

11000001 10010000 01111110 11110011 10100000 10011111 11100011
00100101 10000110 01000110 00100000 01011011 10100000 00110011
01010000 01110010 01001001 01000001 10011111 10101111 01000010
00000110 11001100 01110111 00001010 10000101 01100001 11011100
11110111 01110011 10100101 10110001 00010011 10101000 00010011
10011010 11100110 11100110 01011011 00001001 10101110 10100011
01010100 10111111 11000100 01001000 00011110 01010011 01000011
01111000 00110100 10111110 01100011 01101100 01001101 11001011
10100110 10111010 00010100 10111111 01001010 00010000 00010000
10100011 10100111 11010110 01101100 10001010 10111001 10010010
10001111 01011111 01000101 00001110 11010001 11011001 01101011
10100100 11110010 01111001 01001111 01100001 00011011 01111100
01110011 11001000 10010011 10001111 11101011 10100101 10100000
11101010 10001001 10001010 01001010 01111101 10001100 11010001
10011100 11110100 00110010 11110110 00101010 11010000 11100100
10011011 11110101 01010110 10000001 10000101 10100100 10110011
00001100 01000000 01001111 01101110 00100101 00000011 00100111
00000000 10001101 01110100 01000011 00101011 11001101 10011011
10111111 10010010 11110100 11111000 11001110 00111011 00100010
**RESPONSABILITÉ ALGORITHMIQUE ET
DÉCOUVERTE DE CONTENUS NUMÉRIQUES**
Fenwick McKelvey, Ph. D., et Robert Hunt
Document de réflexion, 7-8 février 2019

Remerciements

Ce rapport a été préparé pour le ministère du Patrimoine canadien et la Commission canadienne pour l'UNESCO par Fenwick McKelvey et Robert Hunt. Les points de vue, les opinions et les recommandations exprimés dans ce rapport sont ceux des auteurs et ne reflètent pas nécessairement la politique ou la position officielle du gouvernement du Canada. La responsabilité de toute erreur, interprétation ou omission incombe uniquement aux auteurs.

TABLE DES MATIÈRES

Sommaire	3
Introduction	4
Algorithmes : De la recherche à la recommandation	4
Impulsion en faveur de la responsabilité algorithmique	7
Obstacles à la responsabilisation	12
Recommandations concernant la responsabilité algorithmique	16
Conclusion : Responsabilité algorithmique mondiale	20
Bibliographie	22

Sommaire

Le mot « algorithme » est mieux compris dans son sens générique pour désigner la prise de décision automatisée. Les algorithmes peuvent être codés par l'humain ou devenir autodidactes grâce à l'apprentissage machine. Les biens culturels et les nouvelles passent de plus en plus par des intermédiaires de l'information connus sous le nom de plateformes, lesquelles s'appuient sur des algorithmes pour filtrer, classer, trier et promouvoir l'information. Ces méthodes de recommandation de contenus représentent un moyen clé dans une ère de découvrabilité, où les plateformes coordonnent les utilisateurs, les créateurs et les logiciels pour rendre les contenus attrayants.

La recommandation de contenu algorithmique joue un rôle important et de plus en plus controversé de contrôleur d'accès. De nombreuses controverses au sujet de la nature du contenu recommandé – qu'il s'agisse de vidéos dérangeantes pour les enfants, de complots ou de désinformation politique – ont miné la confiance dans la neutralité de ces systèmes. Il n'y a pas de réponses claires, et la nature complexe et imprévisible des systèmes algorithmiques rend les explications publiques difficiles.

Le pouvoir de marché des principaux intermédiaires de l'information, ainsi que la complexité et l'ampleur croissantes du problème, a compliqué l'application des mécanismes réglementaires traditionnels utilisés dans la radiodiffusion et le journalisme. Le pouvoir économique croissant des intermédiaires de l'information – souvent appelé capitalisme de plateforme – a mis ces questions au premier plan, surtout en raison du rôle qu'elles jouent dans le succès économique des organisations journalistiques et des créateurs culturels indépendants.

Au milieu d'un défi générationnel pour la politique des médias, la responsabilité algorithmique est apparue comme

un domaine d'innovation en matière de réglementation. La responsabilité algorithmique vise à expliquer la prise de décision automatisée, afin de localiser les responsabilités et d'améliorer le système dans son ensemble.

Le rapport qui suit met l'accent sur les questions techniques et systémiques liées à la responsabilité algorithmique, en soulignant que le déploiement importe autant que le développement lorsqu'on explique les résultats algorithmiques. Le rapport articule la responsabilité algorithmique autour de trois moments : l'entrée, le code et le contexte. La prise en compte de ces moments aide à préciser la nature de la responsabilité algorithmique, les obstacles à sa réalisation ainsi que les moyens pour surmonter ces obstacles.

Les interventions que nous recommandons pour améliorer la responsabilité algorithmique s'articulent autour de trois grands objectifs stratégiques :

- Améliorer la qualité des entrées algorithmiques disponibles dans le domaine public.
- S'assurer que les décisions prises dans les codes peuvent être expliquées et justifiées dans l'intérêt public.
- Investir dans une meilleure compréhension de l'utilisation, des publics et des créateurs des médias numériques.

Ces recommandations fournissent le cadre nécessaire à l'adoption de la responsabilité algorithmique. Le Canada ne fait que commencer à comprendre son rôle et ses responsabilités dans le renforcement de la responsabilité algorithmique à l'échelle nationale et internationale. Ces recommandations visent à stimuler l'innovation politique pour répondre à cette question centrale.

Introduction

Comment les algorithmes influencent-ils la façon dont les Canadiens découvrent la culture en ligne ou se renseignent sur les actualités? De plus en plus de Canadiens utilisent les plateformes en ligne – comme Netflix, Spotify et Steam ainsi que YouTube, Facebook et Google – pour trouver et consommer du contenu, notamment des actualités, de la musique, des films, des jeux vidéo et des émissions télévisées. Mais les processus algorithmiques automatisés que les plateformes utilisent pour filtrer, classer, trier et promouvoir ces contenus restent largement cachés du public.

L'opacité de la logique qui guide les décisions algorithmiques a d'importantes répercussions sur l'engagement du Canada envers la Convention sur la protection et la promotion de la diversité des expressions culturelles de l'UNESCO de 2005. Les processus algorithmiques influent sur l'utilisation à la fois culturelle et économique des œuvres de création. Leur pouvoir exige un examen approfondi de l'influence algorithmique sur les flux de contenu culturel et d'information, la vie privée des consommateurs de contenu, la diversité des recommandations et la viabilité économique du travail créatif. Les algorithmes peuvent être à la fois très complexes et à l'abri d'un examen minutieux, ce qui représente un défi de taille quand vient le moment de mesurer et de qualifier cette influence sur la culture et l'engagement civique. De plus, ces facteurs rendent les systèmes algorithmiques responsables de leurs résultats, ce qui constitue un grave problème qui nécessite l'intervention d'acteurs de divers secteurs.

Notre rapport met l'accent sur la responsabilité en ce qui a trait au contenu culturel et aux actualités. Le présent rapport vise à :

- passer brièvement en revue l'historique de la recommandation algorithmique et anticiper son avenir;
- mettre en évidence certaines préoccupations clés concernant la façon dont les filtres algorithmiques influent sur les actualités et la culture;
- relever les principaux obstacles à la compréhension des systèmes algorithmiques par le public;

- discuter de la notion émergente de responsabilité algorithmique;
- proposer des interventions aux parties prenantes du gouvernement, de l'industrie et de la société civile pour y parvenir.

Algorithmes : De la recherche à la recommandation

Trouver le bon contenu dans Internet, ce qui revient à chercher une aiguille dans une botte de foin en expansion constante, est un problème qui ne date pas d'hier. Les solutions à la « surcharge d'information » remontent à encore plus loin. Ces solutions ont souvent été automatisées par des algorithmes.

Un algorithme peut se définir, dans son sens large, comme un ensemble d'instructions pour résoudre un problème ou exécuter une tâche. Dans les discussions populaires sur les médias numériques, « algorithme » désigne un raccourci pour une procédure d'exploitation qui repose sur des données et des logiciels (Bucher, 2018, 20-28). Ces procédures peuvent être écrites par des codeurs humains, ou les ingénieurs peuvent adopter l'approche de l'apprentissage machine selon laquelle les algorithmes « apprennent » comment exécuter des tâches sous différents niveaux de supervision humaine. Sur les plateformes de découverte de contenu, les algorithmes sont une solution technique pour les questions complexes de réception et d'interprétation culturelles, telles que la pertinence, le goût, le plaisir et la personnalité (Langlois, 2014). Les algorithmes interprètent les utilisateurs et les contenus afin d'idéalement améliorer leurs interactions et de trouver des contenus pertinents pour les utilisateurs.

Les moteurs de recherche sont un exemple clé de l'utilisation d'un algorithme pour trouver des contenus en ligne. Les moteurs de recherche indexent les sites Web et recueillent divers indices, tels que des hyperliens et des éléments du code du site Web, afin d'aider ses algorithmes à déterminer dans quelle mesure chaque site est pertinent pour les requêtes de recherche et dans quelle mesure les sites sont populaires et fiables. L'algorithme PageRank de Google,

leader de l'industrie, a classé les sites Web en fonction du nombre et de la qualité des liens d'autres sites vers eux (connus sous le nom de liens entrants).

Google a fondamentalement modifié son moteur de recherche en 2005 en redéfinissant la pertinence en fonction des intérêts *individuels* de l'utilisateur. Ce passage à des résultats personnalisés a recalibré le processus de recherche sur Internet. Google fournissait désormais des résultats dynamiques filtrés en fonction des préférences de l'utilisateur qui synthétisaient les données sur les caractéristiques de l'utilisateur et son comportement antérieur, comme l'historique de recherche et l'adresse IP, avec des tendances abstraites dérivées d'ensembles de données beaucoup plus importants (c.-à-d. des données massives).

L'évolution de Google vers un contexte personnalisé a ponctué une évolution industrielle plus large vers des systèmes de recommandation. D'un point de vue général, un système de recommandation pourrait être n'importe quel mécanisme qui filtre le contenu pour l'utilisateur, y compris des caractéristiques impersonnelles telles qu'une liste « la plus lue » sur un site Web d'actualités ou des mots-clés tendances sur une plateforme de médias sociaux.

