

Les algorithmes aux commandes

par Jérémie Poiroux

Quel raisonnement aboutit à la recommandation d'un morceau, d'un film ou d'une personne à rencontrer ? Puisqu'un algorithme est un système dont le code est seulement le reflet, il est essentiel de comprendre les ingénieurs et les organisations qui ont contribué à sa conception.

Qu'est-ce qu'un algorithme ?

Les algorithmes¹ sont un objet d'étude grandissant en sociologie et en sciences de l'information et de la communication. Leur utilisation croissante par les entreprises et les administrations publiques (notamment sous le terme « intelligence artificielle ») a suscité la curiosité médiatique et académique. La question n'est plus seulement technique (par exemple, comment faire mieux et plus vite ?) mais également sociale : quels sont les effets des algorithmes ? Qu'implique l'automatisation du recrutement, de la distribution et l'organisation du travail, de la validation de demandes de crédit ou encore de la prédiction des récidives ? De ce fait, les algorithmes sont parfois vus comme des objets obscurs et suscitent craintes, fantasmes et espoirs.

Mais qu'est-ce qu'un algorithme ? Définir le terme dans un cadre sociologique demanderait un article en soi, et certains chercheurs se sont déjà penchés sur cette

¹ Cet essai a été partiellement écrit dans le cadre du projet Algodiv (ANR-15-CE38-0001) financé par l'Agence Nationale pour la Recherche et mené par Camille Roth que je remercie pour ses conseils et relectures.

question complexe². Nous avancerons simplement ici qu'un algorithme ne peut pas être réduit à une recette ou à un programme informatique. En plus d'être du code, des données et des instructions, un algorithme comprend des intentions et des stratégies entrepreneuriales ou politiques³. En 2011, Eli Pariser développe dans son livre le concept de *filter bubble* (« bulle de filtre » ou « bulle filtrante »)⁴ largement repris en 2016 lors du Brexit et de l'élection présidentielle américaine. Les algorithmes de Google ou de Facebook (pour ne citer qu'eux) provoqueraient un phénomène de renforcement de l'existant, voire enfermeraient les personnes dans leur zone de confort informationnelle en se basant sur des critères dits personnalisés tels que l'historique de navigation ou la position géographique. Cette bulle de filtre pourrait être considérée comme un effet indirect d'un des objectifs de Google ou Facebook ; que les résultats correspondent aux personnes et les incitent à continuer à utiliser le service.

Un second exemple a défrayé la chronique en France en 2017 : Admission Post-Bac (APB). Mise en place en 2008, la plateforme d'affectation des lycéens dans des formations du supérieur était critiquée pour son manque de transparence. Malgré la publicisation du code, il était impossible de comprendre les directives du gouvernement derrière les affectations. Ce dernier exemple rappelle que le programme informatique peut être considéré comme une « boîte noire »⁵ (s'il n'est pas libre d'accès et expliqué) qu'il faudrait nécessairement décoder. L'algorithme est un système dont le code est seulement le reflet. Une compréhension des acteurs et des organisations derrière les algorithmes est nécessaire. Cet essai s'efforce de contribuer à une littérature encore émergente : l'étude de la conception des algorithmes et du travail effectué par les ingénieurs.

Les algorithmes de recommandation

L'adjectif « recommandation » apporte assez de précision pour qualifier les « algorithmes de recommandation » comme objet de recherche. Cet essai se base sur

² Rob Kitchin, « Thinking critically about and researching algorithms », *Information, Communication & Society*, 2017, vol. 20, n° 1, p. 14-29.

³ Jenna Burrell, « How the machine 'thinks': Understanding opacity in machine learning algorithms », *Big Data & Society*, 2016, vol. 3, n° 1.

⁴ Eli Pariser. *The filter bubble: What the Internet is hiding from you*, Penguin UK, 2011.

