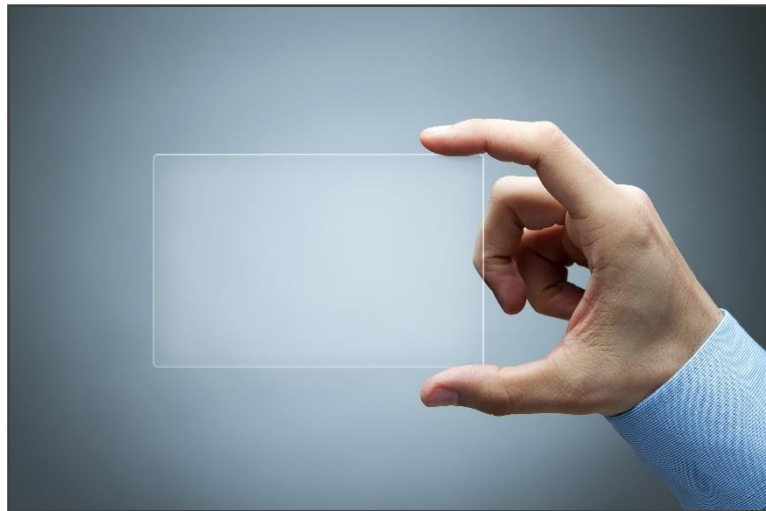


SATISFACTION A L'EGARD DES MOTEURS DE RECOMMANDATION : LA TRANSPARENCE OPTIMALE



Pauline VAUTROT

Chef de projet marketing digital – Agence 1000mercis
pauline.vautrot@1000mercis.com

Christophe BENAVENT

Professeur des universités – Université Paris Nanterre – CEROS-MIL
christophe.benavent@parisnanterre.fr

Pauline de PECHPEYROU

Professeur des universités – Université Paris-Est – IRG (EA 2354)
pauline.de-pechpeyrou@u-pec.fr

RESUME

L'objectif de la recherche est de tester l'impact d'une stratégie de transparence sur la satisfaction des internautes à l'égard des moteurs de recommandation. Une expérimentation longitudinale auprès de 52 répondants affectés de manière aléatoire à l'une des conditions (absence d'information, information simplifiée, information exhaustive) établit l'effet très clair de la modalité « information simplifiée » sur le niveau de confiance. C'est moins le contenu de l'information que le signal qui compte, et l'information simplifiée joue le rôle d'un gage d'autant plus efficace qu'il est contrôlable.

INTRODUCTION

Au début des années 1990, avec la démocratisation de l'usage d'Internet, les systèmes de recommandation basés sur le filtrage collaboratif se sont développés afin « d'aider les utilisateurs à gérer la surcharge informationnelle en construisant des modèles prédictifs destinés à estimer le degré d'intérêt de l'utilisateur pour chacun des items d'une immense base de données » (Konstan et Riedl, 2012). Ils s'appuient généralement sur les opinions d'une communauté d'utilisateurs pour aider chacun d'entre eux à identifier le contenu pertinent (Resnick et Varian, 1997).

Les travaux en marketing soulignent les bénéfices de ces recommandations pour les internautes. La mise en avant d'articles susceptibles de les intéresser permet aux utilisateurs de gagner en temps et en effort. Ce faisant, l'internaute sera tenté de revenir sur le même site (De Pechpeyrou, 2009). Toutefois, l'internaute peut se sentir « manipulé » ou être

déçu par le manque de pertinence de ces recommandations. Enfin, une limite également évoquée concerne le manque de variété et l'absence de sérendipité qu'engendrent les algorithmes prédictifs. Les plus réfractaires y voient même un risque d'enfermement dans ces choix et d'appauvrissement de la culture de la société. C'est le cas du militant d'Internet Eli Pariser qui a développé une théorie sur l'isolement culturel et intellectuel engendré par les recommandations qu'il a baptisée : The Filter Bubble.

La transparence des algorithmes est une réponse couramment avancée à ce problème. Selon Diakopoulos (2014), le « big open data push » aurait commencé dans l'administration américaine en 2009 suite à une requête du président Obama. Les grandes sociétés ont alors suivi, avec, en première ligne, Google qui publie désormais un rapport de transparence biannuel montrant à quelle fréquence il supprime ou divulgue des informations aux gouvernements. En France, la question de la transparence des données est toutefois apparue dès 2007 avec la première édition du Baromètre de l'intrusion.

