



HAL
open science

Vers des médias numériques en sciences humaines et sociales

Marin Dacos

► **To cite this version:**

Marin Dacos. Vers des médias numériques en sciences humaines et sociales : Une contribution à l'épanouissement de la place des sciences humaines et sociales dans les sociétés contemporaines. Tracés : Revue de Sciences Humaines, 2012, pp.A venir. halshs-00700238

HAL Id: halshs-00700238

<https://shs.hal.science/halshs-00700238>

Submitted on 22 May 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Vers des médias numériques en sciences humaines et sociales : une contribution à l'épanouissement de la place des sciences humaines et sociales dans les sociétés contemporaines

Marin Dacos

directeur du Cléo (Centre pour l'édition électronique ouverte, CNRS)

Je propose l'idée selon laquelle la presse en ligne se comporte comme un prédateur du Web en général, et des sciences humaines et sociales (SHS) en particulier¹. Si les SHS veulent pleinement jouer leur rôle dans l'interprétation et la compréhension de notre société, elles ne peuvent pas se permettre de le faire seulement dans le confort et l'isolement des murs de l'université. Elles ont intérêt à se doter de leur propre force de projection des idées, c'est-à-dire de leur propre média, au sens noble du terme de passeur entre deux mondes. En effet, la naissance du Web et son développement prodigieux constituent une opportunité historique et crédible pour un tel projet. Par provocation, mais également par conviction, je considérerai Wikipédia comme un exemple et comme un levier permettant l'émergence de véritables contre-propositions médiatiques de la part des SHS. Celles-ci doivent se doter d'une stratégie particulièrement efficace pour ne pas rester à l'état de déclarations d'intentions ou d'initiatives isolées. Elles devront notamment s'appuyer sur le principe du libre accès aux résultats de la recherche, dans la droite ligne des déclarations de Budapest et de Berlin². En s'orientant ainsi résolument vers le public, elles ne renonceront pas à leur dimension scientifique, et même elles profiteront, par effet de levier, de perspectives méthodologiques nouvelles. L'ensemble de ces développements constitue une branche très riche des humanités numériques, ou *Digital Humanities*, qui sont dotées de leurs organisations³, de leurs rencontres⁴ et de leur Manifeste⁵.

« Noindex, nofollow » : le comportement prédateur des médias généralistes

¹ Je tiens à remercier Pierre Mounier, Anne Durand, Serge Noiret, Emma Bester et Hubert Guillaud pour leur aide lors de la rédaction de cet article.

² La déclaration de Budapest est disponible à l'adresse suivante : <http://www.soros.org/openaccess>, consulté le 2 mai 2012 ; la déclaration de Berlin est disponible à l'adresse suivante : <http://oa.mpg.de/lang/en-uk/berlin-prozess/berliner-erklarung>, consulté le 2 mai 2012.

³ Notamment l'Alliance of Digital Humanities Organizations (ADHO), [URL : <http://digitalhumanities.org>], consulté le 2 mai 2012.

⁴ Notamment les « non conférences » THATCamp, [URL : <http://thatcamp.org>], consulté le 2 mai 2012 ; ainsi que le colloque annuel de l'ADHO, [URL : <http://digitalhumanities.org/conference>], consulté le 2 mai 2012.

⁵ Le *Manifeste des Digital Humanities*, produit en mai 2010 à THATCamp Paris, est disponible à l'adresse suivante : <http://www.humanistica.eu/manifesto>, consulté le 2 mai 2012.

Pour bien comprendre la bataille que livrent actuellement les médias en ligne, il faut mesurer le rôle des hyperliens dans l'économie générale du Web. Les hyperliens (ou liens hypertextes) constituent l'ossature cognitive du Web, dans la mesure où ils assurent la connexion entre deux pages Web différentes, à la manière utilisée naguère par les notes de bas de page. Ces liens permettent une navigation individuelle (*browsing*) passant de lien en lien, comme de liane en liane. Plus encore, ils documentent le Web en permettant à Google, moteur de recherche dominant, de hiérarchiser les pages entre elles. En effet, son algorithme considère qu'un lien vers une autre page constitue un vote. Plus une page Web est citée par d'autres pages Web, plus Google considère que son importance est grande et plus il mettra en avant cette page dans ses résultats (Ertzscheid *et al.*, 2009). Ce système est renforcé par la prise en compte du poids connu de chaque page : être cité par *Lemonde.fr* apportera un pourcentage de son propre *PageRank* à la page concernée. Or, le Web est régulé par l'économie de l'attention : les documents sont en concurrence entre eux, et une des rémunérations principales de ces contenus est liée à la publicité, elle-même proportionnelle à la fréquentation. Pour toutes ces raisons, les médias s'appuyant sur un modèle économique publicitaire ont intérêt à conserver, voire à capturer l'internaute, ce qui commence par éviter de l'inciter à partir ailleurs. Or, citer explicitement un contenu en ligne permet au lecteur d'accéder directement à l'information, donc à quitter le site d'origine, ce qui n'est pas le cas lorsque l'on cite un ouvrage papier, qui se situe à une distance physique et temporelle de plus d'un clic.

Théoriquement, le journalisme est une profession qui pourrait mettre à profit l'édition électronique pour citer plus finement et plus explicitement ses sources. Mais les impératifs économiques l'emportent, et les articles sont avares de citations externes. Il suffit de consulter une page du *monde.fr* et une autre de Wikipédia pour voir la différence de comportement entre les deux types d'écritures en ce domaine. Pire : la plateforme de blogs du *Monde* a longtemps eu un comportement contraire aux bonnes pratiques du Web. Dans la mesure où les blogueurs du monde ne sont pas assujettis aux mêmes règles d'écriture que les journalistes, leurs textes comportent un nombre important de liens pointant vers l'extérieur. De ce fait, le *PageRank* des sites cités par les blogueurs pouvait être positivement influencé par l'ensemble de ces liens, et affaiblir, par contrecoup, le *PageRank* du *Monde*. Pour éviter ce « défaut », la plateforme de blogs du *Monde* superposait sur les hyperliens des métadonnées invisibles par le novice : « noindex, nofollow »⁶. Ces instructions techniques sont destinées aux moteurs de recherche pour leur indiquer de ne pas suivre le lien (« no follow ») et de ne pas indexer (« no index ») le contenu de la page web cible. Ce type d'instruction a été mis au point pour permettre aux webmasters de guider les moteurs de recherche dans leur site, en leur indiquant, notamment, les pages qui sont redondantes avec d'autres pages du même site ou qui ont une vocation confidentielle, voire privée, afin d'optimiser l'indexation de leur propre site web. Déclarer en « noindex, nofollow » revient à cacher ces pages à Google et à refuser de voter pour elles. Officieuse et découverte par hasard, cette politique est choquante et non conforme aux bonnes pratiques de la profession. Elle a fait l'objet de contestations de la part des blogueurs eux-mêmes, qui ont finalement obtenu gain de cause, avec le retrait de ces métadonnées supplémentaires – ces pratiques avaient été constatées en février 2009 ; en février 2012, elles avaient disparu.

