

Référence à citer : SMYRNAIOS Nikos, Rebillard Franck, 2007, « L'information via Google. Les relations entre éditeurs de presse et infomédiaires », communication au colloque Online services: networks, contents, usages, Paris, 12-14 décembre.

**L'information via Google.
Les relations entre éditeurs de presse et infomédiaires sur l'internet.**

Nikos Smyrnaiois, MCF, Université Toulouse 3, LERASS et MSH Paris-Nord
Franck Rebillard, MCF, Université Lyon 2, ELICO et MSH Paris-Nord

Mots-clés :

internet ; information ; journalisme ; infomédiation ; socio-économie

Résumé :

A l'occasion de leur diversification sur l'internet, les éditeurs de presse sont confrontés à des acteurs occupant une position centrale sur les réseaux informatisés : les infomédiaires. De leur côté, les infomédiaires sont réciproquement dépendants de ces producteurs de contenus éditorialement légitimes et parmi les plus porteurs d'audience. La relation d'interdépendance entre ces deux catégories d'acteurs est analysée en détail, au niveau de la France, pour l'infomédiaire majeur de l'internet qu'est devenu Google.

Il s'agit tout d'abord de décrire comment la production de l'information, pour le web, a été repensée au sein des entreprises de presse. Afin d'optimiser son référencement sur l'internet, l'information est conçue pour satisfaire aux critères algorithmiques de Google. Corollairement, l'attitude de Google est observée, et plus particulièrement ses efforts pour passer d'un rôle de simple distributeur à celui de méta-éditeur.

Un point de vue diachronique permet de suivre l'évolution de ces relations de coopération. Les stratégies des éditeurs et de Google s'adaptent les unes aux autres au fil du temps, en même temps que se reconfigure plus largement la filière de la production d'information journalistique avec l'internet.

Keywords :

internet, information, journalism, infomediatioin, political economy

Abstract :

Since they have engaged themselves on the internet, press editors are been confronted to infomediaries, such as portals and search engines, who occupy a central position in the sector as powerful audience agregators. On their side, infomediaries are also depending on the content industry and especially dominant media in order to attract large publics. This complex relation of interdependance is analysed in detail concerning the case of Google News in France.

First of all we describe the way in which production of information processes were modified inside french press groups in order to satisfy Google's pertinence criteria. At the same time we examine Google's strategy and particularly its effort to move on from simple information distributor to meta-editor. Our diachronic perspective analyses the evolution of these *coopetitive* relations, illustrating the process of reombination that characterizes the online information sector.

La diversification de l'industrie de la presse sur l'internet implique de nombreuses mutations dans l'activité du secteur. Le nouvel environnement se caractérise par une concurrence très forte ainsi que par une certaine interdépendance entre les médias traditionnels et les nouveaux acteurs en provenance des secteurs de l'informatique et des télécommunications.

Une facette très importante de ces mutations est l'émergence sur l'internet des nouveaux vecteurs de diffusion de l'information journalistique que sont les moteurs de recherche et autres services de veille spécialisés dans le domaine de l'actualité. Les infomédiaires tels que Google prennent en charge la sélection, l'organisation, la hiérarchisation et la distribution de l'information éditée par des tiers, sur la base de critères thématiques, linguistiques ou géographiques [Boure, Smyrnaio, 2006]. De ce fait, ils constituent en quelque sorte une « industrie de l'accès » – se surajoutant aux fonctions exercées par les FAI et les routeurs aux niveaux successifs des infrastructures et des infostructures de réseau – dans le sens où leur métier est d'organiser l'accès du grand public à la production des industries du contenu. Etant donnée la puissance de certains de ces acteurs, l'hypothèse a été formulée selon laquelle ces derniers seront amenés à prendre une place de plus en plus significative dans les circuits de distribution des biens informationnels et culturels [Bustamante, 2004].

La mise en ligne de l'information de presse a été très tôt considérée comme une opportunité offerte aux éditeurs de s'affranchir de la contrainte lourde et particulièrement coûteuse que représentent l'impression et la distribution des supports papier. Aujourd'hui encore, cette idée reste présente dans certaines études, qui insistent sur les facilités techniques de publication assistée par ordinateur. Ainsi le dernier rapport ministériel français consacré aux défis numériques de la presse affirme-t-il que « *la création d'un site internet ne nécessite pas une infrastructure importante. Bien au contraire, il suffit de disposer d'un ordinateur, des logiciels adéquats et d'une connexion à internet pour pouvoir être présent sur le réseau et présenter et diffuser ainsi ses idées. Les contraintes techniques sont donc très limitées.* » [Tessier, 2007, p. 47]. Dans ce même rapport, les quelques rares réserves émises à propos de la visibilité des informations sur le web sont levées en faisant appel à l'omniprésence des moteurs de recherche : « *Certes, il ne suffit pas d'être en ligne pour être visible et, compte tenu de la masse d'informations et de données disponibles sur le réseau, il peut paraître particulièrement difficile de s'y faire entendre. Cependant, le recours à des moteurs de recherche constitue une aide précieuse et permet à un internaute de se repérer et de trouver les avis et les positions correspondant à tout l'éventail des courants d'expression démocratique.* » [ibid, p. 47].

Cette vision de moteurs de recherche offrant à n'importe quel producteur d'informations une visibilité satisfaisante sur l'internet est très contestable car elle fait fi de la complexité des relations qui s'établissent entre les acteurs concernés. Dès lors il convient de s'intéresser d'une part au fonctionnement des infomédiaires et, d'autre part, aux stratégies que mettent en œuvre les éditeurs afin d'exploiter au mieux ce nouveau canal de diffusion. Du point de vue de l'économie politique, une telle analyse permet d'aborder deux problèmes particulièrement cruciaux : d'une part la question du pluralisme de l'information sur l'internet, en s'intéressant aux effets éventuels de la sélection et de la hiérarchisation des sources opérées par les infomédiaires ; d'autre part l'élaboration de modèles socio-économiques adaptés à l'internet, en pointant les conséquences possibles de cette nouvelle intermédiation (l'infomédiation) sur la filière de la production d'information journalistique.

Notre texte vise à apporter une contribution sur ces deux sujets, en se retranchant toutefois à analyser dans le détail les relations, parfois complexes et ambiguës, qu'entretiennent les

éditeurs de presse, acteurs historiques majeurs de la production de l'information journalistique, avec les nouveaux acteurs de l'internet. Comment s'articulent ces relations de coopération, mi-coopératives, mi-compétitives ? Comment sont-elles intégrées dans les stratégies éditoriales des acteurs de la presse ? Quels sont les problèmes et les conflits qui émergent, notamment sur la question des droits d'auteur ? Quelle est l'importance des infomédiaires en termes d'apport d'audience, et pour quelles catégories des sites d'information ?

Afin de répondre à ces questionnements, nous avons privilégié une approche socio-économique en prenant pour terrain le cas français. Cette recherche s'appuie principalement sur des entretiens semi-directifs avec des acteurs du marché de l'infomédiation ainsi qu'avec des acteurs du secteur de l'information journalistique engagés dans la diversification numérique¹. Ces entretiens sont complétés par l'étude d'un corpus de documents (décisions de justice, articles de presse, rapports publics, travaux scientifiques) en lien avec l'objet de la recherche.

Un infomédiaire majeur : Google

Dans le domaine de la presse, des sociétés comme l'américain Google assurent une fonction *d'infomédiation de l'information en ligne*, autrement dit d'organisation et de mise à disposition de documents numériques relatifs à l'actualité – triés et structurés par des algorithmes simulant une maîtrise de la sémantique – sur la base de requêtes par mots clés mais aussi de critères thématiques, linguistiques ou chronologiques.