Selon Diakopoulos (2014), la question de la transparence des algorithmes doit être abordée sous trois angles : celui de l'implication humaine (Quel est le but, l'intention de l'algorithme ? Qui en a le contrôle et qui en est responsable ?), celui des données (Quelles données ? Comment sont-elles récoltées et utilisées ? A quel point sont-elles

précises ?) et celui du modèle en lui-même (Quelles sont les variables utilisées par l'algorithme ? Comment sont-elles pondérées ?). Pour Diakopoulos (2014), une alternative est en train d'émerger, qu'il appelle « *algorithmic accountability reporting* ». Il s'agit de décortiquer le système par un examen rigoureux afin de pouvoir en dévoiler le fonctionnement. En France, l'initiative NosSystèmes se fixe pour objectif de « rendre les systèmes techniques (programmes, automates, modèles...) plus intelligibles au profit du dialogue et de la négociation entre les utilisateurs, les systèmes techniques et ceux qui les créent ou les exploitent ». Il ne s'agit pas tant de donner accès aux utilisateurs aux catégories qui ont été créées pour eux que de leur expliquer le fonctionnement du moteur de recommandation.

Le challenge est de trouver une façon intelligible de communiquer sur ces aspects des algorithmes. Comme le souligne Bénavent (2017) à propos du dernier service développé par Google (myactivity.google.com) qui restitue à l'internaute l'ensemble des informations que le moteur de recherche enregistre sur lui, la transparence n'est pas seulement formelle, elle doit être aussi cognitive : « il ne s'agit pas simplement de dire aux gens qu'il n'y a pas de secret dans le produit en leur donnant un plan technique détaillé qu'ils sont incapables de lire. Il faudrait expliquer le pourquoi du comment de manière compréhensible ».

Cette recherche vise à explorer l'impact d'une stratégie de transparence quant au fonctionnement du moteur de recommandation sur la qualité de la relation que les clients entretiennent avec le site marchand. Ce faisant, nous nous inscrivons dans la perspective de la « chaîne relationnelle » développée par Aurier et al. (2001), qui, partant de la qualité perçue, « cherche à développer la fidélité à long terme en créant et développant la valeur perçue, la satisfaction, la confiance et l'engagement ». Notre hypothèse centrale est celle d'un effet curvilinéaire de la transparence sur la qualité relationnelle, une transparence modérée constituant le niveau optimal à privilégier.

MOTEURS DE RECOMMANDATION ET TRANSPARENCE : BILAN DES RECHERCHES

Les premiers travaux sur les moteurs de recommandation se sont attachés à comparer la précision statistique des algorithmes sous-jacents aux recommandations, accordant peu d'intérêt au point de vue de l'utilisateur. Il faut attendre les

travaux d'[Herlocker et al. \(2000\)](#) et des chercheurs qui ont suivi pour véritablement prendre en compte l'impact des explications données autour des recommandations sur les jugements et sur la satisfaction des utilisateurs. Comme le soulignent [Tintarev et Masthoff \(2012\)](#), les explications fournies à l'utilisateur servent des objectifs divers (Tableau 1). Par exemple, les explications peuvent viser à aider le consommateur à faire le bon choix (« Cet item est susceptible de vous plaire pour telle raison »). De manière alternative, elles peuvent fournir de la transparence, lorsqu'elles exposent à l'utilisateur les raisons pour lesquelles un item lui est recommandé (« Les clients qui ont acheté cet article ont aussi acheté » sur Amazon.com). C'est cet objectif de transparence qui sera le cœur de notre recherche.

Tableau 1 – Objectifs assignés aux explications ([Tintarev et Masthoff, 2012](#))

Objectif	Définition
Transparence	Expliquer comment le système fonctionne
Examen	Permettre aux utilisateurs de donner un feedback au moteur lorsqu'il se trompe
Confiance	Accroître la confiance des utilisateurs dans le système
Efficacité	Aider les utilisateurs à prendre de bonnes décisions
Persuasion	Convaincre les utilisateurs d'essayer ou d'acheter
Efficience	Aider les utilisateurs à prendre des décisions plus rapidement
Satisfaction	Accroître la facilité d'usage ou le plaisir

L'une des premières applications du filtrage collaboratif aux produits culturels est le projet MoviLens conduit par l'équipe de recherche GroupLens de l'University of Minnesota. [Herlocker et al. \(2000\)](#) démontrent qu'une recommandation couplée à une explication autour de celle-ci accroît la probabilité d'aller voir le film mis en avant. Parmi les 21 variantes d'explications testées, ce sont les notes des autres utilisateurs qui se révèlent le plus persuasives. Une représentation graphique sous forme d'histogramme est plus efficace qu'un tableau de données.