Les bonnes pratiques du Web

Ces bonnes pratiques sont les héritières de la netiquette. Celle-ci est née au milieu des années 1990, au moment de la multiplication de l'usage du courrier électronique, avec son lot de chauffards et de maladresses, et régule les échanges entre les personnes dans les espaces de conversation, essentiellement le chat et l'email⁷. Sont ensuite apparues les bonnes pratiques du Web, dédiées aux personnes réalisant

⁶ « Using the robots meta tag », *Webmaster Central Blog. Official news on crawling and indexing sites for the Google index*, 5 mars 2007, [URL : <http://googlewebmastercentral.blogspot.com/2007/03/using-robots-meta-tag.html>], capturé le 27 février 2012, [URL : <http://www.webcitation.org/651Ci5Y1D>].

des sites web. Ceux-ci se développant de façon anarchique, un mouvement en faveur d'un web ouvert⁸, qui milite notamment pour une standardisation des formats, a émergé et a rapidement été complété par des bonnes pratiques, qui dépassent le problème des formats. En France, les professionnels du Web ont rédigé, de façon collaborative, plus de 200 bonnes pratiques qui font consensus et qui permettent de disposer d'un Web fluide, agréable, non ambigu⁹.

Comment expliquer qu'un journal comme *Le Monde*, visant a priori l'information de ses lecteurs, puisse avoir promu des pratiques de lecture limitant indirectement mais volontairement l'accès libre à l'information par les internautes ? Les usages de lecture sur le Web se développent avec une force tellement considérable que l'on peut imaginer que les intermédiaires culturels, pédagogues et journalistes, comme *Le Monde*, se sentent menacés par ce qu'ils pourraient considérer comme des concurrents. Tous les signaux convergent pour montrer que les médias écrits traditionnels doivent désormais cohabiter, en ligne, avec d'autres contenus textuels sérieux.

Vers l'émergence de médias SHS ?

Dans cette partie, je souhaite évoquer l'émergence d'usages massifs de contenus qui ne sont pas issus de la presse traditionnelle... pas plus que des maisons d'édition historiques. Pour mesurer les usages, il est possible de collecter, ici et là, quelques chiffres de fréquentation donnés par les sites eux-mêmes, essentiellement à des fins de promotion. Ces chiffres sont toujours incomplets, ne s'appuient pas sur les mêmes méthodologies, les mêmes définitions, couvrent souvent des périodes différentes et sont finalement difficilement comparables. Mais nous ne disposons pas d'un accès facile à ce type d'information, que seuls des fournisseurs d'accès pourraient confier à des chercheurs. Dans l'attente de données de meilleure qualité, peut-on utiliser des indicateurs accessibles au grand public et construits dans une perspective commerciale et non scientifique, comme *Google Trends for Websites* ?

Comparer la fréquentation de sites web : un problème méthodologique

Google Trends est un outil développé par la firme de Mountain View pour les acheteurs de mots clés publicitaires (Helft, 2008). Ce service d'identification des mots clés les plus souvent interrogés dans le moteur de recherche a pour objectif d'en augmenter le prix de vente. Un tel outil n'a donc aucune prétention scientifique, mais il a pu être détourné de sa vocation initiale pour détecter l'apparition d'épidémies de grippe (Ginsberg *et al.*, 2009). Le sous-service *Google Trends for Websites* (GTW) propose, quant à lui, une évaluation de la fréquentation des plus importants sites de la Toile. Il ne s'agit que d'une estimation, qui agrège des données diverses¹⁰. Google insiste sur le caractère imprécis de ses

⁷ « Nétiquette », *Wikipédia*, [URL : <http://fr.wikipedia.org/wiki/N%C3%A9tiquette>], consulté le 2 mai 2012.

⁸ Voir le site de l'OpenWebGroup, [URL : <http://openweb.eu.org>], consulté le 2 mai 2012.

⁹ 217 bonnes pratiques sont proposées à l'adresse suivante : <https://checklists.opquast.com/opquastv2>, consulté le 2 mai 2012.

¹⁰ Les deux principales sources sont les usages mesurés par Google lors de l'interrogation du moteur par des centaines de millions d'internautes, mais aussi les millions de sites qui ont adhéré au programme *Google Analytics*, qui est un mouchard inséré dans les pages web par leur webmestre afin de collecter des informations d'usages. Ce service, gratuit pour les webmestres, apporte en retour des informations de très haute valeur à Google. La documentation officielle de *Google Trends for Websites* décrit précisément la méthodologie utilisée : « *Trends for*

mesures :

Il est important de garder en tête le fait que les résultats de *Trends for Websites* sont estimés. De plus, les données sont mises à jour périodiquement, de sorte que des modifications récentes dans les flux de données peuvent ne pas être prises en considération. Enfin, gardez à l'esprit le fait que *Trends for Websites* est un produit *Google Labs*, ce qui veut dire qu'il est encore dans ses premiers stades de développement et peut par conséquent contenir certaines erreurs. À l'avenir, nous tâcherons peut-être d'améliorer la précision de *Trends for Websites* en autorisant les propriétaires à contribuer, en ajoutant les données issues de leurs propres sites¹¹.

Francis Chateauraynaud et Josquin Debaz ont, avec raison, mis en évidence les errements méthodologiques, voire les approximations effarantes de Google en bien des domaines statistiques (Chateauraynaud, 2010 ; Chateauraynaud et Debaz, 2010). Il n'est cependant pas interdit de se pencher sur *Google Trends for Websites*, en l'absence d'études de terrain et de données alternatives accessibles. D'ailleurs, ces données seraient-elles accessibles que leur traitement nécessiterait une instrumentation lourde, qui donnera probablement naissance à une branche des humanités numériques et des sciences du Web, dont la complexité dépasse les capacités du présent article. Je considérerai, ici, seulement deux informations comme exploitables : 1/ les oscillations temporelles, non en tant que mesures exactes, mais comme tendances ; 2/ les informations comparatives entre la fréquentation de deux sites, là encore moins en tant qu'informations absolues que relatives. Ce dernier élément est majeur, car unique au monde, à ma connaissance¹².