Cette configuration place les éditeurs de la presse française qui disposent d'un site internet en situation de *coopétition* avec des acteurs tels que Google. Il s'agit de relations mi-coopératives, mi-compétitives : « *Internet est un réseau où se posent avec une acuité particulière des problèmes de coordination entre des firmes qui sont à la fois complémentaires et concurrentes. En effet, la gamme des services offerts ou supportés par le réseau est tellement vaste que les compétences et les actifs nécessaires pour les produire ne sont pas maîtrisables par une seule entité. Dès lors, toute production de service exige une coopération entre des firmes productrices de « briques de base » qui peuvent être par ailleurs concurrentes car il existe des recouvrements entre ces composants de base qu'elles peuvent produire. (...). [Cela signifie] « que des accords verticaux à composante concurrentielle doivent être conclus. Ils sont complexes à mettre au point car ils recèlent une importante conflictualité du fait des divergences d'intérêts entre des concurrents directs »* [Brousseau, 2001, p. 819].

Dans les relations que nouent les éditeurs de presse avec les infomédiaires, la composante coopérative consiste pour les premiers à établir des partenariats en laissant les seconds utiliser une partie des contenus et constituer des liens vers les sources d'origine générant ainsi du trafic. La composante concurrentielle provient du fait que, simultanément, les deux catégories

¹ La présente recherche mobilise douze entretiens effectués en 2007 avec des responsables de Google France, Wikio, Agence France Presse, Syndicat de la Presse Quotidienne Régionale (SPQR), Lagardere Active Media, Les Echos Editions Electroniques. Ces entretiens font partie d'un ensemble empirique plus large ayant donné lieu à la rédaction d'une note sectorielle sur les mutations de la *Filière Presse et Information* (F. Rebillard, B. Cabedoche, B. Damian-Gaillard, N. Smyrniaos – juin 2007), dans le cadre du Programme de recherche *Diversité culturelle et mutations des industries de la culture, de l'information, et de la communication* (coord. : P. Bouquillion et Y. Combès), programme hébergé par la Maison des Sciences de l'Homme Paris-Nord et soutenu par le Ministère de la Culture et de la Communication.

d'acteurs sont en concurrence directe sur le marché publicitaire, les infomédiaires prenant avantage du contenu cédé par les sites-médias afin d'attirer les internautes.

Les éditeurs cherchent à être présents en bonne position sur les résultats que le moteur de recherche généraliste produit en fonction des requêtes par mot clé des internautes. A la recherche d'une audience accrue, ils ont également intérêt à ce que leurs articles se trouvent en bonne position sur la page d'accueil du service spécialisé Google News (<http://news.google.fr>) qui s'avère être un levier puissant de visibilité. Selon les chiffres du panel Médiamétrie/NetRatings, l'audience de la version française de Google News oscille entre 1,5 et 2 millions d'utilisateurs uniques par mois. Si l'on prend en compte le fait que Google News renvoie systématiquement vers les sites d'origine des articles, cela fait entre 50 000 et 70 000 lecteurs potentiels par jour que se disputent les éditeurs, ce qui constitue un enjeu économique non négligeable. La domination de Google en tant que point nodal du réseau canalisant les flux de visiteurs est écrasante. Google est le « domaine » le plus visité de l'internet français avec 22.762.000 visiteurs uniques pour le mois de septembre 2007, celui qui détient la part de marché la plus importante parmi tous les moteurs de recherche en termes de requêtes (85,8%) mais également en parts de visites (90,2%)². Même si ces chiffres doivent être maniés avec précaution, en raison des différences dans les méthodes de mesure, il apparaît que l'apport global d'audience en provenance de Google constitue un enjeu majeur pour les sites d'information³.

Cependant, loin de constituer une collaboration sans heurts, les relations des éditeurs avec Google ont été marquées par des tensions et des conflits. Le point de départ étant la création du service Google News lancé en France en 2003.

Le principe qui anime Google News, dont Google Actualités est la déclinaison française, semble en apparence proche de celui présidant aux rubriques d'actualité des grands « portails » tels que Yahoo, MSN, ou Orange. Il s'agit de constituer une page web sous la forme de Une d'un journal électronique découpée en rubriques thématiques (« International », « France », « Economie », « Science », « Sport »,...). Sous chaque rubrique thématique se trouve un ensemble de titres et de brefs extraits d'articles, mais aussi des images, sous forme de liens hypertextes conduisant vers le texte intégral sur le site de la source. Le service, disponible en 22 versions nationales, comprend plusieurs milliers de sources.

La grande différence avec les pages d'information des portails généralistes, c'est que les articles ne se trouvent pas sur Google, mais sur les sites d'origine. Autrement dit, alors que les autres grands « portails » établissent des partenariats avec les producteurs de contenus afin de disposer du droit de les publier - le plus souvent contre rémunération -, Google ne fait que « pointer » vers des contenus externes pour lesquels il ne dispose d'aucun droit de publication. En exploitant sa technologie de recherche, il peut sonder en continu un grand nombre de sources présélectionnées et proposer des pages d'information à jour sans s'entendre avec les producteurs d'informations, ni a fortiori les rémunérer. Ce système est censé être profitable

² Sources : Médiamétrie/Netratings, *L'Observatoire des usages Internet*, septembre 2007 ; ComScore, *qSearch 2.0*, septembre 2007 ; XitiMonitor, *Baromètre des moteurs*, septembre 2007. La part de marché est le pourcentage de requêtes qui ont été faites sur Google par rapport à celles effectuées sur l'ensemble de moteurs. La part de visites est le pourcentage de visites que les sites français reçoivent de la part de Google par rapport à l'ensemble du trafic en provenance de tous les moteurs.

³ En raison de la variété des méthodes de mesure et de la multiplicité des sources, il est difficile de connaître exactement la part d'audience que Google apporte aux sites d'information journalistique. En croisant les données existantes, nous pouvons néanmoins estimer que *lemonde.fr* et *lefigaro.fr*, sites de la presse généraliste les plus visités, voient entre 20% et 30% de leur audience totale provenir des divers services de Google.

aux deux parties : Google exploite les contenus afin de maximiser son audience globale, tandis que les éditeurs bénéficient d'un « trafic » supplémentaire.

Or, la manière dont le service a été lancé en Europe a généré des conflits avec les éditeurs notamment en France et en Belgique. Le conflit provoqué par le lancement de Google News en France en 2003 a abouti au retrait temporaire du service d'un certain nombre d'éditeurs français, notamment ceux qui sont regroupés au sein du GESTE (Groupement des Editeurs de Services En Ligne) comme *Le Monde* et *Les Echos*. Hormis les différends secondaires (accès en cache à des articles payants, reproduction d'images sans autorisation des auteurs et/ou avec des références erronées) l'essentiel du désaccord portait sur la reconnaissance des droits d'auteur des éditeurs sur leurs contenus y compris les titres et les chapôs des articles reproduits par Google News. Il faut savoir que, en raison de « l'organisation matricielle » de la société Google, les décisions stratégiques concernant le lancement et l'implémentation des services au niveau mondial se prennent directement depuis les Etats-Unis sans aucune consultation préalable ni avec ses représentants locaux, qui ne disposent que des prérogatives commerciales, ni avec les acteurs impliqués, en l'occurrence les éditeurs.

