

Avant-propos

Les moteurs (et systèmes) de recommandation

Les travaux de recherche sur les médias de masse ont permis de mettre en avant l'activité du récepteur dans le décodage des informations portées par la sociabilité primaire [HAL 75]. Cette approche a été développée [PAS 99, CAR 99, CEF 03] pour proposer des méthodologies qui prennent en compte des publics de fans et ainsi comprendre la réception des contenus informationnels. D'autres chercheurs ont proposé l'approche dite *two-step-flow* [KAT 08, RAM 11] qui n'est pas basée seulement sur le modèle de l'émission et de la réception des contenus, mais introduit aussi le rôle d'intermédiaire du « leader d'opinion » dont l'influence est prépondérante sur le choix de produits par les consommateurs. Cette approche est basée sur des personnes qui diffusent l'information mais également sur d'autres qui ajoutent une valeur sociale et économique. K. Mellet [MEL 09] a développé ces recherches dans le cadre du marketing viral pour le choix de produits ou services.

« La croissance exponentielle des informations disponibles [sur le Web], ainsi que leur forte hétérogénéité » [PIC 11, p. 38] et l'augmentation rapide de nouveaux services de e-business et de e-commerce « ont très souvent débordé les utilisateurs, les amenant à prendre de mauvaises décisions. » [PIC 11, p. 38] et à se perdre dans la masse d'informations. En effet, cette croissance rapide des sites Web a créé chez les consommateurs un besoin d'aide pour trouver le choix le plus approprié parmi de multiples produits ou informations disponibles, en leur fournissant des recommandations issues du filtrage des alternatives possibles.

L'arrivée du *Web 2.0* vers l'année 1999 a bouleversé la relation entre les producteurs et les consommateurs de contenus en trois aspects [BAS 11] :

1) *la diminution des frontières de production des contenus* : toute personne peut produire des contenus (vidéo, musique, texte, etc.). Ceci implique une surabondance des contenus, produits par des amateurs et des professionnels ;

2) *l'accessibilité des contenus sur des plates-formes* : toute personne peut rendre accessibles ces contenus en les publiant sur des plates-formes éditoriales (Youtube, Dailymotion, blog personnel, etc.). Ceci implique une profusion de contenus, enrichis par des fonctions de tri, de sélection et de qualification de l'information. Ces informations peuvent être produites par l'informatique (comme la trace sur le Web) ou par l'activité des publics. Actuellement, il est possible de qualifier un contenu par un nombre de vues, de visites ou par sa durée d'existence utilisant les outils d'agrégation, les flux RSS, les *folksonomies*, etc. ;

3) *l'expression sociale* : les plates-formes sociales comme Facebook incitent les internautes à « recommander », « coopérer », « partager » et « afficher » les informations sur leurs goûts, leurs passions, leurs activités, conduisant à une surabondance d'avis sur une information ou un objet.

Les recommandations peuvent être utilisées pour résoudre les problèmes de surcharge cognitive et de surabondance des contenus sur le Web. Pour traiter ces problèmes, le système basé sur les recommandations analyse les données inconnues qui peuvent être éventuellement intéressantes pour les utilisateurs [PIC 11, SAL 06], par exemple : historique de la navigation, profil de l'utilisateur, contexte des achats, analyse des actions, trace laissée par l'utilisateur, etc. L'utilisateur consulte les recommandations. « Il peut les accepter ou non et peut fournir, immédiatement ou plus tard, un feedback implicite ou explicite. Toutes les actions et les feedbacks des utilisateurs peuvent être enregistrés dans la base de données et utilisés pour générer de nouvelles recommandations pour les prochaines interactions avec le système » [PIC 11].

Les travaux de recherche de H. Simon [SIM 71] ont par ailleurs montré que « l'abondance d'information engendre une *pénurie d'attention* et la nécessité d'allouer cette attention de manière efficiente parmi la surabondance d'information qui peuvent la consommer ». La valeur dans le contexte numérique ne concerne pas uniquement les contenus informationnels, mais aussi l'attention, vue comme la ressource importante de l'environnement numérique [KES 10]. « Une économie de l'attention envahit le marché de la culture sur Internet plus efficacement que sur n'importe quel autre vecteur informationnel. Salauin déclare que grâce à Internet, le marché de l'attention inverse le sens de l'affichage publicitaire. En effet, le message publicitaire commence à accompagner le cheminement de l'internaute et non celui du document support¹ » [SAL 11].

