



Les usagers et l'information de santé sur Internet

Olivier Le Deuff

Maître de conférences en sciences de l'information et de la communication, laboratoire Mica, université Bordeaux-Montaigne

La recherche d'information de santé sur Internet devient une pratique courante. La facilité d'accès, couplée à des aspects discrétionnaires, explique le succès de cette pratique. Une étude au niveau européen montrait d'ailleurs que les Européens étaient généralement satisfaits de l'information qu'ils trouvaient sur le web en matière de santé¹. Pourtant, plusieurs enquêtes montrent que les niveaux de connaissance en matière de santé et de compétence numérique s'avèrent insuffisants. Le besoin de développer une littératie digitale (ou numérique) de santé apparaît désormais essentiel, tant la complexité des environnements du web n'est pas comprise par une grande partie des usagers, et ce d'autant qu'il est toujours facile d'obtenir une réponse à une requête en la matière car les moteurs de recherche donnent toujours des résultats, quelle que soit la qualité de la recherche d'information. Par conséquent, il faut s'interroger sur la qualité et la pertinence de l'information que les usagers peuvent rencontrer sur Internet.

Le public qui consulte le web est hétérogène, à la fois dans le type d'informations qu'il recherche et dans les compétences qu'il possède. Certains publics n'ont reçu quasiment aucune formation en matière de santé, mais également aucune formation pour rechercher l'information et utiliser les technologies du web. C'est souvent le cas du public senior.

Pourtant, la recherche d'information en ligne n'est pas aussi intuitive qu'elle peut le paraître. Ainsi, une des premières difficultés est celle qui consiste à mettre des mots sur ses maux. Avant même l'utilisation du moteur de recherche, il y a de fortes inégalités dans la prise de conscience du besoin d'information, son identification et son expression. Les résultats dépendent fortement de cette capacité à exprimer une requête appropriée selon que celle-ci comporte ou non des concepts scientifiques maîtrisés. Il est probable que certaines requêtes pourraient potentiellement paraître totalement incohérentes au corps médical. Les coquilles, les fautes d'orthographe influent fortement sur les résultats des requêtes, mais ce sont surtout les connaissances médicales basiques qui s'avèrent importantes pour exprimer correctement d'éventuels symptômes. Pour des symptômes proches, ce sont des centaines de parcours de recherche différents qui peuvent être envisagés. Toutefois, il apparaît que l'optimisation en matière de référencement de certains sites, comme le célèbre Doctissimo, recentre les résultats autour d'acteurs dominants dans l'offre informationnelle en matière de santé.

1. Flash Eurobarometer 404. European Citizen's digital health literacy. Novembre 2014 http://ec.europa.eu/public_opinion/flash/fl_404_en.pdf

S'il existe des usagers experts au point d'aller chercher de l'information directement dans les publications scientifiques en consultant des sites comme Pubmed², nous n'avons pas rencontré de tels cas en France, notamment dans notre enquête sur les seniors (dans le cadre du programme de recherche Ricsa, Risque informationnel chez les seniors et automédication³). Le réflexe majoritaire est l'utilisation du moteur de recherche Google, qui est le leader du marché en France avec plus de 90 % d'usagers, une tendance confirmée dans les résultats de notre enquête en ligne auprès des seniors, qui l'utilisent à plus de 92 %.

S'il est difficile de pouvoir mesurer pleinement quelles sont les requêtes effectuées en matière de santé (les outils du type Google Trends⁴ ne sont pas suffisants pour mesurer justement la diversité des interrogations possibles), il est toutefois opportun d'étudier le fonctionnement des moteurs de recherche, notamment en constatant la forte présence publicitaire.

L'information publicitaire associée au moteur de recherche

L'inclusion de publicités débute dès les premiers résultats des requêtes, avec des publicités contextuelles, c'est-à-dire qui apparaissent en fonction des mots clés tapés dans la barre de recherche. Bien qu'elles soient distinctes dans l'emplacement de la page de résultats, elles ne s'avèrent pas si simples à distinguer visuellement du résultat plus classique (dit « naturel »), pour le public non initié. Certes, Google signale les publicités par la mention « annonces » et les place généralement sur le côté, au-dessus, voire en dessous des résultats, mais beaucoup d'usagers cliquent dessus sans avoir conscience qu'il s'agit d'une annonce publicitaire. Le système est basé sur une mise aux enchères des mots clés, de façon à ce que chaque fois qu'une personne clique sur le lien un versement de quelques centimes soit effectué à Google. Il existe donc un marché des mots clés et, par conséquent, un marché lié à la santé.

L'information traitée par Google

Le mode de fonctionnement de Google et les processus par lesquels il trouve les résultats de façon efficace en un clic reste inconnu pour la plupart des usagers. Or, le système de Google est basé sur des mécanismes de

2. <http://www.ncbi.nlm.nih.gov/pubmed>

3. <https://ricsa.hypotheses.org/>

4. Google Trends permet de connaître la fréquence d'une requête dans le moteur de recherche Google. Il est possible de visualiser ces données par région et par langue.

popularité : un site est bien classé s'il se trouve mentionné par un grand nombre d'autres sites, eux-mêmes déjà bien cités par d'autres. Ce système rentre donc en contradiction avec les systèmes de classification des savoirs qui privilégient l'autorité pour attester de la qualité et de la véracité des informations. Pour ce qui concerne l'information de santé, le problème vient du fait que l'on peut trouver de l'information qui n'émane pas nécessairement de spécialistes diplômés. Si cela peut s'avérer opportun lorsqu'il s'agit d'avis et de retours d'expériences de patients, cela peut être néfaste quand l'information est produite par un non-spécialiste qui outrepassé ses compétences.

