Citations des ressources électroniques dans les publications scientifiques : analyse comparée et stratégie des liens

Ghislaine Chartron, Elisabeth Caillon

▶ **To cite this version:**

Ghislaine Chartron, Elisabeth Caillon. Citations des ressources électroniques dans les publications scientifiques : analyse comparée et stratégie des liens. Document numérique et société, Nov 2008, France. pp. sic_00359488

HAL Id: sic_00359488

https://archivesic.ccsd.cnrs.fr/sic_00359488

Submitted on 9 Feb 2009

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.

Manuscrit auteur, publié dans "Dans Actes de la Conférence "Document numérique et Société, vers un changement de paradigme ?" - Document numérique et société, France (2008)"

sic_00359488, version 1 - 9 Feb 2009

Citations des ressources électroniques dans les publications scientifiques : analyse comparée et stratégie des liens[1]

Ghislaine CHARTRON
Professeur, INTD-CNAM
Laboratoire : ELICO
ghislaine.chartron@cnam.fr

Elisabeth CAILLON
Documentaliste
INTD-CNAM
caillone@yahoo.fr

Résumé :
Cette communication s'intéresse aux nouvelles dimensions de la citation dans les textes scientifiques dans le contexte numérique. A l'appui d'une étude exploratoire quantitative menée fin 2007 et début 2008 dans le domaine de la biologie moléculaire, de la physique appliquée et de l'économie-gestion, les résultats montrent une faible intégration de ressources Web dans les textes (en dehors des versions en ligne des vecteurs traditionnels), sauf pour la biologie moléculaire. L'étude met en évidence une typologie des mécanismes de liens développés par les éditeurs et les agrégateurs. Un ensemble de recommandations finales sont formulées pour accompagner cette économie du lien devenue aujourd'hui centrale pour les citations scientifiques.

Mots clés :
Citation, référence bibliographique, publication scientifique, ressource numérique, éditeur, agrégateur, biologie moléculaire, physique appliquée, sciences économiques, économie des liens.

Introduction

Dans le contexte du fort développement des ressources numériques, cette communication s'intéresse à l'évolution des pratiques de citations dans les publications scientifiques. Pour la documentation-papier, les préoccupations s'étaient essentiellement focalisées sur l'écriture de ces références, les règles précises et non ambigües de leur rédaction dans le souci d'une recherche ultérieure facilitée et de codes communs partagés (Boulogne, 2005). Cette problématique reste bien sûr d'actualité.

Mais la dimension numérique introduit un nouveau contexte marqué notamment par la très grande diversité des ressources pouvant être citées et par la possibilité d'accéder directement aux textes par des mécanismes de liens à partir de ces citations/références.
La citation de ressources électroniques dans les publications scientifiques renvoie donc à une diversité d'enjeux : enjeux cognitifs (continuum du travail intellectuel), à des enjeux stratégiques (reconnaissance et mise en visibilité de toutes les ressources, de tous les acteurs œuvrant aujourd'hui à la transmission des idées), à des enjeux de mémoire (stabilité et pérennité d'accès à ces ressources).

[1] Ces travaux ont été menés dans le cadre d'une étude financée par le Bureau de l'Information Scientifique et Technique à la Direction Générale de la Recherche et de l'Innovation du Ministère de l'Enseignement Supérieur et de la Recherche. Rapport remis en octobre 2008.

sic_00359488, version 1 - 9 Feb 2009

Cette communication résulte d'une étude menée pour le Ministère de l'Enseignement Supérieur et de la Recherche de septembre 2007 à juin 2008[2].
L'analyse des citations de ressources électroniques dans les publications nous conduira tout d'abord à éclairer l'appropriation de ces ressources par les auteurs. Cet indicateur a certes des limites, il vient en tout cas compléter les autres indicateurs souvent utilisés pour rendre compte de l'usage, à savoir les analyses quantitatives exploitant les traces de téléchargement des fichiers selon des protocoles en voie de normalisation, notamment par le standard Counter (Boukacem, Schöpfel, 2005), d'autre part les analyses qualitatives basées sur des entretiens avec les usagers (Mahé, 2002).Les travaux ont conduit dans un second temps à apprécier l'état d'interconnexion des ressources, à l'appui des mécanismes de liens parfois complexes à l'œuvre et de façon transparente pour l'usager.

Par contre, cette communication ne s'intéressera pas à la dimension bibliométrique de la citation et à son renouvellement dans le cadre numérique qui font l'objet de nombreux autres travaux ces dernières années (voir en particulier les travaux de S. Harnad et de son équipe).

