

Les algorithmes de la librairie en ligne Amazon amplifient la désinformation

Description

Sur des sujets comme la santé, l'immigration, le changement climatique ou le droit des femmes, la plateforme en ligne américaine promeut des ouvrages aux contenus mensongers. Publiée en décembre 2023, une étude portant sur les marchés belge et français confirme les risques liés aux failles des algorithmes.

Les recommandations de livres sur Amazon « *façonnent le discours public* », et la plateforme en ligne ne parvient pas à corriger les biais de ses algorithmes. Tel est le constat d'une étude menée par AI Forensics, association européenne à but non lucratif qui enquête sur les algorithmes afin de défendre les droits des internautes, et par la société CheckFirst, éditrice de logiciels, spécialisée dans la lutte contre la désinformation.

Considérant le livre comme une source d'information, notamment au regard du statut du secteur de l'édition en termes de crédibilité, la place occupée par Amazon sur le marché des produits culturels justifiait, selon les auteurs de l'étude, que soient examinés les résultats de ses algorithmes de recherche et de recommandation.

Avec 181 millions d'utilisateurs actifs par mois dans l'Union européenne, Amazon a été désignée « très grande plateforme en ligne » en vertu du règlement sur les services numériques (Digital Services Act, DSA). La plateforme est donc tenue d'évaluer les risques systémiques potentiellement liés à ses activités, notamment le fonctionnement de ses algorithmes de recommandation et de modération des contenus, afin de parer à la diffusion massive de contenus préjudiciables ou illicites.

Intitulée « The Amazing Library : An Analysis of Amazon's Bookstore Algorithms within the DSA Framework », cette enquête porte sur le fonctionnement du système de recommandation d'Amazon. Conduite à partir des sites belge et français de la librairie en ligne, elle s'appuie sur une sélection d'environ 1 700 best-sellers consacrés à des sujets de société (hors fiction), selon un classement en dix-huit catégories : « Actualité, politique et société », « Sciences de la terre et de l'environnement », « Affaires et Bourse », etc.

À partir de ce corpus d'ouvrages, plus de 60 000 recommandations de livres ont été collectées et cela, sans utiliser un profil utilisateur (donc sans historique de navigation, ni cookies), entre le 28 octobre 2023 et le 4 novembre 2023. Les suggestions algorithmiques de lecture du genre « Les clients qui ont acheté cet article ont aussi acheté » représentent une partie significative du parcours de l'internaute sur Amazon, expliquent les auteurs. Les recommandations de lecture sur Amazon peuvent engendrer jusqu'à 30 % des pages consultées. L'étude s'inspire du travail d'un de ses coauteurs, Paul Bouchaud, enseignant chercheur au

Centre d'analyse et de mathématiques sociales (CAMS) à l'École des hautes études en sciences sociales, qui a publié un article intitulé « Browsing Amazon's Book Bubbles » (HAL Id : hal-04308081, novembre 2023).

Afin de comprendre les effets de la curation par l'intelligence artificielle sur la diffusion de l'information, les chercheurs ont formulé des requêtes sur la librairie en ligne sur des thèmes clivants et donc motifs à la désinformation : les vaccins et le Covid, l'avortement, l'immigration, le changement climatique et le féminisme. « *Bien que les algorithmes soient conçus pour améliorer l'expérience de l'utilisateur en fournissant des recommandations personnalisées, il existe des inquiétudes quant aux risques potentiels de renforcer l'exposition à des contenus similaires et d'amplifier des contenus liés à la désinformation. Ce phénomène est particulièrement évident dans le domaine des recommandations de livres, où le choix des lectures suggérées peut influencer de manière significative l'opinion publique et la connaissance en général* », précisent les auteurs.