Comme pour les résultats de recherche personnalisés, les recommandations algorithmiques nécessitent la création et la tenue de profils utilisateurs. Les profils se composent de données sur les préférences de chaque utilisateur, qu'elles soient glanées implicitement – comme le suivi des livres qu'un utilisateur achète, des émissions de télévision qu'il regarde ou des sites Web qu'il a visités – ou sollicitées directement par le biais d'évaluations et de revues. Ces données ne peuvent révéler que la façon dont les utilisateurs ont interagi avec le contenu dans le passé, de sorte que pour prédire ce qu'ils pourraient apprécier dans l'avenir, les systèmes de recommandation doivent s'appuyer sur au moins une autre source de données.

Il existe quatre approches traditionnelles pour corréler les données des utilisateurs avec une autre source afin de formuler des recommandations (Jannach, Zanker, Felfernig et Friedrich, 2010), tandis que de nouveaux progrès révèlent

le potentiel d'un changement radical dans le paysage de la recommandation par l'apprentissage machine. La section suivante résume ces approches :

1. fondées sur le contenu
2. fondées sur les connaissances
3. fondées sur la collaboration
4. en fonction du contexte
5. apprentissage machine

Fondées sur le contenu

Ce type de recommandations vise à faire correspondre le profil de goûts d'un utilisateur à des éléments particuliers que l'utilisateur aimera, selon le système. Par exemple, si un utilisateur écoute des albums de Gil Evans, Oscar Peterson et Diana Krall sur un service de diffusion de musique en continu, son système de recommandation pourrait raisonnablement supposer que l'utilisateur apprécierait d'autres albums ou listes de lecture jazz.

Fondées sur les connaissances

Ces recommandations sont généralement mises en œuvre sur des plateformes où le comportement des utilisateurs est peu fréquent ou le comportement passé est un mauvais facteur de prédiction. En l'absence d'information existante, ces systèmes demandent des conseils directement à l'utilisateur pour établir ses préférences. Par exemple, un concessionnaire automobile en ligne ne saura probablement pas quelle voiture vous avez achetée dans le passé, de sorte que le système vous demandera à l'avance le prix et les caractéristiques que vous préférez pour recommander un article pertinent. Pour les opérateurs de plateformes, l'un des inconvénients est que les approches fondées sur les connaissances exigent qu'ils obtiennent et gèrent une quantité importante de données sur les articles de leur catalogue (cet inconvénient s'applique également à l'approche fondée sur le contenu.) Pour revenir à notre exemple de jazz, le système de recommandation pourrait suggérer un album du pianiste Herbie Hancock, mais si notre utilisateur hypothétique est un puriste du genre, le système devrait éviter de recommander les incursions de Hancock dans le funk ou le disco. En d'autres termes, pour

faire une recommandation vraiment utile, un système fondé sur le contenu devrait pouvoir s'appuyer sur des données détaillées et précises sur les genres, les artistes et les catalogues.

Fondées sur la collaboration

Souvent appelé filtrage collaboratif, ce système évite le dilemme d'avoir besoin de renseignements de catalogue en associant les utilisateurs à d'autres utilisateurs plutôt qu'à des éléments. Cette approche a été popularisée par Amazon et sa fameuse approche « les clients qui ont acheté cet article ont également acheté » pour faire des suggestions (Smith et Linden, 2017). Le filtrage collaboratif exploite les données contenues dans une masse de profils individuels pour trouver des intérêts communs entre les utilisateurs, puis recherche les éléments « manquants ». Par exemple, l'utilisateur A, l'utilisateur B et l'utilisateur C ont tous acheté les mêmes romans, mais, contrairement à l'utilisateur B et à l'utilisateur C, l'utilisateur A n'a jamais acheté *The Handmaid's Tale*. Étant donné le chevauchement entre ces trois utilisateurs, le système d'Amazon peut raisonnablement prédire que l'utilisateur A appréciera également le roman d'Atwood. Il est important de noter que le système n'a pas besoin de savoir que ces utilisateurs partagent un intérêt pour un genre ou un auteur particulier (il est agnostique sur le contenu); il doit seulement enregistrer que leur comportement d'achat a des points communs. Le système de recommandation d'Amazon a eu une grande influence et a été adopté ou adapté par d'autres grandes plateformes, dont Netflix et YouTube (Smith et Linden, 2017).

En fonction du contexte

Au moment où les recommandations précises en sont venues à être vues comme une source de valeur pour les plateformes, d'autres progrès ont modifié la donne. L'adoption généralisée des appareils mobiles a rendu possible l'élaboration de recommandations contextuelles (Portugal, Alencar et Cowan, 2018). Alors que les informations contextuelles telles que l'heure de la journée ou l'adresse IP étaient déjà détectables par la navigation normale sur le Web, les appareils mobiles ont permis d'inclure dans les calculs des systèmes de recommandation des données

plus précises telles que l'emplacement géographique exact de l'utilisateur.

Apprentissage machine

L'apprentissage machine est de plus en plus utilisé dans les systèmes de recommandation. L'apprentissage machine « a recours à des ordinateurs pour simuler l'apprentissage humain et permet aux ordinateurs de cibler et d'acquérir des connaissances du monde réel, et d'améliorer l'exécution de certaines tâches en fonction de ces nouvelles connaissances » (Portugal, Alencar et Cowan, 2018, p. 206) (traduction libre). L'apprentissage machine permet aux systèmes de recommandation de « tirer les leçons » de leurs erreurs, en évaluant indépendamment le succès de leurs recommandations et en s'adaptant rapidement en réponse à un retour d'information implicite ou explicite. Les progrès accomplis en apprentissage machine pourraient permettre aux systèmes de recommandation de concevoir de nouvelles façons de calculer les recommandations, ce qui permettrait de découvrir des modèles de comportement des utilisateurs que les ingénieurs humains n'ont pas encore envisagés.

L'apprentissage machine pourrait avoir un impact majeur sur la recommandation fondée sur le contenu, car un algorithme pourrait être formé pour analyser, trier et classer le contenu, y compris les nouveautés dans le catalogue, ainsi que pour chercher des modèles à l'intérieur et à travers de celui-ci. Auparavant, seuls les humains pouvaient le faire, mais si les machines pouvaient, métaphoriquement parlant, relever chaque élément de la base de données de Netflix et les indexer de manière inédite, la recommandation fondée sur le contenu pourrait évoluer dans des directions imprévisibles.

Comment l'apprentissage machine pourrait-il modifier les systèmes de recommandation tels que nous les connaissons actuellement? Une possibilité prometteuse est que la popularité pourrait avoir moins d'importance. Par exemple, si un produit n'a jamais été acheté sur Amazon ou si une vidéo YouTube n'a presque pas été visionnée, il est peu probable qu'un système collaboratif recommande ces éléments à quiconque à moins que le producteur ne paie

pour obtenir un résultat commandité. Les algorithmes de l'apprentissage machine pourraient résoudre ce problème aberrant – à condition qu'ils ne soient pas programmés pour donner la priorité à la popularité – et recommander des éléments qui manquent de visibilité ou d'adeptes. Une autre possibilité serait d'ouvrir de nouvelles voies de découverte. Versé dans les archives d'enregistrements musicaux en continu d'une plateforme et de données sur nos habitudes d'écoute, un algorithme d'apprentissage machine conçu pour « apprendre » les attributs musicaux pourrait discerner dans nos historiques de consommation (accords privilégiés, rythmes, motifs lyriques, etc.) des schémas que nous ne sommes pas capables de trouver nous-mêmes sans formation ou par manque de temps. En ayant à l'esprit ces schémas musicologiques plutôt que sociaux, les auditeurs pourraient recevoir des recommandations de chansons, d'artistes ou de genres qui ne leur seraient jamais venus à l'esprit ou de pièces qui ont peine à être découvertes en raison d'un manque de promotion.

Par ailleurs, ces caractéristiques peuvent être de mauvais prédicteurs du plaisir, et une recommandation algorithmique pourrait ne jamais être en mesure de calculer pourquoi quelqu'un préfère Carly Rae Jepsen à Justin Bieber. Malgré l'enthousiasme populaire que suscite l'apprentissage machine, les logiciels continueront probablement à lutter contre les dimensions sociales et affectives plus larges et plus profondes de l'engagement envers la culture – l'influence non quantifiable des sentiments, de l'amitié et des passions. Alors que nos plateformes et appareils continueront à recueillir des données de plus en plus détaillées sur notre comportement, l'expression « recommandations en fonction du contexte » masque la nature irréductiblement complexe de ce qui constitue un « contexte » lorsqu'il s'agit de nos expériences culturelles.

Impulsion en faveur de la responsabilité algorithmique

De plus en plus de travaux de recherche ont commencé à étudier les moyens de déterminer l'influence des algorithmes. Ces préoccupations suivent de près l'essor des moteurs de recherche. Tout comme les moteurs de recherche ont accru la possibilité de découvrir certains contenus, ces mêmes algorithmes ont relégué d'autres contenus dans l'ombre (Introna et Nissenbaum, 2000). Selon les spécialistes du journalisme, la façon dont les algorithmes priorisent, classent, associent et filtrent l'information s'apparentait aux décisions éditoriales dans la production d'actualités (Diakopoulos, 2014). Mais à mesure que l'optimisation des moteurs de recherche s'est développée en tant qu'industrie prometteuse pour améliorer le classement crucial des moteurs de recherche, la valeur économique des décisions algorithmiques en a éclipsé les valeurs démocratiques.

D'autres applications des domaines de la finance, de l'embauche, de la manipulation des prix et de l'évaluation des risques ont suscité une réflexion plus large sur la nature de la prise de décision automatisée au sein du gouvernement et du secteur privé (Eubanks, 2018; O'Neil, 2016; Pasquale, 2015). Les organes de réglementation ont commencé à s'interroger sur l'aspect anti-concurrentiel et les parts de marché des entreprises dominantes dans ce domaine. Par exemple, en avril 2018, l'UE a proposé un règlement visant à « promouvoir l'équité et la transparence pour les utilisateurs professionnels des services d'intermédiation en ligne » dans le cadre de sa stratégie pour le marché unique numérique. En fin de compte, la responsabilité algorithmique est devenue étroitement liée aux questions du pouvoir de marché dans l'économie numérique et à l'incertitude sur le rôle des médias dans la vie quotidienne.