⁵ Frank Pasquale, *The black box society*. Harvard University Press, 2015.

une enquête de trente entretiens avec des ingénieurs menée entre mars 2017 et juin 2018. Les algorithmes développés par ceux-ci sont notamment utilisés pour nous faire découvrir de la musique, des vidéos, des livres, nous conseiller des « objets similaires » à acheter, à consommer de l'information pertinente et de qualité ou encore pour nous faire rencontrer des personnes. Grâce aux données de navigation et aux préférences des utilisateurs, les algorithmes de recommandation automatisent la prescription. Leur but est de prédire nos goûts et envies pour s'y adapter le mieux possible. Dans ce cadre, une des problématiques majeures est, comme nous l'avons vu, de formuler des recommandations qui soient convaincantes pour l'utilisateur sans tomber dans un enfermement informationnel (ce que nous consommons) ou interactionnel (avec qui nous échangeons)⁶. Avant de s'intéresser à ses méthodes, prenons une situation illustrant les subtilités de la recommandation musicale impliquant soit un disquaire, soit un algorithme. La seule différence notable est ici la taille du catalogue connu par le disquaire (les vinyles en rayon) et l'algorithme (la base de données du service de streaming musical). À l'instar du disquaire, l'algorithme va au bout d'un certain temps d'écoute vous connaître musicalement et commencera à vous prescrire des albums ou des chansons. L'art de la recommandation musicale consiste, entre autres, à proposer d'une part de la musique que l'on connaît, que ce soit du même artiste, du même album, voire les mêmes chansons à nouveau (ce qui est moins valable lorsqu'on recommande des films) et d'autre part de la musique que l'on ne connaît pas mais qui nous plairait, sans s'éloigner trop de vos goûts musicaux. Pour le dire autrement et à l'aide d'un raisonnement par l'absurde, la meilleure recommandation est par définition une chanson qui a déjà été appréciée.

Des recommandations intrinsèques et extrinsèques

Dans le cadre d'une plateforme en ligne, un système de recommandation propose des objets à des utilisateurs en fonction des propriétés de ces objets et/ou de la consommation des utilisateurs. Gardons l'exemple d'une plateforme de streaming musical. Pour retenir les utilisateurs, un de ses objectifs est de faire découvrir de la musique à ses utilisateurs, comme le rappelle cet ingénieur : « Une entreprise qui fait du streaming doit faire découvrir de la musique aux gens. Pas seulement pour une raison d'éthique

⁶ Frederik Zuiderveen Borgesius, et al., « Should we worry about filter bubbles? », *Internet Policy Review. Journal on Internet Regulation*, 2016, vol. 5, n° 1.

mais aussi, car si les gens tournent sur leurs quatre albums favoris, ils n'ont aucune raison de payer dix euros par mois. Ils ont juste à acheter leurs mp3 ».

Pour cela, la plateforme peut recommander des albums ou des titres grâce à deux méthodes.

La première est la recommandation intrinsèque qui repose sur les qualités des objets. Vous recommander un album des Rolling Stones après l'écoute des Beatles fait sens, car les deux groupes étaient actifs dans les années 1960 et peuvent être classés en pop/rock. La plateforme pourrait aussi vous proposer d'écouter *Dancing In The Street*, une chanson qui réunit Mick Jagger et David Bowie (1985). La méthode *content-based* peut aller plus loin, et trouver des similarités (car c'est de cela qu'il s'agit) dans le rythme, les instruments ou encore les émotions véhiculées par les musiques. Pour cela, les ingénieurs analysent le signal audio pour quantifier et catégoriser la musique. Avec les dernières évolutions du *machine learning* (et plus précisément des techniques utilisant des réseaux de neurones), le traitement du signal a gagné en performance et peut de nouveau faire concurrence à la seconde méthode, le « filtrage collaboratif » qui permet des recommandations extrinsèques.

Comme son nom l'indique, le filtrage collaboratif permet de tirer parti des coopérations explicites (mettre une note, aimer la musique) et implicites (passer à la piste suivante, effectuer une recherche dans le catalogue) ainsi que des schémas de consommation des utilisateurs pour effectuer des recommandations. Avec cette méthode, il est possible de recommander un album de Lady Gaga après l'écoute des Rolling Stones (dont la musique a intrinsèquement peu de propriétés communes), simplement parce qu'une certaine part d'utilisateurs écoutent Lady Gaga et les Rolling Stones dans la même session (peu importe la raison). Certaines habitudes d'écoute sont d'ailleurs liées à des événements ou des jours particuliers et peuvent servir à proposer des recommandations aux utilisateurs. Pour l'anecdote, Paul Lamere, *data scientist* chez Spotify, a identifié ce qu'il appelle l'anomalie Aerosmith⁷ : une des chansons les plus célèbres du groupe, *I don't want to miss a thing*, est écoutée significativement plus à certains jours précis (sur une période de deux ans, voir image 1). Après avoir écarté la piste de « la boom de la Saint-Valentin » (qui se trouve être une coïncidence sur les dates), il a identifié (voire image 2) que l'important volume d'écoute correspondait à chaque fois à une actualité impliquant une comète. Pourquoi ? C'est que la chanson fait partie de la bande originale du film *Armageddon*