Plusieurs recherches établissent que la présence d'explications à côté des recommandations est positivement corrélée à la confiance ressentie par l'utilisateur envers le système ([Sinha et Swearingen, 2002](#), [Chen et Pu, 2005](#)). Le format d'explication à privilégier dépend du domaine d'application, une courte explication étant préférable pour l'achat d'un livre alors qu'une longue le sera pour celui d'une voiture ([Chen et Pu, 2005](#)). La confiance accrue envers les recommandations accompagnées d'explications influence positivement les intentions

comportementales. La compétence perçue du moteur de recommandation incite les utilisateurs à retourner sur le site, même si elle ne les amène pas nécessairement à y effectuer un achat ([Chen et Pu, 2005](#)). A contrario, l'expérimentation de [Cramer et al. \(2008\)](#) sur la base d'un moteur de recommandation dans le domaine de l'art révèle que les explications génèrent une meilleure acceptation des recommandations mais pas une augmentation de la confiance envers le système lui-même.

Les explications accompagnant les recommandations peuvent avoir des effets contrastés selon les objectifs assignés à ces explications ([Tintarev et Masthoff, 2012](#)). La présence d'explications augmente l'efficacité des recommandations. Mais contrairement aux hypothèses formulées, la personnalisation des explications en fonction du profil de l'utilisateur réduit leur efficacité. Enfin, une recherche récente met en évidence l'effet d'interaction entre la source de la recommandation (proche vs. distante) et la justification de la recommandation. Ainsi, une recommandation émanant d'une source proche est

plus persuasive lorsqu'elle est accompagnée d'une justification, tandis qu'une recommandation émanant d'une source distante le sera plus si elle n'est pas accompagnée de justification. L'intrusion perçue est le mécanisme sous-jacent à cet effet d'interaction (Balbo et al., 2017).

Les recherches existantes soulignent donc l'intérêt pour les sites internet d'accompagner les recommandations d'items par des explications. En effet, elles augmentent la confiance vis-à-vis des recommandations et leur efficacité et conduisent, in fine, à une meilleure persuasion. Notre recherche souhaite prolonger ces travaux selon deux directions. La première concerne la mise en œuvre de la transparence. Contrairement aux recherches existantes pour lesquelles les explications sont données pour chaque item recommandé, nous souhaitons tester l'impact d'une information générale à propos du fonctionnement du moteur sur la satisfaction du client. Nous opérons en cela la distinction suggérée par Cramer et al. (2008) entre l'acceptation d'une recommandation et l'acceptation du système de recommandation en lui-même. La deuxième contribution attendue est la prise en compte de la chaîne relationnelle décrite par Aurier et al. (2001). Conformément aux recherches passées, l'une des variables mesurées est la satisfaction à l'égard des recommandations formulées. Cette variable est complétée par la confiance à l'égard de l'entreprise, qui est présentée dans la thèse relationnelle comme l'un des déterminants majeurs de l'engagement du consommateur à l'égard de l'entreprise. La confiance est abordée sous ses trois dimensions, à savoir la compétence, l'intégrité et la bienveillance.

Nous faisons l'hypothèse d'une relation curvilinéaire entre le degré de transparence et la qualité relationnelle, mesurée au travers de la confiance et de la satisfaction à l'égard de l'entreprise. Conformément aux recherches passées, on peut s'attendre à ce que la communication à l'internaute d'explications quant à la manière dont ses données sont utilisées pour formuler des recommandations augmente la confiance. En revanche, viser l'exhaustivité en matière de transparence augmenterait la complexité pour l'internaute et réduirait alors la confiance.

METHODOLOGIE

La nature des hypothèses formulées nécessite une démarche expérimentale en quatre étapes : une première phase d'apprentissage des préférences des consommateurs, une phase d'information des

consommateurs sur le fonctionnement du moteur de recommandation, une nouvelle phase de visite du site marchand et enfin une phase d'administration du questionnaire.