Pour résumer l'attention considérable accordée aux algorithmes, nous établissons trois moments essentiels à l'explication de la prise de décision algorithmique :

1. Entrées, signaux, réalités de terrain et données d'apprentissage qui informent le fonctionnement d'un système algorithmique;
2. Code écrit, algorithmes à boîte noire, bibliothèques et autres agents techniques fonctionnant en tandem dans un système;
3. Contextes dans lesquels le fonctionnement attendu d'un système rencontre son application dans le monde réel.

Ces moments se produisent simultanément, mais les distinguer aide à préciser quels sont les agents qui prennent des décisions, qui ou quoi décide si ces décisions sont de haute qualité, et comment ces pratiques pourraient être plus transparentes.

Ces moments se conjuguent pour produire des expériences de recommandation de contenu qui ont suscité des appels à une plus grande responsabilité. Dans les sections suivantes, nous abordons certaines des questions les plus essentielles en ce qui concerne la diffusion des filtres algorithmiques et des recommandations.

Moment	Définition	Exemples
Entrée	Données qui éclairent les décisions algorithmiques	Courriels d'Enron utilisés comme données de formation dans l'apprentissage machine, profils d'utilisateurs sur Google, évaluations sur Netflix
Code	Code écrit et programmes en cours d'exécution qui traitent les entrées en sorties	PageRank de Google, fil d'actualité de Facebook, recommandations de Netflix
Contexte	Applications réelles de la prise de décision algorithmique avec des publics et des objectifs définis	Recherche d'information en politique, divertissement aux heures de grande écoute, substitution des tâches ménagères par les médias

Tableau 1 – Moments de la responsabilité algorithmique

Solipsisme ou heureux hasard?

Les recommandations algorithmiques peuvent accroître l'isolement social et diminuer la culture publique en limitant notre régime d'information et de culture. L'omniprésence de ce solipsisme algorithmique conduit à une compréhension polarisée et intéressée du monde (Sunstein, 2007). Les détracteurs de ce point de vue soutiennent que les algorithmes personnalisés piègent les utilisateurs dans des bulles d'information partisans lorsqu'il s'agit des actualités, tout en leur proposant un régime restrictif de contenu homogène lorsqu'il s'agit de culture (Pariser, 2011). Certains chercheurs ont affirmé que ces « bulles filtrantes » peuvent amplifier les points de vue extrémistes. Zeynep Tufekci (2018), chercheur et écrivain, a fait une mise en garde contre la possibilité d'une radicalisation algorithmique, où les recommandations automatisées poussent insidieusement les utilisateurs en marge du discours politique.

L'hypothèse de la bulle filtrante a été contestée et ses effets sont exagérés (Dubois et Blank, 2018). Par exemple, les spécialistes du journalisme Fletcher et Nielsen (2018) ont récemment constaté que l'utilisation de Google pour chercher des articles d'actualité mène à ce qu'ils appellent un « heureux hasard automatisé », où les utilisateurs de moteurs de recherche sont en fait exposés à un plus large éventail de sources qui reflètent une gamme diversifiée de positions politiques.

Bien que la bulle filtrante soit surestimée, elle présente un attrait en raison d'une revendication moins litigieuse : il y a un manque de transparence autour des entrées algorithmiques et du code. La sociologue Francesca Tripodi a observé que certaines personnes principalement motivées à utiliser Google pour s'éloigner de leurs positions idéologiques ignorent que leurs requêtes retourneront simplement des données qui réaffirment leurs croyances selon l'expression qu'elles choisissent de taper dans le moteur de recherche (2018, p. 47). En d'autres termes, les résultats du tri algorithmique de Google sont si difficiles à prévoir que, comme le montre l'étude de Fletcher et Nielsen, certains utilisateurs lisent des sources d'information qu'ils auraient pu ignorer autrement, tandis que, comme dans les travaux de Tripodi, d'autres utilisateurs préfiltrent idéologiquement leurs résultats même en tentant explicitement de faire le contraire.

Gouvernance par algorithme

Qu'ils soient bénéfiques ou nuisibles aux utilisateurs, les filtres algorithmiques et les recommandations constituent une forme de politique culturelle, même s'ils sont présentés comme une science objective. Mike Ananny, spécialiste des communications, soutient que les algorithmes « soulèvent des préoccupations éthiques dans la mesure où ils signalent la certitude, découragent les explorations alternatives et créent une cohérence entre des objets disparates – ce qui réduit de façon catégorique l'ensemble des réponses socialement acceptables à la question de savoir ce qui doit être fait » (2016, p. 103) (traduction libre). Autrement dit, la façon dont les ingénieurs d'une plateforme programment leur système

de recommandation fait en sorte qu'il devienne le seul qui compte, et certaines valeurs seront inévitablement exclues.

La gouvernance par algorithme peut masquer la politique culturelle en tant que décision technique. Des décisions apparemment techniques peuvent avoir d'importantes répercussions sur les résultats des politiques, surtout si leur conception ne tient pas activement compte de l'iniquité systémique dans le statu quo. Par exemple, l'apprentissage machine nécessite habituellement des ensembles massifs de données pour que l'algorithme puisse découvrir des modèles significatifs, mais ces modèles peuvent finir par refléter l'iniquité systémique contenue dans les données. Par exemple, un algorithme d'apprentissage machine formé sur un corpus de textes en anglais associait le mot « génie » aux hommes et « mannequin » aux femmes (Bolukbasi, Chang, Zou, Saligrama et Kalai, 2016). Ce résultat discriminatoire est-il la faute de l'algorithme, des développeurs ou d'un ensemble de données qui reflète un historique d'hypothèses sexistes? De plus, l'idée qu'il faut « fixer les préjugés » dans les systèmes algorithmiques repose sur l'hypothèse selon laquelle tout problème social a une solution technique et évite les questions urgentes sur la justesse de technologies comme la reconnaissance faciale (Powles et Nissenbaum, 2018).

Pour y réfléchir plus concrètement, il convient de se demander si un système de recommandation de contenu devrait tenir compte de la race ou du sexe dans ses calculs (plutôt que de prétendre de l'ignorer). Comme le cheminement vers la vérité et la réconciliation l'a démontré – de même que les campagnes de sensibilisation telles #JunosSoWhite et #JournalismSoWhite – le travail visant à éliminer l'exclusion et la discrimination dans le secteur culturel exigera une réflexion sur les inégalités historiques que des catégories protégées comme le « contenu canadien » peuvent reproduire. Le manque de représentation et d'occasions pour les acteurs, réalisateurs, écrivains et autres travailleurs des communautés marginalisées est une préoccupation majeure au sein des industries culturelles et de l'information ainsi que pour le public de leurs créations influentes.

Au fur et à mesure que les recommandations algorithmiques gagnent en influence, il convient d'examiner de plus près le rôle qu'elles jouent dans la promotion de l'iniquité. Les plateformes de vidéo en continu ou les recommandations relatives aux actualités devraient-elles changer leur code pour promouvoir un contenu qui aide à accroître la visibilité des personnes et des histoires sous-représentées? Ou la saisie de ce type de données pourrait-elle avoir des conséquences néfastes? Quoi qu'il en soit, ni les données ni les algorithmes ne sont neutres. Les plateformes qui conçoivent et mettent en œuvre ces algorithmes doivent reconnaître les iniquités systémiques et mettre au point des mesures correctives, de peur qu'ils ne reproduisent et n'amplifient l'injustice dans le statu quo.

Demandes de revenu

La recommandation algorithmique comporte un problème majeur : les modèles d'affaire des plateformes de découverte de contenu peuvent fournir la logique sous-jacente plutôt que de tenir compte de ce qui est le plus avantageux pour les utilisateurs. La volonté de démontrer une croissance aux actionnaires peut inciter les entreprises à élaborer des algorithmes visant essentiellement à ce que les utilisateurs restent collés à leur écran et exposés à la publicité. Ou bien, dans le cas de plateformes comme Netflix qui produisent leur propre contenu, les recommandations peuvent être biaisées pour mettre l'accent sur leurs propres produits plutôt que sur d'autres résultats, peut-être plus pertinents (McKelvey et Hunt, à paraître).

Examinons les effets qu'ont les agrégateurs de nouvelles programmés pour optimiser leurs filtres afin de maximiser l'engagement des utilisateurs. En valorisant des contenus qui attirent l'attention et provoquent une réaction intense, sommes-nous en train de remplacer les normes journalistiques traditionnelles comme l'équilibre et l'objectivité (même si ces normes ne sont pas réalisées)? Le but du classement et de la recommandation, en général, est de plaire aux lecteurs et de les maintenir immergés dans le flux du contenu. Quels sont les effets de la transposition de ce modèle au journalisme sur la démocratie? Par exemple,

les reportages sur la politique américaine peuvent intéresser les lecteurs canadiens, mais la promotion de ces reportages au détriment des actualités locales moins sensationnelles les laisse-t-elle mal informés? En ce qui concerne les productions culturelles, quels sont les pièges à éviter pour rendre le contenu « attirant »? Qui gagne et qui perd quand c'est l'engagement qui prime?

Données de l'utilisateur et protection de la vie privée

La façon dont les plateformes mesurent l'engagement nous amène à nous préoccuper de la protection de la vie privée. La plupart des moyens établis pour mesurer l'engagement en ligne sont relativement simples, comme l'affichage d'une page Web, le partage sur les médias sociaux ou le temps passé à lire, regarder ou écouter les contenus. Ces moyens nécessitent déjà une surveillance des utilisateurs, mais les ingénieurs des systèmes de recommandation continuent de chercher des façons d'aller plus loin. Dans un article paru récemment sur l'algorithme de filtrage collaboratif révolutionnaire d'Amazon, deux de ses créateurs, Brent Smith et Greg Linden, émettent des hypothèses quant à l'avenir de la recommandation de contenus; ils affirment que « la découverte devrait revenir à parler avec un ami qui vous connaît, qui sait ce que vous aimez, travaille avec vous à chaque étape, et anticipe vos besoins » (2017, p. 17) (traduction libre). Les assistants numériques qui atteignent ce niveau d'attention intime seraient-ils dans l'intérêt fondamental d'Amazon ou de ses clients? Les détaillants en ligne ou les plateformes de contenu devraient-ils acquérir une connaissance approfondie de notre vie intérieure pour nous donner de meilleures recommandations? Les recommandations ne visent-elles que l'exploitation lucrative des données des utilisateurs?