⁷ RailsConf 2016 - Closing Keynote: Paul Lamere : <https://www.youtube.com/watch?v=f7XN3RuDzmc>

dans lequel Bruce Willis doit sauver la Terre d'un impact de comète. Il est très probable, pour en revenir au filtrage collaboratif, que les algorithmes aient détecté ce schéma d'écoute et recommandent *I don't want to miss a thing* lors d'une actualité « comète » (Paul Lamere ne précise pas si c'est effectivement le cas).

Pour la première méthode basée sur les contenus, la plateforme a besoin d'un maximum d'informations sur les musiques de son catalogue. Pour la seconde basée sur le filtrage collaboratif, la plateforme doit en savoir le plus possible sur ses utilisateurs et leur consommation. Cette course aux signaux explicites (ce que l'on demande aux utilisateurs) et implicites (leurs traces) soulève un certain nombre de problématiques que nous allons présenter ci-après.

Pourquoi parler de représentation ?

Plus l'algorithme en sait sur les utilisateurs, meilleures seront ses recommandations. On peut considérer qu'une bonne recommandation mène à un acte de consommation de la part de l'utilisateur. Par conséquent, pour notre plateforme de streaming musical, l'utilisateur continuera son abonnement. Ainsi, rares sont les algorithmes de recommandation reposant uniquement sur la méthode dite *content-based* : un des objectifs des ingénieurs est d'améliorer la quantité et la qualité de la connaissance des utilisateurs. S'opère alors une représentation qui recouvre plusieurs réalités. La représentation⁸ se définit comme (i) « donner une illustration, constituer un exemple de, être représentatif de » ou être « l'action, [le] fait d'agir ou de parler au nom de quelqu'un ». Elle peut également être (ii) « l'action de rendre quelque chose ou quelqu'un présent sous la forme d'un substitut ou en recourant à un artifice ; ce substitut ». Enfin, c'est (iii) « l'action, [le] fait de se représenter quelque chose ; [la] manière dont on se représente quelque chose ». Dans le cadre de la fabrique des algorithmes de recommandation, nous avons affaire aux représentations suivantes : la représentation de tous les utilisateurs par un groupe d'utilisateurs (représentativité)⁹ et la représentation comme image que se font les ingénieurs des utilisateurs (spéculation)¹⁰. Pour finir, nous nous pencherons sur ce qui nous semble le plus novateur avec les algorithmes de recommandation : la catégorisation implicite et en temps réel des utilisateurs. Dans ces trois cas, il est important de souligner dès à

⁸ Les définitions proviennent du dictionnaire en ligne Ortolang : ww.cnrtl.fr/definition/representation

⁹ Hanna F. Pitkin, *The concept of representation*, University of California Press, 1967.

¹⁰ Pierre Mannoni. *Les représentations sociales, Que sais-je? n° 3329*, PUF, 2016.

présent que les utilisateurs des plateformes de streaming musical délèguent une partie de leur découverte musicale aux algorithmes. S'installe ici une relation de « représentation » des utilisateurs par les algorithmes, en revisitant un extrait de « La représentation politique » dans lequel Hanna F. Pitkin explique du représentant que : « Son action [à l'algorithmes] doit exiger de la sagesse et du jugement ; il doit être celui qui agit [recommande]. Les représentés [les utilisateurs] aussi doivent être [...] capables d'agir et de juger de façon indépendante. »¹¹

