MSI-NET

Comité d'experts sur les intermédiaires d'internet

COUNCIL OF EUROPE

CONSEIL DE L'EUROPE

6 octobre 2017 **MSI-NET (2017)06**

MSI-NET 4e réunion
18-19 septembre 2017
(Strasbourg, Agora, salle G05)

Rapport de réunion

1. Le professeur Wolfgang Schulz, président du MSI-NET, ouvre la réunion. Jan Kleijssen, directeur de la Direction de la société de l'information et de la lutte contre la criminalité, souhaite la bienvenue aux membres et aux participants. Il les félicite pour le travail qu'ils ont effectué jusqu'à présent et les invite à se consacrer à la finalisation des deux projets pendant cette dernière réunion. M. Kleijssen rappelle qu'au cours de la 12e réunion du CDMSI (Comité directeur du Conseil de l'Europe sur les médias et la société de l'information), qui s'est tenue en juin 2017, les délégués ont été informés par le vice-président du MSI-NET des progrès réalisés sur les deux textes, et étaient convenus de lancer des consultations publiques sur le projet de recommandation pendant l'été. Au cours des consultations, 23 séries de commentaires détaillés ont été reçues de la part de représentants des États membres, de comités conventionnels, de la société civile et du milieu universitaire. Il informe également les membres et participants du MSI-NET de la mise en œuvre en cours de la Stratégie du Conseil de l'Europe sur la gouvernance d'internet et des évolutions récentes dans le contexte de l'initiative du Conseil de l'Europe de créer une plateforme pour favoriser le dialogue entre les États membres et les entreprises de l'internet et ainsi améliorer le respect des droits de l'homme, de la démocratie et de l'État de droit sur internet.

M. Kleijssen rappelle par ailleurs l'importance de la conférence intitulée « Le rôle et les responsabilités des intermédiaires de l'internet », organisée conjointement par la Présidence autrichienne de l'OSCE et la Présidence tchèque du Comité des Ministres du Conseil de l'Europe le 13 octobre à Vienne, au cours de laquelle le MSI-NET sera largement représenté et qui offrira une tribune idéale pour attirer l'attention sur les travaux du Comité et sur le projet de recommandation. Souhaitant aux membres et aux participants un débat fructueux sur les points particulièrement pertinents à l'ordre du jour, il les encourage également à se

fixer des objectifs de clarté et de concision pour parvenir à des textes capables d'orienter les actions des États membres.

2. L'ordre du jour (annexe 1) est adopté sans modification. La liste des participants figure à l'annexe 2. La répartition hommes-femmes des 30 participants est la suivante : 12 femmes (40 %) et 18 hommes (60 %).

Conclusions et décisions

3. En ce qui concerne *le projet de recommandation du Comité des Ministres sur les intermédiaires d'internet*, le MSI-NET examine la version révisée du document tel que présenté par le rapporteur et le Secrétariat, qui a intégré les commentaires recueillis au cours des consultations publiques. Le MSI-NET souscrit à la plupart des modifications, telles que présentées, et développe le texte plus avant en :

a) clarifiant les normes relatives aux procédures rapides pour restreindre l'accès aux contenus sans passer par la voie judiciaire ;

b) étoffant les paragraphes dévolus aux procédures reposant sur la notification dans les deux chapitres des lignes directrices (obligations des États et responsabilités des intermédiaires) ;

c) introduisant des garanties plus spécifiques en lien avec les mécanismes de contrôle et de plainte judiciaires et non judiciaires instaurés par les intermédiaires,

et en

d) ajoutant du contenu lié à la responsabilité des grands intermédiaires pour élaborer des normes et des codes de conduite visant à prévenir les propos injurieux, les images offensantes, les discours de haine et l'incitation à la violence

Un certain nombre d'observations, de remarques et de propositions supplémentaires sont formulées pour modifier des points spécifiques du texte et le **projet de recommandation est finalisé pour approbation par le CDMSI à l'occasion de sa 13e réunion (5 – 8 décembre 2017)** (annexe 3).

4. The MSI-NET poursuit ses discussions concernant les parties IV et V du *projet d'étude sur les dimensions droits de l'homme des techniques de traitement automatisé des données (en particulier les algorithmes) et sur de possibles implications en matière de réglementation*, tel que soumis par le rapporteur à la fin du mois d'août. Les membres et les participants formulent divers commentaires ainsi que des suggestions et propositions de modification supplémentaires, qu'il est décidé d'un commun accord d'intégrer dans **la version finale de l'étude** (annexe 4). Il est convenu de conclure le chapitre final par des propositions d'action concrètes pour les États membres, et de suggérer que le Conseil de l'Europe approfondisse ses travaux sur les thèmes étudiés, en vue de l'élaboration prochaine d'instruments normatifs capables de guider l'action des États membres.

Questions diverses

6. Le Secrétariat est chargé de préparer un projet de rapport de réunion à envoyer au président et au vice-président pour examen. Il enverra ensuite le projet de rapport au MSI-NET avec un délai de cinq jours ouvrables pour formuler ses commentaires. En l'absence de commentaires, le rapport sera considéré comme définitif et transmis au CDMSI pour information avec les versions finalisées des deux projets. Il est donc jugé inutile d'établir un rapport de réunion abrégé.

ANNEXE I

ORDRE DU JOUR[1]

1. **Ouverture de la réunion**

2. **Adoption de l'ordre du jour**

3. **Information du Secrétariat**

4. **Finalisation de la 3ème version révisée du projet de recommandation sur les rôles et responsabilités des intermédiaires d'internet**
   Doc MSI-NET(2016)05rev4

5. **Finalisation de la 2ème version révisée du projet d'étude portant sur les dimensions des droits de l'homme dans l'application des techniques de traitement des données informatiques (en particulier les algorithmes) et leurs implications éventuelles sur le plan réglementaire**
   Doc MSI-NET (2016)06rev3

6. **Autres points**

MANDAT MSI-NET

[1] Tel que produit sous doc MSI-NET(2017)05

ANNEXE 2

LISTE DES PARTICIPANTS

MEMBRES DU COMITE

M. Bertrand De la CHAPELLE, Co-fondateur et Directeur de « Internet & Jurisdiction », France

Mme Julia HÖRNLE, Professeur des lois dans le domaine d'Internet, Queen Mary University of London

Mme Tanja KERŠEVAN-SMOKVINA, Conseillère principale auprès du directeur général - Agence pour les réseaux et services de communication – Slovénie (Rapporteur pour l'égalité de genre)

M. Matthias KETTEMANN, Postdoc Fellow, Cluster of Excellence "Normative Orders" Université de Francfort-sur-le-Main (Rapporteur Recommandation)

M. Arseny NEDYAK, Directeur adjoint, Service des politiques nationales des médias, Ministère de la télécommunication – Fédération de Russie

Mme Dörte NIELANDT, division VI A3 (Cadre juridique pour les services numériques, l'industrie des médias), Ministère Fédéral de l'Economie et de l'Energie – Allemagne

M. Pēteris PODVINSKIS, Ministère des affaires étrangères, Direction Organisations Internationales, Service des Politiques publiques dans le domaine de l'Internet – Lettonie

M. Thomas SCHNEIDER, Directeur adjoint des affaires internationales, Coordinateur de la société d'information internationale, Service fédéral de l'environnement, transport, énergie et communication DETEC, Office fédéral des communications (OFCOM) – Suisse

M. Wolfgang SCHULZ, Professeur, Faculté de droit, Université de Hambourg / Institut de Hans-Bredow (président)

Mme Sophie STALLA-BOURDILLON, Professeur agrégée en technologie d'information / droit de la propriété intellectuelle, Directrice de ILAWS, Faculté de droit de Southampton, Université de Southampton

Mme Karmen TURK, Trinity Tallinn – Estonie (vice-présidente)

M. Dirk VOORHOOF, Professeur de droit européen des media, UCPH (Université de Copenhague) / Professeur à l'université de Gand / membre du comité scientifique du CMPF (Centre pour le pluralisme des médias et la liberté de la presse)

M. Benjamin WAGNER – Professeur assistant, Institute for Management Information Systems, Vienna University of Economics and Business / (Rapporteur Étude)

ETATS MEMBRES DU CONSEIL DE L'EUROPE

Autriche – M. Gerhard HOLLEY, Chancellerie fédérale d'Autriche, service constitutionnel

REPUBLIQUE TCHEQUE- M. Jakub SVAB, Direction des medias et de l'audiovisuel, Ministère de la Culture

TURQUIE – M. İrfan Dündar ERENTÜRK, Spécialiste Médias, Conseil Supérieur de l'Audiovisuel (RTÜK) - Ankara

OBSERVATEURS

COMMISSION EUROPEENNE - DG CONNECT - Mme Irene ROCHE LAGUNA, juriste, DG des réseaux de communication, du contenu et des technologies

OBSERVATOIRE EUROPEEN DE L'AUDIOVISUEL - Mme Maja CAPPELLO, Chef du service des Informations juridiques (18.09.2017)

UER - UNION EUROPEENNE DE RADIO-TELEVISION – M. Michael WAGNER, Chef du droit des médias et de la communication, Service des Affaires Juridiques

SOCIETE CIVILE, COMMUNAUTES ACADEMIQUES ET SECTEUR PRIVE

Mme Christina ANGELOPOULOS, Institut de recherches sur le droit de la propriété intellectuelle et l'accès à l'information, Université de Cambridge (Royaume-Uni)

M. Allon BAR, Consultant indépendant

M. Giancarlo FROSIO - Centre d'études internationales de la propriété intellectuelle (CEIPI) - Université de Strasbourg

Mme Gabrielle GUILLEMIN, Article 19 (18.09.2017)

M. Martin HUSOVEC, Professor adjoint; Faculté de Droit, Technologie et Société, Tilburg (Pays-Bas)

Mme Catherine KENT – Université d'Essex (excusée)

Mme Aleksandra KUCZERAWY, chercheuse en droit, Centre de Droit des TI&PI, Université de Louvain, Belgique

M. Joseph McNAMEE, Directeur Exécutif, European Digital Rights (EDRi), Bruxelles, Belgique

M. Michael ROTERT, EuroISPA - European Internet Services Providers Association

ETATS NON MEMBRES

MAROC

M. El Mahdi AROUSSI IDRISSI, Directeur des affaires juridiques, Haute Autorité de la Communication Audiovisuelle (HACA)

Mme Chanaz El AKRICHI, Chef de la division de la Coopération, Ministère de la Communication

Mme Meriem KHATOURI, Directrice des études et du développement des médias, Ministère de la Communication

SECRÉTARIAT

M. Jan KLEIJSSEN, Directeur, Direction de la Société de l'information et de la lutte contre la criminalité

M. Patrick PENNINCKX, chef du service de la Société de l'information (excusé)

Mme Silvia GRUNDMANN, chef de la division médias et internet, service de la Société de l'information

Mme Charlotte ALTENHÖNER-DION, Secrétariat MSI-NET, division médias et internet, service de la Société de l'information

Mme Francesca MONTAGNA, Administratrice, division médias et internet, service de la Société de l'information

Mme MAETZ Elisabeth, Assistante, division médias et internet, service de la Société de l'information

INTERPRETERS / INTERPRETES

M. Grégoire DEVICTOR, M. PEDUSSAUD Jean-Jacques, M. Nicolas GUITTONNEAU

ANNEXE 3

PROJET FINAL[2]

PROJET DE RECOMMANDATION DU COMITE DES MINISTRES AUX ÉTATS MEMBRES SUR LES ROLES ET LES RESPONSABILITES DES INTERMEDIAIRES D'INTERNET

Préambule

1. Conformément à la jurisprudence de la Cour européenne des droits de l'homme (ci-après « la Cour »), les États membres du Conseil de l'Europe sont tenus de reconnaître à toute personne relevant de leur juridiction les droits et libertés définis dans la Convention de sauvegarde des droits de l'homme et des libertés fondamentales (STE n° 5, ci-après « la Convention »), tant en ligne qu'hors ligne. L'accès à internet est un préalable indispensable à l'exercice en ligne des droits et libertés protégés par la Convention dans un environnement numérique.

2. En améliorant la possibilité pour le public de chercher, de recevoir et de communiquer des informations sans ingérence et sans considération de frontière, internet joue un rôle particulièrement important pour la liberté d'expression. C'est en outre un élément clé permettant l'exercice d'autres droits protégés par la Convention et ses protocoles, tels que le droit à la liberté de réunion et d'association, le droit à l'éducation, l'accès à la connaissance et à la culture, ainsi que la participation au débat public et politique et à la gouvernance démocratique.

3. La protection de la vie privée et des données à caractère personnel est l'un des éléments constitutifs de la jouissance et de l'exercice de la plupart des droits et des libertés garantis dans la Convention. Cependant, internet a également contribué à une augmentation des risques d'atteinte à la vie privée et de non-respect de la vie privée, et a favorisé la propagation de certaines formes de harcèlement, de haine et d'incitation à la violence fondées en particulier sur le genre, la race et la religion, qui restent insuffisamment signalées et donnent rarement lieu à réparation ou à des poursuites. En outre, l'essor d'internet et les évolutions technologiques connexes sont à l'origine de défis considérables pour le maintien de l'ordre public et de la sécurité nationale, pour la prévention et la répression de la criminalité, ainsi que pour la protection des droits d'autrui, y inclus les droits de propriété intellectuelle.

4. Une large diversité d'acteurs, communément appelés « intermédiaires d'internet » et dont le nombre ne cesse de s'étendre, facilite les interactions entre les personnes physiques et morales sur internet en exerçant des fonctions diverses et en proposant des services divers. Certains connectent les utilisateurs à internet, assurent le traitement d'informations et de données ou hébergent des services en ligne, y compris pour du contenu généré par les utilisateurs. D'autres agrègent des informations et permettent de faire des recherches ; ils

[2] Tel que produit sous doc MSI-NET(2016)05rev4 daté du 19 septembre 2017

donnent accès à des contenus et des services conçus ou gérés par des tiers, les hébergent et les indexent. Certains facilitent la vente de biens et de services, notamment de services audiovisuels, et rendent possibles d'autres transactions commerciales, y compris les paiements.

5. Les intermédiaires sont susceptibles de remplir plusieurs fonctions en parallèle. Il arrive également qu'ils contrôlent les contenus et les classent au moyen, par exemple, de techniques de traitement automatisé des données et, partant, peuvent exercer certaines formes de contrôle qui influencent l'accès des utilisateurs aux informations en ligne, à l'instar des médias, ou encore qu'ils assurent d'autres fonctions qui se rapprochent de celles des éditeurs. Les services d'intermédiaires peuvent aussi être fournis par les médias traditionnels, par exemple, lorsque de l'espace pour les contenus générés par les utilisateurs est proposé sur leurs plateformes. Le cadre réglementaire régissant la fonction d'intermédiaire ne porte pas atteinte aux cadres applicables aux autres fonctions proposées par la même entité.

6. L'État de droit est une condition indispensable à la protection et la promotion de l'exercice des droits de l'homme en ligne ainsi qu'à une démocratie pluraliste et participative. Les États membres ont l'obligation négative de s'abstenir de violer la liberté d'expression et d'autres droits fondamentaux sur internet. Ils ont également l'obligation positive de protéger les droits de l'homme dans l'environnement numérique. Cette obligation positive de garantir l'exercice et la jouissance des droits et des libertés comprend, de par les effets horizontaux des droits de l'homme, la protection des individus contre les actes de parties privées, en s'assurant du respect des cadres légaux et réglementaires applicables. Il est par ailleurs indispensable de mettre en place des garanties procédurales et de faciliter l'accès à des recours effectifs à la fois contre les États et contre les intermédiaires au regard des services en question.

7. Il est en outre essentiel de soutenir les initiatives visant au développement des compétences en termes de compréhension des médias et de l'information pour l'accès à l'espace numérique et sa gestion. Ces mesures devraient être mises en œuvre par des moyens divers, notamment via l'éducation formelle et non formelle, afin de promouvoir la jouissance effective, dans des conditions d'égalité des droits consacrés par la Convention sans aucune distinction. Eu égard au nombre particulièrement élevé d'enfants et de jeunes utilisateurs d'internet, l'importance particulière de permettre, protéger et soutenir l'accès en toute sécurité des enfants à leurs droits dans l'environnement numérique doit être reconnue en permanence. A cet effet, un engagement continu est nécessaire afin de renforcer parmi les enfants, parents et éducateurs les compétences relatives à la manière de faire face à un environnement comprenant toute sorte d'informations et de messages et rendant accessible un contenu dégradant et potentiellement dommageable de nature sexuelle ou violente.

8. Le cadre réglementaire régissant les services assurés directement ou indirectement par les intermédiaires est divers, complexe et en constante évolution. Les États font face au défi complexe que représente la réglementation d'un environnement dans lequel des acteurs privés jouent un rôle essentiel dans la prestation de services qui ont une importante valeur de service public. La nature mondiale des réseaux et services d'internet, la diversité des intermédiaires, le volume des communications internet et la vitesse à laquelle elles sont produites et traitées compliquent encore davantage le travail de réglementation. Étant

donné que les intermédiaires opèrent dans de nombreux pays, y compris dans un environnement de l'informatique en nuage, leurs activités peuvent aussi relever simultanément de plusieurs lois, parfois en conflit, de différentes juridictions.

9. Les intermédiaires d'internet développent également leurs propres règles, généralement établies sous la forme de conditions de service ou de standards de la communauté, et incluant fréquemment des politiques en matière de restriction des contenus. De plus, les intermédiaires collectent, génèrent, conservent et traitent une quantité considérable d'informations et de données émanant des utilisateurs et les concernant. Ces activités peuvent porter atteinte à d'autres droits des utilisateurs, notamment leurs droits à la liberté d'expression et au respect de leur vie privée. Les mécanismes effectifs de contrôle et de plainte sont parfois inexistants ou insuffisamment transparents et efficaces ou encore limités à des procédés automatisés.

10. Conformément aux Principes directeurs relatifs aux entreprises et aux droits de l'homme et au cadre de référence «protéger, respecter et réparer» des Nations Unies, les intermédiaires devraient, dans toutes leurs actions, respecter les droits de l'homme des utilisateurs et des parties affectées. Cela inclut la responsabilité d'agir dans le respect des lois et des cadres réglementaires applicables. En raison du caractère multifonctionnel des intermédiaires, leurs responsabilités et devoirs correspondants ainsi que leur protection en vertu de la loi doivent être définis en fonction des services spécifiques qu'ils fournissent et des fonctions spécifiques qu'ils exercent.

11. De multiples effets de réseau et de fusions ont conduit à l'existence d'entités plus grandes et en nombre plus restreint qui dominent le marché au risque de priver de débouchés les petits intermédiaires ou les « start-ups » et qui sont en position d'influencer, voire de contrôler, les principaux modes de communication publique. Le pouvoir de ces intermédiaires en tant que protagonistes de l'expression en ligne impose de clarifier leur rôle et leur impact sur les droits de l'homme, ainsi que leurs devoirs et responsabilités correspondants.

12. Compte tenu des considérations ci-dessus et dans le but de donner des orientations à tous les acteurs concernés qui sont confrontés à la tâche complexe que représentent la protection et le respect des droits de l'homme dans l'environnement numérique, le Comité des Ministres, agissant en vertu de l'article 15.b du Statut du Conseil de l'Europe, recommande aux États membres :

- de mettre en œuvre les lignes directrices figurant dans la présente recommandation, lors de l'élaboration et de l'application de cadres législatifs concernant les intermédiaires d'internet conformément à leurs obligations découlant de la Convention, de la Convention pour la protection des personnes à l'égard du traitement automatisé des données à caractère personnel (STE n 108, ci-après dénommée « la Convention 108 »), la Convention sur la cybercriminalité (STE n° 185, « la Convention de Budapest »), la Convention du Conseil de l'Europe sur la protection des enfants contre l'exploitation et les abus sexuels (STCE n° 201, ci-après « la Convention de Lanzarote ») et la Convention sur la prévention et la lutte contre la violence à l'égard des femmes et la violence domestique (STCE n° 210, « la Convention d'Istanbul »), et de les promouvoir dans les enceintes internationales et régionales qui traitent des rôles et responsabilités des intermédiaires d'internet ;

- de prendre toutes les mesures nécessaires pour que les intermédiaires d'internet assurent leurs responsabilités en matière de respect des droits de l'homme, conformément aux Principes directeurs de l'ONU relatifs aux entreprises et aux droits de l'homme et à la Recommandation CM/Rec (2016)3 du Comité des Ministres aux États membres sur les droits de l'homme et les entreprises ;

- de prendre en compte, pour la mise en œuvre des lignes directrices, les recommandations suivantes du Comité des Ministres : la Recommandation 2016/5 sur la liberté d'internet, la Recommandation 2016/3 sur les droits de l'homme et les entreprises, la Recommandation 2016/1 sur la protection et la promotion du droit à la liberté d'expression et du droit à la vie privée en lien avec la neutralité du réseau, la Recommandation 2015/6 sur la libre circulation transfrontière des informations sur internet, la Recommandation 2014/6 sur un Guide des droits de l'homme pour les utilisateurs d'internet, la Recommandation 2013/1 sur l'égalité entre les femmes et les hommes et les medias, la Recommandation 2012/3 sur la protection des droits de l'homme dans le contexte des moteurs de recherche, la Recommandation 2012/4 sur la protection des droits de l'homme dans le cadre des services de réseaux sociaux, la Recommandation 2011/7 sur une nouvelle conception des médias, la Recommandation (2010)13 sur la protection des personnes à l'égard du traitement automatisé des données à caractère personnel dans le cadre du profilage, la Recommandation 2007/16 sur des mesures visant à promouvoir la valeur de service public de l'Internet, les Lignes directrices sur la protection des personnes à l'égard du traitement des données à caractère personnel à l'ère des mégadonnées (2017), les Lignes directrices pour la coopération entre organes de répression et fournisseurs de services internet contre la cybercriminalité (2008) et les Lignes directrices visant à aider les fournisseurs de services internet, élaborées en 2008 par le Conseil de l'Europe en coopération avec l'Association européenne des fournisseurs de services Internet, qui, s'agissant des responsabilités des fournisseurs de services internet, sont renforcées par la présente recommandation.