Émergence récente d'un média pédagogique en SHS : Wikipédia

GTW permet de se concentrer sur la population d'un pays, ici la France. On peut demander une information relative aux grands organes de presse quotidienne, qui est, bien entendu, inférieure à la fréquentation mondiale. En janvier 2012, *Lemonde.fr* et *Lefigaro.fr* faisaient presque jeu égal, autour de 700 000 visiteurs uniques par jour¹³. De son côté, *Liberation.fr* jouissait d'une fréquentation inférieure

Websites combine des informations provenant de sources diverses, telles que les données agrégées des recherches Google, les données agrégées des abonnés anonymes de *Google Analytics*, les données de panels de consommateurs abonnés, et d'autres études de marché. Les données sont agrégées à partir de millions d'utilisateurs, traitées grâce à des algorithmes informatiques, et ne contiennent aucune information personnelle identifiable. De plus, *Google Trends for Websites* ne fournit de résultats que pour des sites qui reçoivent un nombre significatif de visites, et définit des seuils minimaux d'inclusion dans l'outil. », *Google Trends for Websites*, [URL : <http://www.google.com/intl/en/trends/websites/help/index.html#2>], capturé le 25 février 2012, [URL : <http://www.webcitation.org/65idXew1S>] ; toutes les traductions sont de la rédaction.

¹¹ « *It's important to keep in mind that all results from Trends for Websites are estimated. Moreover, the data is updated periodically, so recent changes in traffic data may not be reflected. Finally, keep in mind that Trends for Websites is a Google Labs product, so it's still in its early stages of development and may therefore contain some inaccuracies. In the future, we may consider ways to improve the accuracy of Trends for Websites by allowing owners to contribute additional data from their sites.* », *Google Trends for Websites*, [URL : <http://www.google.com/intl/en/trends/websites/help/index.html#15>], capturé le 25 février 2012, [URL : <http://www.webcitation.org/65idXew1S>].

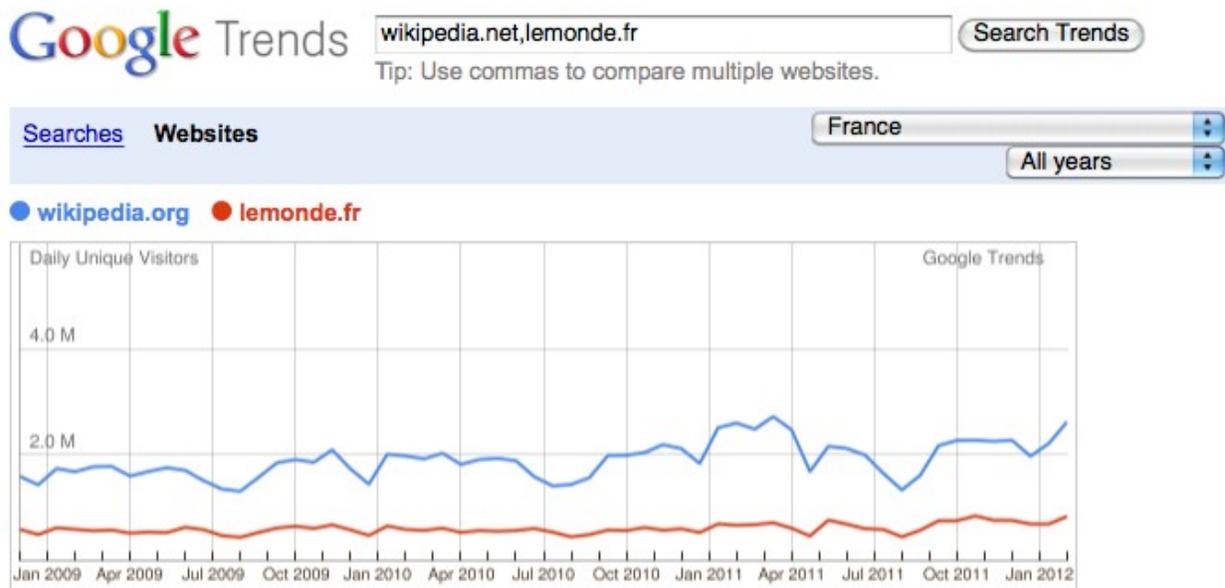
¹² La firme Alexa offre des services similaires, mais encore moins précis, [URL : <http://www.alexa.com>], consulté le 2 mai 2012.

¹³ Un visiteur unique est une personne ou un poste de consultation, qui réalise au moins une visite du site dans la journée.

d'environ 50 %¹⁴.



Comparons à présent le site *Lemonde.fr* avec *Wikipedia.org*, là encore en nous restreignant à la consultation française.



En janvier 2012, *Wikipedia.org* attirait environ 2 millions de visiteurs uniques. On pourra longuement commenter un tel résultat, notamment en soulignant qu'il est difficile de comparer un quotidien national avec une encyclopédie collaborative multilingue. Je me contenterai de constater que les français utilisent désormais une source d'information non journalistique, née au XXI^e siècle, alors que la presse est née au XIX^e siècle. Ce qu'un chercheur en sciences humaines et sociales pourrait tirer comme enseignement de ce nouveau rapport de forces est qu'il existe désormais de nouveaux champs d'expression du savoir qu'il pourrait/devoir investir pour faire progresser le savoir collectif. Certes,

¹⁴ On notera que le graphique utilise une échelle logarithmique.

Wikipédia est organisé par des règles collaboratives très différentes de celles de l'édition scientifique. Pour devenir auteur dans *Wikipédia*, le chercheur doit accepter ces règles qui peuvent produire un choc culturel, d'abord parce que la contribution de chacun est complexe à identifier dans le texte final, ensuite parce que des interventions de personnes moins établies scientifiquement se produiront sur le texte, enfin parce que le savoir à transmettre doit être strictement neutre et encyclopédique. L'exercice de style est donc rarement récompensé en retour, car il ne sera pas reconnu par l'université, pas considéré comme une publication, et ne contribuera pas à la renommée de l'auteur – qui décortique l'historique d'une page Wikipédia pour identifier les contributions les plus significatives ? Ce que nous perdons du point de vue de la reconnaissance, nous le gagnons en liberté, car Wikipédia est un média inscriptible, c'est-à-dire ouvert en permanence à des modifications collaboratives (Dacos éd., 2010), alors que la presse est par nature beaucoup plus restrictive. À titre personnel, j'utilise Wikipédia comme un support pédagogique, demandant à mes étudiants en Master¹⁵ d'apporter leur contribution aux pages relatives aux cours que je leur donne. En effet, Wikipédia n'est pas un espace scientifique, mais peut être exploité comme espace pédagogique, aux deux sens du terme. Pédagogie pour nos étudiants, qui apprennent à suivre une ligne éditoriale dans un contexte de formation intellectuelle. Pédagogie pour le public, avide de compréhension et de savoirs, qui reçoit grâce à ce type d'instrument un ensemble de connaissances qui prolonge la mission de l'Éducation nationale au-delà de la période de scolarité. Pour l'instant, l'Éducation nationale n'a pas investi officiellement ce support, mais une réflexion et un débat sur cette question devraient avoir lieu publiquement, afin de réfléchir à la modernisation des missions, des méthodes et des finalités de l'école et de l'université.