Les limites de l'intelligence artificielle

Il peut être trompeur de croire qu'en interagissant avec les systèmes de recommandation, nous avons affaire à une forme d'intelligence. La plupart des gens ont déjà cliqué sur un élément de contenu et ont vu leurs recommandations

devenir étrangères ou irrationnelles : vous cliquez sur une vidéo virale de façon irréfléchie et maintenant, YouTube pense que vous êtes obsédé par les chatons. Les systèmes de recommandation ne peuvent pas porter le genre de jugements sensibles et multidimensionnels que peuvent les humains; c'est pourquoi les plateformes utilisent un terme ambigu et neutre comme « engagement » pour définir leur but. Ces systèmes ne peuvent jamais comprendre pleinement notre contexte personnel, social et culturel en raison de leur dépendance par rapport aux mandataires (cliquer sur un lien) pour deviner les attributs corrélés (aime les chatons).

Ainsi, l'algorithme d'apprentissage profond de YouTube ne peut pas savoir si vous avez regardé une vidéo parce que vos amis s'en moquaient, parce que vous effectuez des recherches sur des théories de conspiration pour l'école, parce que vous êtes endormi avec la lecture automatique ou parce que votre sœur a emprunté votre téléphone – il enregistre seulement que vous avez fait jouer cette vidéo. Comme le dit Meredith Broussard, ancienne journaliste de données et conceptrice de logiciels, on dit que les ordinateurs sont capables de faire n'importe quoi, mais ce n'est que de la rhétorique parce qu'en fin de compte, ce sont des machines et ce qu'elles font, c'est calculer, et donc tout ce que vous pouvez transformer en mathématiques, un ordinateur peut le faire (Chen, 2018). Les plateformes comptent et mesurent les signaux pour déduire nos expériences ou nos sentiments à l'égard du contenu, mais toutes les entrées ne sont pas aussi claires que lorsque l'on clique sur le bouton J'aime, et tous les sentiments et toutes les expériences ne peuvent être traduits en mathématiques.

Les systèmes de recommandation à intelligence artificielle peuvent également se comporter de manière imprévisible, et l'automatisation complète de la découverte en ligne pourrait avoir des résultats aberrants. Par exemple, YouTube a été critiqué pour avoir recommandé du contenu dérangeant à des enfants ou, comme l'a dit *Wired UK*, pour avoir « répandu le sang, le suicide et le cannibalisme » (Orphanides, 2018). L'algorithme fondé sur l'apprentissage machine semble

avoir jugé le succès de ses recommandations entièrement en fonction de l'engagement des utilisateurs à l'égard du contenu (c.-à-d. combien de temps ils ont passé à le regarder, ce qui, pour de très jeunes enfants, pourrait être une mesure insignifiante). Le coupable de ce phénomène reste un mystère, illustrant certains des obstacles à la responsabilité algorithmique dont nous parlerons plus loin, tels que la difficulté de retracer les problèmes à leurs sources et la vulnérabilité de ces systèmes à la manipulation par les humains et les robots.

Même dans des situations plus quotidiennes, les utilisateurs peuvent être perplexes ou mal à l'aise face aux algorithmes, ce qui les incite à invoquer des théories pour expliquer pourquoi l'algorithme a échoué ou mal compris certains choix (Bucher, 2016; Toff et Nielsen, 2018). Dans les situations où les utilisateurs interagissent avec des filtres algorithmiques et des recommandations au fil du temps, certains peuvent en venir à accepter les catégories algorithmiques comme une détermination scientifique de leurs goûts ou intérêts (Cheney-Lippold, 2017). Au fur et à mesure que les plateformes filtrent le contenu dans les fils d'un utilisateur et à l'extérieur de ceux-ci, le goût personnel devient quelque chose qui se développe en collaboration avec la recommandation algorithmique.

Un manque d'information

Les utilisateurs individuels ne sont pas les seuls à devoir composer avec une pénurie de données concernant les résultats algorithmiques sur les plateformes. En effet, avant même de songer à l'opacité des processus algorithmiques, nous manquons actuellement de moyens pour vérifier les affirmations des plateformes sur de l'information aussi simple que le nombre de visionnements ou de lectures. Le problème, c'est que la plupart des données sur le contenu qui attire l'attention proviennent des plateformes elles-mêmes ou des courtiers en données affiliés. L'agence publicitaire Nielsen, par exemple, recueille des données fournies par Facebook, YouTube et Hulu lors de la compilation de ses statistiques de l'effectif de l'auditoire en ligne (Tran, 2017). Il

est problématique que les mêmes entreprises qui génèrent des revenus en vendant de la publicité mesurent leur propre auditoire; cela suscite des préoccupations continues au sujet de la surdéclaration (Sutton, 2018; Vranica et Marshall, 2016) ou des mesures d'auditoire trompeuses (Keller, 2018; Montgomery, 2018). Entre-temps, certaines plateformes qui ne dépendent pas de la publicité refusent de divulguer des données sur l'auditoire.

Dans l'industrie de la publicité numérique, des préoccupations semblables ont mené à des vérifications indépendantes au sujet des impressions publicitaires par des entreprises comme le Media Ratings Council. Il n'en a pas été de même

pour d'autres paramètres de mesure en ligne, même s'il existe un précédent historique. Les estimations d'auditoires pour la télévision et la radio ont été établies par des tiers comme Nielsen ou des maisons de sondage qui ont complété les données divulguées par les industries culturelles.

À l'ère des plateformes, les efforts visant à développer des sources de données indépendantes se heurtent au défi de violer potentiellement les conditions de service ou les accords de confidentialité des plateformes et, de plus, ils exigent que soient mises au point des méthodes complexes pour enquêter sur des objets de recherche hautement instables (McKelvey, 2015; Rieder, Matamoros-Fernández et Coromina, 2018).

Obstacles à la responsabilité

Les processus algorithmiques peuvent poser des défis uniques aux efforts visant à accroître leur responsabilité; cette section résume certains des principaux obstacles.

Moment	Obstacles	Exemples
Entrées	Données exclusives Inégalité systémique et données compromises Personnalisation Localisation	Application Street Bump de Boston; comparaison des résultats de Google; articles sur les tendances de Facebook
Code	Boîte noire Traçabilité Instabilité Imprévisibilité Absence de diversité	Découvertes de la semaine de Spotify; recommandations YouTube pour les enfants; débrayages des employés de Google
Contexte	Utilisation ou objectifs prévus Changement des habitudes des médias Intégration verticale Manipulation	Campagnes de partisans; agent conversationnel Tay de Microsoft; Netflix Originals; consommation omniprésente des médias

Tableau 2 – Obstacles à la responsabilité

Entrées

Données exclusives

Les données utilisées dans la prise de décisions algorithmiques sont souvent confidentielles et les établissements publics n'y ont pas facilement accès. Elles constituent un actif stratégique connu, et procurent un avantage concurrentiel aux entreprises ayant un accès direct et la capacité de regrouper des données (Bureau de la concurrence du Canada, 2018a). Les données des utilisateurs sont une forme de capital particulièrement précieuse, d'après les récentes révélations concernant l'accès spécial de Facebook à des partenaires privilégiés (Waldie et Kiladze, 2018). L'importance cruciale des données pour l'apprentissage automatique augmente encore leur valeur et incite les entreprises à en restreindre l'accès.

Inégalité systémique et données compromises

Bien qu'elles soient souvent présentées comme politiquement neutres, les données sont la représentation d'un monde inéquitable. Si elles sont utilisées naïvement, les données refléteront et reproduiront le statu quo injuste. Par exemple, Kate Crawford (2013), cofondatrice d'AI Now, a traité de l'utilisation par Boston de l'application Street Bump, qui se sert des données des téléphones intelligents pour détecter automatiquement l'état des routes et relever les nids-de-poule à réparer. Cependant, comme le taux d'adoption des téléphones intelligents est aussi bas que 16% dans la ville et que l'application est inévitablement moins accessible aux communautés à faible revenu, il en est résulté une certaine discrimination dans l'attribution des services municipaux essentiels.

Personnalisation

Les filtres algorithmiques sont presque toujours personnalisés à l'aide de l'une des nombreuses techniques décrites ci-dessus, de sorte que toutes les perspectives concernant un système donné ne sont que cela : une seule perspective. On ne peut supposer qu'un seul résultat soit représentatif de l'ensemble du système, et il est impossible, même pour un utilisateur expert, de déduire comment l'algorithme fonctionne dans l'abstrait à partir de résultats

individuels. Lorsque vous partagez des anecdotes sur les partis pris politiques obtenus dans les résultats de Google Actualités, par exemple, vous devez garder à l'esprit que tous les utilisateurs ne verront pas les mêmes résultats pour la même requête de recherche.

Localisation

Comme pour le problème de la personnalisation, les algorithmes contextuels donnent des résultats variés en fonction de la géographie, en particulier les différentes façons dont ils peuvent être signalés par la localisation, les préférences des utilisateurs, la position GPS mobile ou l'adresse IP. Facebook, par exemple, avait l'habitude de restreindre ses publications vedettes à certaines régions linguistiques, montrant aux Canadiens d'expression anglaise la liste sans faire de même pour les Canadiens d'expression française. De plus, les systèmes développés en anglais fonctionnent-ils aussi bien en français ou dans d'autres langues?

Code

Boîte noire

Les systèmes algorithmiques sont souvent appelés « boîtes noires » parce que les utilisateurs ne voient pas comment ils parviennent à leurs résultats (Pasquale, 2015). Avec l'avènement de l'apprentissage machine, nous sommes confrontés à une situation où ces décisions sont difficiles à comprendre non seulement pour les utilisateurs, mais aussi pour le développeur de l'algorithme. L'utilisation de la boîte noire est souvent délibérée. La connaissance du fonctionnement des algorithmes de recherche de Google ou des Découvertes de la semaine de Spotify représente un avantage concurrentiel pour ces entreprises et ne sera pas facilement divulguée. Le Code criminel (pour des actes tels que les méfaits informatiques) et la nouvelle loi commerciale, comme l'ACEUM en instance, peuvent restreindre l'accès au code source, établissant ainsi des obstacles juridiques considérables à la surveillance.

Traçabilité

Les algorithmes fonctionnent comme les composantes de systèmes distribués où les responsabilités sont partagées. Les recommandations et autres formes de filtrage ne résultent pas d'une seule partie, mais des interactions entre plusieurs éléments. Lorsque les processus algorithmiques conduisent à des résultats néfastes, la traçabilité – découvrir la source de l'échec, en attribuer la responsabilité et punir ou du moins éduquer les responsables – devient une préoccupation éthique clé (Mittelstadt, Allo, Taddeo, Wachter et Floridi, 2016). Comme dans le cas des recommandations de YouTube pour les enfants, le traçage des problèmes jusqu'à leur source peut être particulièrement décourageant dans les systèmes d'apprentissage machine, où le développeur, les données de formation ou les utilisateurs malveillants pourraient tous être la source.