Pour aller plus loin, rappelons ce qu'avance Michael Saward¹², en rupture avec le propos de Hanna F. Pitkin : les utilisateurs d'un service (ici, de streaming musical) existent uniquement car ils sont « représentés » par l'algorithme de recommandation. Le parallèle avec les différents idéaux-types de la représentation politique proposés par Virginie Dutoya et Samuel Hayat¹³ prend alors tout son sens avec l'opposition entre la « représentation comme imposition » et la « représentation comme proposition ». Celle-ci nous permet de développer la deuxième partie de cet essai avec les éléments suivants en tête : dans le cadre d'une « représentation comme imposition » ce qui est effectivement le cas de la majorité des services utilisant un algorithme de recommandation, « les prétentions à la représentation imposent au représenté [à l'utilisateur] son identité [son profil créé à partir des traces d'utilisation] » tandis qu'une « représentation comme proposition » prétend proposer « au représenté [à l'utilisateur] des représentations de lui-même [son profil] qu'il peut accepter, modifier ou contester », ce qui fait écho aux questions de discrimination et de transparence que nous aborderons dans la dernière partie.

Différentes représentations

La représentation comme représentativité

¹¹ Hanna F. Pitkin, trad. Samuel Hayat, « La représentation politique », *Raisons politiques*, 2013, n°2, p. 35-51.

¹² Michael Saward, « The representative claim », *Contemporary political theory*, 2006, vol. 5, n°3, p. 297-318.

¹³ Virginie Dutoya, Samuel Hayat, « Prétendre représenter », *Revue française de science politique*, 2016, vol. 66, n°1, p. 7-25.

Le sens le plus évident de la représentation dans le cadre de la fabrique des algorithmes de recommandation est la représentativité, au sens où représenter, c'est donner une illustration, constituer un exemple de, être représentatif de.

À l'instar des méthodes traditionnelles d'étude de marché et de marketing, les ingénieurs se servent de panels qui sont constitués d'un nombre limité d'utilisateurs. Ceux-ci sont, de différentes manières, représentatifs de tous les utilisateurs. La première manière est celle des *focus groups*. Les ingénieurs invitent physiquement des utilisateurs sélectionnés pour tester le produit et/ou l'algorithme de recommandation. Sans se déplacer, ces derniers peuvent également contribuer en tant que « bêta-testeurs », en remplissant des questionnaires ou – d'après un ingénieur – en envoyant des commentaires via un chat sur WhatsApp. Ces méthodes peuvent donner des résultats quantitatifs et qualitatifs explicites et approfondis, mais elles ne sont pas systématiquement mises en place : seulement neuf ingénieurs (sur trente) les mentionnent.

Outre ces méthodes traditionnelles, les ingénieurs se servent d'une approche très répandue dans le milieu du développement web : le test. Contrairement au *focus group* et au questionnaire, le test peut être déployé en temps réel¹⁴. Il va d'ailleurs plutôt servir à implémenter des fonctionnalités, des améliorations à la volée ou à valider des choix, tandis que les deux premières méthodes vont permettre de faire évoluer le cœur du produit ou de l'algorithme. Lorsque les ingénieurs parlent de « test », ils évoquent le plus souvent le test A/B, qui consiste à vérifier laquelle de la version A ou B permet d'atteindre les objectifs (*Key Performance Indicators*) fixés. Les tests A/B permettent par exemple de confronter une ancienne contre une nouvelle version d'un algorithme ou deux nouvelles versions (d'une nouvelle fonctionnalité).

Si, pour une plateforme de streaming de musique, l'objectif est par exemple d'augmenter la durée moyenne d'une session d'écoute, un test A/B peut permettre d'évaluer deux nouvelles versions d'un algorithme de recommandation grâce à deux panels d'utilisateurs distincts, ou encore de proposer « *différentes playlists qui sont des variations du même algorithme* ». À cheval entre les méthodes explicites et les statistiques d'utilisations implicites (que nous aborderons dans la dernière partie), le test est la pièce maîtresse du développement et de l'amélioration d'un algorithme de recommandation – certains ingénieurs doivent nécessairement réaliser un test pour procéder à un « *changement qui touche les consommateurs* ». En pratique, cela peut se

¹⁴ Zeynep Tufekci, « Engineering the public: Big data, surveillance and computational politics », *First Monday*, 2014, vol. 19, n°7.

traduire par un « *premier test sur 1 % des utilisateurs* », et un lancement qui « *ne se fait jamais sans un test sur 10 % des utilisateurs* ». Ces usages révèlent une place surprenante du test dans le processus de développement. En effet, il semble moins mobilisé par les ingénieurs afin les aider à prendre les décisions les plus difficiles ou pour lancer les améliorations les plus importantes, qu'intégré à part entière dans ledit processus. D'une part, le test semble jouer le rôle d'un acteur-décisionnaire par les ingénieurs. D'autre part, il a dans certains cas le « dernier mot » sur les modifications proposées ; du point de vue des ingénieurs, il applique la décision finale des utilisateurs.