Émergence de portails spécialisés

Revenons aux publications strictement issues de la recherche, en comparant les usages de *Revues.org*¹⁶, portail dédié aux revues spécialisées en sciences humaines et sociales, avec les usages en ligne de la presse. Le tirage moyen du *Monde* est d'environ 320 000 exemplaires par jour, alors qu'une revue de sciences humaines et sociales moyenne a environ 300 abonnés – moins de 3 000 pour les plus connues –, essentiellement des institutions. Dans l'univers du papier, l'usage du *Monde* est au moins 100 fois supérieur à celui de l'ensemble des revues de *Revues.org*¹⁷. Cette situation est logique, au regard des formes d'écriture et de la spécialisation des sujets publiés dans une revue scientifique. Pourtant, *Google Trends for Websites* semble montrer une évolution de cette situation, car le rapport de forces évolue avec un écart de 1 à 25 seulement lorsque l'on passe au format électronique¹⁸. Cette mesure reste grossière et

¹⁵ Le blog du Master « Monde du livre » est accessible, sur *Hypothèses.org*, à l'adresse suivante : <http://mondedulivre.hypotheses.org>, consulté le 2 mai 2012.

¹⁶ *Disclaimer* : je suis le fondateur de *Revues.org* et je dirige le Centre pour l'édition électronique ouverte, qui développe la plateforme *Revues.org* au sein du portail *OpenEdition.org*.

¹⁷ Avec deux numéros en moyenne par an, la diffusion annuelle d'une revue est donc d'environ 600 exemplaires. La diffusion des 265 revues de *Revues.org* aujourd'hui en ligne, si elles étaient au format papier, serait au moins 100 fois inférieure à celle du journal *Le Monde*. En poursuivant la méthode proposée, on peut évaluer à environ 159 000 exemplaires diffusés pour les 265 revues en ligne sur *Revues.org* en janvier 2012, contre environ 15 millions d'exemplaires par an pour *Le Monde*. Ces calculs ne sont qu'indicatifs, d'abord parce que les revues ont une ancienneté variable, ensuite parce qu'environ 50 % seulement ont également une existence papier, les autres étant exclusivement électroniques.

¹⁸ Les évaluations de GTW indiquent que *Lemonde.fr* aurait eu, en janvier 2012, 1 million de visiteurs uniques par jour. Plus de précisions sont apportées par les données publiques du portail *Revues.org* : celui-ci a reçu en moyenne, en janvier 2012, 41 000 visiteurs uniques par jour. Pour plus de détails, consulter les statistiques sur le site du Cléo,

gagnerait à être précisée à l'aide d'enquêtes qualitatives comme celle menée autour du lectorat de *Revue.org* en 2009 (Bester et Dacos, 2010). Cependant, nous pouvons en retirer une information : réunies sous la forme de portails, les productions en SHS peuvent devenir des lieux d'usages de lecture quantitativement significatifs. Cela signifie que peuvent se construire des médias spécialisés crédibles et intéressants au-delà du seul périmètre des acteurs de l'enseignement supérieur et de la recherche¹⁹.

D'un point de vue plus qualitatif, il reste important de savoir, d'une part, si ces fréquentations sont dépendantes d'affluents, c'est-à-dire de sources de fréquentation (*referer*) issus des industriels de l'information que sont Google, Yahoo! et Bing, et d'autre part si la presse contribue à cette fréquentation.

Les statistiques confirment le rôle mineur joué par la presse dans la fréquentation de *Revue.org*. Sur 108 millions de visites comportant un affluent identifiable en 2011, seuls *Lemonde.fr* (53^e), *Rue89.com* (71^e), *Laviedesidées.fr* (76^e) et *Liberation.fr* (95^e) sont classés parmi les 100 premiers affluents. Et le premier quotidien national n'apporte que 5 300 visites, soit 0,03 % des visites. Un examen de détail montre que seuls six articles ont apporté des visiteurs à *Revue.org* au cours de cette période²⁰. Les 167 autres liens proviennent des blogs du *Monde*²¹.

À lui seul, le site personnel de Laurent Mucchielli apporte presque autant de visites (62^e), et *Ruedesfacs.fr* (11^e, avec 144 000 visites) en apporte plus. Bien classés, également, apparaissent des portails partenaires de *Revue.org* : *Liens-socio.org* (4^e), mais aussi *Cairn.info* (6^e), *College-de-france.fr* (21^e) et *Persee.fr* (26^e). De son côté, Wikipédia est classé 9^e, avec 154 000 visites, soit 29 fois plus de visites que *Lemonde.fr*, bien plus que le différentiel de fréquentation entre l'encyclopédie collaborative et le quotidien. Ceci confirme que *Lemonde.fr*, et la presse quotidienne en général, répugne à citer des sources extérieures et à conforter ce qu'elle considère comme de la concurrence.

Rang	Site	Nombre de visites Janvier-	
		Décembre 2011	% des visites
1	revues.org	8 328 539	46,41%
2	google.com	2 856 668	15,92%
3	hypotheses.org	2 130 215	11,87%
4	liens-socio.org	1 861 888	10,37%
5	cybergeu.eu	345 008	1,92%
6	cairn.info	206 419	1,15%

[URL : <http://statistiques.cleo.cnrs.fr>], consulté le 2 mai 2012. *Google Trends for Websites* propose d'ailleurs une évaluation cohérente avec ces chiffres, [URL : <http://trends.google.com/websites?q=revues.org&geo=all&date=all&sort=0>], consulté le 2 mai 2012. Les évaluations de GTW confondent en fait le journal et la plateforme de blogs du *Monde.fr*, alors qu'il les dissocie pour la plateforme de carnets de recherches d'OpenEdition, qui utilisent un autre nom de domaine, *Hypotheses.org*. Pour parfaire la comparaison, il faudrait ajouter 7 300 visiteurs uniques quotidiens d'*Hypotheses.org*.

¹⁹ Les fréquentations de *Cairn.info* et *Persee.fr* sont relativement proches de celle de *Revue.org*, qui n'est donc pas un cas isolé ou une exception. *Scielo.org* et *Jstor.org* ont des visibilitées supérieures, sur des aires linguistiques dont le potentiel est très supérieur (espagnol et portugais pour Scielo, anglais pour Jstor).

²⁰ Ces articles portent sur les sujets suivants : Alassane Ouattara, le mot « crise », l'univers virtuel des néo-maoïstes chinois, le sondage organisé par Mattel pour désigner le métier d'un nouveau modèle de Barbie, le séisme du Sichuan, et la présence de parabène dans des médicaments. Il s'agit donc de sujets pointus, correspondant à une actualité précise, sur laquelle la recherche en sciences humaines et sociales apporte un éclairage quasiment exclusif.

²¹ Principalement le blog du correspondant du quotidien du soir à Istanbul, Guillaume Perrier, *Au fil du Bosphore. Observations d'Istanbul et de la Turquie d'aujourd'hui*, [URL : <http://istanbul.blog.lemonde.fr>], consulté le 2 mai 2012 ; ainsi que le blog *La feuille, l'édition à l'heure de l'innovation*, [URL : <http://lafeuille.blog.lemonde.fr>], consulté le 2 mai 2012, et le blog *Combats pour les droits de l'homme* [URL : <http://combatsdroitshomme.blog.lemonde.fr>], consulté le 2 mai 2012.