Instabilité

La plupart des plateformes élaborent et publient constamment des mises à jour de leur code (Neff et Stark, 2004; Henderson, 2017). Le nuage informatique permet à ces changements de se produire à l'échelle et souvent sans en donner d'indication à l'utilisateur. Par conséquent, toutes les observations des systèmes algorithmiques doivent tenir compte du moment et de l'état de leur objet de recherche instable. Pour détecter et consigner les changements, il faut assurer une surveillance continue de l'instabilité.

Imprévisibilité

Les systèmes algorithmiques, en particulier les systèmes d'apprentissage machine et les systèmes multi-agents, présentent un comportement émergent qui ne peut pas toujours être anticipé. Les bogues et la complexité des logiciels entraînent une imprévisibilité supplémentaire qui accentue l'incertitude quant à l'intention. La création d'un algorithme pour résoudre un problème peut en créer de nouveaux par inadvertance, comme des algorithmes d'application des droits d'auteur ou de lutte contre le piratage qui bloquent les utilisateurs ou les contenus qu'ils n'étaient pas destinés à surveiller (McKelvey, 2018).

Absence de diversité dans les équipes de développement

Les processus algorithmiques peuvent échouer ou marginaliser davantage les personnes et les communautés s'ils perpétuent le statu quo. Un obstacle important à l'amélioration de ce problème est l'absence de diversité dans l'industrie technologique, où 76 % des emplois techniques sont occupés par des hommes et où la main-d'œuvre totale est blanche à 95 % (Winning, 2018), ce qui signifie que les voix sous-représentées peuvent ne pas être entendues et les perspectives importantes, exclues.

Contexte

Utilisation prévue ou objectifs

Quels sont les objectifs des systèmes algorithmiques? Quels sont les comportements optimaux qu'ils tentent d'inciter chez les utilisateurs? Comment ces systèmes automatisés déterminent-ils s'ils ont réussi? Ces questions concernent l'optimisation, et ces principes ne sont généralement pas publics et ne peuvent être connus par des méthodes numériques. Les processus algorithmiques peuvent sembler énigmatiques ou étranges aux utilisateurs parce que ceux-ci ignorent quelles conditions optimales l'algorithme essaie d'atteindre ou de maintenir (Bucher, 2016). En outre, pour s'assurer que les systèmes algorithmiques fonctionnent de manière optimale, il faut procéder à des essais, notamment pour évaluer les effets de la prestation de services sous-optimaux, ce qui peut miner davantage la confiance des utilisateurs et des créateurs (Rosenblat, 2018).

Changement des habitudes des médias

Pour évaluer la force des effets des nouvelles technologies, comme les filtres algorithmiques et les recommandations connexes, il faut se poser des questions fondamentales sur l'utilisation des médias : quels médias les gens utilisent-ils? Quand, comment et pourquoi? Ces préoccupations traditionnelles demeurent pertinentes à mesure que l'utilisation des médias s'inscrit dans de nouveaux contextes.

Alors que la plage d'horaire en soirée, après le travail, demeure une période de pointe pour l'utilisation de Netflix (Deeth, 2015), le recours aux médias en ligne peut avoir lieu à tout moment, et la plupart des Américains consultent leurs téléphones plusieurs fois par jour (Perrin et Jiang, 2018). En raison de l'évolution des modes d'utilisation des médias, il est plus difficile de suivre l'influence algorithmique que, par exemple, d'estimer le nombre de personnes exposées à une émission télévisée.

Intégration verticale

Lorsque les entreprises de plateformes produisent ou vendent leur propre contenu, les objectifs de leur algorithme de recommandation peuvent ne pas être en harmonie avec les intérêts de leurs utilisateurs, et les créateurs de contenu peuvent avoir du mal à atteindre un public. L'intégration verticale amène un conflit potentiel dans la fonction de l'algorithme, en délaissant les avantages pour les utilisateurs au profit des avantages pour les entreprises, et représente un cas particulièrement problématique de manque de transparence quant aux objectifs d'un système automatisé. Amazon, par exemple, a été accusé de manipuler son algorithme d'établissement des prix pour amener les utilisateurs à penser que les produits d'Amazon étaient l'option la moins chère alors qu'ils ne l'étaient pas (Angwin et Mattu, 2016).

Manipulation

Comment le groupe pop coréen BTS s'est-il hissé en tête des palmarès américains? La base mondiale d'admirateurs du groupe a coordonné une campagne pour diffuser les chansons de BTS en continu sur des comptes Spotify américains partagés (Montgomery, 2018). La façon de mesurer la popularité des chansons dans l'ère après l'achat en continu est un enjeu non réglé (Andrews, 2018), et le triomphe fabriqué de BTS est un bon exemple de faille des systèmes automatisés : ils sont vulnérables aux utilisateurs antagonistes et à la manipulation coordonnée. De plus, la valeur des données – en particulier les données utilisées pour former les algorithmes d'apprentissage machine – diminue rapidement si elles sont empoisonnées par des trolls, manipulées par des admirateurs ou influencées par

de nouvelles formes cachées de marketing. Des théories profanes sur « la sagesse de la foule » peuvent obscurcir l'émergence d'une nouvelle classe de gardiens – des utilisateurs très actifs qui savent comment influencer les systèmes algorithmiques (Kalogeropoulos, Negrodo, Picone et Nielsen, 2017). Au fur et à mesure que de nouvelles méthodes de manipulation voient le jour, la responsabilité algorithmique doit être en mesure d'aborder les systèmes dans des environnements d'information défavorables.

Trois questions pour la responsabilité algorithmique

Nous pouvons résumer les enjeux précédents sous forme de trois questions et de sous-questions connexes :

1. Qu'entend-on par entrées pour les systèmes algorithmiques?

- a. Qui décide ce qui compte et ce qui ne compte pas? Comment les données peuvent-elles reproduire les inégalités systémiques et les préjudices historiques? Qui possède ou a un accès privilégié aux données?

2. Les algorithmes devraient-ils être responsables de certaines décisions?

- a. Dans quelle mesure l'automatisation devrait-elle jouer dans les recommandations culturelles ou journalistiques? Quelles logiques justifient ou optimisent la prise de décisions algorithmiques?

3. Les utilisateurs comprennent-ils l'opération aussi bien que ses concepteurs?

- a. Que se passe-t-il lorsque les utilisateurs se comportent de manière inattendue? Comment peut-on résoudre les attentes conflictuelles des utilisateurs? Quels utilisateurs ont la priorité quand vient le temps de décider de la performance optimale?

Pour répondre à ces questions et améliorer la responsabilité algorithmique, il faut s'attaquer à certains obstacles majeurs.

Recommandations concernant la responsabilité algorithmique

Ces obstacles compliquent la recherche de réponses claires aux questions de responsabilité algorithmique que nous avons posées précédemment. Dans la présente section, nous proposons des objectifs stratégiques pour surmonter les obstacles et améliorer la responsabilité algorithmique.

La responsabilité exige des agents qu'ils « répondent de leur travail, témoignent de l'estime qu'ils portent à un travail de grande qualité et encouragent des pratiques diligentes et responsables » (Nissenbaum, 1994, p. 73) (traduction libre). La responsabilité conduit à des systèmes plus performants qui répondent à des normes plus élevées. Toutefois, il existe de sérieux obstacles à la réalisation de cet objectif. Dans son code d'éthique pour la prise de décisions algorithmiques, l'Association for Computing Machinery US Public Policy Council avertit que « certains algorithmes et certaines analyses peuvent être opaques, ce qui fait qu'il est impossible de déterminer quand leurs résultats peuvent être biaisés ou erronés » (USACM, 2017, p. 1) (traduction libre).

Il n'y a pas de consensus clair quant à la forme que devrait prendre la responsabilité algorithmique. Un point de départ commun est de supposer que la responsabilité signifie une plus grande transparence, souvent sous la forme de « données ouvertes » qui permettraient aux parties intéressées de trouver les problèmes et de les régler. Ces appels à la transparence ont toutefois été critiqués parce qu'ils sont potentiellement inefficaces et qu'ils ne tiennent pas compte du travail à effectuer afin de rendre les données brutes significatives (Ananny et Crawford, 2016; McKelvey, 2015; Schrock, 2016; Tkacz, 2012). En revanche, le Règlement général sur la protection des données (RGPD) de l'UE exige que les entités décisionnelles automatisées fournissent « des informations significatives sur la logique de leurs décisions », ce qui a été interprété comme établissant un droit à l'explication (Selbst et Powles, 2017), bien que les avocats et les juristes continuent à débattre pour savoir si

la règle garantit une explication des décisions spécifiques ou simplement une information sur la logique générale du système. Dans l'intervalle, l'explicabilité (ou la possibilité d'interprétation) est devenue un objectif clé au sein de la communauté de l'apprentissage machine, car elle cherche à légitimer des décisions qui, par définition, manquent de surveillance humaine (Gill et Hall, 2018).

Il y a peu d'institutions, à l'échelle mondiale ou nationale, qui s'attachent à expliquer la prise de décisions algorithmiques. Avec l'adoption du Règlement général sur la protection des données (RGPD), l'Union européenne a fait un premier pas vers une réglementation générale des processus algorithmiques. Aux États-Unis, les journalistes ont ouvert la voie en enquêtant sur le fonctionnement des algorithmes et sur les préjudices qu'ils causent (par exemple, les travaux de ProPublica sur le « biais machine »). La responsabilité algorithmique peut être systématiquement minée par la nature concurrentielle des systèmes à boîte noire (Lewis, 2014; Pasquale, 2015) et la nécessité pour les entreprises de protéger les secrets commerciaux, au point même d'élaborer des programmes pour échapper au contrôle réglementaire (Isaac, 2017). Toutefois, la nécessité d'accroître la confiance des consommateurs peut amener les entreprises à adopter la responsabilité sociale en divulguant plus facilement et plus complètement leurs pratiques (Caplan, Donovan, Hanson et Matthews, 2018). En fin de compte, il faudra assurer la coordination entre ces acteurs ainsi qu'entre les utilisateurs, les créateurs et les institutions culturelles si l'on veut aller de l'avant.