Les éditeurs ont fait valoir le droit d'autoriser ou pas la reprise de leur contenu par Google News ainsi que de contrôler les modalités de cette reprise si elle a lieu. De leur côté, les représentants de la société américaine ont dans un premier temps soutenu qu'une telle approche rendrait l'activité de moteur de recherche impossible et, par extension, nuirait également aux intérêts des éditeurs. La principale crainte des dirigeants de Google a été le bouleversement de son modèle économique que causerait une demande de rémunération de la part des producteurs de contenus en cas de reconnaissance des droits d'auteur sur les extraits des pages web reproduits par le moteur. Surtout dans le cas où une telle demande déborderait du secteur de la presse vers l'ensemble des éditeurs de sites. Cette argumentation apparaît dans les décisions de justice rendues suite au procès qui a opposé la société américaine aux éditeurs belges, regroupés au sein de Copiepresse, et dans lequel Google a été condamné pour violation des droits d'auteur en 2006 et en 2007 par le Tribunal de Première Instance de Bruxelles⁴. Une autre procédure judiciaire qui n'a finalement pas abouti à un procès est celle qui a été initiée par l'AFP en 2005 auprès d'un tribunal de Washington. Comme nous le verrons plus loin cette procédure, qui a finalement abouti à la signature d'un accord annoncé en septembre 2007, s'est avéré un tournant important dans la stratégie de Google mais aussi dans celle de l'AFP.

Ces conflits, qu'ils aient pris la forme d'un procès ou non, démontrent les difficultés à coopérer que connaissent Google et les représentants du secteur de la presse, difficultés largement dues à des logiques différentes : le métier des éditeurs se fonde sur la propriété des contenus et consiste à organiser les modalités de leur exploitation, tant avec le niveau amont de la création qu'avec le niveau aval de la diffusion ; Google, tout comme nombre d'infomédiaires tels que Wikio ou Netvibes, emprunte à l'inverse son mode de fonctionnement, aussi bien technique qu'économique, à l'idéologie « originelle » de l'internet d'un accès libre aux données disponibles, sans prendre spécialement en compte les particularités de l'information journalistique. Si le processus de normalisation des relations entre ces acteurs a été long et difficile, nous pouvons affirmer qu'aujourd'hui celles-ci semblent s'être stabilisées. D'une part les éditeurs ont consenti des efforts dans ce sens, ce qui transparaît

⁴ Les décisions du Tribunal de Première Instance de Bruxelles sont disponibles en ligne aux adresses suivantes : < http://www.copiepresse.be/labo/jugement_copiepresse_google_fr.pdf > et < http://www.copiepresse.be/copiepresse_google.pdf >.

dans leurs stratégies respectives et, d'autre part, Google a progressivement intégré les logiques éditoriales dans le fonctionnement de son service d'actualité.

La traduction algorithmique des logiques éditoriales

L'un des principaux problèmes auxquels ont été confrontés les ingénieurs de Google dans leurs tentatives d'améliorer le fonctionnement de Google News a été le fait de pouvoir traduire les logiques complexes qui régissent le domaine de l'actualité à un langage algorithmique compréhensible par le moteur. Autrement dit, la question qui s'est posée a été de savoir comment et sur la base de quels critères il fallait hiérarchiser l'information recensée par le service.

Deux brevets déposés par Google auprès du *US Patent & Trademark Office* nous éclairent sur les pistes suivies par la société américaine, également confirmées par les représentants de Google France⁵. Le premier brevet déposé en 2005 ne l'a pas été par Google en tant que tel mais par une société de gestion de propriété intellectuelle pour le compte du premier. Parmi les noms des dépositaires nous retrouvons celui de Krishna Bharat, chef du centre Recherche & Développement de Google à Bangalore et connu pour être l'inventeur de Google News [Kramer, 2003]. La méthode consiste à transformer un certain nombre de critères d'évaluation sous forme des variables chiffrées. Le résultat des calculs complexes entre ces variables est une sorte d'indice appelé *source rank* qui pondère la place des articles en provenance d'une source donnée en fonction de sa « qualité ».

Les treize critères proposés peuvent être classés en deux catégories : ceux qui sont centrés sur la source et ceux qui sont centrés sur la production. Dans le premier cas, l'objectif semble être la traduction en langage algorithmique du « capital journalistique » d'un média, « correspondant à une reconnaissance qui s'obtient auprès de ses pairs, mais aussi auprès du public et de pouvoirs externes » [Duval, 2004, p.107]. Il en va ainsi du nombre de liens qui pointent vers un site d'information – ce qui constitue une composante importante de l'indice *PageRank* utilisé par le moteur de recherche généraliste⁶ – ; de son audience, y compris à l'international ; de la taille de sa rédaction et du nombre de bureaux à l'étranger dont il dispose ; et enfin de « l'opinion humaine sur la source d'information »⁷ qui peut être mesurée par sondage auprès des internautes ou, alternativement, en fonction de son âge et des prix journalistiques reçus par le passé. A ce jour, rien n'indique qu'en dehors du *PageRank* les autres méthodes de mesure centrées sur la source ont été effectivement intégrées dans l'algorithme de Google News.

En revanche, les critères centrés sur la production semblent avoir été privilégiés par les ingénieurs de Google. Le premier d'entre eux est basé sur le nombre d'articles produits par une source. Une production quotidienne élevée contribue à ce que les articles d'une source donnée soient mieux placés sur la page d'accueil de Google News. La satisfaction de ce critère implique une organisation particulière des rédactions, se rapprochant davantage de celle d'une agence de presse, dans le sens d'une production en flux tendu d'un nombre élevé d'articles par jour.

⁵ Curtiss M., Bharat K., Schmitt M., « Systems and methods for improving the ranking of news articles », United States Patent Application 20050060312, mars 2005, et Charikar M. S., « Methods and apparatus of estimating similarity », United States Patent 7,158,961, janvier 2007.

⁶ Sur le fonctionnement du moteur généraliste de Google et notamment la « loi du plus grand nombre de liens », voir Hindman et alii, 2003.

⁷ Curtiss M., Bharat K., Schmitt M., 2005, *op. cit.*

Le deuxième critère est la taille des articles, autrement dit le volume de texte qu'ils représentent. Il s'agit là d'un critère utilisé également par le moteur généraliste qui privilégie les sites à contenu volumineux dans ces résultats. Ce critère est complété dans le brevet par une variable qualitative liée à la « qualité » de l'écriture mesurée en fonction des fautes d'orthographe et de grammaire repérées de manière aléatoire dans le corpus des articles recensés.

Le troisième critère de production est une mesure appelée *story size*. Cela consiste à évaluer l'importance d'un sujet en fonction du nombre de sources qui le traitent. Plus le *story size* des articles produits par une source donnée est élevé, mieux celle-ci sera référencée par le service. Ce facteur est également primordial dans le processus algorithmique d'« émergence » de la Une de Google News. Les articles qui traitent du même thème sont rassemblés au sein de *clusters* ou « grappes d'information ». La taille du *cluster* définit l'importance d'un sujet (*topic*) et fait apparaître celui-ci en bonne position sur la page d'accueil du service. Ce fonctionnement de Google News matérialise un phénomène déjà bien connu pour les médias traditionnels, vulgarisé et dénoncé par la formule de « circulation circulaire de l'information » [Bourdieu, 1996]. Il tend à privilégier les sujets d'actualité qui font l'objet d'un traitement médiatique par un grand nombre de sources au détriment des informations exclusives et peu reprises. De ce point de vue, Google News participe d'un plus large mouvement de « retraitement » des informations journalistiques, amplifié et systématisé avec l'internet [Rebillard, 2006].