Les travaux en marketing montrent que plus le choix est important, plus il rend le consommateur hésitant, ceci le conduit à recourir au mimétisme ou à la recommandation [MEL 11]. La recommandation devient un élément central dans l'environnement numérique de la surabondance de contenus informationnels. Il est intéressant de

1. <http://www.dicen-idf.org/evenement/journee-etude-moteurs-recommandation>.

s'interroger sur la part de la recommandation dans la réception d'un contenu et la différence entre la recommandation en ligne par rapport à la recommandation hors ligne ? Comment juger le degré de pertinence de la recommandation ? Le public construit-il une hiérarchie des recommandations ? « Le partage [et la coopération] permettent-ils l'extension ou la réduction de la diversité des informations ? Comment qualifier la méta-information [ou la métadonnée] ajoutée au contenu lors d'un partage, issue de l'identité de la personne [...], voire son avis personnel » [BAS 11, p. 6] ?

Il est certainement aussi important d'étudier « le nouvel équilibre induit par le numérique entre recommandation et réception d'information » [BAS 11, p. 4] à partir d'une analyse sociologique, en interrogeant qui sont les « leaders d'opinion » sur le Web, leurs motivations, leurs démarches, leurs influences et leurs rôles dans la sphère publique ou privée. Par ailleurs, la recommandation pose aussi un problème déontologique dans la mesure où les utilisateurs partagent leurs goûts, leurs préférences, leurs compétences et leur savoir-faire, c'est-à-dire l'exposition de soi dans un monde ouvert sans véritable régulation [BAS 11, GEO 11a et 11b, CAR 11, MER 14]. A l'inverse, l'acte de recommandation peut être vu comme une pratique sociale d'entre-aide généralisée tout à fait inédite dans l'histoire des médias.

Cet ouvrage fait écho au séminaire qui fut organisé sur le sujet au CNAM en 2012² réunissant bon nombre des interventions de cette journée ainsi que des contributions complémentaires venant enrichir l'analyse.

Quel est l'historique des moteurs de recommandation ? Quels sont les méthodes et les modèles de recommandation les plus efficaces ? Quels impacts sur les habitudes des usagers ? Qui sont les principaux acteurs de ces services ? Quels sont les développements majeurs dans le domaine des plates-formes de contenu ? Telles sont des dimensions qui sont ici abordées et que les contributions réunies voudraient contribuer à éclairer.

A l'heure où les *Big Data* transforment complètement les métiers du marketing et du e-commerce en analysant en profondeur les comportements des consommateurs, l'intérêt pour la recommandation ne cesse de grandir notamment dans sa dimension prédictive afin d'optimiser « l'adéquation des produits et des services » [BER 14a]. Les *Big Data* apportent des solutions intéressantes en analysant l'audience Web (pages visitées, recherches, métadonnées, etc.), les profils des consommateurs, les traces laissées pour proposer des recommandations d'achat ciblé. Les *Big Data* sont dès lors utiles « pour déterminer le meilleur emplacement pour un point de vente en analysant les *Open Data*, les données géographiques, les données socio-économiques [...] et le marché et la concurrence » [BER 14b]. Le croisement des *Big Data* et

2. <http://www.dicen-idf.org/evenement/journee-etude-moteurs-recommandation>.

des moteurs de recommandation font sans doute partie des enjeux majeurs de ces prochaines années.

Remerciements

Les coordonnateurs de l'ouvrage Gérald Kembellec, Ghislaine Chartron et Imad Saleh tiennent à remercier chaleureusement l'ensemble des membres du comité de lecture. Grâce à leurs efforts et à leur implication sans faille dans les corrections, l'ouvrage s'inscrit dans une démarche de qualité scientifique garantie par l'évaluation en double aveugle.