Nous résumons nos recommandations pour améliorer la responsabilité algorithmique dans le tableau suivant. Chaque recommandation porte sur l'un des grands enjeux cernés précédemment; à la suite du tableau, nous approfondissons nos recommandations.

Moment	Problème	Objectifs
Entrée	Qu'entend-on par entrées pour les systèmes algorithmiques? Qui décide ce qui compte et ce qui ne compte pas? Comment les données peuvent-elles reproduire les inégalités systémiques et les préjudices? Qui possède et a un accès privilégié aux données?	Améliorer la qualité des entrées algorithmiques disponibles dans le domaine public.
Code	Les algorithmes devraient-ils être responsables de certaines décisions? Dans quelle mesure l'automatisation devrait-elle jouer dans les recommandations culturelles ou journalistiques? Quelles sont les logiques qui justifient ou optimisent la prise de décisions algorithmiques?	S'assurer que les décisions prises en vertu du code peuvent être expliquées et justifiées dans l'intérêt public.
Contexte	Les utilisateurs comprennent-ils l'opération aussi bien que ses concepteurs? Que se passe-t-il lorsque les utilisateurs se comportent de manière inattendue? Comment résoudre les attentes conflictuelles des utilisateurs? Quels utilisateurs ont la priorité pour décider de la performance optimale?	Investir dans une meilleure compréhension de l'utilisation, des publics et des créateurs des médias numériques.

Tableau 3 – Résumé des recommandations pour la responsabilité algorithmique

Nous formulons des recommandations par objectif pour les gouvernements (organisations gouvernementales nationales ou internationales), les institutions culturelles (c.-à-d., la société civile, les médias publics indépendants ou les syndicats de travailleurs culturels) ainsi que les entreprises technologiques et médiatiques (c.-à-d., les plateformes de découverte de contenu et les entreprises de médias sociaux).

Améliorer la qualité des entrées algorithmiques disponibles dans le domaine public

Court terme : Les institutions culturelles devraient s'efforcer de rendre le contenu accessible en ligne. Elles devraient contribuer à faire en sorte que les métadonnées et le format respectent les normes des moteurs de recherche et que le contenu soit accessible à la fois sur les principales plateformes commerciales (lorsque c'est possible) et sur les autres plateformes nationales, comme CBC Music, l'Office national du film et Radio-Canada. Par exemple, la chaîne Encore+ sur YouTube, fruit d'une collaboration entre le Fonds des médias du Canada, Bell Media, Deluxe Toronto, Google Canada et Téléfilm Canada, veille à ce que les productions culturelles canadiennes soient accessibles sur la plus grande plateforme de diffusion vidéo en continu du monde.

Moyen terme : Le gouvernement devrait diffuser davantage de données ouvertes afin de fournir de meilleures entrées pour les systèmes de recommandation (par exemple, Patrimoine canadien ne dispose actuellement que de vingt-trois ensembles de données sur le Portail canadien de données ouvertes). À court terme, davantage de données culturelles devraient également être disponibles, y compris des listes de projets financés, de réalisateurs, d'auteurs et d'autres créateurs. En outre, des études devraient être entreprises en consultation avec l'industrie pour élaborer et diffuser des données de formation neutres sur le contenu culturel afin de former et d'évaluer la performance des systèmes de recommandation pour les indicateurs culturels et de la diversité.

Moyen terme : Les plateformes de découverte de contenu et les autres entreprises, dans le cadre de leur responsabilité sociale d'entreprise et de leurs codes de conduite, devraient vérifier si leurs données de formation sont biaisées, investir dans la recherche pour améliorer l'équité de leurs systèmes et communiquer ces résultats (Courtland, 2018). Des représentants de Google et de Microsoft Research, par exemple, ont participé à la conférence 2018 intitulée « Fairness, Accountability, and Transparency in Machine Learning » et ont aidé à soutenir les « Principles for Accountable Algorithms and a Social Impact Statement for Algorithm » (Diakopoulos et coll., s.d.).

Long terme : Toutes les parties prenantes devraient collaborer à l'élaboration de normes pour les mesures d'auditoires et les métadonnées. Les effets du filtrage algorithmique sont difficiles à évaluer en l'absence d'un système public commun de mesure des auditoires en ligne. Les trois secteurs devraient collaborer pour s'attaquer à ce problème et élaborer des mesures qui correspondent à leurs valeurs. Par exemple, la montée des pièges à clics peut être attribuée à l'utilisation des pages vues comme mesure par laquelle les éditeurs en ligne vendent de la publicité, et des efforts ont été faits pour mettre au point de nouvelles mesures qui servent mieux les créateurs, les auditoires et les annonceurs (Cherubini et Nielsen, 2016). Cependant, ces tentatives se sont limitées pour la plupart à des organisations médiatiques individuelles, et les données

qu'elles ont générées sont restées privées. De même, les producteurs culturels et leurs publics bénéficieraient de l'élaboration et de l'adoption de normes communes pour les métadonnées associées au contenu. De cette façon, le sort des types de contenu (par exemple : local, national ou international) pourrait être correctement évalué et, au besoin, des solutions particulières pourraient être mises à l'essai.

S'assurer que les décisions prises dans les codes peuvent être expliquées et justifiées dans l'intérêt public

Court terme : Le gouvernement devrait suivre la diversité de la main-d'œuvre des entreprises technologiques au Canada. L'absence de ces données est un problème connu qui est exacerbé par le nombre croissant de travailleurs contractuels « fantômes » nécessaires au bon fonctionnement des systèmes algorithmiques (Glaser, 2018; Taylor, 2018). De meilleures données sur les équipes de conception et les travailleurs de soutien pourraient donner un aperçu du contexte du développement des produits tout en éclairant le travail humain qui se fait derrière les systèmes automatisés.

Court terme : Les institutions culturelles, en particulier l'ONF et CBC/Radio-Canada, travaillent déjà à la recommandation algorithmique sur leurs plateformes de découverte de contenu, comme ici.tou.tv, CBC Music et CBC TV. Idéalement, ces projets pilotes donneront lieu à des pratiques exemplaires en matière de protection de la vie privée et de responsabilisation dans les systèmes de recommandation. Dans la mesure du possible, ces institutions publiques devraient communiquer les résultats de leurs recherches et, à moyen terme, s'orienter vers l'adoption et le soutien de normes et de codes de logiciels libres, conformément à la Directive sur la gestion des technologies de l'information du gouvernement du Canada. Ces organisations devraient également chercher à mettre sur pied des équipes multidisciplinaires intersectorielles afin d'élaborer des pratiques exemplaires en matière de conception algorithmique.

Moyen terme : Les plateformes de découverte de contenu devraient fournir des explications claires sur les recommandations de contenu, comme c'est le cas pour la publicité numérique. Au Canada, l'Alliance de la publicité numérique du Canada coordonne Choix de pub, un programme d'autoréglementation qui explique aux consommateurs pourquoi ils ont reçu une publicité. Le triangle bleu de Choix de pub pourrait procurer aux entreprises un modèle à suivre pour expliquer comment les résultats ou les recommandations ont été personnalisés pour un utilisateur donné.

Long terme : Le gouvernement devrait se doter de la capacité à l'interne et à l'externe d'établir la responsabilité algorithmique au Canada. À court terme, les ministères clés devraient investir dans de la formation, des cours d'été ou des ateliers en collaboration avec l'industrie et le milieu universitaire pour s'attaquer aux répercussions culturelles de la recommandation algorithmique. Ces initiatives éducatives pourraient inclure l'apprentissage de méthodes numériques pour étudier les plateformes de découverte de contenu ou des explications sur la conception des systèmes fournies par ces plateformes. À moyen terme, des ministères comme Mesures Canada pourraient financer, faire connaître ou maintenir des projets de transparence algorithmique visant à diffuser des données publiques. À long terme, les ministères devraient élargir leur mandat afin d'inclure la responsabilité algorithmique, ou il faudrait créer de nouveaux ministères pour s'en occuper. Ces recommandations mettraient en œuvre une partie de la proposition de vérifications algorithmiques transmise par le Comité permanent de l'accès à l'information, de la protection des renseignements personnels et de l'éthique du 42^e Parlement.

Moyen terme : Tous les secteurs devraient créer un groupe de travail sur l'optimisation qui examine les logiques d'évaluation de la performance algorithmique. Souvent appelées optimisation, ces logiques décident comment la recommandation algorithmique doit s'exécuter et quels résultats doivent être priorités. Les préoccupations

concernant l'engagement étant le critère prépondérant de l'accroissement de la performance algorithmique, le Canada devrait créer une tribune chargée de mieux définir les objectifs optimaux qui conviendront aux divers utilisateurs, de s'attaquer aux inégalités et injustices systémiques ainsi que d'étudier les définitions de l'optimisation dans les systèmes d'information essentiels. Ce groupe de travail pourrait également établir des pratiques exemplaires et des définitions de la neutralité, de la diversité et de la qualité de l'information afin d'améliorer le contrôle démocratique de la performance algorithmique (Overdorf et coll., 2018; Weinberger, 2018).

Long terme : Tous les secteurs devraient suivre l'exemple de l'UE et travailler à l'élaboration de lignes directrices sur le droit à l'explication et la possibilité de se retirer de certains processus décisionnels algorithmiques. Si les plateformes donnent suite à la divulgation de l'influence algorithmique, il faudra tout de même s'assurer que les utilisateurs comprennent ces divulgations et qu'ils puissent agir en fonction de l'information qu'ils reçoivent. À cette fin, les plateformes pourraient fournir des outils permettant aux utilisateurs d'ajuster manuellement leurs filtres personnalisés, notamment en excluant certains comportements passés ou certaines catégories démographiques. Le moteur de recherche DuckDuckGo, par exemple, met à la disposition des utilisateurs un simple bouton qui leur permet de choisir si les résultats de recherche tiendront compte de leur localisation.

Investir dans une meilleure compréhension de l'utilisation des médias numériques, des publics et des créateurs

Court terme : Le gouvernement devrait mener ou soutenir des recherches sur l'évolution des habitudes culturelles et médiatiques. Les questions de l'Enquête canadienne sur l'utilisation d'Internet devraient porter sur les niveaux de connaissances des médias et les attitudes à l'égard de l'utilisation des recommandations algorithmiques dans

divers aspects de la vie culturelle; cette collecte de données devrait se faire de façon continue. Dans la mesure du possible, le gouvernement devrait coordonner et commander des recherches comparables aux études exemplaires du Pew Research Center sur l'utilisation d'Internet. Cela aiderait à expliquer les contextes et les cas d'utilisation des filtres et des systèmes de recommandation.