La première page du questionnaire, commune aux trois conditions expérimentales, invite les répondants à faire une première visite sur le site Amazon et leur donne des instructions de navigation. Ces consignes ont pour but de s'assurer que les participants aient « suffisamment » navigué sur le site pour que des recommandations pertinentes puissent leur être proposées lors de leur seconde visite (celle qu'ils devront évaluer).

La deuxième page fait suite à la première visite des répondants sur le site et à leur déconnexion. Elle leur donne des informations concernant le fonctionnement et les objectifs du moteur de recommandation d'Amazon. La longueur, la précision et la technicité de ces informations varient en fonction des trois conditions expérimentales, correspondant chacune à un niveau de transparence donné. Les notices ont fait l'objet d'une validation d'experts.

Les répondants sont exposés de manière aléatoire à l'une des trois versions suivantes :

- **Version 1** (absence de transparence) : aucune information n'est donnée sur le moteur de recommandation. Les répondants sont simplement invités à effectuer une seconde visite sur le site puis à répondre au questionnaire.
- **Version 2** (transparence modérée) : des informations générales et simplifiées sont partagées sur le moteur de recommandation d'Amazon telles que : son fonctionnement, l'utilisation qu'il fait des données et les objectifs qu'il poursuit en faveur du site et des visiteurs.
- **Version 3** (transparence forte) : cette version reprend les informations de la version 2 sous une forme plus technique et détaillée (et donc plus longue).

La dernière étape consiste en un questionnaire, permettant de mesurer sous forme d'échelles de Likert en 6 points les principaux concepts de la thèse relationnelle mobilisée dans le cadre de cette recherche. La confiance est mesurée par le biais des items de Bartikowski, Chandon et Müller (2009). Les items reflètent les trois dimensions traditionnelles de la confiance : la compétence (par ex : « Je pense que le site Amazon a les compétences et l'expertise nécessaires pour effectuer des

recommandations »), l'intégrité (par ex : « Globalement, le site Amazon semble avoir des principes éthiques ») et enfin la bienveillance (par ex : « Mes besoins et désirs semblent être importants pour Amazon »). La satisfaction est mesurée par le biais d'items inspirés des travaux d'[Oliver \(1980\)](#), par exemple : « Les recommandations du site semblent-elles répondre à vos besoins ? ». La compréhension perçue des recommandations est mesurée par le biais d'items empruntés à [Cramer et al. \(2008\)](#) et à [Sinha et Swearingen \(2002\)](#). Des questions sur la familiarité avec le site et la fréquence d'achat en ligne sont posées en fin de questionnaire afin de pouvoir contrôler ces différences individuelles lors du test des hypothèses.

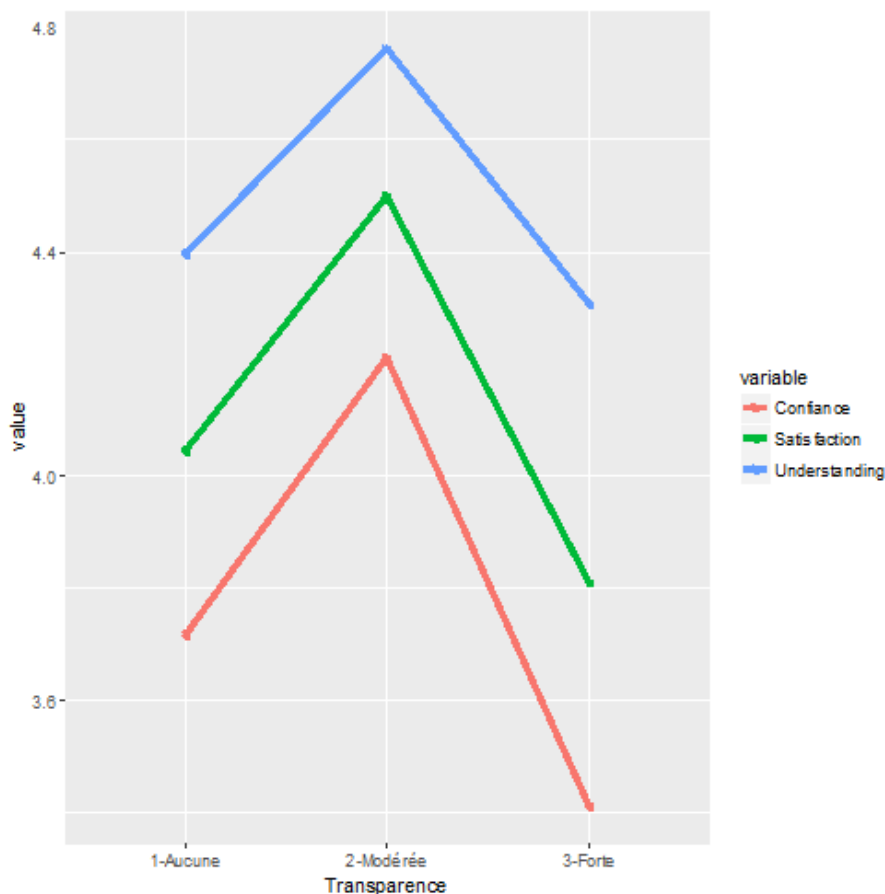
Au vu de la longueur de l'expérience menée, l'échantillon sollicité est un échantillon de convenance. Les 52 répondants (30 femmes, 22