Objectifs, méthodologie, données

Les objectifs de cette étude s'inscrivent dans la continuité de nombreux travaux sur l'évolution des pratiques de citations (Clarke, M.E., Oppenheim, C., 2006), (Dalbello & all, 2006) par exemple et sur les mécanismes de liens (Jacso, 2004a, 2004b), (Apps, 2005).
Son originalité est de fournir une analyse actualisée de ces développements (données collectées en 2007-2008), dans une dimension comparative des citations de ressources électroniques.

Les éléments de comparaison concernent ici le domaine scientifique, le type de publication, la sphère linguistique, le type de ressources électroniques citées, la localisation dans les textes. Trois domaines ont été choisis relativement à la richesse des ressources en ligne: la biologie moléculaire, la physique appliquée, l'économie-gestion.

Les objectifs visés étaient de rendre compte :

- de la présence de citations de ressources électroniques dans les bibliographies, les notes, les textes des publications,
- de caractériser les ressources citées : articles de revues, communications, sites, textes d'archives ouvertes, rapports en ligne, données…,
- de pointer les éventuelles différences entre les pratiques relevées dans les actes de colloques (ou/et thèses) et les articles de revues ? L'hypothèse sous-jacente, étant que la communication pouvait être plus libre dans certaines publications moins contraintes par des prescriptions éditoriales,
- de relever d'éventuelles différences entre des publications de la sphère anglophone et de la sphère francophone,
- de caractériser la valeur ajoutée des intermédiaires (éditeurs, agrégateurs) sur les citations électroniques.

[2] Le rapport sera prochainement en ligne sur le site du Ministère.

sic_00359488, version 1 - 9 Feb 2009

Méthodologie retenue :
Analyse des citations des articles d'un échantillon de 3-5 revues « phares », par domaine, d'un échantillon de thèses et de colloques dans les domaines choisis. Cette analyse s'est effectuée sur l'année 2007 de publication. Les données sont recueillies dans les collections numériques rétrospectives des revues accessibles par un éditeur ou un agrégateur.
Aucune automatisation n'a semblé possible pour analyser ces données. Compte tenu de l'ampleur de la tâche et des contraintes de temps, l'échantillon est nécessairement réduit et cette étude ne pourra être considérée que comme une étude exploratoire concernant la problématique visée.

Le recueil des données s'est déroulé entre octobre2007 et février 2008 pour l'ensemble des publications analysées.4 revues ont été retenues pour les 2 disciplines des sciences de la nature, revues faisant partie des 20 premières du Journal Citation Report (JCR). Ces revues sont anglophones. L'objectif a été aussi de diversifier entre éditeur commercial international et sociétés savantes. Pour l'économie-gestion : 8 revues (4 anglophones, 4 francophones) ont été retenues dans la liste de la section 37 du CNRS. *(Les revues sont listées dans les tableaux ci-dessous).*

Normativité, faiblesse des ressources web

Des ressources web quasi-absentes

Un premier constat général (Cf. les tableaux ci-dessous) concerne le développement important des liens automatiques associés aux citations dans les 3 échantillons traités. Cette automatisation concerne en moyenne 90% des citations dans les revues de biologie moléculaire, entre 70 et 90% pour les revues en physique appliquée. Celle-ci est un peu moins marquée pour les revues anglophones en économie/gestion où le rôle de l'agrégateur Ebsco est toutefois majeur et en cours de développement avec la technologie OpenUrl.
Concernant les revues francophones en économie-gestion, l'automatisation n'avait pas débuté à la date de l'étude. La situation est en évolution depuis.

A l'inverse, les ressources électroniques originales (URL explicitement fournies par les auteurs) sont quasi-absentes. Selon l'analyse des URL des numéros de revues anglophones: elles sont très faibles (<1%) dans les bibliographies en sciences exactes, un peu plus élevées en économie/gestion pour une revue (3%). Elles sont observables aussi dans les notes en économie gestion mais restent en faible quantité (0,3 à 4%). Elles ne sont présentes, de façon notable, que dans le texte en biologie moléculaire (136 et 208 URL relevées par exemple).
Concernant les numéros des revues francophones d'économie-gestion analysés, on note une plus grande quantité et diversité d'URL (jusqu'à 7,4% dans la bibliographie et 1,7% dans les notes), constat certainement lié au sujet de la revue (économie publique) pour lequel de nombreux rapports sont en ligne, et certainement lié à l'absence de contrainte éditoriale, nous reviendrons sur ce point ultérieurement.

Concernant les thèses et les actes de colloques, contrairement à ce qu'on aurait pu penser, on n'observe pas plus de citations de ressources électroniques informelles (URL) que dans les revues, et ce, dans toutes les disciplines. Les thèses analysées sont particulièrement pauvres en URL, plus pauvres que les numéros de revues étudiés. De même, en économie/gestion, les URL y sont presque absentes, alors qu'elles sont présentes dans les revues (et les actes de

colloque) surtout au niveau des notes. Les thèses de physique appliquée ne contiennent quasiment pas d'URL (bibliographie et texte). On peut se demander si les jeunes thésards ne s'autocensurent pas pour les citations d'URL. Mais, l'échantillon de thèses est trop petit pour généraliser ce résultat, résultat qui pourrait aussi être lié aux sujets des thèses observées.