Moyen terme : Les institutions culturelles et les organismes subventionnaires devraient reconnaître le travail et l'argent nécessaires à la promotion, à la recommandation et à la découverte du contenu. Par exemple, les entreprises de jeux vidéo emploient souvent des gestionnaires communautaires pour travailler avec les adeptes au développement et à la réception de leurs jeux. Ce travail est un exemple des nouveaux rôles requis dans la production culturelle, et de nouvelles options et catégories de financement devraient être envisagées pour l'appuyer.

Moyen terme : L'industrie devrait s'engager à effectuer des études d'impact algorithmique lorsqu'elle déploie des mises à jour essentielles des systèmes de recommandation. Le Conseil du Trésor du Canada (Karlin, 2018), Nesta au Royaume-Uni (Copeland, 2018) et AI Now aux États-Unis (Reisman, Schultz, Crawford et Whittaker, 2018) ont proposé et élaboré des outils pour évaluer l'impact social probable et les risques des algorithmes. Les études d'impact algorithmique peuvent être simples, par exemple des enquêtes, ou plus complexes, par exemple des études d'impact environnemental. La réalisation d'études d'impact algorithmique avant l'apport de changements aux systèmes de recommandation permettra de rendre les algorithmes plus responsables et transparents.

Moyen terme : Coordonner la recherche sur les possibilités de découverte en tant que problème médiatique commun qui doit retenir l'attention de tous les secteurs. Le Fonds des médias du Canada, l'Office national du film et le Conseil de la radiodiffusion et des télécommunications canadiennes ont popularisé l'expression « découvribilité » pour décrire comment les auditoires ont accès au contenu. Les créateurs de contenu supposent souvent que la facilité de découverte est une question d'efficacité des campagnes de marketing et

de publicité, mais une meilleure compréhension de la façon dont les auditoires découvrent le contenu doit aussi tenir compte du rôle des systèmes algorithmiques. L'industrie, le gouvernement et la société civile devraient convoquer d'autres sommets et groupes de travail pour élaborer un programme de recherche sur la découvribilité, à l'instar des travaux antérieurs sur le système de radiodiffusion, qui se traduirait par une meilleure sensibilisation de la réglementation à la consommation culturelle et médiatique ainsi que de nouveaux outils pour l'éducation aux médias.

Conclusion : Responsabilité algorithmique mondiale

La responsabilité algorithmique s'inscrit en fin de compte dans l'enjeu juridique, politique et social plus large des responsabilités des entreprises d'Internet et des plateformes numériques. Jusqu'à présent, les préoccupations relatives à la responsabilité algorithmique se sont surtout concentrées sur les applications à risque élevé susceptibles de nuire directement à la société dans des domaines comme l'immigration, la justice pénale, la sécurité nationale et les finances (AI Now Institute, 2018). À l'échelle internationale, la responsabilité algorithmique est souvent incluse dans les lois sur la protection des données qui limitent ce type de prise de décisions automatisée.

Au Canada, la loi nationale sur la protection de la vie privée limite l'utilisation des renseignements personnels dans la prise de décisions automatisée, mais le Canada n'a pas adopté de règlements particuliers en lien avec une responsabilité algorithmique élargie. Le Commissariat à la protection de la vie privée surveille activement les questions algorithmiques, bien qu'il mette l'accent sur les renseignements personnels et le traitement des données. D'autres recherches gouvernementales sur les algorithmes ont eu tendance à cibler davantage leur potentiel dans l'économie numérique que les questions de responsabilité (Bureau de la concurrence du Canada, 2018b). La recherche universitaire est également moins active au Canada qu'ailleurs; seul le Citizen Lab de l'Université de Toronto surveille indirectement les questions de responsabilité

algorithmique. De façon prometteuse, les organisations de la société civile ont commencé à combler la lacune en créant de nouveaux groupes, comme le Digital Justice Lab, Tech Reset et OpenMedia, chargés de se pencher sur l'intérêt du public pour la prise de décisions algorithmiques.

En comparaison, l'Union européenne est devenue un leader mondial évident en matière de responsabilité algorithmique depuis l'adoption du Règlement général sur la protection des données (RGPD) en 2018. Le RGPD limite le recours à la prise de décisions algorithmiques en accordant aux individus « le droit de ne pas être soumis à une décision fondée uniquement sur un traitement automatisé » lorsque ces décisions ont des effets significatifs sur eux. Comme nous l'avons vu précédemment, le RGPD peut aussi donner aux particuliers le droit d'obtenir une explication des décisions automatisées.

Les États-Unis continuent d'étudier la responsabilité algorithmique, bien qu'il n'y ait pas encore eu de changements législatifs majeurs. La Federal Trade Commission examine les effets des technologies émergentes pour la concurrence et la protection des consommateurs, et sa série d'audiences en cours sur la concurrence et la protection des consommateurs au XXI^e siècle a récemment prévu une séance sur les algorithmes, l'intelligence artificielle et l'analyse prédictive (voir www.ftc.gov/policy/hearings-competition-consumer-protection). À l'échelon municipal, la Ville de New York a créé un groupe de travail sur les systèmes de décision automatisés en 2018. À l'extérieur du gouvernement, des établissements de recherche telles que Data and Society, AI Now, le Center for Information Policy de l'Université de Princeton, le Center for Internet and Society de la Stanford Law School et le Berkman Klein Center for Internet and Society de l'Université Harvard contribuent à un volume croissant de travaux sur la responsabilité algorithmique.

Bien que bon nombre de ces progrès soient encourageants, dans l'ensemble, on a accordé moins d'attention à la responsabilité algorithmique dans les actualités et la découverte culturelle. L'UE a lancé un programme axé sur le site Web AlgoAware.eu afin de faire connaître les algorithmes et d'améliorer les connaissances des médias concernant les algorithmes. Depuis 2018, l'UE exige également que les services de vidéo à la demande, tels qu'Amazon et Netflix, veillent à ce que le contenu européen représente au moins 30 % de leur catalogue. Bien qu'il ne s'agisse pas d'une question directe de responsabilité algorithmique, cette règle aura nécessairement une incidence sur les contenus recommandés par ces plateformes. Dans l'intervalle, la désinformation en ligne fait l'objet d'un examen minutieux dans le monde. Des solutions ont été élaborées (telles que le code de bonne pratique de l'UE sur la désinformation, un accord d'autorégulation entre les acteurs de l'industrie), mais étant donné leur nouveauté, elles doivent encore être réellement mises à l'essai. Bien que certains de ces éléments ne touchent qu'indirectement au concept de responsabilité algorithmique, ils n'en offrent pas moins des exemples dont on peut tirer des leçons et des outils à utiliser.

En prenant des mesures pour améliorer la responsabilité algorithmique dans les secteurs de la culture et des actualités, le Canada peut contribuer de façon significative à ce projet de politique mondiale.

Bibliographie

- AI Now Institute. (2018). Algorithmic accountability policy toolkit. Consulté à l'adresse <https://ainowinstitute.org/aap-toolkit.pdf>
- Ananny, M. (2016). Toward an ethics of algorithms. *Science, Technology & Human Values*, 41(1), p. 93-117.
- Ananny, M. et K. Crawford (2016). Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media & Society*, 20(3), p. 973-989.
- Andrews, T. M. (9 juillet 2018). Billboard's charts used to be our barometer for music success. Are they meaningless in the streaming age? *Washington Post*. Consulté à l'adresse <https://www.washingtonpost.com/news/arts-and-entertainment/wp/2018/07/05/billboards-charts-used-to-be-our-barometer-for-music-success-are-they-meaningless-in-the-streaming-age/>
- Angwin, J. et S. Mattu (20 septembre 2016). Amazon says it puts customers first. But its pricing algorithm doesn't. *ProPublica*. Consulté à l'adresse <https://www.propublica.org/article/amazon-says-it-puts-customers-first-but-its-pricing-algorithm-doesnt>
- Bolukbasi, T., K.-W.Chang, J. Zou, V. Saligrama, et A. Kalai (2016). Man is to computer programmer as woman is to homemaker? Debiasing word embeddings. *ArXiv*. Consulté à l'adresse <https://arxiv.org/abs/1607.06520>
- Bucher, T. (2016). The algorithmic imaginary: Exploring the ordinary affects of Facebook algorithms. *Information, Communication & Society*, 20(1), p. 1-15.
- Bucher, T. (2018). *If... then: Algorithmic power and politics*. New York: Oxford University Press.
- Bureau de la concurrence Canada (2018a). Mégadonnées et innovation : les grands thèmes de la politique en matière de concurrence au Canada. Consulté à l'adresse <http://www.competitionbureau.gc.ca/eic/site/cb-bc.nsf/fra/04342.html>
- Bureau de la concurrence Canada (2018b). Plan annuel 2018-2019 : Instaurer la confiance pour promouvoir la concurrence sur le marché. Consulté à l'adresse <http://www.bureaudelaconcurrence.gc.ca/eic/site/cb-bc.nsf/fra/04356.html>
- Caplan, R., J. Donovan, L. Hanson et J. Matthews (2018). *Algorithmic accountability: A primer*. Data & Society Research Institute. Consulté à l'adresse <https://datasociety.net/output/algorithmic-accountability-a-primer/>
- Chen, A. (23 mai 2018). How computers misunderstand the world. *The Verge*. Consulté à l'adresse <https://www.theverge.com/2018/5/23/17384324/meredith-broussard-artificial-unintelligence-technology-criticism-technochauvinism>
- Cheney-Lippold, J. (2017). *We are data: Algorithms and the making of our digital selves*. New York: New York University Press.
- Cherubini, F. et R. K. Nielsen (2016). Editorial analytics: How news media are developing and using audience data and metrics. Reuters Institute for the Study of Journalism. Consulté à l'adresse <http://reutersinstitute.politics.ox.ac.uk/publication/editorial-analytics-how-news-media-are-developing-and-using-audience-data-and-metrics>
- Copeland, E. (20 février 2018). 10 principles for public sector use of algorithmic decision making. Nesta. Consulté à l'adresse <https://www.nesta.org.uk/blog/10-principles-for-public-sector-use-of-algorithmic-decision-making/>
- Courtland, R. (2018). Bias detectives: The researchers striving to make algorithms fair. *Nature*, 558, 357. <https://doi.org/10.1038/d41586-018-05469-3>
- Crawford, K. (1^{er} avril 2013). The hidden biases in big data. *Harvard Business Review*. Consulté à l'adresse <https://hbr.org/2013/04/the-hidden-biases-in-big-data>
- Deeth, D. (7 décembre 2015). Over 70% of North American traffic is now streaming video and audio. Consulté à l'adresse <https://www.sandvine.com/press-releases/blog/sandvine-over-70-of-north-american-traffic-is-now-streaming-video-and-audio>
- Diakopoulos, N. (2014). Algorithmic accountability reporting: On the investigation of black boxes. Tow Center for Digital Journalism. Consulté à l'adresse <http://towcenter.org/algorithmic-accountability-2/>