hommes) sont affectés de manière aléatoire à l'une des trois conditions expérimentales. Au niveau de l'âge, on observe une surreprésentation des 18-25 ans (60% de l'échantillon). Enfin, la majorité des participants est relativement familière avec le site Amazon (62% d'entre eux possèdent un compte depuis plus d'un an).

RESULTATS

Le résultat principal apparaît dans la figure 1 qui représente la comparaison des moyennes des trois indicateurs de réponses des utilisateurs (compréhension, confiance, satisfaction). Au premier aspect, les mesures semblent parallèles et indique un résultat compatible avec notre hypothèse principale d'un effet curvilinéaire.

Figure 1 – Réponses des utilisateurs selon le degré de transparence



Un premier test est effectué par des analyses de variance sur chacune des variables, ainsi qu'une Manova pour tenir compte de leurs corrélations. Le résultat est intéressant, dans la mesure où le test Manova échoue au seuil habituel de 5% de risque,

alors que l'on obtient des F élevés pour la confiance et la satisfaction. On en conclut rapidement que l'effet n'est pas homogène et affecte de manière spécifique chacune des dimensions de la réponse des utilisateurs (Tableau 2).

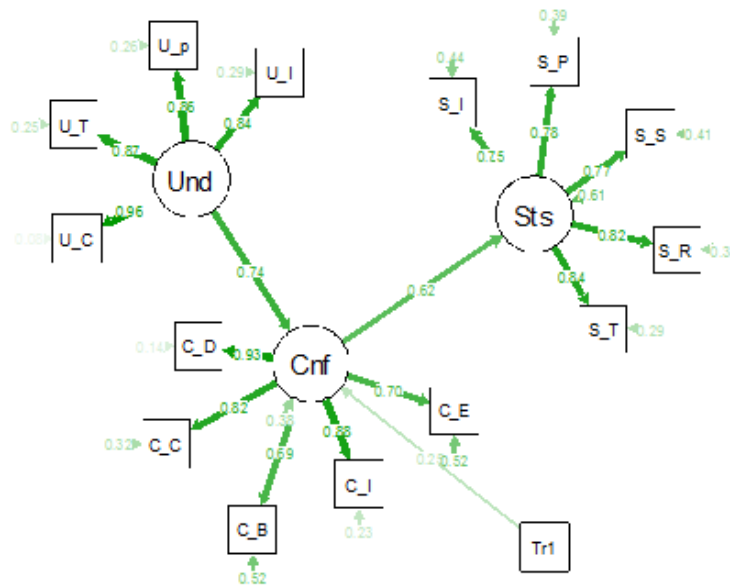
Tableau 2 – Résultats des analyses de la variance

		F	p
Anova	Compréhension	0,08	0,282
	Confiance	4,40	0,012
	Satisfaction	4,81	0,017
Manova		1,98	0,077

Voilà qui mérite une analyse plus approfondie qui rende compte plus précisément de la structure des données. Le modèle structurel a la forme d'une chaîne où la compréhension induit la confiance qui elle-même conditionne la satisfaction. On introduit dans un second temps le degré de transparence au travers de deux dummy dont on teste les effets sur les trois variables.

Le modèle final apparaît dans la figure 2. Avec un chi2 d'ajustement de 126,8 pour 88 degrés de liberté, un CFI de 0,934 et un TLI de 0,921, un RMSEA de 0,092 (avec un intervalle de confiance compris entre 0,053 et 0,126), on obtient un ajustement satisfaisant et de bonnes qualités psychométriques du modèle.