Le quatrième critère centré sur la production est une mesure appelée *breaking score* qui représente la capacité d'une source à produire un article peu de temps après que l'événement traité s'est produit. Autrement dit, plus l'apparition d'une information sur un site est rapide, par rapport aux autres sources, mieux celui-ci sera noté par l'algorithme. Il s'agit là d'évaluer la réactivité d'une rédaction en fonction des événements. La satisfaction de ce critère implique une activité de veille considérable pour les journalistes et une organisation du travail permettant de diffuser un contenu extrêmement rapidement. *In fine* ce critère tend à privilégier la source qui se trouve à l'origine d'une information puisque, par définition, le moment de la première publication sur un sujet donné constitue la référence à partir de laquelle est calculé le *breaking score* des articles suivants.

Le cinquième critère est lié à l'éventail des thèmes traités par une source. Cette dernière est mieux référencée si elle traite un nombre élevé des sujets différents, ce qui tend à privilégier les sources généralistes au détriment des sites d'information spécialisée.

Enfin, un dernier critère proposé dans le brevet est celui de l'*original reporting*, autrement dit de la production d'information originale. Il s'agit pour Google de pouvoir hiérarchiser les articles qui traitent du même thème, et qui se trouvent ainsi dans le même *cluster*, en mesurant la capacité des sources qui en sont à l'origine à produire des « scoops » ou à enrichir une information déjà parue. La méthode proposée consiste à mesurer le nombre de noms propres contenus dans un article (noms de personnes, de villes, d'entreprises etc.). Quand le nombre de noms propres est élevé, la source à l'origine de l'article est considérée comme ayant apporté des informations supplémentaires par un travail de reportage, ce qui entraîne un bonus dans le classement.

Le service Google News évolue de manière quasi-ininterrompue depuis son lancement en 2003 et son fonctionnement exact constitue un secret bien gardé. De ce fait il est impossible

de savoir avec exactitude si tous les critères mentionnés sont effectivement utilisés dans l'algorithme de Google News, et dans quelle mesure le cas échéant. Cependant, on peut observer que le profil de source privilégiée, tel qu'il est esquissé par ces critères, correspond à celui exposé sur la page d'accueil de Google News : un site d'information généraliste, très ancré dans l'actualité « dominante », disposant d'une rédaction de taille relativement importante, avec une production journalière élevée en termes de volume et diffusée en flux tendu. En effet, parmi les sites dont les articles sont repris le plus souvent sur la page d'accueil de Google News, nous retrouvons des sources qui correspondent à ce profil⁸. Comme nous le verrons par la suite, il s'agit à la fois de la traduction des stratégies mises en place par les éditeurs afin d'augmenter leur visibilité sur l'internet, y compris sur les différents services de Google, et réciproquement de l'option suivie par ce dernier de répondre aux demandes des médias dominants.

En effet, la mise en avant des sources qui sont à l'origine des informations et qui ne se bornent pas à un simple travail de réécriture est l'une des demandes formulées de manière répétée par certains membres du GESTE, l'organisme professionnel rassemblant les entités multimédias des principaux éditeurs de presse français. Les responsables de Google News cherchent donc un moyen de les repérer afin de leur attribuer une « prime à l'originalité ». Ainsi, l'une des directions importantes du développement de Google News est la recherche d'une formule automatisée qui permettrait de reconnaître les doublons, c'est-à-dire les articles quasi identiques qui proviendraient de sources différentes mais qui seraient en grande partie basés sur les mêmes informations, en provenance par exemple d'une dépêche d'agence. L'objectif à terme est que le moteur puisse distinguer l'information originale des reprises, et « récompenser » ainsi le média qui en est à l'origine en plaçant son article en bonne position sur la page d'accueil. Cette technique appelée *duplicate detection* existe déjà pour la recherche traditionnelle, mais doit être adaptée pour le domaine de l'actualité. Il s'agit de vérifier, sur la base des critères algorithmiques, le degré de similitude de deux documents numériques et d'éliminer ou de cacher l'un d'entre eux à partir d'un seuil de ressemblance, considérant qu'il s'agit d'une quasi-copie. C'est l'objet du second brevet, d'une portée plus générale, déposé par Google en janvier 2007. Les applications envisagées sont très larges, Google News ne constituant qu'une d'entre elles.

Enfin une dernière évolution concernant Google News est la mise en place d'un service de *sitemaps* spécifiquement dédié aux sites d'information. Le principe des *Google Sitemaps* consiste pour le webmaster à fournir aux moteurs de recherche un plan du site, afin de gérer l'indexation des pages par les robots des moteurs qui scrutent en permanence le web. Il s'agit en quelque sorte de traduire une « politique éditoriale » en un langage compréhensible par le moteur, en offrant par exemple la possibilité au webmaster de supprimer de l'index de Google certaines pages de son site. Cette fonctionnalité faisait déjà partie, en 2003, des revendications du GESTE lors des négociations avec les représentants de la société américaine. Parallèlement, le service *Google Sitemaps* fournit un certain nombre d'informations utiles pour la gestion d'un site internet comme le nombre de liens entrants à une page web, leur origine, le nombre d'abonnés aux fils RSS, les erreurs d'indexation du moteur etc. Pour la société américaine il s'agit, tout en se rendant indispensable, de rassurer les éditeurs de sites

⁸ La liste des sources les mieux représentées sur la page d'accueil de la version française du service (>2.000 articles cités pour la période allant de janvier 2007 à octobre 2007 toutes rubriques confondues) sont dans l'ordre : *Le Monde* (8.419 articles soit environ 23 par jour), *Le Figaro* (7.165), *Nouvelobs.com* (5.622), *Libération* (4.302), *L'Express* (4.161), *TF1* (3.566), *Challenges* (2.693), *Boursorama* (2.535), *Football.fr* (2.502), *20 Minutes* (2.164) *Romandie.com* (2.190), *JDD* (2.182), *Sport365.fr* (2.039), *La Tribune* (2.034).

Source : Google News Report France, Source Rankings, <<http://88.198.49.15/~search/gnews/reports/fr/200709/sources.php>>

en leur offrant une possibilité, même réduite, de contrôler comment leurs pages web sont indexées et rendues accessibles par le moteur de recherche.

La variante de *sitemaps* pour les sites d'information concerne ceux qui sont indexés par le service Google News. Selon la société américaine, l'envoi d'informations spécifiques par le biais d'un *sitemap* dédié permet au moteur de Google News d'indexer plus efficacement et plus rapidement les articles augmentant ainsi leurs chances de générer du trafic. Il s'agit également de « permettre aux webmasters de mieux contrôler le contenu qui s'affiche dans Google Actualités »⁹. Le *sitemap* dédié à l'actualité n'est pas une liste exhaustive de toutes les URL figurant sur un site web, comme le *sitemap* traditionnel, mais un fichier dynamique comprenant uniquement les articles publiés les trois derniers jours. Ce fichier est enrichi avec des métadonnées comme la date et l'heure de publication, des mots-clés qui décrivent la thématique, ainsi que d'autres – qui ne sont pas sans rappeler les critères de *source rank* précédemment cités – comme les noms propres d'entreprises ou les libellés des valeurs en bourse mentionnés dans les articles. De toute évidence, il s'agit là d'une adaptation de Google aux demandes exprimées de la part des éditeurs souhaitant contrôler l'exploitation de leur contenu. Cependant, dans le même temps, il s'agit pour Google de pouvoir ressembler des métadonnées qui sont utilisées dans l'attribution du *source rank* et qui ont donc des répercussions directes sur le processus de hiérarchisation de l'information. On peut voir ici, à travers cet échange de bons procédés à un niveau technique, une illustration de relations plus larges de *coopétition* qui nous semblent caractériser les rapports entre Google et les éditeurs de presse.