- de dialoguer régulièrement, de manière inclusive et transparente, avec tous les acteurs concernés, y compris du secteur privé, des médias de service public, de la société civile, des établissements scolaires et des milieux universitaires, en vue de partager et d'examiner des informations et de promouvoir l'utilisation responsable des dernières évolutions technologiques liées aux intermédiaires d'internet qui ont des répercussions sur l'exercice et la jouissance des droits de l'homme, ainsi que leurs aspects juridiques et politiques ;

- d'encourager et de promouvoir la mise en œuvre de programmes d'éducation aux médias et à l'information, efficaces et différenciés en fonction de l'âge et du genre, afin de permettre aux adultes, aux jeunes et aux enfants de bénéficier des avantages de l'environnement des communications en ligne et de réduire les risques qui y sont associés, en coopération avec les parties prenantes concernées du secteur privé, de la société civile, de l'éducation et des milieux universitaires et techniques.

Lignes directrices sur la protection et la promotion des droits de l'homme et des libertés fondamentales en ce qui concerne les intermédiaires d'internet

I – Devoirs et obligations des États

1.1 LÉGALITÉ

1.1.1. Toute requête, demande ou autre action des autorités publiques adressée à des intermédiaires d'internet qui constitue une ingérence dans l'exercice des droits de l'homme et des libertés fondamentales doit être prévue par la loi et constituer une mesure nécessaire et proportionnée dans une société démocratique. Les pouvoirs des autorités publiques à l'égard des intermédiaires d'internet doivent être clairement définis par la loi et exercés dans les limites légalement fixées. Les États ne devraient pas avoir recours à des moyens informels pour contourner les garanties offertes par les procédures judiciaires.

1.1.2. Indépendamment de leur objectif et de leur champ d'application, étendus ou non aux activités commerciales et non commerciales, les lois, règlements et politiques applicables aux intermédiaires d'internet doivent garantir la protection effective des droits de l'homme et des libertés fondamentales, et maintenir des garanties suffisantes contre une application arbitraire en pratique.

1.1.3. Les États ne doivent pas chercher à se décharger de leur obligation fondamentale de garantir le respect des droits de l'homme et des libertés fondamentales dans l'environnement numérique. Tous les cadres réglementaires, y compris les modes d'autoréglementation ou de corégulation, doivent prévoir des mécanismes de surveillance efficaces pour être conformes à cette obligation et doivent être assortis de possibilités de recours satisfaisantes.

1.1.4. La procédure aboutissant à des dispositions législatives ou réglementaires applicables aux intermédiaires d'internet devrait être transparente et inclusive. Les États devraient consulter régulièrement toutes les parties prenantes concernées pour s'assurer qu'un équilibre approprié est garanti entre l'intérêt général, les intérêts des utilisateurs et des parties concernées, ainsi que l'intérêt de l'intermédiaire. Avant l'adoption d'une loi ou d'un texte réglementaire, les États devraient réaliser des études d'impact du point de vue des droits de l'homme pour en évaluer les effets négatifs potentiels sur les droits de l'homme afin de les prévenir et de les réduire.

1.1.5. Les États doivent veiller à ce que les dispositions législatives et réglementaires ainsi que les politiques relatives aux intermédiaires d'internet soient interprétées, appliquées et mises en œuvre sans aucune distinction, y compris sans forme de discrimination multiple ou croisée. L'interdiction des discriminations pourra amener dans certains cas les intermédiaires à devoir prendre des dispositions spéciales afin de répondre à des besoins spécifiques ou de corriger des inégalités existantes. Dans

l'élaboration, l'interprétation et l'application du cadre législatif, les États devraient en outre prendre en compte les différences notables de taille, de fonction et de structure organisationnelle des intermédiaires afin d'empêcher de potentielles conséquences discriminatoires.

1.1.6. Les États devraient veiller à ce que les dispositions législatives et réglementaires ainsi que les politiques relatives aux intermédiaires d'internet soient effectivement applicables et exécutables, et qu'elles ne restreignent pas indûment le fonctionnement et la circulation des communications transfrontières.

1.2. SÉCURITÉ JURIDIQUE ET TRANSPARENCE

1.2.1. Toute législation applicable aux intermédiaires d'internet et à leurs relations avec les États et les utilisateurs doit être accessible et prévisible. Toutes les lois devraient être claires et suffisamment précises pour permettre aux intermédiaires, aux utilisateurs et aux parties concernées de régler leur conduite en conséquence. La législation devrait créer un environnement en ligne sûr, qui soit propice aux communications privées et au débat public et devrait être conforme aux normes internationales pertinentes.

1.2.2. Toute législation doit limiter clairement les pouvoirs, discrétionnaires ou non, accordés aux autorités publiques à l'égard des intermédiaires d'internet, en particulier lorsqu'ils sont exercés par l'exécutif ou les forces de l'ordre. La loi doit en préciser la portée pour éviter toute application arbitraire.

1.2.3. Les États devraient rendre publiquement disponibles, en temps opportun et de manière régulière, des informations complètes sur le nombre, la nature et la base juridique des restrictions aux droits de l'homme, concernant par exemple une restriction des contenus ou la divulgation de données permettant d'identifier des personnes, qu'ils ont appliquées dans des périodes données par le moyen de demandes adressées à des intermédiaires, notamment celles fondées sur des traités internationaux d'entraide judiciaire et sur des mesures prises à la suite de ces demandes. Les États devraient demander aux intermédiaires de divulguer des données claires (dans un format facilement accessible et lisible par un ordinateur) et utiles sur les ingérences dans l'exercice des droits et libertés dans l'environnement numérique, que ces ingérences soient la conséquence d'ordonnances judiciaires ou administratives, de demandes de plaignants ou de l'application de leurs propres politiques de contrôle des contenus.

1.2.4. En vue d'éviter l'insécurité juridique et les conflits de lois, les États devraient s'engager à coopérer entre eux et avec tous les acteurs concernés dans les situations où des lois différentes s'appliquent ; ils devraient soutenir le développement d'approches et de principes d'attribution de compétences communs, notamment par le biais de structures non étatiques appropriées.

1.3. PROTECTION DE LA LIBERTÉ D'EXPRESSION

1.3.1. Toute requête, demande ou autre action des autorités publiques adressée à des intermédiaires d'internet pour restreindre un accès (y compris le blocage ou la suppression de contenus), ou toute autre mesure qui entraînerait une ingérence dans l'exercice de la liberté d'expression, doit être prévue par la loi, poursuivre l'un des buts légitimes énoncés à l'article 10 de la Convention, être nécessaire dans une société démocratique et proportionnée au but poursuivi. Les autorités doivent évaluer soigneusement les répercussions potentielles, y compris non intentionnelles, de toute restriction avant d'y avoir recours et après les avoir appliquées, tout en cherchant à appliquer la mesure la moins restrictive pour atteindre l'objectif visé.

1.3.2. Pour demander à un intermédiaire la restriction de l'accès à des contenus illégaux, les autorités nationales doivent chercher à obtenir une ordonnance d'une autorité judiciaire ou d'une autre instance étatique indépendante, dont les décisions font l'objet d'un contrôle juridictionnel pour exiger des intermédiaires d'internet qu'ils limitent l'accès à des contenus. Toutes les exceptions doivent être prévues par la loi, poursuivre l'un des buts légitimes visés à l'article 10.2 constituer une mesure nécessaire dans une société démocratique et être proportionnées à l'objectif poursuivi.

1.3.3. Les autorités publiques devraient veiller à ce que les intermédiaires d'internet, lorsqu'ils restreignent l'accès à des contenus de tiers, offrent des mécanismes de recours appropriés et respectent les garanties procédurales. Lorsque les intermédiaires suppriment un contenu sur le fondement de leurs propres conditions de service, les autorités publiques ne devraient pas considérer que cela constitue une forme de contrôle qui les rendrait responsables du contenu d'un tiers auquel ils ont donné accès.

1.3.4 Les autorités nationales devraient envisager d'adopter une législation appropriée pour prévenir les contentieux stratégiques contre la participation du public ou les litiges abusifs et vexatoires utilisés dans le but de restreindre le droit à la liberté d'expression des utilisateurs, des fournisseurs de contenus et des intermédiaires.

1.3.5 Les autorités nationales ne devraient pas imposer aux intermédiaires, directement ou indirectement, une obligation générale de surveiller, par un moyen automatisé ou non, les contenus auxquels ils donnent accès, ou qu'ils transmettent ou stockent. Lorsqu'une demande quelconque est adressée aux intermédiaires d'internet ou quand est encouragée, l'adoption par lesdits intermédiaires de modes de corégulation, les autorités nationales, seules ou avec d'autres États ou des organisations internationales, devraient éviter toute action susceptible d'entraîner une surveillance générale des contenus. Elles devraient, par ailleurs, prendre en considération le fait que la surveillance de contenus est généralement réalisée par des moyens automatisés incapables d'évaluer convenablement les contextes. Tous les modes d'autorégulation doivent être conformes aux principes de l'état de droit et de la transparence.

1.3.6 L'imposition de sanctions disproportionnées aux intermédiaires pour non-respect des cadres réglementaires peut entrainer une restriction excessive de tout contenu légal. Cela a un effet dissuasif sur le droit à la liberté d'expression. La surveillance de contenus risque également de porter atteinte à la jouissance par les utilisateurs de leur droit au respect de la vie privée.

1.3.7. Les États devraient veiller en droit et en pratique à ce que les intermédiaires ne puissent être tenus responsables des contenus de tiers auxquels ils donnent simplement accès, ou qu'ils se contentent de transmettre ou de stocker. Les autorités nationales peuvent tenir les intermédiaires pour coresponsables des contenus qu'ils stockent si ceux-ci n'agissent pas avec la diligence voulue pour restreindre l'accès au contenu ou aux services dès qu'ils ont connaissance de leur caractère illégal, notamment par le biais de procédures reposant sur la notification. Les autorités nationales devraient veiller à ce que les procédures de notification et retrait ne soient pas conçues de telle manière qu'elles incitent les intermédiaires à retirer des contenus légaux, en raison par exemple de délais trop courts. Les notifications devraient contenir suffisamment d'informations pour permettre aux intermédiaires de prendre des mesures. Les notifications soumises par les États devraient reposer sur leur propre évaluation du caractère illégal du contenu signalé. Toute procédure de restriction d'accès à un contenu devrait pouvoir être notifiée dès que possible à son producteur/son émetteur, sauf si cela perturbe des activités en cours des services de répression. Les informations devraient également être accessibles aux utilisateurs qui souhaitent accéder au contenu, conformément aux lois sur la protection des données.

1.3.8. Afin d'empêcher efficacement le nouvel accès à un contenu identique à ce qui a auparavant été identifié comme illégal par une autorité judiciaire ou une autre instance administrative indépendante dont les décisions font l'objet d'un contrôle juridictionnel, les États devraient coopérer étroitement avec les intermédiaires pour assurer la restriction de tels contenus, conformément aux principes de légalité, de nécessité et de proportionnalité. De telles restrictions ne devraient pas empêcher l'utilisation légitime de contenus identiques ou similaires dans d'autres contextes.

1.3.9. Lorsque la fonction des intermédiaires consiste à produire ou gérer des contenus disponibles sur leurs plateformes, ou encore lorsque des intermédiaires exercent des fonctions de conservation ou d'édition, en appliquant notamment des algorithmes, les autorités nationales devraient appliquer l'approche graduelle et différenciée décrite dans la Recommandation CM/Rec(2011)7 du Comité des Ministres aux États membres sur une nouvelle conception des médias. Les États devraient déterminer des niveaux de protection adéquats, ainsi que les devoirs et les responsabilités découlant du rôle que jouent les intermédiaires dans la production et la diffusion de contenus, tout en portant une attention particulière à leur obligation de protéger et de promouvoir le pluralisme et la diversité dans la diffusion en ligne des contenus.

1.3.10. Lors de la détermination des devoirs et responsabilités relevant des intermédiaires qui remplissent des fonctions de conservation ou d'édition, dont la production et la diffusion de contenus, les États devraient encourager l'adoption de mesures

d'autorégulation appropriées ou le développement de mécanismes de corégulation, en tenant dûment compte de la mesure dans laquelle leur action peut porter atteinte au pluralisme et à la diversité des contenus en ligne et affaiblir la capacité des intermédiaires à assurer des prestations à valeur de service public, comme des plateformes propices au discours public et au débat démocratique tels que protégés par l'article 10 de la Convention.

1.4. GARANTIES EN MATIÈRE DE PROTECTION DE LA VIE PRIVÉE ET DE PROTECTION DES DONNÉES À CARACTÈRE PERSONNEL

1.4.1. Toute demande ou requête adressée par les autorités nationales/étatiques à des intermédiaires d'internet sollicitant l'accès à des informations à caractère personnel relatives à leurs utilisateurs, y compris à des fins de justice pénale, la collecte ou l'interception de telles informations, ou toute autre mesure qui entraînerait une ingérence dans l'exercice du droit au respect de la vie privée, doit être prévue par un texte de loi, poursuivre l'un des buts légitimes énoncés à l'article 8 de la Convention et à l'article 9 de la Convention 108, et ne peut être prise que si elle est nécessaire et proportionnée au but poursuivi dans une société démocratique. La garantie du droit au respect de la vie privée et à la protection des données à caractère personnel couvre également les dispositifs utilisés pour accéder à l'internet ou pour conserver des données.

1.4.2. Les autorités nationales devraient s'assurer que leur cadre réglementaire et les politiques et pratiques mises en place sur cette base par les intermédiaires soient conformes aux principes régissant le traitement des données (légalité, équité et transparence, limitation de la finalité, minimisation des données, exactitude, durée limitée de conservation et sécurité des données, notamment intégrité et confidentialité) et protègent les droits de la personne concernée dans le plein respect de la Convention 108, qui prévoit également une supervision exercée par une autorité indépendante au sens de l'article 1 du Protocole additionnel concernant les autorités de contrôle et les flux transfrontières de données.

1.4.3. Les autorités nationales devraient protéger le droit à la confidentialité de toutes les communications privées facilitées par les intermédiaires d'internet et étendre la protection au contenu de la communication et aux métadonnées; elles devraient veiller à ce que des niveaux appropriés de protection des données et de respect de la vie privée soient garantis aussi dans les situations de flux transfrontières de données.

1.4.4. Les mesures de surveillance mises en place par les États, en coopération ou non avec les intermédiaires d'internet, doivent être ciblées, définies de manière précise et conforme aux articles 8 de la Convention et 9 de la Convention 108. Elles doivent en particulier être prescrites par la loi, nécessaires dans une société démocratique et proportionnées au but poursuivi, comporter des garanties de procédure et de contrôle suffisantes et des dispositifs de recours. Toute surveillance doit être autorisée par une autorité judiciaire ou une autre instance étatique indépendante, dont les décisions font l'objet d'un contrôle juridictionnel.

1.4.5. Les autorités étatiques devraient veiller à l'application de garanties complémentaires appropriées, comme le consentement exprès de la personne concernée, lors du traitement automatique de catégories spéciales de données définies à l'article 6 de la Convention 108.

1.5. ACCÈS À UN RECOURS EFFECTIF

1.5.1. Les États devraient garantir des procédures judiciaires et non judiciaires faciles d'accès et efficaces qui assurent un examen impartial de toutes les allégations de violation en ligne des droits protégés par la Convention dans l'environnement numérique, comme le droit à la liberté d'expression, le droit au respect de la vie privée ou le droit de ne pas faire l'objet de discrimination, conformément à l'article 6 de la Convention.

1.5.2. Les États devraient garantir que toute violation des droits de l'homme et des libertés fondamentales de la part des intermédiaires d'internet puisse faire l'objet d'un recours effectif, conformément à l'article 13 de la Convention. Ils devraient en outre veiller à ce que les intermédiaires donnent accès rapidement à un examen transparent et efficace des plaintes formulées par les utilisateurs ou par les parties affectées et des allégations de non-respect des conditions de service, et offrent des voies de recours effectives pouvant être de différentes formes, comme la restauration du contenu, un démenti, une rectification et un dédommagement. Un contrôle judiciaire doit être possible lorsque les mécanismes internes de règlement des litiges et autres systèmes alternatifs s'avèrent insuffisants ou lorsque les parties concernées choisissent cette voie ou font appel.

1.5.3. Les États devraient prendre l'initiative de chercher à éliminer tous les obstacles juridiques, pratiques et autres qui pourraient conduire à priver les utilisateurs, les parties affectées et les intermédiaires d'internet d'un accès à un recours effectif.

1.5.4. Les États devraient soutenir les initiatives de promotion de l'éducation aux médias et à l'information qui soient différenciées en fonction de l'âge et du genre, pour faire en sorte que tous les citoyens aient effectivement connaissance de leurs droits et libertés, en particulier pour ce qui est de leur droit d'accès à un recours effectif contre les autorités nationales et les intermédiaires d'internet. La promotion des compétences en termes de compréhension des médias et de l'information devrait inclure l'éducation aux droits de toutes les parties prenantes, y compris les autres utilisateurs et parties concernées.

II - Responsabilités des intermédiaires d'internet en matière de droits de l'homme et de libertés fondamentales

2.1. RESPECT DES DROITS DE L'HOMME ET DES LIBERTÉS FONDAMENTALES

2.1.1. Dans toutes leurs actions, les intermédiaires d'internet devraient respecter les droits de l'homme et les libertés fondamentales qui sont reconnus internationalement à leurs utilisateurs et aux autres parties concernées par leurs activités. Cette responsabilité, conforme aux Principes directeurs relatifs aux entreprises et aux droits de l'homme de l'ONU, existe indépendamment de la capacité ou de la volonté des États de satisfaire à leurs propres obligations en matière de droits de l'homme.

2.1.2. La responsabilité qui incombe aux intermédiaires de respecter les droits de l'homme et d'utiliser des mesures adaptées vaut quels que soient leur taille, leur secteur d'intervention, leur contexte opérationnel, leur régime de propriété ou encore leur nature. L'ampleur et la complexité des moyens qu'ils mettent en œuvre pour assumer cette responsabilité peuvent cependant varier, compte tenu de la gravité potentielle de l'incidence sur les droits de l'homme des prestations assurées par l'intermédiaire. Plus l'impact et les dommages potentiels causés aux objets de la protection juridique et la valeur de service public sont importants pour l'exercice des droits de l'homme, plus l'intermédiaire doit prendre de précautions dans le cadre de l'élaboration et de l'application de politiques, de standards de la communauté et de codes d'éthique visant notamment à prévenir la diffusion d'abus de langage et d'images, du discours de haine et de l'incitation à la violence.

2.1.3. Toute ingérence des intermédiaires dans les communications et les échanges libres et gratuits d'informations et de données doit reposer sur une politique claire et transparente et être limitée à des buts légitimes spécifiques, par exemple empêcher l'accès à des contenus considérés comme illégaux par une autorité judiciaire ou par une autre instance étatique indépendante, dont les décisions font l'objet d'un contrôle juridictionnel ou conformément à leurs propres politiques de contrôle des contenus ou codes d'éthique.

2.1.4 Les intermédiaires d'internet devraient contrôler régulièrement et avec diligence qu'ils s'acquittent bien de leur responsabilité de respecter les droits de l'homme et les libertés fondamentales et se conforment aux obligations qui leur incombent. Cela implique notamment d'évaluer les incidences directes et indirectes, du point de vue des droits de l'homme, de leurs politiques, produits et services envisagés, tant sur les utilisateurs que sur les personnes qui en sont affectées, et de donner à ces évaluations le suivi qu'elles appellent par des mesures fondées sur les constatations ainsi relevées et en s'attachant à vérifier et jauger l'efficacité des réponses déterminées. Les intermédiaires devraient mener ces évaluations de la manière la plus ouverte possible et encourager les utilisateurs à y participer activement. Dans toutes leurs actions, ils devraient avoir conscience de l'importante valeur de service public des prestations qu'ils assurent et s'efforcer d'éviter ou d'atténuer tout effet négatif sur l'exercice effectif des droits des utilisateurs ou des parties affectées.

2.1.5. Les intermédiaires d'internet devraient s'efforcer de fournir leurs produits et services sans discriminations. Ils devraient veiller à ce que leurs actions n'aient pas de conséquences discriminatoires directes ou indirectes pour leurs utilisateurs ou d'autres parties concernées, notamment ceux qui ont des besoins particuliers ou un handicap ou qui pourraient présenter des désavantages structurels dans leur accès aux droits de l'homme. Les intermédiaires devraient en outre prendre des mesures raisonnables et proportionnées pour s'assurer que leurs conditions de service, normes collectives et codes d'éthique soient appliqués et mis en œuvre de manière cohérente, et conforme aux garanties procédurales applicables. L'interdiction de toute discrimination pourra amener les intermédiaires, dans certaines circonstances, à prendre des dispositions spéciales à l'égard de certains utilisateurs ou groupes d'utilisateurs de façon à corriger les inégalités existantes.

2.2. TRANSPARENCE ET RESPONSABILITÉ

2.2.1. Les intermédiaires d'internet devraient s'assurer que tous les accords relatifs aux conditions de service et les politiques qui précisent les droits des utilisateurs, ainsi que toutes les autres normes et pratiques concernant la modération des contenus, le traitement et la divulgation de données relatives aux utilisateurs soient rédigés en des termes simples et clairs et mis à la disposition du public dans des formats accessibles. Les intermédiaires qui opèrent dans le monde entier devraient traduire ces documents dans les langues que comprennent leurs utilisateurs et les parties concernées. Les utilisateurs devraient être avertis au préalable de toutes les modifications apportées aux politiques concernant leurs conditions de service et de fonctionnement, sans délai et dans des formats aisément accessibles et compréhensibles, y compris les guides explicatifs.