Figure 2 – Résultats du modèle conceptuel



Pour apprécier l'effet du facteur expérimental (niveau de transparence), on compare quatre modèles : un modèle sans effet expérimental (fit1), un modèle avec l'effet des variables auxiliaires sur la confiance et la satisfaction (fit3 – on omet l'effet sur la compréhension qui n'est clairement pas significatif), un modèle où seule la confiance est affectée par l'explication simplifiée (fit2) et une variante où les deux dummy sont prises en compte (fit4).

La comparaison des modèles par un test d'analyse de vraisemblance (tableau 3) indique que prendre en compte tous les effets conduit à une

surparamétrisation ainsi que la dégradation de l'ajustement (AIC, BIC et χ^2) de fit3 à fit2. Le modèle fit4 ne se distingue pas de fit3, c'est donc le modèle fit2 qui est le mieux ajusté.

Le passage de fit2 au modèle sans effet ne montre pas un gain significatif en termes de χ^2 , la règle du rasoir d'Ockham conduirait au plus simple, mais dans notre cas le critère d'intelligibilité conduit à retenir le modèle fit 2 et par conséquent à isoler un effet significatif de la formulation simplifiée sur la confiance (0,15 et $p = 0,01$).

Tableau 3 – Comparaison de la qualité d'ajustement des modèles

	Df	AIC	BIC	Chisq	Chisq diff	Df	diff
fit1	75	1444,1	1502,6	114,35			
fit2	88	1510,4	1570,9	126,84	12,492	13	0,487746
fit3	99	1566,5	1632,9	155,47	28,625	11	0,002595
fit4	101	1564,9	1627,3	157,82	2,3514	2	0,308597

DISCUSSION

Si globalement on confirme l'hypothèse de curvilinéarité, l'analyse structurelle nous conduit à un résultat plus nuancé.

En premier lieu, l'effet ne joue pas sur la compréhension. Ceci signifie que l'information fournie expérimentalement sous une forme simplifiée ou sous celle d'une information intégrale des conditions d'usages n'apporte rien de neuf par rapport à ce que les utilisateurs pensent connaître avant. L'effet n'est donc pas de nature cognitive, d'autant plus que la compréhension explique largement la confiance.

En contrôlant mieux la variance de la satisfaction par la prise en compte des propriétés psychométriques de l'échelle et du rôle de la confiance, on s'aperçoit que notre facteur expérimental n'a pas non plus d'effet. Les tests d'analyse de variance traduisent que l'effet est transporté de la confiance à la satisfaction.

En revanche on observe un effet très clair de la modalité « information simplifiée » sur le niveau de confiance. On en déduit que cet effet joue sur la crédibilité. C'est moins le contenu de l'information que le signal qui compte, l'information simplifiée sur

le moteur de recommandation jouant le rôle d'un gage d'autant plus efficace qu'il est contrôlable. Cet argument ne remet pas en cause l'hypothèse d'un niveau de complexité optimale de l'information fournie mais en précise l'effet. Il s'agit moins d'un problème de compréhension et d'intégration de l'information dans le système de connaissance, qu'un effet rhétorique sur la crédibilité de l'information fournie. En simplifiant on donne le sentiment d'un contrôle possible qui renforce la crédibilité.

Est-ce encore une politique de transparence ? La réponse est négative si on les conçoit comme le fait d'aider les utilisateurs à une meilleure compréhension de l'usage des recommandations. Elle devient positive si on la conçoit comme une politique de signal positif : celui par lequel la plateforme qui recommande des items manifeste sa bonne disposition à l'égard d'un contrôle par l'utilisateur. A long terme, et cela dépasse l'horizon de cette expérience, la rhétorique ne suffit pas et doit s'appuyer sur la possibilité d'accéder à un contrôle effectif. Proposer des CGU simplifiées et des explications pédagogiques doit s'accompagner d'instruments qui permettent à l'utilisateur soupçonneux d'aller plus loin : accéder au texte littéral, à des descriptions précises des algorithmes, éventuellement à une plateforme de contrôles de l'algorithme (pensons par exemple à la possibilité de faire varier certains paramètres : précision,

sérendipité, poids des historiques) ou même à des résultats d'études indépendantes de la performance et des effets secondaires de l'algorithme.

