy. 2015. *What We Can Learn from the Epic Failure of Google Flu Trends*.
- Lazer, David, Ryan Kennedy, Gary King, and Alessandro Vespignani. 2014. 'The Parable of Google Flu: Traps in Big Data Analysis'. *Science* 343(6176):1203–5.
- Mannaro, Katuscia. 2008. 'Adopting Agile Methodologies in Distributed Software Development'. Università degli Studi di Cagliari, Cagliari. Italy. Retrieved (<http://le.uwpress.org/content/87/2/284.short>).
- McCarthy, Daniel R. 2011. 'Open Networks and the Open Door: American Foreign Policy and the Narration of the Internet'. *Foreign Policy Analysis* 7(1):89–111.
- Mueller, Milton. 2010. *Networks and States: The Global Politics of Internet Governance*. MIT Press.
- O'Callaghan, D., D. Greene, M. Conway, J. Carthy, and P. Cunningham. 2015. 'Down the (White) Rabbit Hole: The Extreme Right and Online Recommender Systems'. *Social Science Computer Review Social Science Computer Review* 33(4):459–78.
- O'Neil, Cathy. 2016. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York: Crown.
- Pariser, Eli. 2011. *The Filter Bubble: What the Internet Is Hiding from You*. New York: Penguin Press.
- Pasquale, Frank. 2015. *The Black Box Society: The Secret Algorithms That Control Money and Information*. Harvard University Press.
- Perry, Walt L. 2013. *Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations*. Rand Corporation. Retrieved 9 September 2016 (<https://books.google.com/books?hl=en&lr=&id=ZdstAQAQBAJ&oi=fnd&pg=PP1&dq=Perry,+Walter,+and+Brian+McInnis.+2013.+Predictive+Policing:+The+Role+of+Crime+Forecasting+in+Law+Enforcement+Operations.+Santa+Monica,+CA:+RAND.&ots=924yNa6Vct&sig=N3HnEi1FBr9YyMXV77GsgPbovYc>).
- Rifkind, Malcolm. 2014. *Report on the Intelligence Relating to the Murder of Fusilier Lee Rigby*.
- Rosenblat, Alex, Tamara Kneese, and others. 2014. 'Networked Employment Discrimination'. *Open Society Foundations' Future of Work Commissioned Research Papers*. Retrieved 9 September 2016 (http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2543507).
- Rosnay, Mélanie Dulong de. 2016. 'Algorithmic Transparency and Platform Loyalty or Fairness in the French Digital Republic Bill'. Media Policy Project. Retrieved 1 September 2016 (<http://blogs.lse.ac.uk/mediapolicyproject/2016/04/22/algorithmic-transparency-and-platform-loyalty-or-fairness-in-the-french-digital-republic-bill/>).
- Rubinstein, Ira, Ronald D. Lee, and Paul M. Schwartz. 2008. *Data Mining and Internet Profiling: Emerging Regulatory and Technological Approaches*. Rochester, NY: Social Science Research Network. Retrieved 9 September 2016 (<http://papers.ssrn.com/abstract=1116728>).
- Salamatian, Kavé. 2014. 'From Big Data to Banality of Evil'. Retrieved 9 September 2016 (<https://www.oximity.com/article/Vortrag-Big-Data-und-Ethik-1>).
- Schulz, Wolfgang and Kevin Dankert. 2016. 'Governance by Things' as a Challenge to Regulation by Law'. *Internet Policy Review* 5(2).
- Sills, Arthur J. 1970. 'Automated Data Processing and the Issue of Privacy'. *Seton Hall Law Review* 1.
- Spiekermann, Sarah. 2015. *Ethical IT Innovation: A Value-Based System Design Approach*. CRC Press.
- Staab, Steffen, Sophie Stalla-Bourdillon, and Laura Carmichael. 2016. 'Observing and Recommending from a Social Web with Biases'. arXiv Preprint arXiv:1604.07180. Retrieved 9 September 2016 (<http://arxiv.org/abs/1604.07180>).

- Stanley, JE. 2011. 'Max Mosley and the English Right to Privacy'. Wash. U. Global Stud. L. Rev. 10(3). Retrieved (http://heinonlinebackup.com/hol-cgi-bin/get_pdf.cgi?handle=hein.journals/wasglo10§ion=25).
- Steinbrück, Peer. 2012. *Vertrauen Zurückgewinnen: Ein Neuer Anlauf Zur Bändigung Der Finanzmärkte*. Berlin, Germany: Deutscher Bundestag - German Federal Parliament.
- Tene, Omer and Jules Polonetsky. 2012. 'To Track or "Do Not Track": Advancing Transparency and Individual Control in Online Behavioral Advertising'. Retrieved 9 September 2016 (<http://conservancy.umn.edu/handle/11299/155947>).
- Toor, Amar. 2016. 'Automated Systems Fight ISIS Propaganda, but at What Cost?' The Verge. Retrieved 9 September 2016 (<http://www.theverge.com/2016/9/6/12811680/isis-propaganda-algorithm-facebook-twitter-google>).
- Tufekci, Zeynep, Jillian C. York, Ben Wagner, and Frederike Kaltheuner. 2015. *The Ethics of Algorithms: From Radical Content to Self-Driving Cars*. Berlin, Germany: European University Viadrina. Retrieved (<https://cihr.eu/publication-the-ethics-of-algorithms/>).
- Voorhoof, Dirk and P. Humblet, eds. 2013. 'The Right to Freedom of Expression in the Workplace under Article 10 ECHR'. Pp. 183–208 in *The European Convention on Human Rights and the Employment Relation*. Oxford: Hart Publishing.
- Wagner, Ben. 2016. *Global Free Expression: Governing the Boundaries of Internet Content*. Cham, Switzerland: Springer International Publishing.
- Williamson, Ben. 2016. 'Computing Brains: Learning Algorithms and Neurocomputation in the Smart City'. *Information, Communication & Society* 0(0):1–19.
- Winner, L. 1980. 'Do Artifacts Have Politics?' Daedalus.
- Winner, L. 1986. 'The Whale and the Reactor: A Search for Limits in an Age of High Technology'.
- Woolley, Richard, Charles Livingstone, Kevin Harrigan, and Angela Rintoul. 2013. 'House Edge: Hold Percentage and the Cost of EGM Gambling'. *International Gambling Studies* 13(3):388–402.
- York, Jillian C. 2010. 'Policing Content in the Quasi-Public Sphere'. Boston, MA: Open Net Initiative Bulletin. Berkman Center. Harvard University.
- Zhang, Pei, Sophie Stalla-Bourdillon, and Lester Gilbert. 2016. 'A Content-Linking-Context Model For "notice-and-Take-Down" procedures'. Pp. 161–65 in. ACM Press. Retrieved 9 September 2016 (<http://dl.acm.org/citation.cfm?doid=2908131.2908171>).
- Zuckerman, Ethan. 2013. *Digital Cosmopolitans: Why We Think the Internet Connects Us, Why It Doesn't, and How to Rewire It*. W. W. Norton & Company.