2.2.2. L'élaboration et l'application des accords relatifs aux conditions de service, des standards de la communauté ainsi que des politiques en matière de restriction des contenus devraient se faire de manière transparente, inclusive et avec l'obligation d'en rendre compte. Les intermédiaires devraient chercher à collaborer et négocier avec les associations de consommateurs, les défenseurs des droits de l'homme et autres organismes de défense des intérêts des utilisateurset des parties affectées, ainsi que les autorités chargées de la protection des données, avant d'adopter et de modifier leurs politiques. Les intermédiaires devraient s'efforcer de donner à leurs utilisateurs les moyens d'apprécier, vérifier et réviser, le cas échéant, leurs politiques et pratiques.

2.2.3. Les intermédiaires d'internet devraient clairement et de manière transparente fournir des informations publiques utiles sur la façon dont ils exploitent, dans l'exercice de leurs fonctions, les techniques de traitement automatisé des données, notamment les algorithmes facilitant les recherches fondées sur les profils des utilisateurs ou sur la diffusion de contenus personnalisés qui seraient sélectionnés à l'aide d'algorithmes. Ces informations devraient inclure des informations sur la manière dont les données sont traitées, les critères utilisés et la finalité du traitement des données.

2.2.4. Les intermédiaires devraient régulièrement publier des rapports de transparence rendant compte, à travers des informations claires (facilement accessibles et lisibles par un ordinateur) et utiles, de toute ingérence et demande d'ingérence dans les communications et échanges libres et gratuits des informations et des idées, ainsi que de toute demande liée à l'accès aux données et à leur conservation, qu'elle résulte d'une décision de justice, de traités internationaux d'entraide juridique, d'une requête formée par un plaignant à titre privé ou de la mise en œuvre de leurs propres politiques en matière de restriction de contenus.

2.3. MODÉRATION DES CONTENUS

2.3.1. Les intermédiaires d'internet devraient respecter les droits des utilisateurs de recevoir et de communiquer des informations, des opinions et des idées. Ils ne devraient pas, de manière générale, effectuer de contrôle des contenus auxquels ils donnent simplement accès, qu'ils transmettent ou stockent, sur ordre de l'État ou à sa demande. Toute mesure prise pour restreindre l'accès à un contenu, le supprimer ou le bloquer sur ordre ou requête d'un État devrait être nécessaire et exécutée par les moyens les moins contraignants, après une évaluation minutieuse de leur efficacité et de leur proportionnalité par rapport au but légitime poursuivi.

2.3.2. Lorsqu'ils sont amenés à restreindre l'accès à des contenus conformément à leurs politiques en la matière, les intermédiaires devraient le faire de façon transparente, sans aucune discrimination. Toutes les restrictions d'accès à un contenu doivent être exécutées par les moyens techniques les moins contraignants, et être limitées, dans leur ampleur et leur durée, à ce qui est strictement nécessaire pour éviter toute restriction ou suppression collatérale de contenu légal.

2.3.3 Toute restriction de contenu devrait avoir une portée limitée à l'objet précis de l'ordre ou de la demande et être accompagnée d'une information au public expliquant quel contenu a fait l'objet d'une restriction et quel en est le motif juridique. L'utilisateur et le cas échéant les autres parties concernées devraient être informés, notamment des garanties procédurales, des possibilités de procédures contradictoires pour les deux parties le cas écheant, ainsi que des mécanismes de recours disponibles.

2.3.4. , Les intermédiaires devraient adéquatement former initialement et en continu tous les membres du personnel participant à la modération des contenus aux lois applicables, aux normes internationales des droits de l'homme applicables, à leur lien avec les conditions de service et les normes internes des intermédiaires, ainsi qu'aux mesures à prendre en cas de conflit. Une telle formation peut être fournie de manière interne ou externe, y compris au moyen d'associations d'intermédiaires, et devrait dans son champ d'application correspondre à l'importance du rôle des intermédiaires et des implications que leurs actions sont susceptibles d'avoir sur l'exercice par les utilisateurs de leur droit à la liberté d'expression. Le personnel devrait par ailleurs bénéficier de conditions de travail satisfaisantes. Cela inclut l'allocation d'un temps suffisant pour décider de la légalité des contenus et des possibilités de bénéficier d'un soutien professionnel et de conseils juridiques qualifiés en cas de besoin.

2.3.5. Vu que les moyens automatisés d'identification de contenu utilisés pour empêcher la réapparition d'éléments spécifiques de contenus restreints ont une capacité limitée à évaluer le contexte, les intermédiaires devraient mesurer minutieusement les incidences qu'une gestion automatisée des contenus peut avoir sur le plan des droits de l'homme et procéder à un contrôle humain si nécessaire. Ils devraient tenir compte de la capacité limitée des algorithmes à évaluer le contexte, du risque qui en découle de procéder à des blocages insuffisants ou excessifs, et de ses effets potentiels sur les services fournis pour alimenter le débat public. Les restrictions d'accès à un contenu identique ne devraient pas empêcher l'utilisation légitime d'un tel contenu dans un autre contexte.

2.3.6. Dans le cas où, conformément à leurs propres politiques de restriction des contenus, des intermédiaires limitent le contenu au motif qu'il contient une indication d'une infraction grave au regard du droit international, la restriction doit être accompagnée de mesures appropriées pour que des preuves soient conservées dans un objectif d'efficacité des enquêtes pénales. Si les intermédiaires ont spécifiquement connaissance d'un tel contenu restreint, ils devraient en informer les autorités de répression sans retard injustifié.

2.4. UTILISATION DES DONNÉES À CARACTÈRE PERSONNEL

2.4.1. Les intermédiaires ne devraient pas divulguer de données à caractère personnel, sauf si la loi l'exige ou sur demande d'une autorité judiciaire ou d'une autre instance étatique indépendante dont les décisions font l'objet d'un contrôle juridictionnel, ayant établi que leur divulgation est conforme à la législation et aux normes applicables, nécessaire dans une société démocratique et proportionnée au but légitime poursuivi.

2.4.2. Les intermédiaires d'internet devraient limiter le traitement de données à caractère personnel des utilisateurs aux informations qui leur sont directement nécessaires dans le cadre d'un objectif clairement défini et expressément communiqué à tous les utilisateurs de manière proactive. Le traitement, notamment la collecte, la conservation, la compilation, la mise en relation ou le partage de données à caractère personnel, doit obéir à un intérêt légitime et suppose le consentement libre, spécifique, éclairé et sans équivoque de l'utilisateur sur l'objectif spécifique poursuivi, ou selon un autre fondement légitime prévu par la loi, conformément à la Convention 108. Des garanties supplémentaires, comme le consentement explicite, devraient être appliquées au traitement automatique des catégories de données particulières, conformément à l'article 6 de la Convention 108.

2.4.3. Les intermédiaires d'internet devraient réduire autant que possible le traitement des données à caractère personnel, compte tenu des finalités pour lesquelles elles sont traitées. Les principes de « respect de la vie privée par défaut » et de « respect de la vie privée dès la conception » devraient être appliqués à toutes les étapes afin de prévenir ou de minimiser tout risque d'interférence avec les droits et les libertés fondamentales des utilisateurs. La compilation des données des utilisateurs et leur migration au travers de services ou de dispositifs multiples ne devraient intervenir qu'après obtention du consentement libre, spécifique, éclairé et

sans équivoque des intéressés. Les utilisateurs devraient avoir la possibilité d'utiliser un service sans consentir à la compilation de leurs données.

2.4.4. Les utilisateurs ont le droit d'accès à leurs données à caractère personnel et d'en obtenir la rectification, et devraient en être informés dans un langage clair et simple. Ils devraient en outre être explicitement avertis des conditions dans lesquelles ils peuvent exercer leur droit de supprimer des données à caractère personnel et de s'opposer à leur traitement, ainsi que de leur droit de retirer leur consentement, auquel cas il conviendrait de mettre fin à tout traitement de données personnelles reposant sur ce consentement.

2.4.5. Les intermédiaires devraient opérer conformément aux conditions et garanties juridiques, quel que soit l'endroit où les données ont été collectées et y compris du point de vue des flux transfrontières de données.

2.4.6. Toute action de suivi ou de profilage des utilisateurs menée par des intermédiaires devrait être totalement transparente pour les utilisateurs. Afin de protéger l'identité de leurs utilisateurs, les intermédiaires d'internet ne devraient pas employer de techniques de profilage ou de suivi numérique qui portent atteinte à l'exercice de leurs droits de l'homme. Les intermédiaires devraient s'efforcer de protéger leurs utilisateurs contre de telles pratiques par des tiers. Du personnel dûment formé devrait superviser toutes les questions liées à la divulgation de données des utilisateurs à des tiers, conformément aux responsabilités et obligations des intermédiaires en vertu des normes internationales de protection des données à caractère personnel et de respect de la vie privée. Toute personne faisant l'objet d'une décision fondée sur le profilage ou affectée par les répercussions juridiques découlant de cette décision devrait pouvoir s'opposer à cette décision.

2.5. ACCÈS À UN RECOURS EFFECTIF

2.5.1. Les intermédiaires d'internet devraient mettre en place – en ligne et hors ligne – des voies de recours et des systèmes de règlement des litiges efficaces qui offrent aux utilisateurs, aux fournisseurs de contenus et aux parties concernées la possibilité d'un recours rapide et direct en cas de grief. Si les mécanismes de plainte et leurs procédures de mise en œuvre peuvent varier selon la taille, l'impact et le rôle de l'intermédiaire d'internet, toutes les voies de recours doivent permettre un examen impartial et indépendant des allégations de violation. En fonction de l'infraction, celles-ci devraient entraîner une enquête, des explications, une réponse, une rectification, des excuses, le rétablissement d'un statut, le rétablissement d'une connexion ou une réparation.

2.5.2. Tous les mécanismes de plaintes, y compris les procédures reposant sur la notification, devraient être assortis de garanties procédurales et être accessibles, équitables, compatibles avec les droits, d'un coût abordable et transparents. Ils devraient, en outre, comporter des garanties intégrées (par exemple, une structure de contrôle) en vue d'éviter les conflits d'intérêt lorsque l'entreprise gère directement le mécanisme. Ces mécanismes de plaintes devraient être conduits sans retards injustifiés et ne devraient pas avoir d'incidence négative sur la

possibilité pour les plaignants d'exercer des recours auprès de dispositifs de contrôle nationaux, y compris judiciaires.

2.5.3. Les intermédiaires devraient veiller à ce que tous les utilisateurs ainsi que les autres parties concernées par leurs actions puissent avoir pleinement et aisément accès à des informations transparentes, dans une langue claire et aisément compréhensible, relatives aux mécanismes en vigueur pour la réception et le traitement des plaintes, aux différentes phases de la procédure, à un calendrier indicatif et aux résultats attendus.

2.5.4. Les intermédiaires ne devraient pas prévoir dans leurs conditions de service de possibilité de renonciation aux droits ni de règles entravant l'accès effectif à des voies de recours, telles que l'attribution impérative de compétence dans un pays autre que celui de résidence de l'utilisateur, ou encore des clauses obligatoires de recours à l'arbitrage.

2.5.5. Les intermédiaires devraient chercher à donner accès à des dispositifs de contrôle alternatifs qui puissent faciliter le règlement des litiges pouvant opposer des utilisateurs. Ils ne devraient toutefois pas les rendre obligatoires pour en faire les seuls moyens de règlement des litiges.

2.5.6. Les intermédiaires devraient engager un dialogue avec les associations de consommateurs, les défenseurs des droits de l'homme et autres organismes de défense des intérêts des utilisateurs et des parties concernées, ainsi que les autorités de protection des données personnelles, afin de s'assurer que la conception, la mise en œuvre et l'évaluation de leurs mécanismes de réception et traitement des plaintes reposent sur un processus participatif. Ils devraient, en outre, analyser régulièrement la fréquence, les profils et les causes des plaintes reçues et en tirer les enseignements afin d'améliorer leurs politiques, procédures et pratiques, et d'en éviter la répétition.

2.5.7. Les intermédiaires devraient s'engager dans, et promouvoir, des initiatives prenant en compte les facteurs d'âge et de genre, pour développer chez tous les utilisateurs la connaissance de leurs droits et libertés dans un environnement numérique, vis-à-vis des États comme des intermédiaires, et en particulier l'information sur les dispositifs et procédures de plaintes en vigueur. La promotion des compétences en termes de compréhension des médias et de l'information devrait inclure l'éducation aux droits de toutes les parties prenantes, y compris les autres utilisateurs et les parties affectées.

ANNEXE 4

MSI-NET(2016)06rev3

PROJET FINAL[3]

PROJET D'ÉTUDE SUR LES DIMENSIONS DES DROITS HUMAINS DANS LES TECHNIQUES DE TRAITEMENT AUTOMATISÉ DES DONNÉES (EN PARTICULIER LES ALGORITHMES) ET ÉVENTUELLES IMPLICATIONS RÉGLEMENTAIRES

FINALISÉ LE 6 OCTOBRE 2017

[3] Tel que produit sous doc MSI-NET(2016)06rev3 daté du 6 octobre 2017

Le Comité directeur sur les médias et la société de l'information (CDMSI), dans le cadre de son mandat biannuel 2016 – 2017, a eu pour tâche de préparer une « étude sur les dimensions des droits humains dans l'application des techniques de traitement des données informatiques (en particulier les algorithmes) et leurs implications éventuelles sur le plan réglementaire » par le biais du Comité d'experts sur les intermédiaires d'internet (MSI-NET), structure subordonnée nommée par le Comité des Ministres pour faciliter le travail du CDMSI. Lors de sa première réunion, le 17-18 March 2016, le comité d'experts MSI-NET a décidé de nommer Benjamin Wagner rapporteur de l'étude. D'autres membres du MSI-NET ont exprimé le souhait d'apporter leur soutien au rapporteur dans le cadre d'un petit groupe de travail.

COMPOSITION DU MSI-NET

Wolfgang SCHULZ, Professeur, Faculté de droit, Université de Hambourg / Institut de Hans-Bredow (président)

Karmen TURK, Trinity Tallinn – Estonie (vice-présidente)

Bertrand De la CHAPELLE, Co-fondateur et Directeur de « Internet & Jurisdiction », France

Julia HÖRNLE, Professeur des lois dans le domaine d'Internet, Queen Mary University of London

Tanja KERŠEVAN-SMOKVINA, Conseillère principale auprès du directeur général - Agence pour les réseaux et services de communication – Slovénie (Rapporteur pour l'égalité de genre)

Matthias KETTEMANN, Postdoc Fellow, Cluster of Excellence "Normative Orders" Université de Francfort-sur-le-Main (Rapporteur Recommandation)

Arseny NEDYAK, Directeur adjoint, Service des politiques nationales des médias, Ministère de la télécommunication – Fédération de Russie

Dörte NIELANDT, division VI A3 (Cadre juridique pour les services numériques, l'industrie des médias), Ministère Fédéral de l'Economie et de l'Energie – Allemagne

Pēteris PODVINSKIS, Ministère des affaires étrangères, Direction Organisations Internationales, Service des Politiques publiques dans le domaine de l'Internet – Lettonie

Thomas SCHNEIDER, Directeur adjoint des affaires internationales, Coordinateur de la société d'information internationale, Service fédéral de l'environnement, transport, énergie et communication DETEC, Office fédéral des communications (OFCOM) – Suisse

Sophie STALLA-BOURDILLON, Professeur agrégée en technologie d'information / droit de la propriété intellectuelle, Directrice de ILAWS, Faculté de droit de Southampton, Université de Southampton

Dirk VOORHOOF, Professeur de droit européen des media, UCPH (Université de Copenhague) / Professeur à l'université de Gand / membre du comité scientifique du CMPF (Centre pour le pluralisme des médias et la liberté de la presse)

Benjamin WAGNER – Professeur assistant, Institute for Management Information Systems, Vienna University of Economics and Business / (Rapporteur Étude)

TABLE DES MATIÈRES

1. Introduction

Quelles informations sont mises à la disposition des utilisateurs de Facebook sur leurs fils d'actualité ? Sur quels critères les profils de risque d'une personne reposent-ils ? Quels profils offrent les meilleures chances d'obtenir une assurance santé ou un emploi ? Lesquels risquent au contraire d'être considérés comme étant ceux d'un délinquant ou d'un terroriste en puissance ? Les techniques de traitement automatisé de données, tels que les algorithmes, permettent aux internautes de rechercher des informations et d'y accéder, mais elles sont également – de plus en plus – utilisées dans les processus décisionnels, un domaine qui relevait autrefois exclusivement de la compétence des êtres humains. Les algorithmes servent aussi bien à préparer les décisions humaines qu'à les prendre immédiatement, par le biais de l'automatisation. En fait, la frontière entre une prise de décision humaine et une autre découlant d'un processus automatisé est souvent floue, d'où l'apparition de la notion de « prise de décision quasi automatisée ».

Le recours aux algorithmes pose des problèmes considérables non seulement pour les domaines d'action particuliers dans lesquels ils sont utilisés, mais aussi pour la société dans son ensemble. Comment protéger les droits fondamentaux et la dignité humaine face à des technologies en constante évolution dont les effets se font sentir aussi bien sur le droit à la vie, le droit à un procès équitable et à la présomption d'innocence que sur le droit à la vie privée et à la liberté d'expression, les droits des travailleurs, le droit à des élections libres voire sur l'État de droit ? Comment relever les défis associés à l'utilisation des « algorithmes » par le secteur public et le secteur privé, et en particulier par les plateformes internet ? C'est aujourd'hui l'une des questions suscitant les plus vifs débats.

Les hommes ayant le sentiment de ne pas contrôler et de ne pas comprendre les systèmes techniques qui les entourent, l'idée se répand de plus en plus que les « logiciels mangent le monde » (Andreessen, 2011). Bien que déconcertant, ce constat 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 ? Y a-t-il plus (ou moins) de risques de préjugés sexistes, ethniques ou raciaux dans un système automatisé ? Les inégalités sociales sont-elles seulement reproduites ou sont-elles amplifiées par les techniques de traitement automatisé des données ?

Dans le passé, les entreprises privées décidaient de développer leurs logiciels selon les cadres économique, juridique et éthique qu'elles jugeaient appropriés. Des cadres normatifs pour le développement des systèmes et des processus qui aboutissent à une prise de décision algorithmique, ou à leur mise en œuvre, commencent certes à voir le jour, mais ils en sont encore à un stade précoce et n'abordent généralement pas explicitement de problématiques liées aux droits de l'homme. En réalité, on ignore si et dans quelle mesure les concepts juridiques existants sont capables de décrire les défis éthiques posés par les algorithmes. En outre, de nombreuses technologies fondées sur des algorithmes en sont toujours à leurs balbutiements et il est nécessaire de comprendre davantage leurs

répercussions sociales ; par conséquent, la question de savoir s'il est même possible d'établir un cadre normatif relatif à l'utilisation des algorithmes ou une réglementation efficace des techniques de traitement automatisé des données reste ouverte. Les questions soulevées par l'utilisation des algorithmes dans le processus décisionnel sont multiples et complexes. Dans le même temps, le débat sur les algorithmes et leurs conséquences éventuelles sur les personnes, les groupes et les sociétés n'en est qu'au début. Cela ne doit pas empêcher de chercher à comprendre ce qu'ils font réellement, les conséquences qui en découlent pour la société et comment les éventuelles préoccupations concomitantes liées aux droits de l'homme peuvent être prises en compte.

La présente étude recense un certain nombre d'inquiétudes suscitées en matière de droits de l'homme par le rôle croissant des algorithmes dans les processus décisionnels. Leur impact sur l'exercice des droits de l'homme sera différent selon le type de fonction exécutée. Lorsque les algorithmes violent les droits de l'homme, qui est responsable ? La personne qui a programmé l'algorithme, l'opérateur qui l'utilise ou l'être humain qui a mis en œuvre une décision fondée dessus ? Existe-t-il une différence entre une telle décision et une décision prise par un humain ? Quelles sont les incidences sur l'exercice des droits de l'homme et sur 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 qu'aller croissant, 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 tous les aspects liés aux incidences des algorithmes sur les droits de l'homme. 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 d'en minimiser les effets négatifs ou de promouvoir les bonnes pratiques. Plusieurs thèmes liés nécessiteront des recherches plus détaillées 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 et l'internet des objets.

2. PORTÉE DU RAPPORT

Dans le cadre de l'évaluation des techniques de traitement automatisé des données et des algorithmes qu'elles utilisent, il est important de bien préciser de quels types d'algorithmes on parle. La présente étude s'appuie sur des définitions existantes bien établies et en particulier sur les travaux de Tarleton Gillespie (2014), de Nicholas Diakopoulos (2015) et de Frank Pasquale (2015). Un autre point à ne pas perdre de vue est l'ample et fréquente utilisation du terme « algorithme » et ses diverses significations possibles selon qu'il est utilisé au sein de la communauté des sciences informatiques, par les mathématiciens et les informaticiens, dans des études consacrées à la communication et aux médias culturels ou dans la sphère publique, y compris dans le discours politique et social. 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 sur les algorithmes sont moins axés sur les algorithmes eux-mêmes que sur le rôle de la technologie dans la société (Bucher, 2016).

L'approche adoptée pour la présente étude repose pour l'essentiel sur 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, à partir de calculs spécifiés. Les procédures désignent à la fois le problème et la démarche qui devrait suivie pour le résoudre. » (Gillespie 2014:167). Les algorithmes sont donc perçus comme « une série d'opérations réalisées afin de résoudre un problème particulier ou d'obtenir un résultat défini ». (Diakopoulos 2015:400).