CONCLUSION

Au final, notre recherche souligne que ce sont les signaux de volonté de transparence qui génèrent de la confiance plus que l'explication en elle-même. Ce résultat est comparable à ce que l'on observe sur le plan juridique quant au renforcement des dispositifs d'information du consommateur. Il existe en effet une différence entre une approche de stricte conformité de la règle de l'opt-in qui n'est pas toujours compatible avec l'expérience utilisateur et une approche d'information que l'on pourrait qualifier d'opérationnelle plus orientée vers l'acquisition de compétences chez le client au travers d'outils dédiés. Reste cependant à étudier comment ces signaux de confiance vont évoluer dans les années à venir, avec le développement des privacy

polices et les règles imposées par la GDPR (*General Data Protection Regulation*). Quand toutes les entreprises montreront ces signes, comment susciter la confiance ? Sans doute par la certification par des tiers indépendants.

L'évolution technologique, notamment en matière d'agents intelligents vocaux, pose par ailleurs de nouveaux défis quant à la manière de coupler transparence des algorithmes et expérience utilisateur optimale. La disparition des écrans au profit de la voix portée par des assistants multiples (dédiés à des tâches spécifiques) va en effet créer une situation nouvelle en la matière. Comment la structure de l'offre va-t-elle être modifiée par ce transfert de l'intention de l'humain à la machine ? Comment assurer la transparence des recommandations algorithmiques dans un environnement vocal notamment pour des produits facilement substituables ? Quels impacts sont à prévoir sur les futurs rapports de force entre distributeurs et industriels ■

BIBLIOGRAPHIE

Aurier, P., C. Bénévent, C. et G. N'Goala, G., 2001, Validité discriminante et prédictive des composantes de la relation à la marque, *Actes du Congrès de l'Association Française de Marketing*, Deauville.

Balbo L., F. Jeannot et A. Helme-Guizon, 2017, When recommending a product backfires: The effects of justification and source on user responses to online personalized recommendations, *Systèmes d'Information et Management*, 22, 2, pp. 81-102.

Bartkowski B., J-L. Chandon et B. Muller, 2010, Mesurer la confiance des internautes : adaptation de McKnight, Choudhury et Kacmar (2002), *Journal of Marketing Trends*, 1, 1, pp. 11-25.

Bénévent C., 2017, Comment Google enregistre et stocke secrètement vos conversations, *Atlantico*. [En ligne]
URL : <http://www.atlantico.fr/decryptage/comment-google-enregistre-et-stocke-secretement-vos-conversations-3146068.html>. Consulté le 23 octobre 2017.

Chen L. et P. Pu, 2005, Trust building in recommender agents, *1st International Workshop on Web Personalization, Recommender Systems and Intelligent User Interfaces*, pp. 135-145.

Cramer H., V. Evers, S. Ramlal, M. van Someren, L. Rutledge, N. Stash, L. Aroyo et B. Wielinga, 2008, The effects of transparency on trust in and acceptance of a content-based art recommender, *User Modeling and User-Adapted Interaction*, 18, 5, pp. 455-496.

De Pechpeyrou P., 2009, How consumers value online personalization: a longitudinal experiment, *Journal of Research in Interactive Marketing*, 3, 1, pp. 35-51.

Diakopoulos N., 2014, Algorithmic Accountability Reporting: On the Investigation of Black Boxes, *Tow Center for Digital Journalism*, 33 p.

Herlocker J.L., J.A. Konstan, A. Borchers et J. Riedl, 2000, Explaining collaborative filtering recommendations, *ACM Conference on Computer Supported Cooperative Work*, Philadelphie, PA, pp. 241-250.

Konstan J.A. et J. Riedl, 2012, Recommender systems: from algorithms to user experience, *User Modeling and User-Adapted Interaction*, 22, 1-2, pp. 101-123.

Oliver R.L., 1980, A Cognitive Model of Antecedents and Consequences of Satisfaction Decisions, *Journal of Marketing Research*, 17, 4, pp. 460-469.

Pariser E., 2011, *The Filter Bubble: What The Internet Is Hiding From You*, Penguin, 304 p.

Resnick P. et H.R. Varian, 1997, Recommender systems, *Communications of the ACM*, 40, 3, pp. 56-58.

Sinha R. et K. Swearingen, 2002, The Role of Transparency in Recommender Systems, *CHI '02 Extended Abstracts on Human Factors in Computing Systems*, Minneapolis, Minnesota, pp. 830-831.

Tintarev N. et J. Masthoff, 2012, Evaluating the effectiveness of explanations for recommender systems, *User Modeling and User-Adapted Interaction*, 22, 4-5, pp 399-439.