Voici quelques-unes des « recommandations à risque d'Amazon » révélées par l'étude :

- Une requête « Covid » ou « Vaccination » aboutit, en moyenne sur les dix premiers, à huit résultats affichant des ouvrages « antivaccination » ou « conspirationnistes », notamment par des auteurs identifiés pour avoir diffusé des informations erronées sur ces sujets.
- Sur le changement climatique, la moitié des résultats de recherche sur la plateforme en France induisent en erreur les utilisateurs sur le consensus scientifique. Une part significative de livres sur le climat sont négationnistes.
- Le mot « Féminin » sur Amazon Belgique renvoie principalement vers des livres critiques ou banalisant le mouvement féministe. La section fait également la part belle aux stéréotypes et aux inégalités fondés sur le genre : la moitié des livres présentés sur la première page des résultats font référence au « féminin sacré » ou à l'« énergie féminine ». Selon la Miviludes (Mission interministérielle de vigilance et de lutte contre les dérives sectaires), le concept de « *féminin sacré* » renvoie à « *une appropriation du féminisme par des dérives sectaires* » réduisant les femmes à des facultés reproductives.
- Avortement : 40 % de livres critiques en réponse à une requête concernant ce sujet sur Amazon Belgique affichent une position conservatrice sur la question, privilégiant les préoccupations morales et sociales de l'« enfant à naître » et accordant moins d'importance aux droits et à la santé des femmes.
- Immigration : les résultats de recherche d'Amazon Belgique surreprésentent les ouvrages anti-immigration, contribuant potentiellement à la diffusion de lectures affectant les droits fondamentaux.

En plus de l'absence de pluralisme dans la mise en avant des points de vue sur Amazon, les chercheurs de AI Forensics et de CheckFirst montrent comment les algorithmes amplifient le phénomène de désinformation en enfermant strictement les utilisateurs dans des « bulles de recommandation ». Après avoir cliqué une fois sur un livre « antivaccination », un utilisateur devra ensuite cliquer jusqu'à vingt fois avant de pouvoir échapper à des récits similaires. « *Les systèmes de recommandation d'Amazon créent des communautés de livres très confinées* », constatent les auteurs. Comme une soixantaine de communautés couvrent 90 % des livres sur Amazon, des groupes de livres sont souvent recommandés ensemble,

amplifiant encore davantage le phénomène.

Les systèmes de recherche et de recommandation d'Amazon exposent donc les utilisateurs à « un nombre disproportionné de récits trompeurs ou partiels, quand il ne s'agit pas de faussetés ou de désinformations pures et simples ». En outre, les utilisateurs se retrouvent totalement piégés une fois qu'ils ont sélectionné un livre, au lieu de se voir proposer une pluralité de points de vue.

Par ailleurs, les auteurs reprochent à Amazon de ne pas respecter ses propres règles de modération. Des ouvrages « *sexuellement explicites* » sont recommandés sans filtre dans un contexte inapproprié, en particulier parmi les livres destinés aux adolescents.

« *Compte tenu de l'impact d'Amazon sur la diffusion de l'information par la vente de millions de livres dans l'Union, l'évaluation des risques dans le cadre du Digital Services Act devrait aller au-delà de la protection des consommateurs et englober les risques systémiques posés par les systèmes d'Amazon pour les droits fondamentaux, le discours public et le bien-être de la société* », concluent les auteurs. Enfin, ils soulignent la nécessité d'une plus grande responsabilité de la plateforme, rappelant que si des audits indépendants et obligatoires pour les très grandes plateformes en ligne telles qu'Amazon sont prévus par le DSA, il reste que ces rapports ne seront pas rendus publics.

Sources :

- AI Forensics (Paul Bouchaud, Raziye Buse Çetin, Marc Faddoul), Check First (Amaury Lesplingart, Guillaume Kuster, Nelly Pailleux), « The Amazing Library : An Analysis of Amazon's Bookstore Algorithms within the DSA Framework », aiforensics.org, checkfirst.network, December 2023.
- Croquet Pauline, « Dans la librairie d'Amazon, « les algorithmes façonnent les ventes, mais aussi le discours public » », lemonde.fr, 11 décembre 2023.

Categorie

1. Usages

date créée

28 mars 2024

Auteur

francoise