Ce rapport n'examinera pas les algorithmes qui automatisent les processus de fabrication ou exécutent d'autres tâches de routine. Il semble en effet raisonnable de limiter l'examen aux algorithmes numériques et qui concernent le grand public, et donc de le centrer sur la prise de décision algorithmique ayant des répercussions sur les droits de l'homme. Sans être exhaustifs ni chercher à prévoir toutes les propriétés potentielles des algorithmes et les décisions qu'ils déclencheront à l'avenir, le présent rapport considère comme essentielles du point de vue des droits de l'homme les caractéristiques suivantes des algorithmes intervenant dans le traitement automatisé de données et la prise de décision (semi) automatisée : automatisation, analyse des données et adaptabilité. Par ailleurs, les algorithmes et les techniques de traitement de données sont produits et exploités par l'homme ; il serait donc vain de vouloir comprendre leurs implications sans reconnaître les constructions sociales qui les entourent.

1. AUTOMATISATION

L'automatisation est l'une des principales caractéristiques associées à 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. Le remplacement de l'homme par des systèmes

informatiques automatisés trouve généralement son origine et sa justification dans des problématiques anciennes telles que le traitement de données à grande échelle, la rapidité et le volume des décisions à prendre et procède, dans de nombreux cas, d'exigences pour des taux d'erreur moindres que ceux observés chez l'homme. Différents domaines utilisent des algorithmes de décision automatisée, 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. La prise de décision algorithmique automatisée est généralement difficile à prévoir pour un être humain, et sa logique sera difficile à expliquer après coup.

2. 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 jeux de données sans nécessairement établir de lien de causalité (Grindrod, 2014). Le fait qu'ils utilisent l'exploration de données et la reconnaissance de tendances sans « comprendre » leurs corrélations ou leurs liens de causalité peut conduire à des erreurs et suscite des inquiétudes quant à la qualité des données. Ces algorithmes reproduisent des fonctions auparavant exécutées par des êtres humains mais font appel à une logique décisionnelle quantitativement et qualitativement différente appliquée à des masses beaucoup plus importantes de données d'entrée.

Notons que les effets d'une prise de décision automatisée peuvent être envisagés comme une interaction entre les outils d'analyse appliquée (basés sur les algorithmes) et les ensembles de données utilisés. Il conviendrait de tenir compte de ces deux éléments pour évaluer l'incidence sur les droits de l'homme, dans la mesure où un ensemble de données peut, par exemple, contenir dès le départ une part de partialité cachée que l'analyse des algorithmes pourrait donc ne pas déceler. Au surplus, les performances de l'opérateur de l'algorithme peuvent davantage dépendre de sa connaissance de la structure des données analysées que de sa compréhension du fonctionnement exact de l'algorithme.

3. ADAPTABILITÉ

L'adaptabilité est mise en évidence dans des algorithmes d'auto-apprentissage automatique qui utilisent des données pour créer des schémas et savoirs originaux et générer des règles décisionnelles via des techniques d'apprentissage automatique (Williamson, 2016). En adoptant divers styles d'apprentissage, les algorithmes peuvent modéliser des problèmes à partir d'ensembles de données et proposer de nouvelles solutions qu'il serait impossible à un être humain d'appréhender. C'est essentiellement par des techniques d'approximations successives que les algorithmes d'apprentissage automatique détectent des tendances parmi les données existantes, identifient des tendances similaires parmi les données futures et réalisent des prédictions basées sur les données.

L'apprentissage automatique est utilisé notamment par les moteurs de recherche qui corrigent automatiquement les erreurs d'orthographe, ainsi que dans d'autres domaines plus complexes, comme la prévention des fraudes, l'analyse de risques, l'avancement des renseignements sur le comportement des consommateurs et l'amélioration du matériel médical.

La prévisibilité du résultat d'un algorithme par l'opérateur est importante lorsque sa responsabilité et la conception des structures appropriées de gouvernance sont envisagées. Les progrès des technologies de l'apprentissage profond (« deep learning ») sont susceptibles de conduire à un plus grand nombre de systèmes qui ne peuvent être compris au moyen des modèles mentaux des machines mécaniques. Il existe un important débat au sein de la communauté scientifique au sujet du degré auquel de tels systèmes peuvent être rendus intelligibles pour les êtres humains et quelles conséquences pourrait avoir une telle intelligibilité[4].

La capacité d'un opérateur à prévoir les résultats d'un algorithme peut revêtir une certaine importance, notamment pour la conception de structures de gouvernance adéquates. Les progrès des technologies d'apprentissage profond pourraient conduire à équiper davantage de systèmes d'une « intelligence artificielle » qu'il est impossible à comprendre au moyen du modèle mental utilisé par les machines mécaniques. Au sein de la communauté scientifique, la question de savoir jusqu'à quel point ces systèmes peuvent être rendus intelligibles à l'homme et quelles seraient les conséquences d'une telle intelligibilité suscite un vaste débat.

4. 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, font des erreurs différentes et peuvent produire des résultats différents et donc des conséquences différentes. Si la société et les gouvernements ont une vaste expérience et une connaissance approfondie de la prise de décision humaine et de ses échecs, ils commencent à peine à comprendre les points faibles, les limites et les frontières 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, non discriminatoires et indépendantes concernant des événements futurs. L'effervescence autour de l'exploitation des *Google Flu Trends* en 2011, qui s'est plus tard avérée injustifiée car la capacité de prédiction de ce service était bien moindre que ce qui avait été prétendu, est un exemple du débat permanent autour des affirmations

[4] Voir par exemple Yuan Stevens, 'The Promises and Perils of Artificial Intelligence: Why Human Rights and the Rule of Law Matter': https://medium.com/@ystvns/the-promises-and-perils-of-artificial-intelligence-why-human-rights-norms-and-the-rule-of-law-40c57338e806, 5 septembre, 2017.

relatives à la précision des algorithmes prédictifs (Lazer et al., 2014 ; Lazer et Kennedy, 2015). Néanmoins, ce problème tient moins aux algorithmes en tant qu'outils qu'à leur conception et à la perception et à l'interprétation humaines de leur mise en œuvre et de leurs résultats. Ainsi, pour promouvoir le respect des droits de l'homme dans l'usage des algorithmes, faut-il peut-être bien comprendre ce que ces algorithmes peuvent et ne peuvent pas faire et ne pas laisser de simples considérations d'efficacité ou d'efficience en prescrire à elles seules l'usage.

Traditionnellement, les développeurs programmaient les algorithmes à la main « afin de traiter et de transformer les données d'entrée en un produit souhaité, sur la base de calculs précis » (Gillespie, 2014). Toutefois, avec l'évolution technologique, les systèmes sociotechniques tels que les algorithmes deviennent de plus en plus opaques. Cette opacité ne découle pas tant d'une nécessité technique que d'un choix fréquent de conception conduisant à des systèmes algorithmiques dont le fonctionnement interne ne peut être ni présenté de façon transparente ni expliqué au monde extérieur. 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 3 ci-après), il sera souvent amené à 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 une distinction entre prise de décision entièrement automatisée et prise de décision semi-automatisée, dans la pratique, les frontières sont floues. Dans aucun des deux cas, l'être humain ne sera en mesure d'avancer un argument raisonné expliquant la nécessité de telle ou telle décision dans une situation considérée. 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 5 ci-après).

Il convient de noter que ce problème peut se poser non seulement aux professionnels qui développent les algorithmes, mais aussi à d'autres groupes, tels que les analystes qui les utilisent. Il a souvent été affirmé que l'application des algorithmes dans l'apprentissage automatique se fait en grande partie sans « compréhension » des relations de cause à effet (corrélation au lieu de causalité), ce qui peut entraîner biais et erreurs et susciter des craintes 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), exempts de toutes formes de considérations politiques (Denardis, 2008) est ici au cœur du problème. C'est pourquoi il serait plus utile de susciter une participation plus critique aux débats publics sur les algorithmes que d'essayer de les modifier.

Les algorithmes évoqués ici n'ont pas d'existence qui fasse sens sans interaction avec des êtres humains. Les concepts mathématiques ou informatiques n'ont pas en eux-mêmes d'impacts négatifs sur les droits de l'homme, mais ce n'est pas le cas de leur mise en œuvre et de leur application à l'interaction humaine. Les technologies, dans leur application à l'interaction humaine, sont des constructions profondément sociales (Winner, 1980, 1986) aux conséquences politiques importantes (Denardis, 2012). Si par exemple 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 confère sens et poids.

Il est par conséquent trop simple de blâmer l'algorithme ou de suggérer de ne plus avoir recours aux ordinateurs ou à l'informatique. Ce sont plutôt la construction sociale et les normes et valeurs spécifiques intégrées aux algorithmes qu'il est nécessaire de remettre en question, critiquer et contester. Ainsi, ce ne sont pas les algorithmes eux-mêmes mais les processus décisionnels y afférents qu'il convient d'analyser pour déterminer leurs éventuelles conséquences pour les droits de l'homme.

3. INCIDENCES DES ALGORITHMES SUR LES DROITS HUMAINS

Les réserves émises à l'encontre des algorithmes et des techniques de traitement automatisé des données dénoncent généralement leur opacité et leur imprévisibilité[5]. Au-delà de ces préoccupations générales, il apparaît toutefois de plus en plus clairement que des droits de l'homme spécifiques sont particulièrement concernés. Ils sont recensés ci-après avec des cas pratiques montrant comment et pourquoi l'utilisation des algorithmes peut entraîner des violations de droits humains ou en compromettre la jouissance effective.

1. DROIT À UN PROCÈS ÉQUITABLE ET GARANTIES D'UNE PROCÉDURE RÉGULIÈRE

Dans le domaine de la prévention de la criminalité et dans le système pénal, la tendance est de plus en plus à l'utilisation de techniques de traitement automatisé et d'algorithmes. De tels systèmes peuvent en effet s'avérer bénéfiques en ce qu'ils permettent d'accélérer le traitement d'énormes quantités de données ou d'évaluer plus précisément les risques des vols aériens. En outre, le recours à des techniques de traitement automatisé pour déterminer la durée des peines d'emprisonnement peut permettre des approches plus homogènes d'affaires comparables. Les préoccupations grandissantes en matière de sécurité nationale conduisent pourtant à des applications toujours plus ambitieuses des nouvelles technologies. Après la vague d'attentats terroristes aux États-Unis et en Europe, des responsables politiques ont demandé aux plateformes de réseaux sociaux en ligne d'utiliser leurs algorithmes afin de repérer les terroristes potentiels et d'agir en conséquence (Rifkind, 2014; Toor, 2016). Certaines de ces plateformes en utilisent déjà pour détecter les comptes qui mettent en ligne des contenus extrémistes. Hormis son impact significatif sur la liberté d'expression (voir point 3 ci-après), cette application des algorithmes pose également des problèmes 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, le droit à ce que sa cause soit entendue équitablement et le droit de se défendre soi-même. D'autres questions peuvent aussi se poser en relation avec l'article 5 de la Convention qui protège contre la privation arbitraire de liberté, et l'article 7 (pas de peine sans loi).

En ce qui concerne la prévention de la criminalité, les principaux débats politiques au sujet de l'utilisation des algorithmes portent sur la police prédictive. Cette approche va au-delà de la capacité des êtres humains à tirer des conclusions à partir 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

[5]Voir *Tim O'Reilly, "The great question of the 21st century: Whose black box do you trust?", 13 September 2016,* disponible au lien suivant: 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.

peines plus sévères[6]. Elle intègre aussi des systèmes destinés à prédire le lieu où un délit est susceptible d'être commis à un moment donné, systèmes qui sont ensuite utilisés pour prioriser le temps consacré par les policiers aux enquêtes et aux arrestations. De telles approches peuvent être extrêmement préjudiciables en termes d'origines ethniques et raciales, et nécessitent par conséquent un examen minutieux et des garanties adéquates. Souvent, les systèmes reposent sur des bases de données policières existantes qui contiennent intentionnellement ou non des biais systémiques[7]. En fonction des modalités d'enregistrement des délits, du choix des délits à inclure dans l'analyse et des outils d'analyse utilisés, les algorithmes prédictifs peuvent concourir à une prise de décision biaisée et à des cas de discrimination.

Il est en outre à craindre que la pratique de telles évaluations dans le contexte de la prévention de la criminalité ne crée des chambres d'écho dans lesquelles les préjugés ne font que se renforcer. 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 6). Il en découlerait une normalisation des partis pris qui seraient dès lors moins susceptibles d'être identifiés et dénoncés comme tels. 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 l'homme et du principe du droit à l'égalité des armes et à une procédure contradictoire établi par la Cour européenne des droits de l'homme[8].

On note par ailleurs une utilisation croissante des algorithmes dans la justice civile et pénale où le développement de l'intelligence artificielle vise, *in fine*, à appuyer ou remplacer les décisions de juges humains.

Aujourd'hui de tels systèmes sont testés pour identifier les résultats des décisions, afin de déceler des schémas dans les prises de décision judiciaires complexes. A ce jour, le taux de fiabilité des prévisions relativement faible (79 %). Il semble donc prématuré, à l'heure actuelle, d'imaginer que de tels systèmes puissent remplacer les juges[9]. Cela étant, il est suggéré qu'ils pourraient aider ou épauler les juges (et les avocats)[10]. Compte tenu de la pression qu'une lourde charge de travail conjuguée à l'insuffisance des ressources fait peser

---

[6] Voir également Article 19, *Algorithms and Automated Decision-Making in the Content of Crime Prevention: A Briefing paper,* 2016.

[7] Voir par exemple William Issac, Kristian Lum Kristian Lum et William Isaac (2016), « *To predict and serve? Significance* », 10 octobre 2016 The Royal Statistical Society, disponible au lien suivant: http://onlinelibrary.wiley.com/doi/10.1111/j.1740-9713.2016.00960.x/epdf.

[8] Voir, par exemple, l'affaire *Jespers c. Belgique,* 15 octobre 1980, n° 8404/78, *Salduz c. Turquie* 17 novembre 2008, n° 36391/02et *Blokhin c. Russie,* 13 avril 2016, n° 47152/06.

[9] Nikolaos Altreas et al, « *Predicting judicial decisions of the European Court of Human Rights: a Natural Language Processing perspective* », PeerJ Computer Science Open Access (publié le 24 octobre 2016), https://peerj.com/articles/cs-93.pdf at p.2 ; voir également https://www.lawgazette.co.uk/law/artificial-intelligence-mimics-judicial-reasoning/5056017.article.

[10] Ibid.

sur la plupart des corps judiciaires, on peut craindre que les systèmes d'aide à base d'intelligence artificielle soient utilisés à mauvais escient par les juges pour « déléguer » des décisions à des technologies conçues à l'origine à d'autres fins et perçues à tort comme étant plus « objectives ». Il convient donc d'évaluer avec soin ce que ces systèmes sont capables de produire et dans quelles conditions, afin de ne pas compromettre le droit à un procès équitable. Cette précaution vaut tout particulièrement lorsque le recours à de tels systèmes est obligatoire, comme aux États-Unis, dans le cas de décisions prononcées dans le cadre de libérations conditionnelles. Dans de nombreux États américains, la perspective inquiétante de voir des décisions de liberté conditionnelle entachées de parti pris a conduit à imposer l'utilisation de logiciels pour prédire la probabilité de récidive des délinquants[11]. Une enquête indépendante a toutefois conclu que le « logiciel utilisé […] pour prévoir les futurs délits » était « partial à l'égard des Noirs » (Angwin, Mattu et Kirchner, 2016).

2. VIE PRIVÉE ET PROTECTION DES DONNÉES

Le débat le plus ancien et le plus soutenu concernant les droits de l'homme dans l'application des processus de traitement automatisé de données et des algorithmes porte sur le droit au respect de la vie privée[12]. Les algorithmes facilitent la collecte, le traitement et la réutilisation de grandes quantités de données et d'images. Cela peut à ce titre avoir de graves répercussions sur la jouissance du droit au respect de la vie privée et familiale, qui inclut le droit à la protection des données, garanti par l'article 8 de la Convention européenne des droits de l'homme. Les algorithmes sont utilisés dans le suivi et le profilage en ligne des individus dont les habitudes de navigation sont enregistrées par des « cookies »[13] et des technologies similaires telles que les empreintes numériques, agrégées à des requêtes de recherche (moteurs de recherche/assistants virtuels). Les données de comportement sont en outre traitées à partir d'appareils intelligents, comme les appareils de géolocalisation et autres capteurs via les applications mobiles (Tene et Polonetsky, 2012), soulevant des défis croissants pour la vie privée et la protection des données.

Les applications de suivi et de profilage en ligne sont également utilisées à des fins de publicité ciblée basée sur le profil des intérêts présumés d'une personne. D'un point de vue réglementaire, la question du consentement de l'utilisateur est ici primordiale. Des travaux de recherche menés à Berkeley en 2012 ont établi, par exemple, que l'utilisation de technologies de pistage portant atteinte à la vie privée qui ne peuvent pas être détectées par les utilisateurs (comme les empreintes numériques et les données de comportement

[11] Voir GCN, Kevin McCaney, « Prisons turn to analytics software for parole decisions », 1er novembre 2013, disponible au lien suivant https://gcn.com/articles/2013/11/01/prison-analytics-software.aspx (consulté le 25 septembre 2017).

[12] Voir Sills 1970.

[13] Un « cookie » est une petite quantité de données générée par un site internet et sauvegardée par un navigateur internet 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, la principale 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.

générées par capteurs) a augmenté en réponse à une prise de conscience croissante des consommateurs qui prennent l'habitude de supprimer ou de désactiver les cookies dans le cadre du paramétrage *Do Not Track* (« ne pas me pister ») de leurs navigateurs internet[14]. Par ailleurs, le traitement extensif de données au moyen d'algorithmes pourrait aggraver les atteintes à d'autres droits, dans la mesure où les données personnelles sont utilisées pour cibler des individus, comme dans le contexte des applications d'assurance ou de recrutement.

L'un des enjeux spécifiques du traitement algorithmique des données personnelles est la production de nouvelles données. Lorsqu'une personne partage plusieurs données discrètes, il est souvent possible de les fusionner ; de cette fusion naissent des données de deuxième, voire de troisième génération sur l'individu en question. Deux données anodines, si on les évalue par rapport à un jeu de données nettement plus volumineux, peuvent « engendrer » et générer des « données enfants » d'une nature totalement imprévisible pour l'intéressé. Ceci soulève des questions importantes en termes de consentement, de transparence et d'autonomie personnelle. Une étude réalisée par des chercheurs des universités de Cambridge et de Stanford illustre l'ampleur de ces questions[15].

Les efforts déployés pour moderniser la Convention du Conseil de l'Europe de 1981 relative à la protection des personnes à l'égard du traitement automatisé des données à caractère personnel (Convention STE n° 108) et l'aligner sur l'évolution technologique se poursuivent aujourd'hui afin de mieux définir les droits des personnes concernées, compte tenu des conséquences sur la vie privée des outils maintenant utilisés pour la collecte, le traitement et la réutilisation de données et pour le profilage. L'article 8 du projet de convention modernisée 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 fondamentales. Les propositions de modernisation visent en outre à prévoir des garanties complémentaires en ce qui concerne la transparence (article 7bis) et la nécessité d'évaluer l'incidence potentielle du traitement de données sur les droits et les libertés fondamentales de la personne préalablement à un tel traitement (article 8bis)[16].

---

[14] CJ Hoofnagle, « *Behavioural Advertising: The Offer You Cannot Refuse* », 2012, 6 Harvard Policy & Law Review 273-296.

[15] Voir Stanford news, « New Stanford research finds computers are better judges of personality than friends and family », disponible au lien suivant : http://news.stanford.edu/2015/01/12/personality-computer-knows-011215/ (consulté le 25 septembre 2017).

[16] Voir le projet de Convention modernisée pour la protection des personnes à l'égard du traitement automatisé des données à caractère personnel, septembre 2016, disponible au lien suivant : https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=09000016806a616c

Les « Lignes directrices sur la protection des personnes à l'égard du traitement des données à caractère personnel à l'ère des mégadonnées »[17], récemment adoptées par le Comité consultatif de la convention pour la protection des personnes à l'égard du traitement automatisé des données à caractère personnel, fournissent un cadre général pour appliquer les politiques et les mesures appropriées pour continuer à rendre effectifs les principes de protection des données dans le cadre des mégadonnées.

Les cadres réglementaires de la protection des données à l'échelle de l'Union européenne tels que le Règlement général sur la protection des données d'avril 2016 (Règlement (UE) 2016/679 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données), qui s'appliquera à compter de mai 2018, fixent également des normes pour l'utilisation des algorithmes dans la collecte de données dont, éventuellement un droit limité à l'information, voire un « droit à l'explication » (Goodman et Flaxman, 2016) au regard des processus de prise de décision – même si la portée exacte de ce droit à l'explication est vivement contestée[18] (Wachter, Mittelstadt et Floridi, 2016) – et le droit pour la personne de « connaître la logique qui sous-tend le traitement automatisé des données la concernant »[19].

Le recours à des courtiers en données qui agrègent les informations contenues dans les profils personnels pose en particulier des problèmes. Le profilage consiste en soi à extrapoler des données disponibles sur internet au moyen de processus de collecte automatisée d'informations, puis à construire des profils et à les appliquer. Ces techniques présentent des avantages à la fois pour les individus et la société, en permettant notamment une meilleure segmentation des marchés et une analyse ciblée des risques et des fraudes. Leur application suscite néanmoins de vives inquiétudes. La recommandation du Conseil de l'Europe sur le profilage[20] traite du risque que les profils attribués à une personne permettent de générer de *nouvelles* données, y compris par l'agrégation de données. 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). Ce point de vue est repris par le Conseil des droits de l'homme des Nations Unies qui note avec

---

[17] Voir les lignes directrices sur la protection des personnes à l'égard du traitement des données à caractère personnel à l'ère des mégadonnées du Conseil de l'Europe, 17 janvier 2017, disponible au lien suivant : https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=09000016806f06d0

[18] Voir Wachter, Mittelstadt and Floridi, 2016. Voir aussi Lilian Edvards and Michael Veale: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2972855.