En particulier :

- Evelyne Broudoux, MCF au CNAM, laboratoire DICEN ;
- Camille Prime Claverie, MCF à l'Université Paris Ouest Nanterre ;
- Orélie Desfriches Doria, Dr, Ingénieure de recherche au Labex CAP (Cnam) ;
- Lucile Desmoulin, MCF, Université Paris-Est Marne-la-Vallée ;
- Widad Mustafa el Hadi, PR, Université Lille 3, laboratoire Gériico ;
- Camille Paloque-Berges, Dr, Ingénieure de recherche, Laboratoire histoire des technosciences en société (HT2S) au Cnam ;
- Sahbi Sidhom, MCF, Université de Lorraine et chercheur au laboratoire LORIA ;
- Isis Truck, MCF/HDR à l'Université Paris 8, équipe CHArt.

Imad SALEH

Bibliographie

- [BAS 11] BASTARD I., « Les logiques du partage d'information sur Internet », Projet de thèse, D. Cardon et D. Pasquier (dir.), Orange Labs, 2011.
- [BER 14a] BÉRANGER F., *Big Data Analyse et valorisation de masses de données*, Smile, Livre Blanc Open Sources solutions, 2014.
- [BER 14b] BÉRANGER F., « Big Data Analyse et valorisation de masses de données », Smile, *Livre Blanc Open Sources solutions*, p. 22, 2014.
- [CAR 11] CARDON D., *La décomposition de l'identité numérique*, 2011, disponible à l'adresse : <http://spiralconnect.univ-lyon1.fr/webapp/page/page.html?id=1459331>, consulté le 25 avril 2014.
- [CEF 03] CEFAI D., PASQUIER D. (DIR.), *Les sens du public. Publics politiques, publics médiatiques*, Presses Universitaires de France, Paris, 2003.

- [GEO 11a] GEORGES F., « L'identité numérique sous emprise culturelle. De l'expression de soi à sa standardisation », *Les cahiers du numérique*, vol. 7, n° 1, p. 31-48, 2011.
- [GEO 11b] GEORGES F., « Pratiques informationnelles et identité numérique », *Etudes de communication*, n° 35, p. 105-120, 2011.
- [HAL 99] HALL S., ALBARET M., GAMBERINI M.C., « Codage / Décodage », *Réseaux*, vol. 12, n° 68, p. 27-39, 1994.
- [KAT 08] KATZ E., LAZARSELD P.L., *Influence personnelle. Ce que les gens font des médias*, Armand Colin/Institut national de l'audiovisuel, Paris, 1975/2008.
- [KES 10] KESSOUS E., MELLET K., ZOUINAR M., « L'économie de l'attention entre protection des ressources cognitives et extraction de la valeur », *Sociologie du travail*, vol. 52, n° 3, p. 359-373, 2010.
- [MEL 09] MELLET K., « Aux sources du marketing viral », *Réseaux*, n° 157-158, p. 267-292, 2009.
- [MEL 11] MELLET K., « Le marketing en ligne, "Cultures numériques" », *Communications*, n° 88, p. 103-111, 2011.
- [MER 14] MERZEAU L., « Identity commons : du marquage au partage », dans Coutant A. et Stenger T. (dir.), *Identités numériques*, p. 35-49, L'Harmattan, Paris, 2014.
- [PAS 99] PASQUIER D., *La culture des sentiments : l'expérience télévisuelle des adolescentes*, Les Editions de la MSH, 1999.
- [PIC 11] PICOT-CLEMENTE R., Une architecture générique de Systèmes de recommandation de combinaison d'items: application au domaine du tourisme, Thèse, Université de Franche-Comté, p. 39, 2011.
- [SAL 06] SALEH I., MKADMI A., REYES E., « L'hypermédia au service du travail collaboratif », dans Saleh I. (dir.), *Les hypermédias : conception et réalisation*, Hermès-Lavoisier, Paris, 2006.
- [SAL 11] SALAÜN J.-M., « Economie de l'information : les fondamentaux », *Documentaliste-Sciences de l'Information*, vol. 48, n° 3, p. 24-35, 2011.
- [SIM 71] SIMON H.A., « Designing Organizations for an Information-Rich World », dans Greenberger M. (dir.), *Computers, Communications and the Public Interest*, p. 37-72, John Hopkins Press, 1971.