Les stratégies individuelles des éditeurs

Les stratégies sur l'internet des acteurs de la presse dépendent de facteurs multiples et ne peuvent se réduire à un effort d'adaptation aux exigences des infomédiaires. Cependant, nous pouvons déceler parmi les actions entreprises, que ce soit de manière individuelle au niveau d'un groupe de presse ou par le biais d'organisations interprofessionnelles, certaines décisions qui vont dans le sens d'une telle adaptation. Sans toujours aller jusqu'à constituer des rédactions « asservies au référencement de Google »¹⁰, il s'agit pour les éditeurs d'intégrer dans leur fonctionnement des logiques propres au média internet. Ce virage stratégique a été notamment impulsé par l'arrivée progressive de cadres dirigeants en provenance du monde de l'informatique et des télécommunications à la tête des filiales multimédia de plusieurs éditeurs¹¹.

La première étape dans la quête des éditeurs d'une audience importante sur l'internet, y compris par le biais de moteurs de recherche, consiste à pouvoir assurer en interne une production d'information régulière, en s'affranchissant largement de la temporalité de leurs médias d'origine. D'où la constitution des rédactions dédiées. Ces rédactions numériques ne se substituent pas aux rédactions papier mais les complètent. Cet état de fait témoigne plus largement, de la part des groupes de presse, d'une prise de conscience et d'une appréhension beaucoup plus mûre et maîtrisée des spécificités de l'information sur l'internet. Les personnels oeuvrant dans ces rédactions numériques sont largement acculturés aux

⁹ Aide pour les éditeurs, < http://www.google.fr/support/news_pub/bin/answer.py?answer=74479&topic=11671>

¹⁰ Expression utilisée par le responsable d'un groupe de presse lors de nos entretiens.

¹¹ A titre d'exemple Didier Quillot, président du directoire de Lagardère Active Media, et Julien Billot, responsable de la Direction numérique de Lagardère, ont été récemment recrutés après un parcours professionnel chez l'opérateur de télécommunications Orange. De son côté, la Directrice générale du Monde Interactif Dao Nguyen est diplômée d'informatique et a commencé sa carrière aux Etats-Unis en tant que consultant dans le domaine de l'internet.

spécificités de l'information en ligne. Jeunes, leur recrutement est perçu comme un gain de productivité en comparaison avec les personnels des rédactions papier, car plus productifs au niveau des informations confectionnées à la journée, et moins coûteux au niveau salarial (rémunérations et avantages professionnels moindres).

La deuxième étape dans le développement en ligne de certains éditeurs, notamment ceux qui accusent actuellement un déficit au niveau de leur audience internet, a été l'acquisition par croissance externe des compétences techniques et d'un savoir-faire éditorial spécifiques au web. Il en va ainsi de l'acquisition de *pure players* de l'information sur l'internet par des groupes de presse ces derniers mois : *sport24.com* et *evene.fr* par la Socpresse (sociétés acquises en février 2006 et mai 2007) ; Thotnet (*Infobébés*), Newsweb (*Sports.fr*, *Boursier.com*, *Autonews.fr*), et Nextedia par Lagardère Active Media (sociétés acquises entre l'été 2006 et l'été 2007) ; *programme-tv.net* par Prisma Presse (février 2007) ; *aufeminin.com* par Axel Springer (juin 2007). Quant à Next Radio TV (RMC, BFM), rappelons que son rachat du groupe de presse informatique Tests en février 2007 était en grande partie motivé par l'acquisition des sites *01net* et *01men*.

Si les rédactions numériques peuvent parfois être le lieu de création d'informations originales, elles sont surtout un atelier de réagencement de contenus existants. Aux *Echos*, le travail essentiel des rédacteurs web consiste à alimenter le site en informations d'actualité, principalement à partir d'un travail de *desk* : appui sur les dépêches d'agences de presse, étude des communiqués et des rapports d'activité des entreprises à partir de leur site, et, plus rarement, appels téléphoniques auprès des protagonistes. Les rédacteurs numériques ont une tâche supplémentaire à accomplir, qui renvoie à la recherche d'un référencement efficace par les moteurs de recherche : l'attribution de mots-clés à leurs articles en vue de l'indexation dans la base de données des *Echos*. Les articles des journalistes papiers font eux l'objet d'une indexation par des employés spécifiquement affectés à cette tâche (des documentalistes) qui classent les articles papier au moment de leur mise au format XML et de leur basculement dans la base de données. Les rédactions numériques de Lagardère Active Media présentent les mêmes caractéristiques : tâches techniques supplémentaires, et productivité plus élevée quitte à ce que les contenus soient moins originaux.

Les filiales web de Lagardère Active Media – Newsweb et Thotnet – effectuent un travail de ré-écriture des articles des éditions papier. Ce travail de ré-écriture ne consiste pas seulement en une adaptation aux formats de l'écran d'ordinateur (textes plus courts, plus directs, etc.) comme on le conçoit couramment. Il s'agit aussi d'un travail d'optimisation des contenus en vue de leur référencement sur l'internet, et en particulier auprès des infomédiaires. Les filiales web de Lagardère Active Media sont ainsi chargées de truffer les articles numérisés des termes les plus stratégiques pour apparaître dans les sommets des classements des moteurs de recherche. De façon plus générale, ces *ex-pure players* apportent une valeur ajoutée liée à leur expertise des techniques d'indexation sur l'internet.

Chacune des rédactions numériques du groupe, qui correspond à un support papier, est liée par l'intermédiaire de son « rédacteur en chef numérique » aux filiales web de Lagardère Active Media récemment acquises. En fonction du type de contenu journalistique produit par le groupe - plutôt contenu de flux pour les quotidiens et plutôt contenu de stock pour les magazines - c'est l'une ou l'autre des filiales web de Lagardère Active Media qui sera chargée de son traitement pour l'internet. Pour tout ce qui concerne les informations « chaudes » ou d'actualité, l'activité éditoriale est ainsi élaborée à partir de la société Newsweb. C'est déjà le cas pour la version internet du *Journal du Dimanche*. A terme, tous les sites de

news seront gérés à partir de Newsweb, qui possède déjà des sites dans le domaine de l'automobile, des finances, et du sport. Pour ce qui concerne l'information avec une dimension de base de données importante, c'est la société Thotnet qui est cette fois mobilisée. Elle regroupe les équipes éditoriales de *elle.fr*, *parents.fr*, ou encore *premiere.fr*. Nous retrouvons également une distinction similaire dans l'organisation interne de la rédaction multimédia du Monde Interactif. Une partie de la rédaction est affectée à la production des articles de flux liés à l'actualité immédiate du moment, alors qu'une autre partie se charge elle de produire des articles plus longs à un rythme moins soutenu [Attias, 2007].

La distinction établie ici entre information de flux et information de stock rejaille ainsi sur l'organisation de la production numérique. Une telle distinction s'avère être en phase avec les logiques d'infomédiation de Google, grand pourvoyeur d'audience. D'une part les contenus de l'actualité immédiate, référencés par Google News, impliquent une organisation du travail spécifique (notamment la production d'articles en flux tendu) qui renvoie au fonctionnement des agences de presse ou des chaînes d'information en continu. D'autre part, les contenus « magazine », référencés par le moteur de recherche généraliste, nécessitent quant à eux un savoir-faire davantage documentaire et similaire à celui des gestionnaires de bases de données. Encore une fois, ceci montre que les groupes de presse ont franchi un stade important – un stade « cognitif » – dans leur analyse des développements sur l'internet.