[19] Voir https://edps.europa.eu/data-protection/our-work/ethics_fr pour plus de détails. La Directive (UE) 2016/680 relative à la protection des personnes physiques à l'égard du traitement des données à caractère personnel par les autorités compétentes à des fins de prévention et de détection des infractions pénales, d'enquêtes et de poursuites en la matière ou d'exécution de sanctions pénales, et à la libre circulation de ces données, fournit un cadre pour les traitements des données réalisés au cours d'actions qui ne sont pas régies par le droit communautaire, comme une coopération judiciaire dans une affaire pénale ou une coopération policière.

[20] Recommandation CM/Rec(2010)13 du Comité des Ministres aux États membres sur la protection des personnes à l'égard du traitement automatisé des données à caractère personnel dans le cadre du profilage.

préoccupation le 22 mars que « le traitement automatique des données à caractère personnel pour l'établissement de profils individuels peut aboutir à une discrimination ou à des décisions pouvant avoir des conséquences sur la réalisation des droits de l'homme, notamment les droits économiques, sociaux et culturels »[21].

Le principal problème que pose l'utilisation des données extraites des profils à différentes fins, grâce à des algorithmes, réside dans le fait que ces données perdent leur contexte d'origine. La réutilisation des données à de nouvelles fins est susceptible de nuire à l'autonomie d'information de la personne concernée. Les moteurs de recherche peuvent avoir un effet similaire sur le droit au respect de la vie privée et à la protection des données, dans la mesure où ils facilitent l'agrégation de données sur une personne en particulier.

L'utilisation de données extraites de profils, y compris des profils créés sur la base de données collectées par des algorithmes et des moteurs de recherche, a une incidence directe sur le droit d'une personne à l'autodétermination de son information. La personne concernée n'est généralement pas consciente du profilage lui-même ni de la réutilisation ultérieure des données en dehors de leur contexte originel qui facilitent l'accès à des informations en réduisant la protection que permettent les données anonymes. Par ailleurs, les résultats obtenus par le biais des algorithmes de recherche peuvent être incomplets, inexacts ou dépassés et présenter ainsi une fausse image des individus, ce qui peut leur être préjudiciable[22]. De tels profils peuvent avoir des conséquences particulièrement graves pour les enfants et leur vie future.

Enfin, des éléments de plus en plus nombreux semblent indiquer que les données sont recueillies pour mieux cerner les comportements et ainsi cibler les électeurs voire – *in fine* – manipuler des élections (voir point 8 ci-dessous)[23].

Un autre aspect clé de l'utilisation des algorithmes pour le traitement automatisé des données réside dans le stockage des données dans le nuage *(cloud*). Il s'agit de solutions qui permettent de stocker des fichiers et autres données 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 mises en œuvre ordinairement. 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

---

[21] Résolution du Conseil des droits de l'homme des Nations Unies sur le droit à la vie privée à l'ère du numérique, Doc. A/HRC/34/7, 23 Mar. 2017, para.17.

[22] Voir Solove (2006), p. 547. En ce qui concerne le traitement de données dans le cadre de coopérations judiciaires dans des affaires pénales et de coopération policières non régies par le droit communautaire, la Directive (UE) 2016/680 *relative à la protection des personnes physiques à l'égard du traitement des données à caractère personnel par les autorités compétentes à des fins de prévention et de détection des infractions pénales, d'enquêtes et de poursuites en la matière ou d'exécution de sanctions pénales, et à la libre circulation de ces données* fixe des garanties pour la protection des données.

[23] Voir également The Guardian, « The great British Brexit robberty: how our democracy was hijacked », 7 mai 2017, disponible au lien suivant: https://www.theguardian.com/technology/2017/may/07/the-great-british-brexit-robbery-hijacked-democracy (consulté le 25 septembre 2017).

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 autre[24].

3. LIBERTÉ D'EXPRESSION

L'utilisation d'algorithmes et de techniques de traitement de données a des répercussions considérables sur le droit à la liberté d'expression qui englobe le droit recevoir et de communiquer des informations. Si l'effet positif des algorithmes de recherche et des moteurs de recherche pour le droit de l'homme à la liberté d'expression a été évoqué à maintes reprises[25], le risque qu'ils portent atteinte à la liberté d'information et d'expression des personnes, des groupes et de segments entiers de la société est dorénavant étudié de plus en plus fréquemment[26]. De graves questions se posent concernant non seulement la liberté d'expression, mais également l'objectif inhérent à l'article 10 qui est de créer un environnement propice à un débat public pluraliste, également accessible et ouvert à tous. De surcroît, les inquiétudes précédemment exprimées à propos du respect de la vie privée et de la protection des données entravent considérablement la capacité des individus à s'exprimer librement.

Les moteurs de recherche jouent un rôle crucial de filtres pour les êtres humains qui souhaitent rechercher, recevoir ou communiquer des informations. Les contenus qui ne sont ni indexés ni très bien classés par un moteur de recherche internet ont moins de probabilité d'atteindre un large public ou même d'être consultés. L'utilisation des algorithmes peut par conséquent mener à une fragmentation de la sphère publique et à la création de « chambres d'écho » qui privilégient certains types d'organes de presse et renforcent ainsi la polarisation de la société, ce qui peut gravement menacer la cohésion sociale[27]. Un

---

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

[25] Voir, par exemple, la 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 du Conseil de l'Europe le 4 avril 2012 lors de la 1139e réunion des Délégués des ministres, paragraphe 1, consultable sur la page https://search.coe.int/cm/Pages/result_details.aspx?ObjectId=09000016805caa93 (visité le 25 septembre 2017), qui 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. »*.

[26] 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 32e 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 ».*

[27] Voir aussi Arstecnica, Roheeni Saxena, "The social media "echo chamber" is real", disponible au lien suivant: https://arstechnica.com/science/2017/03/the-social-media-echo-chamber-is-real/ (visité le 25 septembre 2017).

algorithme de recherche peut également privilégier certains types de contenus ou fournisseurs de contenus, avec le risque de mettre en péril des valeurs comme le pluralisme et la diversité des médias[28]. C'est particulièrement le cas des moteurs de recherche dominants (Pasquale, 2016).

Les algorithmes déployés par les plateformes de réseaux sociaux pour prédire les préférences des utilisateurs orientent non seulement les publicités qui leur sont présentées, mais ils personnalisent également les résultats de recherche et dictent la manière dont les fils des réseaux sociaux, y compris les fils d'actualités, sont organisés. 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). A contrario, la personnalisation des informations reçues par les utilisateurs à partir d'une prédiction de préférences et d'intérêts peut créer des « bulles de filtres » et pourrait mettre gravement en péril la liberté d'expression et le droit à l'information.

Bien que les bulles de filtres et les chambres d'écho soient des concepts plausibles et fassent donc à ce titre l'objet de nombreuses études, il convient de noter que, dans ce contexte, les preuves matérielles de leur existence en Europe sont contradictoires (Nguyen et al., 2014 ; Zuiderveen Borgesius et al., 2016). Les individus s'informent généralement eux-mêmes au moyen de sources diverses et pas uniquement par les réseaux sociaux ou les recherches sur internet.

Conformément à 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. Selon 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)[29].

---

[28] D'après la publication de l'UNESCO intitulée « *Tendances mondiales en matière de liberté d'expression et de développement des médias* », les technologies internet ont offert un terrain d'expression à beaucoup de nouvelles opinions. Si le manque de statistiques ventilées par sexe empêche pour l'instant de mieux comprendre, pour chaque sexe, l'incidence des outils de recherche contrôlés par algorithmes sur l'exercice du droit à la liberté d'expression, il semblerait que les modèles de communication hommes-femmes et régionaux continuent à se reproduire dans ce nouveau volume d'opinions. Voir les tendances mondiales en matière de liberté d'expression et de développement des médias de l'UNESCO, disponible au lien suivant : http://unesdoc.unesco.org/images/0022/002270/227025e.pdf. (visité le 25 septembre 2017).

[29] La Cour européenne des droits de l'homme a rappelé dans *Yildirim v. Turkey*, 18 mars 2013, n° 3111/10, que « de telles restrictions présentent pourtant de si grands dangers qu'elles appellent de la part de la Cour l'examen le plus scrupuleux » et que « l'information est un bien périssable et en retarder la publication, même pour une brève période, risque fort de la priver de toute valeur et de tout intérêt ». Par conséquent, bloquer l'accès à internet ou supprimer un contenu publié en ligne nécessite « un cadre légal particulièrement strict quant à la délimitation de l'interdiction et efficace quant au contrôle juridictionnel contre les abus éventuels (...) À cet égard, un contrôle judiciaire de telles mesures opéré par le juge, fondé sur une mise en balance des intérêts en conflit et visant à aménager un équilibre entre ces intérêts, ne saurait se concevoir sans un cadre fixant des règles précises et spécifiques quant à l'application des restrictions préventives à la liberté d'expression ».

Cependant, la suppression de contenus sur les plateformes de réseaux sociaux s'opère souvent à l'aide de processus semi-automatisés ou automatisés. 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 s'agissant de l'État de droit (légalité, légitimité et proportionnalité). Bien que les grandes plateformes de réseaux sociaux comme Google ou Facebook aient fréquemment affirmé que toute les suppressions de contenus sont effectuées par des humains (Buni et Chemaly, 2016), le processus est en grande partie automatisé (Wagner, 2016b) et basé sur des opérations semi-automatisées. Selon un rapport de la Commission parlementaire britannique du renseignement et de la sécurité[30], il existe plusieurs techniques automatisées pour détecter les contenus présumés contraires aux conditions de service du fournisseur concerné – contenus extrémistes, exploitation d'enfants ou 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). Un des problèmes qui se posent dans ce contexte est que les intermédiaires sont encouragés à supprimer d'eux-mêmes ce type de contenus, sans base juridique claire et précise. Cette absence de fondement juridique de la suppression automatisée « volontaire » de contenus complique encore davantage la tâche quand il s'agit de veiller à ce que les garanties juridiques de base, telles que la responsabilité, la transparence et le droit à une procédure régulière soient respectées (Fernández Pérez, 2017).

Aux États-Unis, l'administration Obama a prôné la détection et la suppression automatisées des vidéos et images extrémistes[31]. En outre, il a été proposé de modifier les algorithmes de recherche afin de « cacher » les sites web qui encouragent et soutiennent l'extrémisme. Des mécanismes de filtrage automatisé des vidéos extrémistes ont été adoptés par Facebook et YouTube. Cependant, aucune information n'a été communiquée sur le processus ou les critères retenus afin de déterminer le caractère « extrémiste » ou le « contenu clairement illégal » des vidéos[32]. À la suite d'informations relayées par *The Times of London* et *The Wall Street Journal*, selon lesquelles des publicités épousant des thèses « extrémistes » et diffusant des « discours de haine » apparaissant dans des vidéos sur YouTube, YouTube a réagi par une utilisation plus rigoureuse de son algorithme de détection de « contenus non

[30] Voir UK Intelligence and Security Committee of Parliament report, Privacy and Security: A modern and transparent legal framework, mars 2015, disponibel au lien suivant: http://isc.independent.gov.uk/committee-reports/special-reports (consulté le 25 septembre 2017).

[31]Voir Article 19, « Algorithms and automated decision-making in the context of crime prevention », 2 décembre 2016, disponible au lien suivant : https://www.article19.org/resources.php/resource/38579/en/algorithms-and-automated-decision-making-in-the-context-of-crime-prevention (consulté le 25 septembre 2017).

[32] Voir Reuters, Joseph Menn, Dustin Volz, Exclusive: Google, "Facebook quietly move toward automatic blocking of extremist videos", disponible au lien suivant: http://www.reuters.com/article/us-internet-extremism-video-exclusive-idUSKCN0ZB00M (consulté le 25 septembre 2017).

adaptés aux annonceurs » qui aurait affecté des organes de presse indépendants, mais aussi des comédiens, des commentateurs et des experts politiques[33].

Des initiatives similaires ont été mises en place en Europe où, en réponse à une pression publique et politique, des fournisseurs de services intermédiaires se sont engagés à combattre activement les discours de haine en ligne via des techniques automatisées qui détectent et suppriment tout contenu illégal. Bien que la nécessité de lutter efficacement contre les discours de haine ne soit pas contestée, de tels dispositifs ont été accusés de transférer les responsabilités de l'application de la loi de l'État à des entreprises privées, de créer un risque d'interférence abusive avec le droit à la liberté d'expression et critiqués pour leur manque de conformité avec les principes de légalité, de proportionnalité et de garantie d'une procédure régulière. Le fait d'exiger des intermédiaires de restreindre l'accès à des contenus sur la base de notions vagues comme « l'extrémisme » les oblige à contrôler tous les flux de communication et de données en ligne afin d'être en mesure de détecter ce qui pourrait constituer du contenu illégal. Cela va donc à l'encontre du principe d'absence d'obligation de surveillance des intermédiaires, consacré par le droit de l'UE et par les lignes directrices pertinentes du Conseil de l'Europe[34]. En raison de l'important effet dissuasif qu'une telle surveillance entraine sur la liberté d'expression, ce principe est aussi rappelé dans le projet de recommandation sur les rôles et les responsabilités des intermédiaires d'internet préparé par le Comité d'experts sur les intermédiaires d'internet du Conseil de l'Europe en septembre 2017.

Par ailleurs, en ordonnant aux intermédiaires de décider eux-mêmes ce qu'il convient de supprimer ou non comme étant « extrémiste », les pouvoirs publics délèguent le choix des outils et des mesures à un opérateur privé qui peut alors mettre en œuvre des solutions (comme la restriction d'accès ou la suppression de contenus) que la loi ne les autoriserait pas à ordonner eux-mêmes. Des 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, ces types de demandes faites par des instances publiques à des acteurs privés conduisent à une surveillance et un filtrage automatisés et par trop indiscriminé des contenus.

Un an après sa création en juillet 2015, l'Unité de signalement des contenus sur internet d'Europol a évalué et traité 11 000 messages contenant des matériels à contenu extrémiste violent sur 31 plateformes en ligne dans huit langues, ce qui aurait entraîné la suppression

[33] Voir The New York Times, Amanda Hess, "How YouTube's Shifting Algorithms Hurt Independent Media", 17 avril 2017, disponible au lien suivant:https://www.nytimes.com/2017/04/17/arts/youtube-broadcasters-algorithm-ads.html?_r=0 (consulté le 25 septembre 2017).

[34] Voir l'article 15 de la Directive 2000/31/EC du 8 juin 2000 ("Directive sur le commerce électronique"), et le principe 6 relatif à la responsabilité limitée des fournisseurs de services pour les contenus diffusés sur l'internet de la Déclaration du Conseil de l'Europe sur la liberté de la communication sur l'Internet en date du 28 mai 2003.

de 91,4 % du contenu total de ces plateformes[35]. Des mesures auraient été prises pour automatiser le système avec la mise en place de la plateforme commune annoncée en avril 2016[36].

Si la nécessité impérative de lutter résolument contre la diffusion de discours de haine et l'incitation aux infractions à motivation raciste est indiscutable, cela suscite néanmoins de vives préoccupations quant au caractère prévisible et à la légalité des ingérences induites dans la liberté d'expression. En particulier, les données traitées par Europol recouvrent non seulement les contenus qui sont illégaux dans les États membres du Conseil de l'Europe, mais aussi les matériels qui enfreignent les conditions de service des intermédiaires d'internet. Par ailleurs, il est bien souvent difficile même pour un être humain formé à cet exercice d'identifier un contenu extrémiste ou un matériel incitant à la violence, car il faut réussir la tâche complexe de différencier certains facteurs comme le contexte culturel et l'humour. À ce jour, les algorithmes sont incapables de distinguer un propos ironique d'une analyse critique. Le filtrage par algorithme du discours afin d'en retirer des contenus nocifs fait poindre le risque d'un blocage excessif et d'une suppression de propos non seulement inoffensifs, mais également susceptibles de contribuer positivement au débat public. Selon la Cour européenne des droits de l'homme, l'article 10 protège également les contenus qui heurtent, choquent ou inquiètent[37]. Le blocage, le filtrage ou la suppression de contenu par algorithme peut par conséquent avoir une forte incidence négative sur des contenus légitimes. Le problème déjà très aigu de la suppression de grandes quantités de contenus licites en raison des conditions de service des plateformes internet est encore accentué par la pression exercée sur elles afin qu'elles pratiquent un filtrage actif selon de vagues notions telles que l'« extrémisme », le « discours de haine » ou le « contenu clairement illégal ». Selon la Cour européenne des droits de l'homme, toute obligation de filtrage ou de suppression de certains types de commentaires publiés par des utilisateurs sur les plateformes en ligne impose aux opérateurs une exigence « excessive et irréaliste » qui risque de les obliger à mettre en place un système de suivi « susceptible de porter atteinte

---

[35] Voir Europol Internet Referral Unit One Year On, communiqué de presse, 22 juillet 2016, disponible au lien suivant : https://www.europol.europa.eu/newsroom/news/europol-internet-referral-unit-one-year (consulté le 25 septembre 2017).

[36] Voir la 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, disponible au lien suivant : 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 (consulté le 25 septembre 2017). Voir également Article 19, *Algorithms and Automated Decision-Making in the Content of Crime Prevention: A Briefing paper,* 2016.

[37] Comme le prouvent la jurisprudence des juridictions nationales et de la Cour européenne des droits de l'homme, il est toujours délicat de qualifier des propos de « discours de haine (illégal) ». Plusieurs arrêts de la Cour portant sur la nécessité ou la possibilité de qualifier certains propos de « discours de haine répréhensible » ont donné lieu à des votes divisés : *I.A. c. Turquie,* 13 septembre 2005, n° 42571/98; *Lindon, Otchakovsky-Laurens et July c. France,* 22 octobre 2007, n° 21279/02 and n° 36448/02*; Féret c. Belgique,* 16 juillet 2009, n° 15615/07 et *Perinçek c. Suisse*, 15 octobre 2015, n° 27510/08. Voir également *Vejdeland et autres c. Suède*, 9 février 2012, n° 1813/07.

aux libertés fondamentales et au droit de communiquer des informations sur internet »[38]. La Commission de Venise a également préconisé des mesures pour renforcer les garanties en matière de droits de l'homme et pour éviter d'imposer aux fournisseurs de réseaux et de systèmes électroniques de communication des contraintes excessives[39].

Depuis les élections américaines en 2016, les citoyens européens et américains s'inquiètent de plus en plus des campagnes de désinformation menées à coup d'informations montées de toutes pièces, délibérément fausses et trompeuses (« *fake news* »), y compris par des techniques automatisées et sur les plateformes de réseaux sociaux, et qui peuvent avoir une grande influence sur les processus décisionnels démocratiques (voir également point 8 ci-après)[40]. En conséquence, des appels ont été renouvelés 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 créateurs de 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, impliquent des responsabilités spécifiques pour les médias[41].

4. LIBERTÉ DE RÉUNION ET D'ASSOCIATION

L'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 qui augmentent les possibilités de participation des individus à la vie politique, sociale et culturelle[42]. La liberté d'utiliser des plateformes internet, telles que les réseaux sociaux, pour créer des associations, et s'organiser afin de se réunir pacifiquement, y compris pour manifester,

[38] Affaire *Magyar Tartalomszolgáltatók Egyesülete et Index.hu Zrt c. Hongrie,* 2 février 2016, n° 22947/13.–

[39] Voir l'Avis conjoint de la Commission de Venise et de la Direction de la société de l'information et de la lutte contre la criminalité et de la Direction des droits de l'homme (DDH) de la Direction générale des droits de l'homme et de l'État de droit (DGI) du Conseil de l'Europe sur le projet de loi n° 281 portant révision de la législation moldave sur le « mandat de sécurité » adopté par la Commission de Venise lors de sa 110e session plénière, (Venise, 10-11 mars 2017), disponible au lien suivant : http://www.venice.coe.int/webforms/documents/default.aspx?pdffile=CDL-AD(2017)009-f (consulté le 25 septembre 2017).

[40] Voir par exemple « The Power of Big Data and Psychographics », disponible au lien suivant : https://www.youtube.com/watch?v=n8Dd5aVXLCc (consulté le 25 septembre 2017).
ou Das Magazin, Hannes Grassegger und Mikael Krogerus, « Ich habe nur gezeigt, dass es die Bombe gibt », n° 48, 3 décembre 2016, disponible au lien suivant: https://www.dasmagazin.ch/2016/12/03/ich-habe-nur-gezeigt-dass-es-die-bombe-gibt/ (consulté le 25 septembre 2017)., même si le rôle exact des techniques appliquées par Cambridge Analytica et d'autres entités pendant les élections américaines est fortement contestée.

[41]Voir également http://reutersinstitute.politics.ox.ac.uk/news/editors-vs-algorithms-who-do-you-want-choosing-your-news et le code de conduite relatif aux discours haineux illégaux en ligne approuvé le 31 mai 2016 entre l'UE et Facebook, Microsoft, Twitter et YouTube. Voir également https://www.theguardian.com/technology/2016/dec/12/facebook-2016-problems-fake-news-censorship.

[42] Voir la 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.

conformément à l'article 11 de la Convention européenne des droits de l'homme, a également été mise en avant[43]. Partout dans le monde, on a laissé entendre que les réseaux sociaux et leur potentiel de diffusion d'informations et de mise en réseau – un potentiel dopé par les algorithmes – jouaient un rôle de premier plan dans l'organisation et la motivation des militants et des groupes contestataires[44].

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 et la vaste quantité d'informations à caractère personnel dont on dispose sur les individus peuvent à l'évidence servir à pister et identifier les êtres humains et entraîner l'exclusion automatique de certains individus ou groupes d'appels à des rassemblements, ce qui peut avoir un impact significatif sur la liberté de réunion. Le profilage et le contrôle de la foule de manifestants n'a pas seulement lieu sur l'internet, mais s'étend également aux méthodes de contrôle de la foule hors ligne. Théoriquement, les algorithmes utilisés pour prédire de possibles situations de conflit et contestation peuvent aussi servir d'outils de prévention pour contrecarrer les projets de manifestation ou de contestation en arrêtant certaines personnes avant même qu'elles ne se réunissent[45].