7	pearltrees.com	174 826	0,97%
8	bing.com	155 444	0,87%
9	wikipedia.org	154 071	0,86%
10	conflits.org	153 334	0,85%
11	ruedesfacs.fr	144 025	0,80%
12	yahoo.com	122 578	0,68%
13	facebook.com	108 655	0,61%
14	orange.fr	104 326	0,58%
15	openedition.org	85 942	0,48%
16	netvibes.com	71 359	0,40%
17	search-results.com	66 026	0,37%
18	lodel.org	65 064	0,36%
19	sfr.fr	49 026	0,27%
20	laposte.net	42 580	0,24%
21	college-de-france.fr	29 044	0,16%
22	free.fr	27 920	0,16%
23	longurl.org	24 939	0,14%
24	rechercheisidore.fr	24 613	0,14%
25	cnrs.fr	22 067	0,12%
26	persee.fr	21 999	0,12%
27	bingj.com	21 809	0,12%
28	t.co	20 657	0,12%
29	ebsco.com	17 742	0,10%
30	doaj.org	15 661	0,09%
31	scirus.com	15 157	0,08%
32	ehess.fr	14 356	0,08%
33	uni-regensburg.de	13 237	0,07%
34	univ-provence.fr	12 777	0,07%
35	lo.st	12 312	0,07%
36	123people.fr	10 304	0,06%
37	rezo.net	9 561	0,05%
38	fabula.org	9 550	0,05%
39	cleo.cnrs.fr	8 932	0,05%
40	u-paris10.fr	8 140	0,05%
41	over-blog.com	8 072	0,04%
42	blogariak.net	7 619	0,04%
43	gouv.fr	7 498	0,04%
44	lexilogos.com	7 115	0,04%
45	blogspot.com	6 984	0,04%
46	hardware.fr	6 731	0,04%
47	ifporient.org	6 143	0,03%
48	translendum.net	6 107	0,03%
49	facemoods.com	5 664	0,03%
50	scoop.it	5 571	0,03%

51	twitter.com	5 467	0,03%
52	univ-lyon2.fr	5 459	0,03%
53	lemonde.fr	5 300	0,03%
54	about.com	5 096	0,03%
55	unice.fr	5 057	0,03%
56	numericable.fr	4 843	0,03%
57	ecosia.org	4 829	0,03%
58	youtube.com	4 706	0,03%
59	inist.fr	4 431	0,02%
60	aliceadsl.fr	4 329	0,02%
61	tinyurl.com	4 317	0,02%
62	laurent-mucchielli.org	4 313	0,02%
63	sfp.photographie.com	3 998	0,02%
64	cereq.fr	3 950	0,02%
65	iadah.com	3 815	0,02%
66	aol.com	3 813	0,02%
67	speedbit.com	3 682	0,02%
68	uqac.ca	3 680	0,02%
69	base-juniper.org	3 606	0,02%
70	uquebec.ca	3 584	0,02%
71	rue89.com	3 542	0,02%
72	adf.ly	3 387	0,02%
73	sorbonne.fr	3 254	0,02%
74	univ-paris1.fr	3 193	0,02%
75	umontreal.ca	3 081	0,02%
76	laviedesidees.fr	2 934	0,02%
77	univ-paris8.fr	2 800	0,02%
78	feedly.com	2 790	0,02%
79	mail.ru	2 749	0,02%
80	cafepedagogique.net	2 743	0,02%
81	uqam.ca	2 725	0,02%
82	ciera.fr	2 626	0,01%
83	wordpress.com	2 596	0,01%
84	univ-lille3.fr	2 590	0,01%
85	libertysecurity.org	2 545	0,01%
86	univ-tlse2.fr	2 354	0,01%
87	lusopenedition.org	2 347	0,01%
88	bitly.com	2 304	0,01%
89	ulaval.ca	2 281	0,01%
90	racocatala.cat	2 132	0,01%
91	inrp.fr	2 091	0,01%
92	assoeconomiepolitique.org	2 080	0,01%
93	journalseek.net	2 069	0,01%
94	univ-angers.fr	2 048	0,01%

95	liberation.fr	1 946	0,01%
96	bnf.fr	1 930	0,01%
97	u-cergy.fr	1 899	0,01%
98	ifre.fr	1 899	0,01%
99	unirioja.es	1 849	0,01%
100	sciences-po.fr	1 789	0,01%

Selon la conception la plus répandue, Google caracolait en tête de la fréquentation des sites Web, parfois jusqu'à 80 %, et serait donc en mesure de faire la pluie et le beau temps sur le Web, réduisant potentiellement à néant des années de construction, à partir de réglages de son algorithme complexe et secret, donc arbitraire... Si Google est bien le premier affluent de *Revue.org* (2 856 000 visites en 2011), il est désormais talonné par *Liens-Socio.org* (1 861 000 visites). Il est même devancé par *Revue.org* lui-même (8 320 000 visites) et talonné par *Hypotheses.org* (2 130 000). Dit autrement, l'écosystème d'OpenEdition, regroupé, devient le premier affluent d'une revue de *Revue.org*. OpenEdition est le portail développé par le Centre pour l'édition électronique ouverte (Cléo), centre issu de l'enseignement supérieur et de la recherche, qui déploie une plateforme de revues depuis 1999 (*Revue.org*), un calendrier scientifique depuis 2000 (Calenda), une plateforme de carnets de recherche depuis 2008 (*Hypotheses.org*) et ouvrira en 2012 une plateforme de livres (OpenEdition Books). Le projet a désormais le statut d'équipement d'excellence²².

Ces résultats alimentent l'idée d'une capacité d'indépendance de portails SHS puissants, capables de développer leur propre audience et de proposer des circulations entre eux. Les moteurs de recherche conservent, c'est entendu, un rôle décisif, mais celui-ci n'est pas exclusif. La leçon principale à retirer de ces chiffres parfois difficiles à interpréter est que le Web est un dispositif distribué. Il n'est pas devenu un média verticalisé et centralisé autour de quelques points d'accès exclusifs. Le Web reste une Toile.

Affinons encore l'approche, en nous concentrant sur les affluents d'un carnet de recherche, *Culture et politique arabe*²³, rédigé par Yves Gonzalez-Quijano, chercheur à l'Institut français du Proche-Orient (IFPO). Celui-ci est l'un des carnets les plus fréquentés d'Hypotheses, avec plus de 270 000 visites au cours de l'année 2011 (117 000 visiteurs uniques). Le profil de ses affluents identifiables au cours du mois d'octobre 2011 est significatif²⁴.