- Diakopoulos, N., S. Friedler, M. Arenas, S. Barocas, M. Hay, B. Howe, H. Jagadish, K. Unsworth, A. Sahuguet, S. Venkatasubramanian, C. Wilson, C. Yu et B. Zevenbergen (n.d.). Principles for accountable algorithms and a social impact statement for algorithms. Consulté à l'adresse <http://www.fatml.org/resources/principles-for-accountable-algorithms>
- Dubois, E. et G. Blank (2018). The echo chamber is overstated: The moderating effect of political interest and diverse media. *Information, Communication & Society*, 21(5), p. 729–745.
- Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. New York: St. Martin's Press.
- Even, A. D. et J. Powles (2017). Meaningful information and the right to explanation. *International Data Privacy Law*, 7(4), p. 233–242.
- Fletcher, R. et R. K. Nielsen (2018). Automated serendipity. *Digital Journalism*, 6(8), p. 976–989.
- Gill, N. et P. Hall (2018). *An introduction to machine learning interpretability*. Oakville, ON: O'Reilly Media, Inc.
- Glaser, A. (5 décembre 2018). Protests inside Google are growing—and now its massive “shadow workforce” has demands, too. *Slate*. Consulté à l'adresse <https://slate.com/business/2018/12/google-walkout-protests-contract-workers.html>
- Henderson, F. (2017). Software engineering at Google. *ArXiv*. Consulté à l'adresse <https://arxiv.org/pdf/1702.01715.pdf>
- Introna, L. D. et H. Nissenbaum (2000). Shaping the web: Why the politics of search engines matters. *Information Society*, 16(3), p. 169–185.
- Isaac, M. (22 décembre 2017). How Uber deceives the authorities worldwide. *New York Times*. Consulté à l'adresse <https://www.nytimes.com/2017/03/03/technology/uber-greyball-program-evade-authorities.html>
- Jannach, D., M. Zanker, A. Felfernig et G. Friedrich (2010). *Recommender systems: An introduction*. Cambridge, Royaume-Uni : Cambridge University Press.
- Kalogeropoulos, A., S. Negro, I. Picone et R. K. Nielsen (2017). Who shares and comments on news? A cross-national comparative analysis of online and social media participation. *Social Media + Society*. <https://doi.org/10.1177/2056305117735754>
- Karlin, M. (18 mars 2018). A Canadian algorithmic impact assessment. Consulté à l'adresse <https://medium.com/@supergouvernance/a-canadian-algorithmic-impact-imeessment-128a2b2e7f85>.
- Keller, M. H. (11 août 2018). The flourishing business of fake YouTube views. *New York Times*. Consulté à l'adresse <https://www.nytimes.com/interactive/2018/08/11/technology/youtube-fake-view-sellers.html>
- Langlois, G. (2014). *Meaning in the age of social media*. New York: Palgrave Macmillan.
- Lewis, M. (2014). *Flash Boys*. New York: W. W. Norton & Company.
- McKelvey, F. (2015). Openness compromised? Questioning the role of openness in digital methods and contemporary critical praxis. Dans G. Elmer, G. Langlois et J. Redden (dir. publ.), *Compromised data: From social media to Big Data* (pp. 126–146). New York, NY: Bloomsbury Academic.
- McKelvey, F. (2018). *Internet daemons: Digital communications possessed*. Minneapolis: University of Minnesota Press.
- McKelvey, F. et R. Hunt (à paraître). Discoverability: Toward a definition of content discovery through platforms. *Social Media + Society*.
- Mittelstadt, B. D., P. Allo, M. Taddeo, S. Wachter et L. Floridi (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, 3(2). <https://doi.org/10.1177/2053951716679679>
- Montgomery, B. (13 septembre 2018). Fans are spoofing Spotify with “fake plays,” and that’s a problem for music charts. *BuzzFeed News*. Consulté à l'adresse <https://www.buzzfeednews.com/article/blakemontgomery/spotify-billboard-charts>

- Neff, G. et D. Stark (2004). Permanently Beta: Responsive Organization in the Internet Era. Dans P. N. Howard et S. Jones (dir. publ.) (p. 173-188). Thousand Oaks: SAGE Publications.
- Nissenbaum, H. (1994). Computing and accountability. *Communications of the ACM*, 37(1), p. 72-80.
- Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. New York: New York University Press.
- O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. New York: Crown.
- Orphanides, K. G. (23 mars 2018). Children's YouTube is still churning out blood, suicide and cannibalism. *Wired UK*. Consulté à l'adresse <https://www.wired.co.uk/article/youtube-for-kids-videos-problems-algorithm-recommend>
- Overdorf, R., B. Kulynych, E. Balsa, C. Troncoso et S. Gürses (2018). POTs: Protective optimization technologies. *ArXiv*. Consulté à l'adresse <https://arxiv.org/abs/1806.02711>
- Pariser, E. (2011). *The filter bubble: How the new personalized web is changing what we read and how we think*. New York: Penguin Press.
- Pasquale, F. (2015). *The black box society: The secret algorithms that control money and information*. Cambridge: Harvard University Press.
- Perrin, A. et J. Jiang (14 mars 2018). A quarter of Americans are online almost constantly. Pew Research Center. Consulté à l'adresse <http://www.pewresearch.org/fact-tank/2018/03/14/about-a-quarter-of-americans-report-going-online-almost-constantly/>
- Portugal, I., P. Alencar et D. Cowan (2018). The use of machine learning algorithms in recommender systems: A systematic review. *Expert Systems with Applications*, 97, p. 205-227. <https://doi.org/10.1016/j.eswa.2017.12.020>
- Powles, J. et H. Nissenbaum (7 décembre 2018). Le détournement séduisant du biais de « résolution » dans l'intelligence artificielle. Consulté à l'adresse <https://medium.com/s/story/the-seductive-diversion-of-solving-bias-in-artificial-intelligence-890df5e5ef53>
- Reisman, D., J. Schultz, K. Crawford et M. Whittaker (2018). *Algorithmic impact assessments: A practical framework for public agency accountability*. AI Now Institute. Consulté à l'adresse <https://ainowinstitute.org/aiareport2018.pdf>
- Rieder, B., A. Matamoros-Fernández et Ò. Coromina (2018). From ranking algorithms to 'ranking cultures': Investigating the modulation of visibility in YouTube search results. *Convergence*, 24(1), p. 50-68.
- Rosenblat, A. (2018). *Uberland: How algorithms are rewriting the rules of work*. Berkeley: University of California Press.
- Schrock, A. R. (2016). Civic hacking as data activism and advocacy: A history from publicity to open government data. *New Media & Society*, 18(4), p. 581-599.
- Selbst, A. D., et Powles, J. (2017). Meaningful information and the right to explanation. *International Data Privacy Law*, 7(4), p. 233-242.
- Smith, B. et G. Linden (juin 2017). Two decades of recommender systems at Amazon.com. *IEEE Internet Computing*, 21(3), p. 12-18.
- Sunstein, C. R. (2007). *Republic.com 2.0*. Princeton: Princeton University Press.
- Sutton, K. (19 octobre 2018). Facebook video ad metric lawsuit prompts publishers to revisit the 'pivot to video.' *AdWeek*. Consulté à l'adresse <https://www.adweek.com/digital/facebook-video-ad-metric-lawsuit-prompts-publishers-to-revisit-the-pivot-to-video/>
- Taylor, A. (2 octobre 2018). The automation charade. *Logic*. Consulté à l'adresse <https://logicmag.io/05-the-automation-charade/>
- Tkacz, N. (2012). From open source to open government: A critique of open politics. *Ephemeria: Theory and Politics in Organization*, 12(4), p. 386-405.
- Toff, B. et R. K. Nielsen (2018). « I just Google it »: Folk theories of distributed discovery. *Journal of Communication*, 68(3), p. 636-657.
- Tran, K. (16 août 2017). Nielsen adds Facebook, YouTube, and Hulu to digital ratings. *Business Insider*. Consulté à l'adresse <https://www.businessinsider.com/nielsen-adds-facebook-youtube-and-hulu-to-digital-ratings-2017-8>

Tripodi, F. (2018). Searching for alternative facts. Data & Society Research Institute. Consulté à l'adresse <https://datasociety.net/output/searching-for-alternative-facts/>

Tufekci, Z. (8 juin 2018). YouTube, the great radicalizer. *New York Times*. Consulté à l'adresse <https://www.nytimes.com/2018/03/10/opinion/sunday/youtube-politics-radical.html>

USACM. (12 janvier 2017). Statement on algorithmic transparency and accountability. Association for Computing Machinery US Public Policy Council. Consulté à l'adresse https://www.acm.org/binaries/content/assets/public-policy/2017_usacm_statement_algorithms.pdf

Vranica, S. et J. Marshall (22 septembre 2016). Facebook overestimated key video metric for two years. *Wall Street Journal*. Consulté à l'adresse <https://www.wsj.com/articles/facebook-overestimated-key-video-metric-for-two-years-1474586951>

Waldie, P. et T. Kiladze (5 décembre 2018). Facebook gave RBC, other companies preferential access to users' data, documents show. *Globe and Mail*. Consulté à l'adresse <https://www.theglobeandmail.com/world/article-facebook-gave-rbc-other-companies-preferential-access-to-users-data/>

Weinberger, D. (28 janvier 2018). Optimization over explanation. Consulté à l'adresse <https://medium.com/berkman-klein-center/optimization-over-explanation-41ecb135763d>

Winning, L. (13 mars 2018). It's time to prioritize diversity across tech. *Forbes*. Consulté à l'adresse <https://www.forbes.com/sites/lisawinning/2018/03/13/its-time-to-prioritize-diversity-across-tech/>