La représentation comme spéculation

Bien que les ingénieurs aient à disposition un catalogue de méthodes limitant les ambiguïtés et les erreurs d'interprétations (les *focus groups*, les groupes WhatsApp, les questionnaires, les tests ainsi que les analyses des statistiques d'utilisation), ils reposent également leur développement sur des idées reçues. En d'autres termes, les ingénieurs spéculent sur le comportement et les envies des utilisateurs – et disent d'ailleurs être « *capables d'imaginer ce que les gens ont apprécié ou pas* ».

À l'inverse des tests A/B dont le déploiement aussi précis est propre au numérique (et qui peut s'apparenter à une procédure quasiment scientifique), ces spéculations se retrouvent dans tous les domaines et n'est pas nouvelle. Nous pourrions d'ailleurs également parler de démarche « empathique » consciente ou non, où les ingénieurs se représentent les utilisateurs selon leur activité. Par exemple « *pour évaluer la qualité de l'expérience musicale* », des ingénieurs se basent sur la durée de la session d'écoute. Ce positionnement intuitif n'est d'ailleurs jamais référencé. Il obstrue parfois un développement fondé sur les données et les méthodes citées ci-dessus : « *À vouloir faire plein de choses avec des critères qu'on trouvait à droite à gauche, on s'est retrouvé dans la situation où les gens nous disaient "si vous voulez que ça me corresponde, pourquoi vous ne me demandez pas ?"* ».

Les ingénieurs font finalement souvent confiance à leur propre intuition sur les envies et les comportements des utilisateurs. Puisque « *les utilisateurs ne savent pas non plus toujours ce qui est important pour eux ou pensent le savoir mais ce n'est pas le cas* », les ingénieurs se retrouvent parfois à généraliser des résultats. Ils jouent ici un numéro d'équilibriste ; ils doivent jauger l'importance à donner aux différents résultats, signaux faibles et forts ainsi qu'à leurs intuitions : « *Si on s'aperçoit que ce client est en train de préparer son voyage et qu'il a déjà fait des réservations dans deux endroits, on va lui*

proposer une troisième destination en tête de page. Comme tout le monde a envie de la même chose, dans 9 cas sur 10, ça marche » (1).

Malgré la quantité de données que les ingénieurs ont à leur disposition, la fabrique d'un algorithme de recommandation se base souvent, sous couvert des envies et des besoins des utilisateurs, sur la propre consommation de ces ingénieurs. Le risque étant de tomber dans un cercle vicieux où ceux-ci testent leurs intuitions plutôt que des améliorations à des problèmes identifiés.

La représentation comme catégorisation implicite

En plus des méthodes considérées ci-dessus comme explicites et des idées reçues qu'ils ont a priori, les ingénieurs doivent traiter les statistiques d'utilisation. Celles-ci sont un calcul des traces et fournissent aux ingénieurs des informations sur l'utilisation du service. Ces traces sont des signaux implicites bruts qui peuvent faire appel, ici aussi, à l'intuition des ingénieurs pour les interpréter. C'est par exemple le cas du *feed-back* envoyé par un utilisateur lorsqu'il « *n'écoute pas la piste de musique jusqu'au bout* ».

Pour un service de streaming de musique en ligne, on rencontre deux catégories de statistiques reposant sur les traces d'utilisation, rendant les utilisateurs « présents sous la forme d'un substitut ». La première regroupe les statistiques de navigation et de collecte des données, par exemple sur la durée de connexion, les liens cliqués, les recherches effectuées, les albums, playlists, albums ou catégories parcourues. La seconde catégorie relève davantage des statistiques d'écoute, à savoir par exemple les musiques et albums joués, rejoués, mis en pause, passés, mis en favori ou encore les playlists créées. Il est intéressant de soulever ici que cette collecte de données se fait a priori et en continu, sans nécessairement avoir d'objectif précis : trouver du sens aux données et « raffiner ce pétrole brut » est d'ailleurs une des principales missions d'un *data scientist*.