Affluents du carnet <i>Culture et politique arabe</i>	Apport en nombre de visites en octobre 2011
Hypotheses.org	3 049
Rezo.net	941
Facebook	690
Netvibes	292
Webmail	85

²² « OpenEdition facility of excellence awards (Equipex) », *OpenEdition*, 23 février 2012, [URL : <http://www.openedition.org/10221>], consulté le 2 mai 2012.

²³ Consultable à l'adresse suivante : <http://cpa.hypotheses.org>, consulté le 2 mai 2012.

²⁴ Source : statistiques d'usages d'OpenEdition. Ces statistiques consistent en des calculs, enrichis à l'aide d'informations éditoriales, sur la base des résultats publics disponibles sur <http://statistiques.cleo.cnrs.fr>, consulté le 2 mai 2012.

Wikipédia FR	38
Google translate	31
@rrêt sur images	25
Google	24
Twitter	15

Le principal enseignement de ce tableau est la puissance des réseaux hypertextuels conçus par des humains et, toujours, la faiblesse des médias, alors qu'en 2011 le « Printemps arabe » était au cœur de l'actualité... Seul @rrêt sur images a apporté, au cours de ce mois d'octobre 2011, quelques visites au carnet. Le site qui lui a apporté le plus de visiteurs est la plateforme de carnets de recherches qui l'accueille, *Hypothèses*. Vient ensuite le site alternatif d'actualités *Rezo.net*. Une part importante des affluents connus vient des réseaux sociaux, Facebook et Twitter, ainsi que du courrier électronique (les « webmails »). Cet ensemble d'affluents provient de recommandations individuelles, qu'elles aient lieu sur des réseaux sociaux ou par courrier électronique. La force du « bouche à oreille », propre au grand public, mais aussi de l'échange entre pairs, est donc importante. Au mois d'octobre 2011, Google n'est pas un affluent majeur pour *Culture et politique arabe*, alors qu'il reste un affluent majeur, d'un point de vue plus global, au niveau de l'ensemble des plateformes d'OpenEdition : alors qu'il s'agit du premier affluent de la plateforme Calenda, il n'est déjà que le cinquième pour *Hypothèses*, et seulement le neuvième pour *Culture et politique arabe* sur la totalité de l'année 2011. Autant dire qu'il n'y a pas que Google dans la vie d'un site web...

De 1984 au lecteur inattendu : vers une édition électronique ouverte

Il n'est malheureusement pas possible, ici, de rédiger un *Petit traité pratique de création d'un média nativement SHS*²⁵. En revanche, nous pouvons mettre en avant quelques intérêts scientifiques de ce type d'initiative.

Tout d'abord, créer un média en sciences humaines et sociales devrait pouvoir se faire selon les principes énoncés par le Déclaration de Berlin pour le libre accès aux matériaux et aux résultats de la recherche²⁶. Cette déclaration a été signée par des centaines d'établissements universitaires depuis 2003 et continue à se développer. Elle s'appuie sur l'idée que la recherche sur fonds publics doit revenir au public, à la fois pour des raisons éthiques, sociétales et scientifiques. Les vertus heuristiques de l'*open access* font, en effet, partie des arguments en faveur d'une libre circulation des idées, afin de favoriser l'apparition du « lecteur inattendu » (Smith, 2011 ; Mounier, 2012), faisant fi des barrières innombrables à l'accès que constitue un modèle économique à péage aval. De la même façon, l'accès des familles des malades aux mêmes informations que le corps médical a favorisé l'émergence d'un nouvel acteur de la recherche, disposant d'une expertise spécifique en voie de reconnaissance (Akrich *et al.*, 2009). Il n'est cependant pas aisé de trouver une solution qui combine libre accès et capacités de financement régulier pour l'édition.

²⁵ Pour une introduction problématique, voir Mounier (2010). Pour une introduction historique sur l'aventure spécifique que constitue *Revue.org*, voir Cavallo (2009).

²⁶ Sur les travers de l'écosystème éditorial pré-*open access*, voir Guédon (2001). Sur les vertus du libre accès, voir Willinsky (2006). Un ouvrage plus récent élargit les enjeux à toutes les dimensions du savoir (VECAM, 2011).

La question se pose en effet. Au-delà du bien-fondé de la démarche, comment financer le libre accès ? Le modèle de la subvention publique, verticale, soumise aux aléas et même aux réductions progressives, est-il le seul possible ? Il existe deux autres voies susceptibles de compléter le financement public classique. La première est celle de l'auteur-payeur, initiée notamment par *Public Library of Science* (PLOS²⁷). L'auteur doit acquitter des frais de publication, qui sont assumés par son laboratoire. Les tarifs sont variables selon les portails, mais ils s'étalent, en général, entre 1 000 et 3 000 dollars par article. Ce modèle pose des problèmes éthiques évidents et est, par ailleurs, inadapté aux disciplines pauvres que sont les sciences humaines et sociales²⁸.

Le deuxième modèle est appelé *freemium*. Ce modèle a été développé par les entreprises du Web qui ne pouvaient ou ne voulaient pas miser sur la publicité et souhaitaient s'inscrire dans la logique du libre accès. Il est appliqué avec succès par Flickr, le célèbre site de partage de photographies, et par Skype, le logiciel de téléphonie par IP. Son modèle est simple : en étant gratuits, les services attirent des millions d'utilisateurs, et un faible pourcentage de ces utilisateurs est intéressé par des fonctionnalités Premium, payantes, qui financent la totalité de la plateforme. C'est le sens du mot *freemium*, qui associe *free* (gratuit) et *premium* (payant pour des fonctionnalités avancées). Le *New York Times* teste, depuis mars 2011, une formule similaire (Mathieu, 2011 ; Peters, 2011 ; Salmon, 2011), s'appuyant sur le principe d'un quota d'articles lisibles gratuitement, chaque mois, par un lecteur. Les premiers résultats attestent d'un succès notable, dont il faudra observer les développements à l'avenir. OpenEdition a lancé OpenEdition *Freemium* le 15 février 2011 (Dacos, 2011). Il propose un abonnement payant aux bibliothèques pour des contenus placés en libre accès. Les bibliothèques abonnées jouissent de fonctionnalités *premium* telles que l'accès aux formats non HTML (PDF, Epub, pour tablettes et liseuses, mais aussi pour consultation hors ligne et construction de bibliothèques personnelles). Par ailleurs, les bibliothèques abonnées ont accès à un ensemble de six services à valeur ajoutée : alertes par courriel pour l'ensemble de leurs usagers, fourniture et services de données spécialisés (statistiques de fréquentation par campus, API Calenda, notices UNIMARC...), assistance et formation, informations et documentation, association à la gouvernance du projet. Cette offre garantit une diffusion maximale des textes scientifiques, grâce au libre accès, tout en offrant une opportunité de revenus à l'édition, grâce aux services *premium*. En effet, 66 % des revenus issus de ce bouquet seront reversés aux revues et éditeurs partenaires. Après moins d'un an d'expérience, cette initiative apporte des résultats plus positifs et clairs que prévu. Trente-sept établissements se sont déjà abonnés à OpenEdition *Freemium* et la réception de la proposition par la communauté des scientifiques, des éditeurs et des bibliothèques est plutôt positive.