5. RECOURS EFFECTIF

L'article 13 de la Convention européenne des droits de l'homme dispose que toute personne dont les droits ont été violés a droit à un recours effectif devant une instance nationale. Le recours disponible doit être effectif en pratique comme en droit. Par conséquent, les États doivent s'assurer que les personnes ont accès à des procédures judiciaires ou autres 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 de défense des droits de l'homme. Étant 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. Il s'agira notamment de veiller à ce que les acteurs du secteur privé respectent les droits de l'homme dans l'ensemble de leurs

[43] 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.

[44] Voir, entre autres, http://www.huffingtonpost.com/pablo-barbera/tweeting-the-revolution-s_b_4831104.html. Voir également Zeynep Tufekci, *Twitter and Tear Gas: The Power and Fragility of Networked Protest*, 2017.

[45] Voir Tim de Chant, « The Inevitability of Predicting the Future », disponible au lien suivant : http://www.pbs.org/wgbh/nova/next/tech/predicting-the-future/ (consulté le 25 septembre 2017).

opérations, en particulier en mettant en place des mécanismes de réclamation effectifs qui remédient rapidement aux griefs des personnes.

Les processus décisionnels automatisés limitent de par leur nature même la capacité des individus à exercer leur droit à un recours effectif. Ces limites peuvent provenir de l'opacité de la décision elle-même, de son fondement, de la question de savoir si l'individu a consenti à l'utilisation de ses données au moment de la prise de décision, ou s'il est même conscient de faire l'objet d'une décision. Compte tenu de la difficulté à imputer la responsabilité de la décision, l'individu ne sait plus vers qui se tourner pour faire valoir ses droits en la matière. Le fait que la décision émane d'un système automatisé, et ait été prise sans intervention humaine ou presque, et pour répondre à des impératifs d'efficience plutôt qu'à une réflexion menée dans un contexte humain, complique un peu plus la tâche des organisations qui recourent à de tels systèmes dès lors qu'elles doivent fournir aux individus concernés un moyen d'obtenir réparation.

La multitude des secteurs dans lesquels des systèmes décisionnels automatisés sont employés peut avoir une incidence grave sur les droits de l'homme, pour peu qu'ils touchent aux soins de santé, aux perspectives d'emploi, à la police prédictive ou autre, rendant la capacité à obtenir réparation de manière effective d'autant plus essentielle dans chacun de ces domaines.

De plus en plus d'entreprises, les plus grandes d'entre elles notamment, utilisent des algorithmes et des techniques de traitement automatisé des données pour gérer leurs procédures de traitement des réclamations. Dans le cadre des processus automatisés de suppression de contenu sur les plateformes de réseaux sociaux (voir point 3 ci-dessus), l'utilisation d'algorithmes se ressent particulièrement dans les réponses apportées 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 requis pour qu'un contenu soit révisé. Tout laisse à penser que les systèmes mis en place par les plateformes internet telles que Facebook, Google ou Microsoft pour répondre aux demandes des utilisateurs sont automatisées de bout en bout pour de nombreux types de demandes et de réclamations (Wagner, 2016b ; Zhang, Stalla-Bourdillon et Gilbert, 2016). Bien souvent, il faut attendre qu'un grand nombre d'internautes se soient plaints d'un type de contenu spécifique pour qu'un algorithme automatisé juge utile de le soumettre au contrôle d'un opérateur humain. Ces opérateurs sont réputés travailler souvent sous une pression de temps considérable et avec des consignes minimales quant « règles de suppression » internes[46].

Le droit à un recours effectif implique le droit à une décision motivée et individuelle. Traditionnellement, toutes ces décisions ont été prises par des êtres humains qui, dans l'exercice de leurs fonctions, après avoir suivi une formation approfondie et conformément aux processus décisionnels applicables, bénéficient d'une liberté d'appréciation. En principe, il appartient à un juge, à un ministre du gouvernement ou à un responsable administratif de

[46] Voir Süddeutsche Zeitung, Till Krause and Hannes Grassegger, « Inside Facebook », disponible au lien suivant : http://international.sueddeutsche.de/post/154513473995/inside-facebook (consulté le 25 septembre 2017).

décider, en conformité avec les critères et la jurisprudence établis par la Cour, comment l'équilibre des droits individuels, tels que la liberté d'expression et la protection contre la violence ou la protection des droits d'autrui, doit être mis en pratique. La décision doit se fonder sur un examen minutieux du contexte spécifique, en tenant compte de l'éventuel «effet dissuasif» d'une interférence et de la proportionnalité de l'interférence. Aujourd'hui, ce sont néanmoins de plus en plus les techniques de traitement algorithmique des données qui préparent et influencent les décisions dans les procédures de réclamation.

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 en 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'attention[47]. Au contraire, les algorithmes empêchent souvent d'accéder à une explication motivée concernant les mesures qui ont été prises dans une affaire donnée.

Dans tous les cas, le droit à un recours effectif exige un système de résolution des litiges progressif. Si la première étape peut être réalisée au moyen d'un système automatisé, il faut qu'existe une possibilité de recours contre la décision prise auprès d'un mécanisme de réexamen interne de niveau supérieur. Si le plaignant n'est pas satisfait de la décision, il doit pouvoir la contester en introduisant un recours judiciaire, conformément à l'article 13 de la Convention européenne[48]. Il semble toutefois qu'un mécanisme de recours judiciaire à lui seul ne suffise pas et qu'une « supervision des négociations collaboratives entre consommateurs et entreprises » par l'État soit nécessaire (Loo, 2016).

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 mesures[49].

6. INTERDICTION DE LA DISCRIMINATION

Un autre droit fondamental fréquemment cité en relation avec l'application des algorithmes et des autres techniques de traitement automatisé est le droit de jouir de tous les droits de l'homme et de toutes les libertés fondamentales sans discrimination.

En termes de rapidité de traitement et de volume de données traitées, les décisions algorithmiques peuvent présenter des avantages considérables par rapport à certains types

[47] Voir Marietje Schaake, « When You Tube took down my video », disponible au lien suivant : https://www.marietjeschaake.eu/en/when-youtube-took-down-my-video (consulté le 25 septembre 2017).

[48] Voir, entre autres, *O'Keefe c. Irlande,* 28 janvier 2014, n° 35810/09,

[49] Voir *Roman Zakharov c. Russie,* 4 December 2015, n° 47143/06..

de décisions humaines. Cela étant, les algorithmes peuvent également présenter une partialité intrinsèque difficile à déceler et/ou à corriger (Sandvig et al., 2016). C'est notamment le cas lorsque les variables individuelles utilisées dans les algorithmes de traitement de mégadonnées (« *big data* ») servent d'indicateurs pour les catégories protégées telles que la race, le genre ou l'âge. Un algorithme peut choisir de discriminer un groupe d'utilisateurs corrélé à 80 %, 90 %, 95 %, voire 99 % avec une variable telle que la race, le genre ou l'âge, sans toutefois le faire systématiquement.

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 classement (voir également le point 3)[50]. En raison de l'agrégation de données et du profilage, 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, ce qui peut les désavantager d'un point de vue commercial. Les moteurs et les algorithmes de recherche ne traitent pas non plus 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 différenciée, en proposant par exemple des prix différents pour les mêmes biens ou services à différents consommateurs en fonction de leur profil (voir point 2 ci-dessus)[51].

Un algorithme tendancieux 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 utilisateurs finaux concernés par ces décisions, mais aussi pour la société dans son ensemble. Certains auteurs ont même suggéré que les services en ligne utilisant des systèmes de notation personnalisés sont susceptibles, de par leur conception même, de mener à des pratiques discriminatoires (Rosenblat et al., 2016). On peut dès lors faire valoir que les personnes devraient avoir le droit de voir une version « impartiale » et non individuellement ciblée des résultats de leurs recherches. Cela pourrait permettre à un internaute de sortir de sa « bulle de filtres » et d'avoir accès à une version non ciblée du contenu de sa recherche, de sa *timeline* sur les réseaux sociaux ou de tout autre service ou produit internet qu'il utilise. En théorie, les algorithmes pourraient être des outils utiles pour réduire la partialité dans des domaines où elle est fréquente, comme dans les processus de recrutement. Pourtant, les experts ont mis en garde contre l'automatisation et l'apprentissage automatique (*machine learning*) qui sont

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

[51] Voir également les dispositions du Règlement (UE) 2016/679 relatives au profilage, au traitement automatisé des données et aux droits de la personne concernée.

susceptibles de renforcer les préjugés car, à la différence des humains, les algorithmes ne sont pas à même de contrer consciemment des préjugés appris[52].

Pour savoir si un algorithme favorise ou permet au contraire de prévenir un traitement discriminatoire, il peut être utile de se référer à la distinction faite sur le plan juridique entre discrimination directe et discrimination indirecte. On parle de discrimination directe lorsqu'une personne fonde sa décision sur des critères ou des facteurs considérés comme étant illégaux (tels que la race, l'origine ethnique, la religion, le genre, l'orientation sexuelle, l'âge ou le handicap). Ces préjugés illégaux se forment souvent inconsciemment et à partir d'informations externes à l'ensemble de données qui *devrait* constituer la base de la prise de décision (par exemple, lors d'un entretien le fait pour le recruteur de relever l'âge ou l'origine raciale du candidat). Les systèmes algorithmiques sont sans doute plus performants pour exclure ces préjugés directs. On parle de discrimination indirecte lorsqu'une caractéristique ou un facteur donné s'observe plus fréquemment dans les groupes de population que la loi interdit de discriminer (une personne d'une certaine origine raciale ou ethnique vivant dans une certaine zone géographique ; les femmes ayant accumulé moins d'annuités de retraite en raison d'interruptions de carrière). Puisque les algorithmes décisionnels se basent parfois sur la corrélation entre des ensembles de données et des considérations d'efficacité, il y a lieu de craindre qu'ils reproduisent ou accentuent la discrimination indirecte par le biais des stéréotypes. Il n'y a discrimination indirecte que si la différence de traitement est impossible à justifier.

En cas d'utilisation d'algorithmes décisionnels, il importe donc d'écarter tout risque de différence de traitement injustifiée et de développer les algorithmes en conséquence. En particulier, une différence de traitement sera injustifiée et illégale si elle se fonde sur des données partiales pour générer une évaluation de risque. Dans une telle éventualité, la décision elle-même n'est pas directement discriminatoire ; elle l'est indirectement dans la mesure où elle repose sur des données et des informations qui peuvent être empreintes de parti pris raciste. C'est le cas, par exemple, lorsque le système pénal utilise des outils d'évaluation des risques pour décider d'accorder ou non une libération sous caution. Le système génère des profils de risque à partir de données policières telles que le nombre d'arrestations répétées pour le même délit. Ces arrestations répétées peuvent néanmoins être la conséquence d'une discrimination directe (préjugé racial)[53]. Si les systèmes décisionnels algorithmiques sont basés sur des décisions humaines antérieures, ils

[52] Voir, par exemple, The Guardian, « AI programs exhibit racial and gender biases, research reveals », disponible au lien suivant : https://www.theguardian.com/technology/2017/apr/13/ai-programs-exhibit-racist-and-sexist-biases-research-reveals (consulté le 25 septembre 2017) ; et The Guardian, « How algorithms rule our working lives », disponible au lien suivant : https://www.theguardian.com/science/2016/sep/01/how-algorithms-rule-our-working-lives (consulté le 25 septembre 2017).

[53] Voir entre autres, Laurel Eckhouse, "Big data may be reinforcing racial bias in the criminal justice system", disponible au lien suivant: https://www.washingtonpost.com/opinions/big-data-may-be-reinforcing-racial-bias-in-the-criminal-justice-system/2017/02/10/d63de518-ee3a-11e6-9973-c5efb7ccfb0d_story.html?utm_term=.720084735d73 (consulté le 25 septembre 2017)et ProPublica, Angwin, Julia, Surya Mattu, and Lauren Kirchner, "Machine Bias: There's Software Used Across the Country to Predict Future Criminals. And It's Biased Against Blacks", 2016, disponible au lien suivant: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing (consulté le 25 septembre 2017.

reproduiront et multiplieront vraisemblablement les mêmes préjugés susceptibles de fausser la prise de décision humaine, la seule différence étant qu'ils seront alors plus difficiles à identifier et à corriger.

7. 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 et Hensel, 2016). Lorsqu'ils sont exécutés par des humains, les processus décisionnels sont loin d'être parfaits. La partialité liée à la race (Bertrand et Mullainathan, 2004), à la catégorie sociale et au sexe (Altonji et Blank, 1999; Goldin et Rouse, 1997) a été démontrée 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 plateformes internet. Permettre à la « sagesse des foules » de prendre des décisions à propos de l'emploi d'individus n'est pas seulement hautement discutable d'un point de vue éthique, mais limite en outre la possibilité pour les travailleurs de contester ces décisions, car elles paraissent constituer des mesures « objectives » de leur performance (Tufekci et al., 2015).

À mesure que les plateformes d'emploi « transforment les gens en intelligence collective » (« *human computation* ») (Irani, 2015:227), des questions se posent sur les droits des travailleurs, l'autodétermination des employés et la façon dont les sociétés dans leur ensemble pensent que les êtres humains devraient être traités au travail[54]. En particulier, l'automatisation accrue du travail pose des problèmes considérables en ce qui concerne les droits au respect de la vie privée des employés (Hendrickx et van Bever, 2013) et la protection de ces droits dans le cadre professionnel. 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 menacés, même s'ils ne sont pas directement ciblés par des mesures générales de collecte des données (voir point 2 ci-dessus). Il existe enfin d'autres problèmes liés à l'utilisation des algorithmes par des organisations des secteurs public comme privé pour contrôler les communications du personnel ou pour procéder à des « notations » internes des employés, ce qui n'entrent pas dans le cadre du processus d'évaluation formel mais peut s'avérer plus déterminant en termes d'opportunités de carrière pour les intéressés. Ces pratiques sont généralement utilisées pour veiller à ce que les employés représentent bien l'entreprise ou

[54] Voir F. Dorssemont, K. Lörcher et I. Schömann (éds.), *The European Convention of Human Rights and the Employment Relation,* Hart Publishing, Oxford, 2013.

l'administration ; elles ont à l'évidence des incidences sur 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 3 ci-dessus).

Les organismes et les services publics automatisent de plus en plus leur prise de décision au moyen d'algorithmes (van Haastert, 2016). Alors que la question de savoir si ces systèmes peuvent ou non accroître l'efficacité est encore vivement débattue, 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 grand risque de tri social dans les données médicales, étant donné que les algorithmes peuvent sélectionner des groupes de citoyens ou des profils humains particuliers 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 l'octroi de prestations sociales.

8. DROIT À DES ÉLECTIONS LIBRES

L'utilisation d'algorithmes et de systèmes de recommandation automatisés capables de créer des « bulles de filtres », c'est-à-dire 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 d'entrainement très importants sur les processus démocratiques des sociétés. Si les effets concrets sur la formation d'une opinion politique des « bulles de filtres » et des fausses informations ciblées sont difficiles à mesurer précisément[55], les 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). Cela peut avoir des incidences déterminantes, en particulier dans le contexte d'élections.

[55] Voir Nguyen, Tien T., Pik-Mai Hui, F. Maxwell Harper, Loren Terveen et Joseph A. Konstan (2014), « *Exploring the Filter Bubble: The Effect of Using Recommender Systems on Content Diversity* », p. 677–686 dans Proceedings of the 23rd International Conference on World Wide Web, WWW '14, New York, NY, États-Unis: ACM, consulté (http://doi.acm.org/10.1145/2566486.2568012) et Zuiderveen Borgesius, rederik J. et al (2016), « *Should We Worry About Filter Bubbles*? », Internet Policy Review, Journal on Internet Regulation 5(1), consulté le 1er septembre 2016 (http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2758126)

Si, depuis l'avènement d'internet, il a été débattu que les campagnes en ligne et les réseaux sociaux étaient susceptibles de changer la façon de mener la politique et les élections, ce n'est que plus 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 » un vote. 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 lors des élections de mi-mandat au Congrès le 2 novembre 2010 (Bond et al., 2012)[56]. Un comportement similaire ayant été observé lors des élections locales au Royaume-Uni, en 2016 (Griffin, 2016), des indications sérieuses montrent que depuis lors Facebook vend des services de publicité politique aux partis politiques du monde entier. L'essentiel n'est pas tant de savoir si Facebook et d'autres plateformes dominantes utilisent, délibérément ou non, leur pouvoir pour influencer le vote humain que de comprendre qu'elles ont, en principe, la capacité d'influer sur des élections.

Selon de récentes études, les élections peuvent être remportées non pas par les candidats présentant les meilleurs argumentaires politiques, mais par ceux qui recourent aux technologies les plus performantes pour manipuler les électeurs, parfois sur un mode affectif et irrationnel[57]. Si le phénomène n'est pas totalement nouveau, sa portée et ses effets se sont certainement amplifiés au point de mener à un changement de paradigme qui pourrait compromettre la démocratie en soi. Les données qui sont discrètement amassées, récoltées et stockées au moyen de technologies algorithmiques ont été assimilées à la nouvelle « monnaie du pouvoir politique », dans la mesure où elles peuvent être directement utilisées à des fins de micro-ciblage des électeurs, et avoir ainsi des incidences décisives sur des élections. En effet, des candidats moins connus ne disposent pas toujours des moyens nécessaires pour s'offrir les technologies de manipulation les plus efficaces qui aident à prédire les préférences des électeurs[58]. Si les publicités politiques télévisées sont dorénavant réglementées et les diffuseurs publics se voient imposer des obligations d'impartialité, il n'existe pas de dispositions équivalentes pour l'utilisation des algorithmes

---

[56] Dans le cadre d'une expérience, des chercheurs de Facebook ont présenté à certains utilisateurs, 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 savoir que leurs amis avaient voté avait accru de 0,39 % la probabilité que les utilisateurs aillent eux-mêmes 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 directement augmenté 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).

[57] Voir également *« The great British Brexit robberty: how our democracy was hijacked »* in the Guardian, 7 mai 2017 https://www.theguardian.com/technology/2017/may/07/the-great-british-brexit-robbery-hijacked-democracy, qui soutient que le référendum sur le Brexit s'est finalement joué à quelque 600 000 votes, soit à peine plus de 1 % du total des électeurs enregistrés qui ont été ciblés par une entreprise ayant « introduit la collecte de données en masse dans ses techniques de guerre psychologique », combinant « psychologie, propagande et technologie de manière inédite et puissante ».

[58] https://motherboard.vice.com/en_us/article/big-data-cambridge-analytica-brexit-trump

de prédiction des préférences et comportements des électeurs qui peuvent influencer ces derniers tout autant, si ce n'est davantage.

Dans ce contexte, le rôle particulier joué par les « bots sociaux » dans l'orientation du débat politique et public précédant des élections a été examiné, notamment dans le cadre des élections américaines de 2016 et du référendum sur le Brexit. Les bots sociaux sont des comptes contrôlés par algorithme qui imitent l'activité d'utilisateurs humains, mais opèrent à un rythme nettement supérieur (par exemple, en produisant du contenu ou en prenant part à des interactions sociales), tout en parvenant à ne pas révéler leur identité artificielle. Des études consacrées à l'influence des bots sociaux dans le débat politique lors des élections américaines de 2016 semblent indiquer que leur présence est plus susceptible de nuire au débat démocratique qu'à l'enrichir, ce qui peut aussi alors modifier l'opinion publique et menacer l'intégrité du processus électoral[59]. 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 comme principe fondamental d'un régime politique véritablement démocratique. Et surtout, comme l'indique l'étude de faisabilité relative à l'utilisation d'internet dans le cadre des élections réalisée 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 appropriée. Comme le souligne l'étude, « [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. »[60].

9. AUTRES RÉPERCUSSIONS ÉVENTUELLES

La liste ci-dessus des droits de l'homme sur lesquels l'utilisation de techniques de traitement automatisé et d'algorithmes est susceptible d'avoir une incidence n'est pas exhaustive. Elle vise plutôt à mettre en avant les droits qui sont les plus manifestement concernés et appartiennent déjà à des degrés divers au débat public. Les droits de l'homme et les libertés fondamentales sont universels, indivisibles, interdépendants et corrélés et sont, par conséquent, tous potentiellement visés par l'utilisation de technologies algorithmiques. Compte tenu de sa portée limitée, la présente étude n'a pas abordé la question du droit à la vie dans le contexte des « armes intelligentes » et des drones pilotés par algorithmes ou dans celui de la santé et des recherches s'y rapportant. Elle n'a pas exploré non plus les effets que pourrait avoir la systématisation des points de vue et des opinions via les algorithmes sur la liberté d'opinion et sur le droit à la liberté de pensée, de conscience et de religion.

---

[59] Bessi, Alessandro et Emilio Ferrara (2016), « *Social bots distort the 2016 U.S. Presidential election online discussion* », FIRST MONDAY, Volume 21, Numéro 11, 7 novembre 2016, http://journals.uic.edu/ojs/index.php/fm/article/view/7090/5653.

[60] Voir l'étude de faisabilité sur l'utilisation d'internet dans les élections (MSI-MED(2016)10rev (en cours de publication)).

De fait, le recours croissant à l'automatisation et aux algorithmes décisionnels dans toutes les sphères de la vie publique et privée constitue une menace potentielle pour le concept même de droits de l'homme considérés comme remparts contre l'ingérence des États, dans la mesure où l'on passe progressivement de l'asymétrie traditionnelle du pouvoir et de l'information existant entre structures d'État et les êtres humains à une asymétrie entre opérateurs d'algorithmes (publics ou privés) et celles et ceux qui sont influencés et gouvernés.