Si la première catégorie sert plutôt à améliorer le service au sens large (le design d'interface, l'expérience utilisateur, les vitesses d'exécution, etc.), la seconde est primordiale pour les algorithmes de recommandation qui peuvent les utiliser à la volée, proposant par exemple du contenu sur « *les dix dernières destinations à la mode ou Neymar* », car ils ont « *détecté que vous cliquiez souvent sur les articles foot et voyage* ».

Ce maniement des traces par les algorithmes se fait, à l'inverse des autres méthodes, sans intervention humaine significative et leur permet de créer de nouvelles

catégories implicites ¹⁵ éloignées des catégories socio-démographiques et géographiques traditionnelles¹⁶ :

« Plus tu rajoutes de la data, plus tu rajoutes de la personnalisation. Notre but c'était de recommander aux gens des lieux qui leur correspondaient, donc nos algorithmes catégorisaient l'utilisateur en fonction de qui il est, du quartier et des endroits qu'il aime bien pour conditionner l'apparition des lieux ».

Cet extrait est intéressant par son paradoxe sur lequel nous revenons peu après : la collecte (massive) de données est censée améliorer la recommandation par une personnalisation toujours plus fine, autrement dit un « hyper-ciblage »¹⁷, mais en même temps la catégorisation des utilisateurs reste basique (leurs goûts et caractéristiques socio-démographiques).

Historiquement et avant la montée en puissance des algorithmes de recommandation, les techniques de vente se basent majoritairement sur les catégories explicites pour augmenter le chiffre d'affaires. Les cibles sont clairement identifiées en fonction du produit vendu et surtout se recourent avec les catégories socio-démographiques comme le sexe, l'âge ou la profession et la catégorie socioprofessionnelle (PCS). La segmentation du marché se fait en fonction de ces catégories explicites renseignées par les personnes elles-mêmes via l'INSEE (en France)¹⁸.

Aujourd'hui, les systèmes de recommandation permettent (et pas seulement pour des services en ligne) de non plus partir de ces informations explicites offrant une « vision stable et rassurante », mais de créer des profils éphémères en continu¹⁹ en fonction des traces de consommation, à l'instar des catégories d'intérêt dans le marketing. Un site de streaming de musique ne recommandera pas Lady Gaga à une personne « par convention », car celle-ci est une jeune femme habitant dans une grande ville, mais parce que sa consommation (ou sa catégorie d'intérêt) induit une probabilité forte qu'elle aime Lady Gaga. Les consommateurs-types ne sont plus

¹⁵ Cathy O'Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, Crown, 2016.

¹⁶ T.B Beane, D.M. Ennis, « La segmentation des marchés : une revue de la littérature », *Recherche et Applications en Marketing*, 1989, vol 4, n° 3, p. 25-52.

¹⁷ Thomas Beauvisage et al., « Fabriquer des cibles dans la publicité programmatique. Enquête sur l'audience planning », 2019, Rapport de recherche, Orange Labs.

¹⁸ Olivier Sautory, « L'échantillon démographique permanent de l'INSEE », *Courrier des Statistiques*, 1987, n°41, p. 1-4.

¹⁹ Dominique Boullier, « Pour des sciences sociales de troisième génération (SS3G) », in Pierre-Michel Menger et Simon Paye (dir.), *Big Data et Traçabilité Numérique : Les Sciences Sociales Face à La Quantification Massive Des Individus*, Paris, Open Editions Press (Collège de France), 2017, p. 163-184.

pertinents, car les données qualifiant un utilisateur (ses écoutes et ses favoris) sont constamment modifiées.

Depuis leur utilisation à grande échelle, les algorithmes soulèvent la question de la personnalisation (ou de l'individualisation) des contenus accessibles et consommés par les utilisateurs²⁰. S'ils sont plutôt encensés pour la recommandation de biens de consommation standards, leur rôle est régulièrement discuté dans le cadre de la personnalisation de l'accès à l'information et aux contenus culturels²¹. Sans rentrer dans le débat, nous souhaitons relativiser cette personnalisation qui n'en porterait que le nom : nous avons plutôt affaire à différents degrés de catégorisation, avec d'une part une catégorisation macro explicite (les caractéristiques socio-démographiques et géographiques) et de l'autre une catégorisation micro implicite au niveau de l'utilisateur (ses interactions avec le service). La personnalisation parfois redoutée ne serait qu'une catégorisation utilisant d'autres critères. Peu importe si nous sommes des femmes ou des hommes, des citadins ou des ruraux, c'est notre consommation de contenus qui nous classifie et au prisme de celle-ci que les services basent leurs recommandations et leur marketing.