Aux questions d'accès s'ajoute la question des conditions d'accès. Les plateformes d'édition électronique commerciales ont tendance à imposer des quotas de téléchargements aux bibliothèques abonnées ainsi que des DRM – *Digital Rights Management*, systèmes de gestion des droits (Benghozi, 2006). Appliqués aux livres, les DRM sont des procédés cryptant les ouvrages qui sont achetés. À chaque ouverture de l'ouvrage, ces logiciels sont chargés de vérifier que l'utilisateur a bien le droit de le lire... Le décryptage du livre n'a lieu qu'après la vérification de ce droit. Cela signifie, concrètement, qu'il n'est pas possible de lire un livre sur toutes les machines existantes : le principal fournisseur de DRM, Adobe, ne propose d'outil de décryptage que pour Mac OS et pour Windows (donc pas pour Linux, mais également pas pour tous les systèmes d'exploitation de smartphones). Par ailleurs, cette vérification donne des droits nouveaux aux éditeurs qui les promeuvent. Il n'est plus possible de prêter, de revendre, de donner un livre

²⁷ Voir le site internet de PLOS, [URL : www.plos.org], consulté le 2 mai 2012.

²⁸ Pour une synthèse générale et à jour sur le libre accès, lire l'*Open Access Directory*, [URL : http://oad.simmons.edu/oadwiki/Main_Page], consulté le 2 mai 2012. On trouvera également des points de vue et des informations de premier plan dans *The Scholarly Kitchen*, [URL : <http://scholarlykitchen.sspnet.org>], consulté le 2 mai 2012.

(Dacos et Mounier, 2010). Il faut être connecté à Internet pour confirmer son identité, qui est stockée, avec une liste d'ordinateurs autorisés, dans une base de données distante. Il est obligatoire d'utiliser le logiciel de lecture qui vous est associé au concepteur du verrou numérique (Adobe Digital Edition, en général). Certains DRM sont dits chronodégradables, car ils ne constituent en fait qu'une location, et le livre s'autodétruit littéralement après une durée fixée, par exemple à huit jours. Les DRM peuvent également être considérés comme des intrus introduits par les éditeurs dans nos machines, dans nos maisons. Amazon en a fourni la magnifique illustration en retirant des milliers de copies de *1984* de George Orwell au cours de l'été 2009. Le libraire en ligne s'est rendu compte, un peu tard, qu'il n'avait pas le droit de vendre cet ouvrage. Au lieu de simplement cesser la commercialisation du livre, il a pris une mesure radicale, que permet facilement le mécanisme des DRM : il s'est introduit dans les liseuses (Kindle) de ses clients et a détruit le fichier qui s'y trouvait, annotations personnelles incluses²⁹... Le fait que ce soit précisément l'ouvrage de George Orwell est pure coïncidence, mais il éclaire sur le potentiel liberticide de certaines technologies lorsqu'elles échappent au citoyen. Les DRM offrent *de facto* au vendeur un contrôle total des usages de la marchandise vendue, y compris après la vente, ce qui vise à contrôler et à altérer l'autonomie du consommateur, qui avait plus de libertés dans l'écosystème papier... Il est donc nécessaire d'abandonner les DRM, qui soupçonnent l'honnête citoyen de vouloir devenir un voleur, ainsi que tous les dispositifs technico-commerciaux du même type.

Nous avons évoqué le potentiel heuristique lié au lecteur inattendu. Il faut également évoquer celui que permet le caractère inscriptible du Web (Dacos éd., 2010 ; Dacos et Mounier, 2011). Se multiplient, en effet, les exemples d'utilisation du Web par les scientifiques pour appeler les internautes à contribuer à un apport de connaissances nouvelles, ce que l'on appelle en anglais le *crowdsourcing*. Le programme Flickr Commons, par exemple, a été initié par la société Flickr (filiale de Yahoo!) et par la Bibliothèque du Congrès. Flickr est une plateforme de partage, d'exposition et de stockage de photographies, qu'elles soient professionnelles ou amateur. La Bibliothèque du Congrès a décidé de déposer des milliers de photographies issues de ses fonds sur Flickr. Elle a ainsi fait le choix d'aller vers le public amateur de photographie, là où il se trouve. Il s'agit donc d'une stratégie d'exposition virtuelle, qui est légitime d'un point de vue culturel. À ce dispositif s'est ajoutée une stratégie scientifique, consistant à faire appel aux internautes pour contribuer à l'identification et à la datation de photographies sur lesquelles la Bibliothèque avait peu ou pas d'information (Casemajor Loustau, 2011). Ce type d'approche a été initié également par Patrice Peccatte (2012) pour des photographies de la seconde guerre mondiale mais ne se limite pas à la collecte d'informations sur la photographie. Le *Roy Rosenzweig Center for History and New Media*³⁰ (CHNM) de Washington a développé plusieurs programmes de collectes d'informations en ligne, sur des sujets contemporains (le 11 Septembre³¹, le Programme Bracero (1924-1964)³², ou les ouragans Katrina et Rita³³). Initié en Italie et ouvert à tous les continents, le projet *Memoro* est un instrument de collecte de la mémoire des individus³⁴. Ce type de programme peut porter sur des matières moins contemporaines et plus difficiles d'accès, comme le projet *Transcribe Bentham*, qui a pour objet la retranscription collaborative de manuscrits de Bentham³⁵, ou la collecte d'archives historiques par l'Open Society Archives (OSA) de la Central European University, dans le cadre de *Parallel Archive*³⁶. Ces

²⁹ « Amazon jette *1984* dans le trou de mémoire » , *Le Monde*, 22 juillet 2009.

³⁰ Voir le site internet du CHNM, [URL : <http://chnm.gmu.edu>], consulté le 2 mai 2012.

³¹ The September 11 Digital Archive, [URL : <http://911digitalarchive.org>], consulté le 2 mai 2012.

³² Bracero History Archive, [URL : <http://braceroarchive.org>], consulté le 2 mai 2012.

³³ Hurricane Digital Memory Bank, [URL : <http://www.hurricanearchive.org>], consulté le 2 mai 2012.

³⁴ Memoro : the Bank of Memories, [URL : <http://www.memoro.org>], consulté le 2 mai 2012.

³⁵ Transcribe Bentham : a Participatory Initiative, [URL : <http://www.ucl.ac.uk/transcribe-bentham>], consulté le 2 mai 2012.

³⁶ Parallel Archive, [URL : <http://www.parallelarchive.org>], consulté le 2 mai 2012.

programmes d'enquête en ligne sont les dignes héritiers de l'enquête de Jacques Ozouf (1983) sur les maîtres d'école à la Belle Époque et sont prometteurs. Ils doivent leur force à la nouvelle visibilité en ligne des sciences humaines et sociales, car de telles enquêtes n'auraient guère de sens sur les chemins vicinaux du Web. Sur des routes goudronnées et fortement fréquentées, en revanche, le *crowdsourcing* s'appuie sur un effet de réseau qui est en mesure de contribuer aux avancées de la recherche.