Un cas particulier est celui de l'AFP qui n'est pas un éditeur de la presse à proprement parler et qui, à l'inverse des acteurs précédemment mentionnés, n'a pas *a priori* de vocation à toucher directement le grand public. A partir des dernières évolutions, nous pouvons affirmer que l'AFP est passée du statut d'opposant affirmé de Google à celui de partenaire privilégié. En effet, comme nous l'avons mentionné précédemment, l'AFP a été la première grande organisation médiatique à assigner Google en justice pour violation des droits d'auteur. La plainte déposée aux Etats-Unis visait la reproduction par Google News des titres et courts extraits de dépêches ainsi que d'images produites par l'AFP et publiées sur les sites de ses clients. Après deux années de procédure préliminaire (*discovery*) le différent juridique entre l'AFP et Google s'est soldé par la signature d'un accord entre les deux sociétés, accord annoncé fin août 2007. L'annonce a suivi des accords similaires signés entre Google et les agences Associated Press, Canadian Press et Press Association.

Cet accord commercial, dont les dimensions stratégiques pour les deux acteurs impliqués sont importantes, prévoit l'achat d'une licence par Google, du même type que celle dont disposent Yahoo et MSN, pour exploiter le Journal Internet de l'AFP. Détail supplémentaire : l'accord prévoit également que Google se charge d'héberger dans ses propres pages les dépêches de l'agence française. Finalement, Google s'aligne dans ses tous récents rapports avec l'AFP sur le modèle d'intégration rémunérée de contenus propriétaires à l'œuvre depuis plusieurs années au travers des rubriques d'actualité des portails comme Yahoo ou MSN. Il est encore trop tôt pour dire s'il s'agit là pour Google d'une exception à sa règle de distribution sans contrepartie ou d'un virage stratégique de la part de cet infomédiaire si particulier jusqu'ici. Toujours est-il que la priorité semble pour Google de maîtriser totalement le référencement de ces articles et leur distribution par son service d'actualité.

Pour l'AFP, cet accord avec Google présente un double intérêt. En dehors de son intérêt financier évident, il participe de la prolongation d'une stratégie amorcée par ailleurs qui consiste à descendre dans la chaîne d'information vers le consommateur final et qui trouve avec Google une plateforme puissante de distribution grand public. Sur ce dernier point, la position de l'AFP est ambivalente. Le responsable de l'accord avec Google nous a réaffirmé

que le métier de l'AFP est celui d'un grossiste en informations, orienté principalement vers le BtoB et non le BtoC. Une preuve de cette volonté de ne pas court-circuiter les éditeurs de presse serait illustrée par un lien hypertextuel vers les sites de ces derniers, sur chaque page AFP placée au sein de Google Actualités. Néanmoins, les expériences menées par l'AFP à l'étranger (application *user-generated-content* de photographies au Japon ou encore blogs depuis les Etats-Unis) laissent penser que l'agence française ne perd pas de vue ce marché du BtoC déjà bien occupé par son principal concurrent Reuters, qui dispose d'un site grand public et bénéficie déjà d'une audience significative.

Pour Google, l'accord avec l'AFP constitue également une occasion pour « nettoyer » son service d'actualité. En effet, il y a une véritable lutte entre les éditeurs visant à conquérir les places privilégiées sur Google News. Cette situation est équivalente à celle qui voit les webmasters essayer d'améliorer la position de leurs sites dans l'index global de Google en utilisant parfois des méthodes abusives pensées pour tromper le robot d'indexation et appelées communément *spamindexing*. L'équivalent pour l'éditeur d'un site d'information est par exemple la pratique, mentionnée par les responsables de Google lors de nos entretiens, consistant à publier la même information plusieurs fois, souvent basée sur une dépêche d'agence, sous un titre différent et avec un texte légèrement modifié. L'intégration des dépêches de l'AFP directement au sein de Google News signifie la relégation à des positions moins visibles des articles qui sont largement inspirés de ces mêmes dépêches. Ainsi, l'annonce de l'accord entre Google et l'AFP a dans un premier temps soulevé des réactions de la part de nombreux éditeurs inquiets de voir se tarir la source d'audience que constitue Google News. Cependant, les changements qu'a apportés cet accord dans la hiérarchie des sources recensées par Google News se révèlent mineurs.

En effet, en comparant le classement¹² des sources les plus citées dans Google News avant et après l'accord Google-AFP nous observons que les changements intervenus sont mineurs. Outre l'entrée de l'AFP en bonne position dans Google News, il y a l'arrivée des quelques sources relativement marginales mais productrices d'information originale comme *Journal Chrétien* et *AgoraVox*. Celles-ci supplantent dans les premières places du classement des sources étrangères dont la couverture de l'actualité française est basée essentiellement sur des dépêches de l'AFP comme Radio Canada, Cyberpresse et RTL Belgique. Parallèlement, des sources d'information financière comme *Boursier.com* et *Boursorama.com*, qui apportent une valeur ajoutée journalistique minime aux dépêches de l'AFP qu'elles diffusent, semblent être les plus pénalisées. D'autres facteurs sont à prendre en compte pour évaluer ces changements comme la variation normale d'une période à une autre, le fait pour certaines sources d'utiliser le fil de dépêches de Reuters, ou les différences qui existent au niveau du contenu entre le produit grand public qu'est le Journal Internet de l'AFP et le fil de dépêches vendu aux professionnels de l'information. En effet, selon les dires des responsables de Google France et de l'AFP, l'accord entre ces deux acteurs ne doit pas nuire aux relations qu'ils entretiennent respectivement avec les éditeurs de presse. Il s'agit de trouver un nouvel équilibre entre agences, infomédiaires et éditeurs sans pour autant bouleverser le modèle d'interdépendance qui régit leurs relations tel qu'il s'est stabilisé au fil du temps.

¹² L'outil qui produit ce classement visite à intervalles réguliers la Une de Google News et recense les articles qui y sont présents. A partir de cet élément mais également d'autres (la fréquence et l'heure d'apparition, la position sur la page etc.) l'algorithme calcule une note censée représenter l'apport de l'audience de Google News pour les sources présentes dans le service. Plus la note est élevée, plus l'audience potentiellement renvoyée par Google l'est aussi. Si ce critère peut être sujet à précaution en raison de l'arbitraire du calcul, l'outil fournit également le nombre d'articles repris par source ce qui constitue un élément d'évaluation plus objectif. Google News Report est disponible à l'adresse :
< <http://88.198.49.15/~search/gnews/reports/fr/index.php> >

Les actions collectives des éditeurs

En dehors des stratégies individuelles que les éditeurs et agences de presse adoptent afin d'optimiser les nouveaux canaux de distribution que constituent les dispositifs d'infomédiation, il existe d'autres actions qui s'organisent au niveau collectif. En effet, un certain nombre d'organisations interprofessionnelles prennent des initiatives afin d'encadrer les relations que leurs membres établissent avec Google. Le principal problème auquel sont confrontées ces organisations afin d'établir des stratégies cohérentes à ce sujet est le fait que souvent elles représentent des acteurs concurrents. L'obtention d'un consensus est de ce fait souvent difficile.