Dans le cadre de la recommandation, les seules catégories explicites sont finalement celles qui qualifient les objets recommandés (le rythme d'une musique, par exemple). Les algorithmes de filtrage collaboratif font abstraction de ce que les utilisateurs « sont » pour s'intéresser à ce qu'ils « font » : « On a des pourcentages, on a des ratios de clics et on va calculer la pertinence d'une chanson pour les gens à partir de ça ».

Questions soulevées

Bien que les ingénieurs soient les « fabricants » des algorithmes, leur implication, voire leur responsabilité dans les recommandations effectuées peuvent dans certains cas être limitées. Un ingénieur va même jusqu'à dire qu'il ne peut et ne doit pas forcer l'algorithme à agir d'une manière ou d'une autre (soit à coder en dur une variable), « *l'algorithme décide tout seul en fonction de ses objectifs* ».

²⁰ Antoinette Rouvroy, Thomas Berns, « Gouvernamentalité algorithmique et perspectives d'émancipation », *Réseaux*, 2013, n°1, p. 163-196.

²¹ Seth Flaxman, Sharad Goel, Justin M. Rao, « Filter bubbles, echo chambers, and online news consumption », *Public opinion quarterly*, 2016, vol. 80, n° 1, p. 298-320.

Sans aller jusqu'à aborder les problématiques soulevées par les algorithmes de *deep learning* dont les résultats sont parfois incompréhensibles pour leurs ingénieurs, nous observons que ces derniers laissent volontairement aux algorithmes de recommandation le dernier mot, considérant que « *l'algorithme est lui-même un test, sa fonctionnalité principale étant d'apprendre et de modifier ses paramètres lui-même pour lui permettre d'être plus performant* ».

En s'affranchissant des catégories explicites traditionnelles finalement trop « humaines », les algorithmes proposent leur propre réalité socio-économique. Pourtant, plusieurs études montrent que les algorithmes (ici plus largement, d'affectation, de prise de décision, de prédiction) ont un pouvoir discriminant sur les catégories explicites traditionnelles²². D'ailleurs, Beauvisage et al. soulignent dans leur enquête que les acteurs de la publicité programmatique en ligne utilisent peu (*a fortiori* moins qu'avant) les catégories d'intérêt et considèrent les « micro-cibles comme des exceptions », leur préférant « le socio-démographique » qui a fait ses preuves.

En d'autres termes, malgré la liberté donnée aux algorithmes pour créer des catégories implicites (par exemple, les fans de Lady Gaga ou les experts de tel logiciel de comptabilité), ils reproduisent des biais humains (recommander Lady Gaga à de jeunes femmes ou recruter des hommes blancs).

Pour élargir la réflexion sur ce propos, nous souhaitons proposer deux pistes de discussion. La première est de donner la possibilité aux utilisateurs d'un service de connaître leur « profil algorithmique » : dans quelles catégories implicites, et par ricochet catégories explicites sont-ils classés par l'algorithme ? La seconde est de permettre à ces utilisateurs d'une part de ne pas alimenter l'algorithme avec leurs traces et d'autre part de pouvoir bénéficier, à l'instar du droit à l'oubli entré en vigueur en Europe en 2012, d'un droit à donner un consentement éclairé et explicite pour l'exploitation de leurs traces²³.

Publié dans lavedesidees.fr, le 28 septembre 2021

²² Joseph Turow, *The daily you: how the new advertising industry is defining your identity and your worth*, Yale University Press, 2012. Benjamin G. Edelman, Michael Luca, « Digital discrimination: the case of Airbnb.com », *Harvard Business School NOM Unit Working Paper*, 2014, n° 14-054.

²³ Lilian Edwards and Michael Veale, « Slave to the algorithm: why a right to an explanation is probably not the remedy you are looking for », *Duke L. & Tech Rev.*, 2017, vol. 16, p. 18.