La puissance du nombre

Il est évident que l'ensemble de ces programmes ne fonctionne que parce qu'il bénéficie d'un très fort effet de réseau, permis par les millions d'hyperliens établis entre les projets, d'une part, et par la masse critique apportée par les regroupements autour de portails, qui bénéficient d'effets de plateformes très importants, d'autre part. Ces effets de plateforme ont vocation à s'étendre avec l'émergence d'outils de détection automatisée des références bibliographiques, permettant d'offrir des références cliquables (Kim *et al.*, 2012), des dispositifs de recommandation adaptés à la diversité des usages, des annotations collaboratives proposant des parcours de lecture aux étudiants. Toutes ces perspectives s'ouvrent à peine et nécessiteront des travaux de recherche et développement mutualisés qui s'articulent, tous, autour de la notion de *Digital Humanities*. Dans le domaine de la mise au point de puissantes plateformes d'édition électronique comme dans celui de l'émergence des humanités numériques, l'alliance des acteurs de la recherche s'impose et l'isolement serait une stratégie vouée à l'échec. En l'espèce, l'union fait la force.

Bibliographie

- AKRICH Madeleine, MÉADEL Cécile, RABEHARISOA Vololona, 2009, *Se mobiliser pour la santé. Des associations de patients témoignent*, Paris, Presses des mines.
- BENGHOZI Pierre-Jean, « DRM et nouveaux modèles d'affaires pour les contenus numériques ». [URL : <http://hal.archives-ouvertes.fr/hal-00263105/en/>], consulté le 7 mai 2012.
- BESTER Emma et DACOS Marin, 2010, « Que savons-nous de l'identité, des comportements et des attentes des lecteurs de Revues.org en 2008 et 2009 ? » [en ligne], 1^{er} février 2010, [URL : http://halshs.archives-ouvertes.fr/sic_00452002], consulté le 2 mai 2012.
- CASEMAJOR LOUSTAU Nathalie, 2011, « La contribution triviale des amateurs sur le Web : quelle efficacité documentaire ? », *Études de communication*, n° 36, p. 39-52.
- CAVALLO Delphine, 2009, « Revues.org, l'invention de l'édition électronique scientifique, entre libre accès et modèle économique pérenne » [en ligne], *Mémoires du livre*, vol. 1, n° 1, [URL : <http://www.erudit.org/revue/memoires/2009/v1/n1/038638ar.html>], consulté le 2 mai 2012.
- CHATEAURAYNAUD Francis, 2010, « Les humanités numériques sont-elles solubles dans Google ? » [en ligne], *Socio-informatique et argumentation*, 28 juillet 2010, [URL : <http://socioargu.hypotheses.org/1095>], consulté le 2 mai 2012.
- CHATEAURAYNAUD Francis et DEBAZ Josquin, 2010, « Prodiges et vertiges de la lexicométrie » [en ligne], *Socio-informatique et argumentation*, 23 décembre 2010, [URL : <http://socioargu.hypotheses.org/1963>], consulté le 2 mai 2012.
- DACOS Marin éd., 2009, *Read/Write Book. Le livre inscriptible*, Marseille, Cléo.
- , 2010, « OpenEdition Freemium » [en ligne], *OPENEDITION*, 18 JANVIER 2011, [URL : <http://www.openedition.org/8699>], consulté le 2 mai 2012.
- DACOS Marin et MOUNIER Pierre, 2010, *L'édition électronique*, Paris, La Découverte.
- 2011, « L'édition électronique », *Communications*, n° 88, p. 47-57.
- ERTZSCHEID Olivier, GALLEZOT Gabriel, BOUTIN Éric, 2009, « PageRank : entre sérendipité et logique marchande », *L'Entonnoir. Google sous la loupe des sciences de l'information et de la communication*,

- G. Gallezot et B. Simonnot éd., Caen, C&F Éditions, p. 113-136.
- GINSBERG Jeremy *et al.*, 2009, « Detecting influenza epidemics using search engine query data », *Nature*, vol. 457, n° 7232, p. 1012-1014.
- GUÉDON Jean-Claude, 2001, *In Oldenburg's Long Shadow. Librarians, Research Scientists, Publishers, and the Control of Scientific Publishing*, Washington, ARL.
- HELFT Miguel, 2008, « Google's new tool is meant for marketers », *The New York Times*, 6 Août 2008.
- KIM Young-Min, BELLOT Patrice, FAATH Élodie et DACOS Marin, 2012, « Machine Learning for Automatic Annotation of References in DH scholarly papers », *DH2012*, Université de Hambourg, 16-22 septembre.
- MATHIEU Gaétan, 2011, « Le *New York Times* sauvé par son site internet » [en ligne], *Owmi*, 24 novembre 2011, [URL : <http://owni.fr/2011/11/24/le-new-york-times-sauve-par-son-site-internet>], consulté le 2 mai 2012.
- MOUNIER Pierre, 2010, « L'édition électronique : un nouvel eldorado pour les sciences humaines ? », *Read/Write Book*, M. Dacos éd., Marseille, OpenEdition Press, p. 149-156.
<http://press.openedition.org/169>
- 2012, « Impressions d'automne 1 : le cas Meredith » [en ligne], *Blogo-numericus*, 28 janvier 2012, [URL : <http://blog.homo-numericus.net/article11010.html>], consulté le 2 mai 2012.
- OZOUF Jacques, 1983, *Nous les maîtres d'école. Autobiographies d'instituteurs de la Belle époque*, Paris, Gallimard.
- PECCATE Patrick, 2012, « PhotosNormandie a cinq ans – un bilan en forme de FAQ » [en ligne], *Déjà vu*, 27 janvier 2012, [URL : <http://culturevisuelle.org/dejavu/1097>], consulté le 3 mai 2012.
- PETERS Jeremy W., 2011, « *The Times* announces digital subscription plan », *The New York Times*, 17 mars 2011.
- SALMON Félix, 2011 « The NYT paywall is working », *Reuters*, 26 juillet 2011.
- SMITH Kevin, 2011, « The unexpected reader » [en ligne], *Scholarly Communications @ Duke*, 15 novembre 2011, [URL : <http://blogs.library.duke.edu/scholcomm/2011/11/15/the-unexpected-reader>], consulté le 2 mai 2012.
- VECAM, 2011, *Libres savoirs. Les biens communs de la connaissance ; produire collectivement, partager et diffuser les connaissances au XXIe siècle*, Caen, C&F éditions.
- WILLINSKY John, 2006, *The Access Principle. The Case for Open Access to Research and Scholarship*, Cambridge (Mass.), MIT Press.