Un exemple de cette situation est le GESTE. Son directeur Philippe Jannet, par ailleurs directeur des Editions électroniques des Echos, est parmi les premiers à avoir réagi au lancement de Google News en France. Appuyé par des membres du groupement comme les responsables du Monde Interactif, il a très vite tenté de prendre contact avec les représentants de la société américaine afin de leur faire part des exigences des éditeurs. Cette initiative a constitué la suite logique des actions du GESTE pour mieux encadrer l'activité d'information en ligne d'un point de vue juridique, comme la rédaction d'une charte d'édition électronique dès 2001¹³. Par ailleurs, des actions ont été entreprises par le GESTE directement en réaction à la création de Google News et la prolifération des services d'infomédiation, comme la mise en place d'un lecteur des fils RSS appelé AlertInfo¹⁴.

Cependant, les démarches du GESTE au sujet de Google News n'ont pas été soutenues activement par tous ses membres et notamment par ceux qui avaient réussi dans un premier temps à exploiter au mieux ce service et à accéder ainsi à une audience nouvelle. Globalement, les exigences formulées par le GESTE – comme par exemple le fait que Google News accorde une « prime à l'originalité » aux sources productrices d'information – représentent davantage les intérêts des médias généralistes dominants, en provenance notamment de la presse nationale, que ceux des éditeurs de sites d'information spécialisés dont le fonctionnement se base davantage sur le retraitement d'informations existantes. Le GESTE constitue actuellement un interlocuteur privilégié des principaux infomédiaires comme Google, Wikio ou Netvibes qui en sont devenus membres. Il a notamment imposé le principe d'un accord explicite de la part d'un éditeur pour toute reprise de contenu de la part de Google News.

Une autre instance représentative qui a mis en place des actions collectives au sujet de Google News est le SPQR (Syndicat de la presse quotidienne régionale) notamment par le biais de sa Commission de développement. Le principe qui a régi les actions entreprises par le SPQR est l'idée que les éditeurs doivent rester maîtres de la distribution de leur production, y compris sur l'internet. La position du SPQR est encore plus radicale que celle du GESTE au sujet des infomédiaires, puisque ses représentants considèrent que toute exploitation des articles de la presse régionale doit se faire uniquement dans le périmètre des sites propriétaires des éditeurs. De ce fait, ils sont hostiles au principe même d'infomédiation sans contrepartie financière. Lors du lancement de Google News, après une période d'observation de trois mois, le SPQR a demandé le retrait du service de tous les articles de ses adhérents, ce qui a été fait début 2004. L'accord trouvé avec Google a entériné le principe d'une permission explicite et *a priori* de la part d'un éditeur pour toute reprise de son contenu.

¹³ Accessible à l'adresse : <http://www.geste.fr/9_commissions/commi_pres1.htm>

¹⁴ <<http://www.geste.fr/alertinfo/home.html>>

Dans les faits, aucun éditeur de la presse régionale n'est actuellement présent dans Google News. Après avoir examiné le fonctionnement du service, notamment par une présentation de Nathan Stole chef de projet chez Google, les représentants du SPQR ont considéré que l'algorithme de Google News n'est pas adapté à la nature de l'information régionale. En effet, celle-ci est caractérisée par une certaine exclusivité dont dispose chaque titre sur l'actualité de sa région. Or, la visibilité d'une information sur Google News dépend du *story size*, autrement dit de l'occurrence de cette information sur un nombre élevé de sources. De ce fait, pendant la période d'essai de trois mois, les informations de la presse régionale n'ont émergé que très rarement sur la Une de Google News, ce qui a abouti à une visibilité moindre pour les titres. Les éditeurs de la presse régionale ont considéré alors qu'ils n'avaient pas intérêt à être présents dans le service d'actualité de Google.

Les discussions actuelles entre le SPQR et Google, qui rejoignent celles qui ont lieu aux Etats-Unis, portent sur la mise en place d'un service dédié à l'actualité régionale avec un algorithme adapté et sur la base d'un accord commercial comme celui que la société américaine a signé avec l'AFP. Parallèlement, le SPQR travaille également sur l'idée de mise en place d'un portail dédié à l'actualité régionale, géré conjointement par les éditeurs à l'image de l'outil de développement en vue de l'offre de panorama de presse régionale.

Enfin, une dernière initiative collective des éditeurs face à la prolifération des services d'infomédiation, cette fois-ci à un niveau international est celle entreprise par la World Association of Newspapers (WAN), le European Publishers Council (EPC) et la International Publishers Association (IPA). Il s'agit du projet ACAP (Automated content access protocol)¹⁵. L'objectif de ce projet conjoint est la mise en place d'un protocole universel pouvant traduire une politique éditoriale dans un langage algorithmique. Les documents produits par les responsables du projet placent cette initiative dans le cadre plus large des relations entre moteurs de recherche et éditeurs. Prenant acte du rôle-pivot qu'occupent les moteurs de recherche sur le réseau pour les acteurs qui sont à la recherche d'une audience élevée, les représentants de ACAP s'efforcent d'y apporter une réponse satisfaisante. Il s'agit, tout en permettant l'indexation automatique de leurs sites par les moteurs, de donner la possibilité aux éditeurs de contrôler cette indexation et l'exploitation qui en est faite.

Le concept se base sur le principe d'intégration des métadonnées aux documents numériques définissant les conditions d'utilisation de manière compréhensible pour une machine. Le protocole se décline sur deux niveaux, d'une part la permission accordée à un moteur pour son indexation et d'autre part la permission que celui-ci accorde à son tour aux utilisateurs pour accéder à un contenu tiers. Les principes qui régissent ACAP (fonctionnalité, flexibilité, universalité, automaticité) traduisent le souci des éditeurs de concevoir un outil en osmose avec le fonctionnement des moteurs de recherche tout en gardant une certaine maîtrise des conditions de diffusion. Nous pouvons noter par ailleurs que l'imposition d'un protocole universel d'indexation des sites d'information constitue un enjeu de première importance également pour les infomédiaires. Ainsi, le service Google News Sitemaps remplit des fonctionnalités comparables à celles d'ACAP et constitue de ce point de vue un protocole concurrent. Après un an de travaux, la première version de ACAP a été présentée à New York en novembre 2007.

¹⁵ < <http://www.the-acap.org/> >

Conclusion et perspectives

A l'issue de cette observation, le constat peut être établi d'une maturation des relations entre Google et les éditeurs de presse français. Chacune des parties en présence semble avoir pris conscience des logiques respectives de ses partenaires, les éditeurs intégrant progressivement la logique d'un accès « libre » à l'information héritée des origines académiques de l'internet sur laquelle se fonde Google, et ce dernier devant composer avec les sources d'information majeures et légitimes que représentent les médias établis. On retrouve finalement ici pour le développement de l'information journalistique sur l'internet un ajustement de positions - non exempt de conflictualité - entre acteurs autour d'une logique commune en gestation, mouvement comparable à celui qui se dessine dans d'autres filières des industries culturelles confrontées à la diversification numérique : *« la mise en ligne de contenus sur des supports de plus en plus diversifiés et l'accès direct aux utilisateurs sont révélateurs de l'entrée de nouveaux acteurs (opérateurs de télécommunications, moteurs de recherche, développeurs d'outils logiciels, etc.) qui ont intérêt à rechercher un autre type de structuration de ce champ. Cependant, ils doivent jouer avec les structures industrielles en place, en particulier celles des contenus (médias, presse, éditeurs). La détermination d'un équivalent doit d'abord être établie par les acteurs, puis devenir ensuite un référent social. »* [Bouquillion, Combès, 2007, p. 26].