4. IMPLICATIONS DE L'UTILISATION DES TECHNIQUES DE TRAITEMENT AUTOMATISÉ ET DES ALGORITHMES DANS LE DOMAINE RÉGLEMENTAIRE

L'utilisation croissante d'algorithmes et de techniques de traitement automatisé ainsi que leurs effets considérables sur l'exercice des droits humains suscitent de plus en plus d'inquiétudes dans le monde, tant au niveau politique que public. Des voix s'élèvent en réponse à cette situation pour demander plus de contrôle et de réglementation[61].

Des gouvernements et contrôleurs indépendants entreprennent déjà, dans de nombreux cas, de réglementer le développement des algorithmes sous une forme ou une autre, en général au préalable de leur exploitation. Les logiciels et les systèmes de traitement des données, qui fonctionnent avec des algorithmes et sont 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). Les développeurs de ces machines sont tenus de soumettre leurs systèmes algorithmiques 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 »[62]. Au Royaume-Uni, les machines de jeu sont contrôlées par un régime d'autorisation spécifique et l'Union européenne a adopté des normes techniques de réglementation précisant les exigences organisationnelles applicables aux entreprises d'investissement recourant au trading algorithmique[63]. 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 algorithme puisse être utilisé pour prendre une décision sur un contrat[64].

Des dispositifs d'octroi de licence de ce type existent pour les algorithmes utilisés dans certains secteurs. Ils s'apparentent aux systèmes de contrôle et d'assurance qualité mis en place dans le secteur de la production et dans l'industrie manufacturière. Ils sont préparés

---

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

[62] 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

[63] Voir http://ec.europa.eu/finance/securities/docs/isd/mifid/rts/160719-rts-6_fr.pdf

[64] Voir la loi fédérale allemande sur la protection des données, promulguée le 14 janvier 2003 (Federal Law Gazette I p. 66), et amendée par l'article 1 de la loi du 14 August 2009 (Federal Law Gazette I p. 2814), disponible au lien suivant : https://www.gesetze-im-internet.de/bdsg_1990/__28b.html.

par des experts compétents qui connaissent et contrôlent les normes de qualité applicables à leurs domaines respectifs. Il n'est pas certain cependant que l'on puisse complètement exporter de telles méthodes de réglementation dans les sphères évolutives et aux multiples dimensions de la vie publique et privée dans lesquelles les techniques de traitement automatisé des données et les algorithmes sont utilisés. Le Centre de la police britannique pour la protection en ligne et la lutte contre l'exploitation des enfants a exigé, par exemple, que son « bouton Facebook » soit proposé par défaut à tous les internautes (Wagner 2016b). Si cette tentative pour faire pression sur Facebook afin que la plateforme change son code par défaut sur la version britannique de son site s'est soldée par un échec, elle laisse entrevoir le type de réponses réglementaires auxquelles on pourrait s'attendre si les États cherchent à déterminer le fonctionnement des algorithmes sur les grandes plateformes internet.

Des questions éthiques et juridiques 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 excuser 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 compliquent considérablement la question de la responsabilité des décisions algorithmiques. Le 16 février 2017, le Parlement européen a adopté une résolution invitant la Commission européenne à élaborer une proposition législative relative aux règles de droit civil sur la robotique. Cette proposition devrait aborder entre autres choses les principes généraux liés au développement de la robotique et de l'intelligences artificielle destinées à un usage civil, les principes éthiques et d'autres questions diverses, dont celles de la responsabilité, des droits de propriété intellectuelle, des flux de données, de la sûreté et de la sécurité[65].

Jusqu'à présent, les problèmes soulevés par le traitement automatisé des données étaient réglés par la législation relative à la protection des données. Cette même législation donne aujourd'hui lieu à des approches adaptées et novatrices telles que la création d'un « droit à l'explication » (Goodman et Flaxman 2016 ; Wachter et al. 2016) et d'autres droits pour les internautes. Cependant, il existe une différence importante entre le droit au respect de la vie privée et la réglementation relative à la protection des données qui, au final, reste un mécanisme de gouvernance conçu pour protéger la vie privée et les droits relatifs à la protection des données à caractère personnel. Ce qui importe est que la vie privée, comme l'exercice d'autres droits de l'homme, nécessite une application effective. Certains des défis les plus importants dans le domaine de la protection des données proviennent du manque de volonté d'octroyer des ressources suffisantes aux autorités de protection des données. 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 soulèvent des questions fondamentales relatives à un large éventail de domaines, il est néanmoins possible de s'inspirer de l'expertise du secteur de la protection

[65] Voir http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+TA+P8-TA-2017-0051+0+DOC+XML+V0//FR.

des données pour essayer d'identifier des réponses réglementaires adaptées à la gouvernance des algorithmes.

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 le 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). Au-delà des preuves à divulgation nulle de connaissance, de nouveaux types de responsabilisation technique peuvent tout à fait venir étayer les notions humaines communes de confiance et de responsabilité. – et être utilisés par la suite en tant qu'approches technologiques au service de l'instauration de la confiance, de la transparence et de la responsabilité.

Quelques états ont développé des stratégies pour la régulation du contenu par lesquelles ils exigent que les intermédiaires d'internet restreignent le contenu de manière notoire à des moyens automatisés plutôt qu'à un signalement de la part des utilisateurs finaux. Cela soulève des problèmes de transparence, de responsabilité et de respect des droits de l'homme.

Puisque les tentatives de réglementation peuvent d'une part susciter en elles-mêmes des inquiétudes sur le plan des droits de l'homme et d'autre part s'avérer problématiques du fait du manque éventuel d'expertise approfondie dont disposent les autorités de réglementation pour élaborer des normes qui soient le reflet non seulement des réalités technologiques mais aussi de considérations juridiques et éthiques, il semble plus approprié de commencer par promouvoir une plus grande transparence et un plus haut niveau de responsabilité vis-à-vis de l'utilisation des algorithmes plutôt que de les réglementer directement[66]. Il conviendrait en outre de combiner les normes élaborées avec des réglementations de haut niveau neutres vis-à-vis de la technologie.

Si un cadrage est parfois nécessaire sur le plan réglementaire à ce stade de la mise en œuvre des algorithmes et des techniques de traitement automatisé, les implications de ces réglementations pour les droits de l'homme et d'un point de vue éthique doivent faire l'objet d'un examen minutieux. En particulier, le discours scientifique actuel s'articule autour de concepts tels que l'autonomie humaine et la capacité d'action individuelle, qui ont tous deux trait au droit au respect de la vie privée et à l'autonomie informationnelle, sans concorder avec la notion même de vie privée. Il convient donc de faire la distinction entre autonomie d'une part et capacité de décision et d'action de l'autre. Ces concepts font respectivement référence à la capacité humaine à fixer ses propres objectifs et à la capacité humaine à

[66] Pour d'autres exemples, voir le chapitre 5 de Pasquale F. (2015), *The Black Box Society: The Secret Algorithms That Control Money and Information*, Harvard University Press.

prendre des décisions et à exercer son libre arbitre. Ils peuvent à ce titre s'avérer incompatibles avec l'utilisation des algorithmes et des techniques de traitement automatisé. En d'autres termes, il pourrait être nécessaire d'élargir la définition des droits de l'homme ou de procéder à leur réinterprétation pour protéger l'autonomie et la capacité de décision et d'action individuelles.

1. TRANSPARENCE

Pour de nombreux consommateurs et régulateurs, les algorithmes ressemblent à des boîtes noires (Pasquale 2015). On entend donc de plus en plus fréquemment des voix s'élever dans le débat public et politique[67] pour réclamer plus de transparence autour des algorithmes. Certains gouvernements demandent ainsi aux entreprises de faire contrôler leurs algorithmes par des auditeurs indépendants, les autorités de réglementation ou le grand public (Diakopoulos 2015 ; Rosnay 2016) avant leur déploiement[68].

Il paraît important de souligner que ces problèmes se posent non seulement aux professionnels qui développent des algorithmes mais aussi à d'autres groupes, comme les analystes (« data scientists ») qui les utilisent. Différents niveaux d'abstraction et de complexité suscitent des défis divers d'opacité et de transparence. Il a souvent été affirmé que l'application des algorithmes dans l'apprentissage automatique se fait en grande partie sans « comprendre » les relations de cause à effet (corrélation au lieu de causalité), ce qui peut provoquer biais et erreurs et susciter des craintes quant à la qualité des données (O'Neil 2016). Le problème, toutefois, tient également dans la manière dont les êtres humains utilisent, perçoivent et interprètent leurs résultats. L'idée selon laquelle les algorithmes informatiques produisent des résultats impartiaux et neutres (Chun 2006) exempts de toute forme de considérations politiques (Denardis 2008) est ici au cœur du problème. C'est pourquoi il serait plus utile d'inciter les gens à participer de manière plus critique aux débats publics sur les algorithmes que de chercher à modifier ces derniers.

La solution consistant à demander la divulgation publique d'algorithmes entiers ou de leur code source est utopique dans ce contexte, les entreprises privées considérant leurs algorithmes comme des logiciels propriétaires stratégiques, et les protégeant donc en conséquence[69]. Il paraît en revanche envisageable d'exiger la publication d'informations

---

[67] Angela Merkel, par exemple, a demandé aux principales plateformes en ligne de divulguer certaines informations sur leurs algorithmes au motif que les internautes ont le droit de savoir sur quels critères les informations qu'ils reçoivent via les moteurs de recherche leur sont envoyées. Voir The Guardian, « Angela Merkel: internet search engines are 'distorting perception' », disponible au lien suivant : https://www.theguardian.com/world/2016/oct/27/angela-merkel-internet-search-engines-are-distorting-our-perception (consulté le 25 septembre 2017).

[68] 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).

[69] 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

partielles mais néanmoins importantes, comme les variables utilisées, les objectifs visés par l'optimisation des algorithmes, les jeux de données d'apprentissage, les valeurs moyennes et les écarts types des résultats obtenus, ou la quantité et le type de données traitées par l'algorithme.

Le plus important en l'occurrence n'est pas tant la publication de toutes les données possibles et imaginables, mais plutôt la notion de « transparence effective ». Ce sont en réalité les données elles-mêmes qui doivent permettre d'atteindre l'objectif sous-jacent d'une transparence accrue. Il n'est par conséquent pas toujours utile d'exiger « plus de données » – dans le pire des cas, cela peut même nuire à l'objectif poursuivi d'une plus grande lisibilité.

Le fait que les algorithmes déployés évoluent fréquemment ne contribue en rien à améliorer la transparence effective des systèmes automatisés. Google, par exemple, modifie son algorithme des centaines de fois par an (Tufekci et al. 2015). Il existe aussi un risque de manipulation et de « bidouillage » des algorithmes s'ils sont rendus publics. Les techniques d'apprentissage automatique compliquent en outre la compréhension des algorithmes à un point tel que la divulgation de tous leurs codes sources pourrait ne pas suffire à les rendre plus transparents. Il serait préférable pour cela que soit expliquée avec précision la façon dont les résultats d'un algorithme sont produits. Puisque les algorithmes peuvent jeter le doute sur le fait qu'ils ont conduit à une prise de décision, et que leur examen n'a d'intérêt que s'il est mené dans leur contexte social et organisationnel, il pourrait être intéressant de faire porter les mesures de promotion de la transparence également sur le processus décisionnel lui-même.

La mise en place de mesures de renforcement de la transparence pourrait *in fine* simplifier les contrôles publics et ceux menés par des experts indépendants, des commissions ou des agences spécialisées – ce qui pourrait dès lors contribuer aux efforts de promotion de la conformité aux normes de protection du consommateur et des droits de l'homme.

2. RESPONSABILITÉ

La responsabilité est le principe selon lequel une personne tenue responsable d'un préjudice sur le plan politique ou juridique a une obligation de justification ou de dédommagement sous une forme ou sous une autre. Cela étant, une personne ne peut être tenue responsable que si elle dispose d'un certain degré de contrôle, dans le sens où elle a facilité ou causé le préjudice ou est en mesure de l'empêcher ou de l'atténuer. D'un point de vue juridique, la responsabilité se manifeste à travers le concept d'obligation d'apporter réparation (par des dommages-intérêts, par exemple). En règle générale, le droit impute la responsabilité à la personne qui est en mesure d'empêcher un préjudice ou d'atténuer un risque (par le biais d'une assurance ou autre). L'imputabilité de la responsabilité d'un processus décisionnel algorithmique est compliquée par le fait que, bien souvent, on ne sait pas avec certitude qui

---

calculer la solvabilité au moyen de l'algorithme ; voir le jugement du Bundesgerichtshof du 28 janvier 2014, Az. VI ZR 156/13, disponible au lien suivant: https://openjur.de/u/677956.html.

dispose du degré de contrôle nécessaire pour être tenu responsable sur le plan juridique ou politique.

Le cas des développeurs d'outils algorithmiques, qui ignorent parfois tout de l'utilisation et de la mise en œuvre future de leurs outils. Inversement, la (ou les) personne(s) qui intègrent les outils algorithmiques dans des applications peuvent ne comprendre que partiellement leur fonctionnement. La responsabilité doit-elle être imputée aux personnes qui ont développé et codé l'algorithme ? Certains auteurs avancent que la responsabilité et la règlementation 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). Un autre domaine à explorer est le fait de savoir si la réglementation existante de la responsabilité du fait des produits devrait être étendue pour inclure les logiciels. Les responsables sont-ils plutôt à chercher du côté des acteurs public et privé, qui achètent l'algorithme et s'en servent pour proposer leurs services sans même comprendre leur fonctionnement ?

L'échec de gouvernance mis au jour par l'affaire des émissions polluantes automobiles illustre également une problématique plus large, qui est celle du renforcement de la responsabilité pour les algorithmes à travers des secteurs divers et variés. Tous les domaines, de la justice pénale aux réseaux sociaux en passant par la santé, l'assurance ou la banque, pour ne citer qu'eux, nécessitent des réponses juridiques spécifiques pour parvenir à des systèmes de traitement automatisé des données et de prise de décision algorithmique plus transparents et pour lesquels la responsabilité est encourue. La responsabilité algorithmique doit être renforcée par des garanties procédurales et la prééminence du droit. Il est également essentiel que quiconque verrait ses droits bafoués par des systèmes décisionnels automatisés puisse bénéficier de mécanismes de recours effectifs.

Une telle approche impose aux opérateurs des algorithmes et des techniques de traitement automatisé des données, qu'ils soient publics ou privés, la tâche ardue de garantir le respect de droits de l'homme élémentaires. Ces principes fondamentaux ne sauraient être transgressés au motif d'une meilleure efficacité présumée de systèmes technologiques opaques (Wagner 2016a). Des questions du même ordre se posent à propos des acteurs privés qui utilisent des algorithmes et des techniques de traitement automatisé des données dans leurs opérations, notamment s'ils occupent une position dominante sur le marché. Compte tenu de la taille et de l'étendue de leurs activités, ces derniers fournissent des services dont la valeur publique est importante – et qui peuvent donc aussi avoir un impact important sur la jouissance des droits de l'homme.

La responsabilité des individus ou des entreprises vis-à-vis des algorithmes qu'ils mettent en œuvre dépend en grande partie de la nature des algorithmes et des résultats que ces derniers produisent. Dans certains 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 que ces résultats sont limités (Staab, Stalla-Bourdillon et Carmichael 2016). Pour autant, de tels mécanismes s'appliquent en règle générale aux seuls résultats des algorithmes, et non aux algorithmes eux-mêmes. Il existe en fait un manque général de cadres réglementaires permettant de faire en sorte que les algorithmes soient

développés en premier lieu pour produire des résultats qui défendent et protègent les valeurs fondamentales ou les principes éthiques et sociétaux de base.

Nous touchons là à des questions éthiques fondamentales concernant l'exploitation des techniques de traitement automatisé de données et d'algorithmes qui sortent du cadre du présent rapport. Comment des systèmes automatisés peuvent-ils intégrer des valeurs normatives ? Certaines des problématiques éthiques entourant les voitures autonomes offrent un aperçu de la complexité de la question : comment l'algorithme est-il censé déduire d'une situation hypothétique la probabilité qu'un accident menace la vie d'un jeune enfant ou d'une personne âgée ? Le nombre de vies potentiellement menacées constitue-t-il un paramètre de calcul ? Quand parle-t-on d'une « bonne » ou d'une « mauvaise » décision dans une telle situation, et quelles sont leurs conséquences juridiques ? Le cas échéant, qui est tenu responsable d'une « mauvaise » décision ?

3. CADRES ÉTHIQUES ET MEILLEURE ÉVALUATION DES RISQUES

Outre les mécanismes réglementaires directs destinés à influencer le code des algorithmes, des mécanismes indirects élaborés à des fins identiques pourraient également être envisagés. Ils concernent le processus de production ou les producteurs d'algorithmes et cherchent à garantir que ces derniers soient conscients des problèmes juridiques, des dilemmes éthiques et des préoccupations en matière de droits de l'homme suscités par la prise de décision automatisée et le traitement des données. Cet objectif pourrait être atteint au moyen de codes professionnels standardisés ou de formes de systèmes de licence pour les ingénieurs en données et les concepteurs d'algorithmes, analogues à ceux qui existent pour les professions telles que les médecins, les juristes ou les architectes. L'idée d'une amélioration des mécanismes existants pour les processus de gestion et de développement de logiciels revient aussi fréquemment (Spiekermann 2015). Elle pourrait s'appliquer plus particulièrement aux techniques de développement de logiciels agiles, pour 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). Étant donné que l'utilisation des algorithmes dans la prise de décision peut potentiellement nuire aux droits des personnes, des mécanismes de contrôle supplémentaires pourraient permettre de garantir que l'algorithme est appliqué de manière juste et viable.

Afin d'apprécier et de comprendre les risques que l'exploitation des systèmes décisionnels automatisés impliquent pour les droits de l'homme, les entreprises peuvent exercer une diligence raisonnable vis-à-vis de ces droits de l'homme. Pour y parvenir, une solution consiste à mesurer l'incidence des systèmes mentionnés sur les droits de l'homme, c'est-à-dire à étudier les répercussions concrètes et potentielles, directes comme indirectes, de leur utilisation sur les individus et à empêcher ou atténuer les préjudices recensés au cours de ces évaluations.

Plusieurs normes portant sur les algorithmes, la transparence, la vie privée, la partialité et plus généralement sur la conception de systèmes éthiques ont été élaborées par des associations professionnelles comme l'IEEE (Institut des ingénieurs électriciens et électroniciens) et par l'Internet Engineering Task Force (IETF) :

- IEEE P7000 : Model Process for Addressing Ethical Concerns During System Design
- IEEE P7001 : Transparency of Autonomous Systems
- IEEE P7002 : Data Privacy Process
- IEEE P7003 : Algorithmic Bias Considerations
- IETF Research into Human Rights Protocol Considerations draft

Les principes d'équité, d'obligation à rendre des comptes et de transparence dans l'apprentissage automatique (ou FAT-ML, pour « Fairness, Accountability, Transparency in machine learning ») en faveur d'algorithmes plus responsables offrent un autre exemple de cadre sectoriel capable de contribuer à une conformité accrue aux droits de l'homme[70].

L'utilisation de plus en plus courante du terme « éthique » en référence aux algorithmes dans le discours des experts comme dans le débat public témoigne peut-être d'une manœuvre tactique de la part de certains acteurs pour éviter certaines réglementations strictes en prônant plus activement certains concepts normatifs informels. Cela étant, elle trahit peut-être également la nécessité d'approfondir la réflexion engagée sur l'interaction des différents types de normes ainsi que le rôle et la responsabilité des différents acteurs dans le but d'organiser la gouvernance de la prise de décision et de « l'éthique » algorithmique sous forme d'un nouvel ensemble de méta-normes applicables.

[70] Voir http://www.fatml.org/resources/principles-for-accountable-algorithms (consulté le 25 septembre 2017).

5. PRINCIPAUX CONSTATS ET CONCLUSIONS

La notion de « traitement et prise de décision algorithmique » est interprétée et comprise différemment selon qu'elle est utilisée dans les milieux de la justice, des technologies et des sciences sociales, ou encore par le grand public. Il s'agit par ailleurs d'un domaine relativement nouveau. Ses répercussions sur l'exercice des droits de l'homme et sur le développement de la société dans son ensemble commencent seulement à susciter une prise de conscience, qu'il reste encore à traduire en un débat politique public plus large et inclusif concernant d'éventuelles implications sur le plan réglementaire.

Les auteurs de la présente étude admettent volontiers le cruel manque d'informations à disposition pour prendre des décisions éclairées sur le sujet, et donc la nécessité de produire de considérables efforts de recherche et d'analyse supplémentaires, y compris en ce qui concerne les caractéristiques des processus décisionnels. Puisque les décisions humaines découlent de processus non pas forcément « meilleurs » mais simplement différents des systèmes décisionnels automatisés, la prise de décision automatisée peut induire des formes de subjectivité, des risques ou des erreurs différentes. Une discussion ouverte doit donc être engagée sur les critères qu'il est nécessaire de définir pour mesurer la qualité de la prise de décision automatisée.

Que les études consacrées au sujet se multiplient est une excellente chose. Cela étant, les travaux de recherche seuls sont insuffisants. Il est primordial de faire en sorte que les professionnels (issus de milieux technologique, scientifique, juridique, médiatique, philosophique et éthique) prennent part aux discussions et aux débats, tout comme le grand public. Afin de donner envie aux gens de participer et de dynamiser le débat public sur un thème qui concerne tous les êtres humains et groupes de population, des activités appropriées d'éducation aux médias et à l'information devraient être organisées pour renforcer l'autonomisation du public, afin qu'il développe une compréhension critique de la logique et du fonctionnement des algorithmes et agisse en conséquence. Les entités publiques et les gouvernements, notamment, doivent avoir accès à des informations suffisamment complètes pour comprendre correctement les systèmes de prise de décision algorithmique, qui font déjà partie intégrante des sociétés du monde entier. L'affaire des émissions polluantes automobiles offre un exemple parmi d'autres des conséquences potentielles du déploiement à grande échelle d'un simple logiciel d'utilisation courante sans contrôle d'une autorité de régulation indépendante. Il n'est pas souhaitable, du point de vue des droits de l'homme, que de puissants systèmes algorithmiques aux répercussions publiques potentielles ne puissent être soumis à une forme appropriée de contrôle public. L'application d'un cadre relatif aux droits de l'homme est primordiale car elle offre, outre des garanties de transparence et responsabilité, l'assurance que tous les droits sont effectivement pris en compte dans les systèmes décisionnels automatisés tels que les algorithmes. Il s'agit d'une tâche ardue qui nécessite un mélange de normes industrielles plus approfondies qui placent les êtres humains et les droits de l'homme au cœur du processus de conception technologique et des mesures réglementaires effectives garantissant qu'en cas de défaillance des normes industrielles, les gouvernements puissent entrer en scène pour promouvoir et protéger les droits de l'homme.