À cette logique du *page rank* (classement des sites Internet sur le moteur de recherche) s'ajoutent les mécanismes d'individualisation des requêtes, à partir de données recueillies par Google (adresse IP, localisation, cookies et historique, etc.). Le moteur de recherche va ainsi proposer des résultats différents en fonction des personnes. Il faut rappeler ici que Google construit un profil de la personne qui effectue des requêtes dans le but de mieux cibler les publicités contextuelles. Il est fort probable que le moteur sera en mesure d'identifier un usager senior en fonction des requêtes et des sites visités par ce dernier⁵. Au final, il apparaît fortement souhaitable de renforcer les actions de médiation et de formation envers les publics qui effectuent des recherches régulièrement.

L'information pas toujours sourcée : quels en sont les véritables propriétaires et auteurs ?

L'origine de l'information de santé est relativement difficile à identifier, ce qui complexifie l'évaluation de sa qualité. L'identification des propriétaires de site s'avère parfois surprenante. Certains acteurs sont ainsi fortement éloignés du domaine de la santé : le site PasseportSanté.net, par exemple, appartient au groupe Oxygem et ses enregistrements en tant que nom de domaine font apparaître l'appellation « cuisine AZ ». Les sites les mieux référencés sont possédés par des structures qui ne sont pas nécessairement de nature médicale. On retrouve ainsi les organismes de presse et de médias, ce qui n'est guère étonnant puisqu'ils avaient déjà investi ce créneau au sein de la presse magazine, de la radio et de la télévision, ainsi que des entreprises de gestion de contenus sur le web, dont la logique repose sur une maximisation des revenus associés aux contenus proposés. La stratégie de ces sites est de privilégier un système de création de liens hypertextes à l'intérieur du site de façon à ce que l'usager passe le plus de temps possible sur le site ou les sites du même groupe. La logique est celle d'une captation de l'attention à des fins publicitaires.

5. Il est possible de vérifier notre profil construit par Google en fonction de nos requêtes et afin d'affiner sa logique publicitaire : <https://www.google.com/settings/u/0/ads/authenticated>

Une logique publicitaire

La publicité est en effet la source principale de revenus de ces sites. Outre des publicités contextuelles de type Google AdSense⁶, les principaux sites cherchent à cibler au mieux l'apparition des publicités par rapport au contenu qui est diffusé et affichent ainsi des publicités ciblées. On trouve alors aisément de la publicité liée au rhumatisme sur une page qui aborde déjà la question. Beaucoup de publicités renvoient à des produits qui sont disponibles sans ordonnance. L'information proposée est donc couplée à des informations publicitaires dans le but non pas d'optimiser le contenu de l'information dispensée au lecteur mais de susciter l'achat en pharmacie ou parapharmacie, voire dans certains cas sur des pharmacies en ligne.

Si la publicité ne suscite pas immédiatement le désir d'achat, l'association entre les symptômes exprimés et le médicament faisant l'objet de la publicité s'inscrit néanmoins dans l'esprit du patient, favorisant par la suite une possible démarche d'automédication. La logique des cookies amplifie le phénomène en faisant réapparaître les mêmes publicités lors de navigations ultérieures, qu'elles soient sur le même site ou sur un site fort différent. Une des recommandations potentielles serait d'inciter l'utilisateur à réaliser ses requêtes en matière de santé ou d'ordre confidentiel *via* le mode « navigation privée » de son navigateur. Même si elle n'assure pas totalement une disparition des traces, elle diminue fortement la conservation des historiques de recherche.

La redondance de l'information

Beaucoup d'éléments d'information sont issus d'autres sites Internet, notamment sur les forums, avec des citations ou incises qui ne mentionnent pas toujours leur source. Le problème de la viralité d'une information répétée est qu'elle peut produire et accroître les risques associés à l'autodiagnostic, en confortant l'usager. Le conseil de croiser les sources pour s'assurer de la véracité d'une information ne garantit pas pour autant que l'information recueillie soit de qualité.

Favoriser une information de santé de type institutionnel

Les acteurs institutionnels ne sont pas en première position des « consultations » sur Internet, au profit d'acteurs qui optimisent les contenus à des fins marchandes. Les sites institutionnels ont donc intérêt à mieux saisir l'écosystème du web pour optimiser les contenus accessibles au grand public, déjà en se citant davantage les uns les autres. Le site de l'assurance maladie <http://www.ameli.fr/> commence à se positionner de façon efficace dans le système. Il serait opportun que les autres acteurs fassent de même. Dans le cadre du programme de recherche Ricsa et à l'aide d'un financement de la fondation Maif, nous avons développé et mis en ligne un test d'autoévaluation sur l'information de santé et l'automédication : <http://www.megatopie.info/testor/enquete.html?e=1>

6. AdSense est la régie publicitaire de Google : une entreprise peut acheter un mot clé, et quand quelqu'un tape ce mot clé, l'annonce de l'entreprise apparaît.