La configuration qui s'élabore dans le cas de l'information journalistique place Google à un niveau méta-éditorial [Chartron, Rebillard, 2004], à mi-chemin entre l'édition et la distribution dans la filière ainsi remodelée avec l'internet. Google n'est pas un simple distributeur car son rôle ne se limite pas à acheminer les contenus jusqu'à l'utilisateur final, mais ce n'est pas non plus un éditeur à proprement parler car Google se contente d'assembler sur ses pages des extraits de contenus édités par des tiers. En fait, sa fonction paraît s'inscrire plus largement dans le modèle du « courtage informationnel » [Moeglin, 2005], consistant à mettre en relation des demandes extrêmement variées avec une offre d'informations toute aussi multiple dans le contexte de l'internet.

Le recours à la notion de modèle socio-économique permet de penser l'articulation, au sein d'une filière donnée, entre la fonction centrale, son mode de valorisation principal, et ses répercussions sur les autres niveaux de la filière. Dans le cas étudié ici, l'infomédiation constitue la fonction centrale. La valorisation économique des services d'infomédiation ne s'effectue pas par le biais d'un paiement par l'utilisateur final en aval mais via des modalités de financement indirect (publicité, liens sponsorisés, exploitation des données collectées auprès des internautes), ce qui explique les négociations en cours avec les éditeurs, plus en amont dans la filière. En effet, dans un tel modèle, les éditeurs de presse voient disparaître l'une de leurs sources de revenus : la vente de contenus au consommateur final¹⁶. Parallèlement, leur autre source de revenus, celle de la vente d'espaces publicitaires, peut bénéficier de l'apport d'audience amené grâce à l'infomédiation, mais sans que le gain ainsi procuré ne parvienne aujourd'hui à compenser le manque à gagner dû à la disparition de la première source de revenus, si l'on en croit les déclarations des éditeurs interrogés¹⁷.

¹⁶Insistons bien sur le fait que notre raisonnement porte ici uniquement sur le modèle de courtage informationnel, qui n'est qu'un des modèles dans lesquels les éditeurs de presse s'inscrivent sur l'internet. Dans ce cas précis, la source de revenus constituée par la vente à l'utilisateur final se tarit pour les contenus accessibles via les infomédiaires tels que Google News, puisqu'ils sont accessibles gratuitement par définition. Cependant, il faut garder à l'esprit que la vente d'archives ou d'abonnements sur le site propre des éditeurs existe par ailleurs et qu'elle peut également être indirectement favorisée par la visibilité des contenus procurée par Google, que ce soit le moteur généraliste ou le service d'actualité.

¹⁷Le gain potentiel en termes publicitaires est aussi à relativiser. Google avec ses « liens profonds » apporte des visiteurs mais peut potentiellement diminuer les pages vues puisque ceux-ci accèdent directement à la page de l'article sans suivre le trajet préconçu du journal en ligne pour une exposition maximale à des publicités. Ce qui amène une réflexion des éditeurs quant à

De ce fait, nous sommes aujourd'hui dans une situation potentiellement lourde de conséquences pour le pluralisme de l'information. D'une part, les principaux éditeurs négocient avec Google pour être prioritairement « pointés » par cet infomédiaire majeur, et en ce sens tentent d'obtenir un traitement de faveur par rapport aux autres producteurs d'information, de plus petite taille ou moins reconnus. D'autre part, pour compenser le tarissement de la source de revenus représentée par la vente de contenus, les éditeurs de presse cherchent à réaliser des économies sur leurs dépenses en matière de création d'informations : les moyens moindres investis dans les rédactions numériques par rapport aux rédactions traditionnelles (ou le couplage entre les deux), tout comme l'appui sur les « contenus générés par l'utilisateur », afin de reporter sur l'utilisateur final les coûts de conception de l'information, illustrent de telles stratégies.

Nous avons pointé ici des problèmes essentiels du point de vue de l'économie politique de la communication, mais qui ne sauraient valoir en dehors du cas précis de l'infomédiation de l'information en ligne opérée par Google. Même si ce dernier constitue l'acteur majeur en la matière, il faudrait pour se prononcer plus largement sur le sujet et avoir une vue d'ensemble, multiplier les approches : s'intéresser à d'autres infomédiaires tels que Wikio qui cherchent explicitement à diversifier les sources pour mettre en avant une information de niche ; considérer également les stratégies globales des éditeurs de presse, qui au-delà du canal de l'infomédiation et de leur inscription dans le modèle du courtage informationnel, adaptent par ailleurs peu ou prou leur modèle éditorial de l'imprimé dans leurs sites web ; considérer enfin des modes de circulation de l'information de nature plus « communautaire » qui amènent à un filtrage éditorial collectif et distribué [Rebillard, 2007]. De tels travaux permettraient d'aboutir à une connaissance véritablement exhaustive du développement de l'information journalistique sur l'internet, édifice colossal à construire auquel notre étude n'apporte qu'une pierre.

Références bibliographiques

- ATTIAS D., 2007, *L'impact d'Internet sur l'économie de la presse : quel chemin vers la profitabilité ?*, Thèse en Sciences économiques, Université Paris X.
- BOUQUILLION P., COMBES Y., (dir.), 2007, *Les industries de la culture et de la communication en mutation*, Paris : L'Harmattan.
- BOURDIEU P., 1996, *Sur la télévision*, Paris : Liber.
- BOURE R., SMYRNAIOS N., 2006, « L'infomédiation de l'information en ligne. Les cas des filiales françaises de Google et de Yahoo », in Chartron G., Broudoux E., (dir.) *Document numérique et société*, Paris : ADBS Éditions, p. 43-55.
- BROUSSEAU E., 2001, « e-Economie : Qu'y a-t-il de nouveau ? », *Annuaire des Relations Internationales*, Bruxelles : Emile Bruylant, p. 813-833.
- BUSTAMANTE E., 2004, « Cultural industries in the Digital Age : some provisional conclusions », *Media, Culture & Society*, vol. 26 (6), p. 803-820.
- CHARTRON G., REBILLARD F., (dir.), 2004, *Modèles de publication sur le web - Rapport final AS 103*, CNRS / Département STIC / RTP 33, <<http://archivesic.ccsd.cnrs.fr>>
- DUVAL J., 2004, *Critique de la raison journalistique. Les transformations de la presse économique en France*, Paris : Seuil.
- HINDMAN M., TSIOUTSIOLIKLIS K., JOHNSON J.A., 2003, « Googlearchy : how a few heavily-linked sites dominate politics on the web », *Annual Meeting of the Midwest Political Science Association* (Chicago), <<http://www.cs.princeton.edu/~kt/mpsa03.pdf>>
- KRAMER S.D., 2003, « Google News Creator Watches Portal Quiet Critics With 'Best News' Webby », *Online Journalism Review*, <<http://www.ojr.org/ojr/kramer/1064449044.php>>
- MOEGLIN P., 2005, *Outils et médias éducatifs. Une approche communicationnelle*, Grenoble : PUG
- REBILLARD F., 2006, « Du traitement de l'information à son retraitement. La publication de l'information journalistique sur l'internet », *Réseaux*, vol. 24, n°137, p. 29-68.
- REBILLARD F., 2007, *Le web 2.0 en perspective. Une analyse socio-économique de l'internet*, Paris : L'Harmattan.
- SCHROEDER R., KRALEMANN M., 2005, « Journalism ex machina. Google News Germany and its news selection processes », *Journalism Studies*, vol. 6 (2), p. 245-247.
- TESSIER M. (et BAFFERT M.), 2007, *La presse au défi du numérique*, Rapport pour le Ministre de la culture et de la communication, <www.culture.gouv.fr/culture/actualites/rapports/tessier/rapport-fev2007.pdf>
- ULKEN E., 2005, « A question of balance : are Google News search results politically biased ? », *USC Annenberg School for Communication Review*, vol. 2 (3), p. 78-99.