Les êtres humains ont le droit d'exercer un contrôle effectif sur les décisions prises par les autorités publiques. Les questions de gouvernance et/ou de réglementation des algorithmes relèvent de la sphère publique et ne devraient pas être laissées aux mains des seuls acteurs privés. Ces derniers peuvent entreprendre des mesures volontaires pour promouvoir la transparence et la responsabilité au cours de leurs activités, et alors qu'ils sont tenus envers leurs utilisateurs d'une obligation de vigilance et d'une obligation de respect des droits de l'homme, c'est aux États qu'il incombe d'élaborer des mécanismes complets et effectifs garantissant la responsabilité algorithmique. Ce point est crucial non seulement en raison des fortes répercussions que les techniques de traitement automatisé des données et que les algorithmes peuvent avoir sur l'exercice et la jouissance des droits de l'homme, mais aussi du fait de la capacité de ces techniques et algorithmes à étendre, à renforcer et à redistribuer le pouvoir, les compétences et les ressources dans la société.

Surtout, il existe certains domaines de l'interaction sociétale et humaine où la prise de décision algorithmique est déconseillée. Il ne faudrait pas se reposer excessivement sur les systèmes de traitement automatisé des données et la prise de décision automatisée pour promouvoir le développement de la société et résoudre les nouveaux défis complexes pour les générations futures, étant donné que cela risque de faire plus de mal que de bien. Il est donc vital de s'assurer que dans les domaines clés où l'automatisation ne semble pas appropriée du point de vue des droits de l'homme, elle ne soit pas utilisée.

Les auteurs de la présente étude considèrent que le débat public qui s'est engagé à propos des multiples aspects des algorithmes ayant trait aux droits de l'homme accuse du retard sur les progrès technologiques et doit le combler au plus vite pour s'assurer que les droits de l'homme et les intérêts des individus soient garantis de manière effective et durable, et en conformité avec les valeurs énoncées dans la Convention européenne et d'autres traités internationaux. L'utilisation des algorithmes et d'autres techniques de traitement automatisé des données peut avoir des répercussions tant positives que négatives sur l'exercice et la jouissance des droits de l'homme. Les responsables politiques doivent se fixer pour objectif de faire en sorte que ces technologies soient utilisées conformément au principe de « primauté de l'être humain »[71], et que nos sociétés de plus en plus technicisées soient conçues – d'abord et avant tout – dans le souci de garantir à chaque être humain l'exercice et la jouissance effectifs de ses droits.

En conséquence, les auteurs de la présente étude tirent les <u>conclusions suivantes</u> :

1. Il serait souhaitable que les entités publiques et que les acteurs indépendants non étatiques lancent et soutiennent des travaux de recherche destinés à mieux comprendre, pour mieux y répondre, les conséquences de la prise de décision algorithmique à la fois pour les droits de l'homme et d'un point de vue éthique et juridique. Pour y parvenir, ces entités et ces acteurs

[71] Voir également *Human rights in the robot age: Challenges arising from the use of robotics, artificial intelligence, and virtual and augmented reality*, rapport de l'Institut Rathenau commandé et financé par l'Assemblée parlementaire du Conseil de l'Europe, adopté par l'APCE le 28 avril 2017.

devraient apporter leur soutien et s'associer non seulement à des recherches transdisciplinaires axées sur les problèmes et fondées sur des données factuelles, mais aussi à des échanges de bonnes pratiques.

2. Les entités publiques devraient être tenues responsables des décisions qu'elles prennent au moyen de processus algorithmiques. Il conviendrait d'encourager l'adoption de mécanismes permettant à quiconque ayant subi un préjudice à la suite d'une décision algorithmique d'obtenir réparation. Des études d'impact sur les droits de l'homme devraient être conduites au préalable de l'utilisation de systèmes de prise de décision algorithmique dans toutes les sphères de l'administration publique.

3. Il serait souhaitable que les évolutions technologiques soient analysées et suivies de près pour examiner leurs conséquences négatives potentielles, en portant une attention particulière à l'utilisation qui est faite des techniques de traitement algorithmique pendant les élections et les campagnes électorales. Parmi les réponses effectives à ces conséquences négatives pourraient figurer des stratégies de réglementation expérimentales définissant les moyens de mieux protéger les droits d'autrui et de garantir certains objectifs réglementaires, sous réserve que ces stratégies s'accompagnent d'un suivi systématique de leurs effets.

4. La sensibilisation de la population et le discours public sont d'une importance cruciale. Tous les moyens disponibles devraient être mis à profit pour informer et associer le grand public afin que les utilisateurs soient mis en mesure de comprendre de manière critique la logique et le fonctionnement des algorithmes et d'agir en conséquence. Cela peut inclure, sans y être limité, des campagnes d'information et d'éducation aux médias. Il conviendrait d'encourager les institutions ayant recours à des processus algorithmiques à fournir des explications facilement accessibles quant aux procédures suivies par les algorithmes et à la manière dont les décisions sont prises. Les sociétés qui développent les systèmes analytiques utilisés par les processus de prise de décision algorithmique et de collecte de données ont tout particulièrement le devoir de sensibiliser et d'informer les gens, y compris sur les risques de biais qui peuvent être induits par la conception et l'utilisation des algorithmes.

5. Des mécanismes de certification et d'audit des techniques de traitement automatisé des données telles que les algorithmes devraient être mis au point pour s'assurer que ces

techniques respectent les droits de l'homme. Il serait souhaitable que les entités publiques et que les acteurs non étatiques encouragent et promeuvent le renforcement des droits de l'homme au moyen de stratégies de conception éthique et de la mise en place de procédures d'évaluation des risques plus solides lors du développement des logiciels.

6. Les États ne devraient pas imposer aux intermédiaires d'internet une obligation générale d'utiliser des techniques automatisées pour surveiller les informations qu'ils transmettent, stockent ou rendent accessibles, étant donné qu'une telle surveillance interfère avec le respect de la vie privée des utilisateurs et provoque un effet dissuasif sur la liberté d'expression.

7. Les entités publiques devraient entamer, en collaboration avec les autorités de régulation de leurs propres secteurs (assurances, établissements de notation de crédit, banques, commerce électronique et autres), un travail d'élaboration de normes et de lignes directrices spécifiques, afin de s'assurer qu'elles sont en mesure de répondre aux défis posés par l'utilisation des algorithmes décisionnels tout en tenant compte des intérêts des consommateurs et du grand public.

8. Compte tenu de la complexité du domaine, la sensibilisation du grand public, aussi importante soit-elle, ne suffira pas. Il est évident que des dispositifs supplémentaires doivent être mis en place au niveau institutionnel. Par conséquent, les entités publiques devraient lancer et soutenir la création de réseaux et d'espaces au sein desquels les parties prenantes concernées pourraient analyser et évaluer différentes formes de prise de décision algorithmique. Il serait souhaitable que toutes les parties prenantes concernées s'associent à une telle entreprise.

9. Le Conseil de l'Europe, en tant qu'organisation des droits de l'homme de pointe sur le continent, est une tribune appropriée pour explorer l'effet d'une utilisation croissante des systèmes de traitements automatisés de données et des processus de décision automatisés (en particulier les algorithmes), tant dans le domaine public que privé, sur l'exercice effectif des droits de l'homme. Le Conseil de l'Europe devrait poursuivre ses efforts à ce sujet en vue de développer des instruments normatifs appropriés visant à fournir des lignes directrices aux états membres.

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).

Angwin, Julia, Surya Mattu, and Lauren Kirchner. 2016. 'Machine Bias: There's Software Used Across the Country to Predict Future Criminals. And It's Biased Against Blacks.' *ProPublica*. Retrieved 31 August 2016 (https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing).

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.

Clay, Kelly. 2012. 'Is Microsoft Spying On SkyDrive Users?' Forbes. Retrieved 31 August 2016 (http://www.forbes.com/sites/kellyclay/2012/07/19/is-microsoft-spying-on-skydrive-users/).

Denardis, Laura. 2008. 'Architecting Civil Liberties'. in *Global Internet Governance Academic Network Annual Meeting*. Hyderabad (Andra 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.

Edwards, Lilian and Veale, Michael. 2017. 'Slave to the Algorithm? Why a 'Right to an Explanation' Is Probably Not the Remedy You Are Looking For', Duke Law & Technology Review (Forthcoming), available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2972855.

Fernández Pérez, Maryant. 2017. 'Parliamentarians Encourage Online Platforms to Censor Legal Content'. *EDRi*. Retrieved 7 June 2017 (https://edri.org/parliamentarians-encourage-online-platforms-to-censor-legal-content/).

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 Rouse, Cecilia. 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).

Grindrod, Peter. 2014. Mathematical Underpinnings of Analytics: Theory and Applications for Data Science in Customer-Facing Industries. Oxford: Oxford Univ. Press.

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. Dorssemont, K. Lörcher, and I. Schömann. Oxford: Hart Publishing.

Hildebrandt, Mireille and Serge Gutwirth. 2008. 'General Introduction and Overview'. Pp. 1–13 in Profiling the European Citizen. Springer, Dordrecht. Retrieved 26 September 2017 (https://link.springer.com/chapter/10.1007/978-1-4020-6914-7_1).

Hoofnagle Chris Jay "Behavioural Advertising: The Offer You Cannot Refuse" (2012) 6 Harvard Policy & Law Review 273-296.

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).

Kim H, Giacomin J and Macredie R (2014) A qualitative study of stakeholders' perspectives on the social network service environment. International Journal of Human– Computer Interaction 30(12): 965–976.

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, Ryan Kennedy, Gary King, and Alessandro Vespignani. 2014. 'The Parable of Google Flu: Traps in Big Data Analysis'. Science 343(6176):1203–5.

Lazer, David and Ryan Kennedy. 2015. What We Can Learn from the Epic Failure of Google Flu Trends.

Loo, Van. 2016. *The Corporation as Courthouse*. Rochester, NY: Social Science Research Network. Retrieved 7 June 2017 (https://papers.ssrn.com/abstract=2872096).

Mannaro, Katiuscia. 2008. 'Adopting Agile Methodologies in Distributed Software Development'. Universita` 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.

Nguyen, Tien T., Pik-Mai Hui, F.Maxwell Harper, Loren Terveen, and Joseph A. Konstan. 2014. 'Exploring the Filter Bubble: The Effect of Using Recommender Systems on Content Diversity'. Pp. 677–686 in *Proceedings of the 23rd International Conference on World Wide Web*, *WWW '14*. New York, NY, USA: ACM. Retrieved (http://doi.acm.org/10.1145/2566486.2568012).

Nikolaos Altreas et al "Predicting judicial decisions of the European Court of Human Rights: a Natural Language Processing perspective" PeerJ Computer Science Open Access (Published 24. October 2016)

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.

Pasquale, Frank A. 2016. *Platform Neutrality: Enhancing Freedom of Expression in Spheres of Private Power*. Rochester, NY: Social Science Research Network. Retrieved 7 June 2017 (https://papers.ssrn.com/abstract=2779270).

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=ZdstAQAAQBAJ&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).

Rosenblat, Alex, Karen EC Levy, Solon Barocas, and Tim Hwang. 2016. 'Discriminating Tastes: Customer Ratings as Vehicles for Bias'. Retrieved 7 June 2017 (https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2858946).

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).

Sandvig, Christian, Kevin Hamilton, Karrie Karahalios, and Cedric Langbort. 2016. 'When the Algorithm Itself Is a Racist: Diagnosing Ethical Harm in the Basic Components of Software'. *International Journal of Communication* 10:19.

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).

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/).

Tufekci, Zeynep, Twitter and Tear Gas: The Power and Fragility of Networked Protest, Yale University Press, 2017.

Tversky, Amos and Daniel Kahneman. 1974. 'Judgment under Uncertainty: Heuristics and Biases'. *Science* 185(4157):1124–31.

Urban, Jennifer M., Joe Karaganis, and Brianna L. Schofield. 2016. 'Notice and Takedown in Everyday Practice'. Available at SSRN 2755628. Retrieved 28 October 2016 (http://papers.ssrn.com/sol3/Papers.cfm?abstract_id=2755628).

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.

Wachter, Sandra, Brent Mittelstadt, and Luciano Floridi. 2016. *Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation*. Rochester, NY: Social Science Research Network. Retrieved 7 June 2017 (https://papers.ssrn.com/abstract=2903469).

Wagner, Ben. 2016a. 'Efficiency vs. Accountability?' *Bureau de Helling*. Retrieved 10 March 2017 (https://bureaudehelling.nl/artikel-tijdschrift/efficiency-vs-accountability).

Wagner, Ben. 2016b. *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).

Zittrain, Jonathan, Engineering an election, Harvard Law Review Forum Vol. 127, 335 – 339 (2014).

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.

Zuiderveen Borgesius, Frederik J. et al. 2016. 'Should We Worry About Filter Bubbles?' *Internet Policy Review. Journal on Internet Regulation* 5(1). Retrieved 1 September 2016 (http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2758126).

RÉFÉRENCES

I) Instruments du Conseil de l'Europe

Convention européenne des droits de l'homme (STE no 5).

Convention pour la protection des personnes à l'égard du traitement automatisé des données à caractère personnel (STE no 108).

Déclaration du Comité des Ministres sur la liberté de la communication sur l'Internet du 28 mai 2003.

Recommandation CM/Rec(2010)13 du Comité des Ministres aux Etats membres sur la protection des personnes à l'égard du traitement automatisé des données à caractère personnel dans le cadre du profilage**,** adoptée le 23 novembre 2010.

Recommandation CM/Rec(2012)3 du Comité des Ministres aux Etats membres sur la protection des droits de l'homme dans le contexte des moteurs de recherche, adoptée le 4 avril 2012.

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

Recommandation CM/Rec(2014)6 du Comité des Ministres aux Etats membres sur un Guide des droits de l'homme pour les utilisateurs d'internet, adoptée le 26 avril 2014.

Recommandation CM/Rec(2016)5 du Comité des Ministres aux Etats membres sur la liberté d'internet, adoptée le 13 avril 2016.

Lignes directrices sur la protection des personnes à l'égard du traitement des données à caractère personnel à l'ère des mégadonnées, 17 janvier 2017.

Le projet de recommandation du Comité des Ministres aux États membres sur les rôles et les responsabilités des intermédiaires d'internet, finalisé par le MSI-NET le 19 septembre 2017, disponible au lien : https://rm.coe.int/draft-recommendation-on-internet-intermediaries-version-4/1680759e67.

II) Instruments de l'Union Européenne

Directive 2000/31/CE du Parlement européen et du Conseil du 8 juin 2000 relative à certains aspects juridiques des services de la société de l'information, et notamment du commerce électronique, dans le marché intérieur («directive sur le commerce électronique»).

Règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données.

Directive (UE) 2016/680 du Parlement Européen et du Conseil du 27 avril 2016 relative à la protection des personnes physiques à l'égard du traitement des données à caractère personnel par les autorités compétentes à des fins de prévention et de détection des infractions pénales, d'enquêtes et de poursuites en la matière ou d'exécution de sanctions pénales, et à la libre circulation de ces données.

Résolution du Parlement européen du 16 février 2017 contenant des recommandations à la Commission concernant des règles de droit civil sur la robotique (2015/2103(INL)), disponible au lien: http://www.europarl.europa.eu.

III) Journaux et articles and ligne

The great question of the 21st century: Whose black box do you trust?, available at: 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.

Article 19, "Algorithms and Automated Decision-Making in the Content of Crime Prevention: A Briefing paper", 2016.

Das Magazin, Hannes Grassegger und Mikael Krogerus, "Ich habe nur gezeigt, dass es die Bombe gibt", no 48, 3 December 2016, available at https://www.dasmagazin.ch/2016/12/03/ich-habe-nur-gezeigt-dass-es-die-bombe-gibt/.

Reuters institute, Emma Goodman, "Editors vs algorithms: who do you want choosing your news?", available at: http://reutersinstitute.politics.ox.ac.uk/news/editors-vs-algorithms-who-do-you-want-choosing-your-news.

GCN, Kevin McCaney, "Prisons turn to analytics software for parole decisions", 1 November 2013, available at: https://gcn.com/articles/2013/11/01/prison-analytics-software.aspx.

Stanford news, New Stanford research finds computers are better judges of personality than friends and family, available at: http://news.stanford.edu/2015/01/12/personality-computer-knows-011215/.

The Guardian, "The great British Brexit robberty: how our democracy was hijacked", 7 May 2017, available at: https://www.theguardian.com/technology/2017/may/07/the-great-british-brexit-robbery-hijacked-democracy.

Roheeni Saxena, arstecnica, "The social media "echo chamber" is real", available at: https://arstechnica.com/science/2017/03/the-social-media-echo-chamber-is-real/.

Article 19, "Algorithms and automated decision-making in the context of crime prevention", 2 December 2016, available at: https://www.article19.org/resources.php/resource/38579/en/algorithms-and-automated-decision-making-in-the-context-of-crime-prevention.

Joseph Menn, Dustin Volz, Reuters, Exclusive: Google, "Facebook quietly move toward automatic blocking of extremist videos", available at: http://www.reuters.com/article/us-internet-extremism-video-exclusive-idUSKCN0ZB00M.

The Guardian, "2016: the year Facebook became the bad guy", available at: https://www.theguardian.com/technology/2016/dec/12/facebook-2016-problems-fake-news-censorship.

Pablo Barberá and Megan Metzger, "Tweeting the Revolution: Social Media Use and the #Euromaidan Protests", available at: http://www.huffingtonpost.com/pablo-barbera/tweeting-the-revolution-s_b_4831104.html.

Tim de Chant, "The Inevitability of Predicting the Future", available at: http://www.pbs.org/wgbh/nova/next/tech/predicting-the-future/.

Till Krause and Hannes Grassegger, Süddeutsche Zeitung, "Inside Facebook", available at: http://international.sueddeutsche.de/post/154513473995/inside-facebook.

Marietje Schaake, "When You Tube took down my video", available at: https://www.marietjeschaake.eu/en/when-youtube-took-down-my-video.

The Guardian, "AI programs exhibit racial and gender biases, research reveals", available at: https://www.theguardian.com/technology/2017/apr/13/ai-programs-exhibit-racist-and-sexist-biases-research-reveals.

The Guardian, "How algorithms rule our working lives", available at: https://www.theguardian.com/science/2016/sep/01/how-algorithms-rule-our-working-lives.

Laurel Eckhouse, "Big data may be reinforcing racial bias in the criminal justice system", available at: https://www.washingtonpost.com/opinions/big-data-may-be-reinforcing-racial-bias-in-the-criminal-justice-system/2017/02/10/d63de518-ee3a-11e6-9973-c5efb7ccfb0d_story.html?utm_term=.720084735d73.

ProPublica, "Machine Bias", available at: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

Hannes Grassegger & Mikael Krogerus, "The Data That Turned the World Upside Down", available at: https://motherboard.vice.com/en_us/article/mg9vvn/how-our-likes-helped-trump-win.

Alessandro Bessi and Emilio Ferrara, "Social bots distort the 2016 U.S. Presidential election online discussion", FIRST MONDAY, Volume 21, Number 11, 7 November 2016, available at: http://journals.uic.edu/ojs/index.php/fm/article/view/7090/5653.

The Guardian, "Angela Merkel: internet search engines are 'distorting perception'", available at: https://www.theguardian.com/world/2016/oct/27/angela-merkel-internet-search-engines-are-distorting-our-perception.

IV) Divers

Tendances mondiales en matière de liberté d'expression et de développement des medias de l'UNESCO, 2014, disponible au lien: http://www.unesco.org/new/en/world-media-trends.

« UK Intelligence and Security Committee of Parliament report, Privacy and Security: A modern and transparent legal framework », mars 2015, disponible au lien: http://isc.independent.gov.uk/committee-reports/special-reports.

Communication de la Commission au Parlement Européen, au Conseil Européen et au Conseil sur la mise en oeuvre du programme européen en matière de sécurité pour lutter contre le terrorisme et ouvrir la voie à une union de la sécurité réelle et effective, disponible au lien : http://eur-lex.europa.eu/

Rapport du Rapporteur special des Nations Unies, David Kaye, à la trente deuxième session du Conseil des droits de l'homme, sur la promotion et la protection du droit à la liberté d'opinion et d'expression, 2016, (A/HRC/32/ 38).

Avis Conjoint de la Commission de Venise et de la Direction de la Société de l'information et de la lutte contre la criminalité et de la Direction des Droits de l'Homme (DDH) de la Direction Générale des Droits de l'Homme et de l'Etat De Droit (DGI) du Conseil de l'Europe sur le Projet de Loi n° 281 portant révision de la législation Moldave sur le « Mandat De Securité », adopté par la Commission de Venise lors de sa 110ème session plénière, (Venise, 10-11 mars 2017), disponible au lien suivant: http://www.venice.coe.int/.

Résolution du Conseil des droits de l'homme des Nations Unies sur le droit à la vie privée à l'ère du numérique, Doc. A/HRC/34/7.

« Human rights in the robot age: Challenges arising from the use of robotics, artificial intelligence, and virtual and augmented reality », rapport du Rathenau Institut commandé et financé par l'Assemblée Parlementaire du Conseil de l'Europe, adopté par l'APCE le 29 avril 2017.