Tableau 1 : tableau comparatif des types de liens associés aux ressources électroniques dans l'échantillon analysé

(Dans ces tableaux, la codification des URL se trouve en Annexe 1)

Publications (revues, thèses, actes de colloque)	**Liens automatiques dans la bibliographie** % / aux références totales	**URL dans la bibliographie** % / aux ressources élect. citées. Typologie	**URL dans le texte** Typologie	**URL dans les notes** % / aux notes totales. Typologie
BIOLOGIE MOLECULAIRE **Revues :** PloS Biology, (Public Library of Science), (n°sept 2007)	93%	<1% 4 URL: 2rap, 1sit, 1acl	136 URL 67%fou, 19% bdd 5% log, 5% out 4% sit	Pas de notes
Genome Research, (CSHLP) (n°sept 2007)	95%	0%	65 URL: 33% bdd ,29% sit 21% log,17% out	0%
Current Biology, (Elsevier) (n°sept 2007)	86%	<1% 1 URL: 1log	1 URL: 1bdd	Pas de notes
Plant Cell, (American society of plant biologists) (n°août 2007)	96%	<1% 6 URL : 4 ebo, 2 art	208 URL: 49% bdd (nac) 16% bdd, 15% sit 12% log ,5% fou 2% out, 1% met	0%
4 Thèses	Pas de liens	-<1%, 1 URL : 1bdd -<1%, 2 URL : 1bdd, 1out -Aucune URL dans 2 thèses	26 URL 21bdd (1nac)3log, 1fou, 1sit Aucune URL dans 3 thèses	Pas de notes dans les 4 thèses
PHYSIQUE APPLIQUEE **Revues** : Journal of Synchrotron Radiation, Blackwell – Wiley, (n°sept 2007)	72%	<1% 1 URL: 1 log	0	Pas de notes
Journal of Applied Physics, American Institute of Physics (n°sept 2007)	87%	<1% 2 URL: 1sit, 1out	0	Pas de note
Applied Physics A: Materials Science & Processing, Springer (n°sept 2007)	77%	0	0	Pas de notes
Nanotechnology, Institute of Physics Publishing (n°sept 2007)	89%	<1%, 5 URL: 2 sit, 1 rap, 1 out, 1pho	0	Pas de notes

sic_00359488, version 1 - 9 Feb 2009

4 Thèses ;	-0%, -92% (l'étudiant a fait sa biblio avec Latex) -0% - 0%	-3%, 4 URL: 3art, 1 out -0% -0% -0%	-0, -1 URL:1 log -0 -0	Pas de notes dans les 4 thèses
Actes de colloque : Journal of physics : conference series vol.85, 22-25/05/2007 (Institute of Physics Publishing) Analyse de 10 articles/37	34%	0%	0	pas de notes

ECONOMIE-GESTION **Revues anglophones** : Canadian Journal of Economics, (Blackwell Synergy), (fév 2007)	65%	3% 16 URL: 3don, 2 sip, 2 nor, 1sit, 3 art, 1pret, 1 rap, bel, 1 dtc, 1tex	7 URL: 2sit, 5don	2% 5 URL: 2sit,2don,1tex
European Economic Review, (Elsevier, diffuseur : Ebsco), (janv 2007)	49%	0	0	0
Quarterly Journal of Economics, (MIT press, diffuseur : Ebsco) (nov 2007)	99% (en construction, OpenUrl)	0	0	0,3% 1 URL: 1art
Journal of Economic, (Literature, American Economic Association, diffuseur : Ebsco) (sept 2007)	99% (en construction, OpenUrl)	0	0	0,5% 1 URL: 1 don
Southern Economic Journal, (Southern Economic Association, diffuseur : Ebsco) (oct 2007)	99% (en construction avec l'OpneUrl)	0	0	4% 8 URL: 3rap,3sit,2don
ECONOMIE-GESTION **Revues francophones** Revue Française de Gestion (Lavoisier diffuseur: Cairn) (août 2007)	0%	0	0	0
Revue économique, (Presses de la Fondation Nationale des Sciences Politiques, diffuseur : Cairn) (mai 2007)	0%	1,8% 5 URL: 1art, 1sip, 2pret, 1jou	0	1,7% 3 URL: 2sit, 1art
Economie publique, (Institut d'économie publique, diffuseur : revues.org)	0%	7,4% 17 URL : 5art,5rap,4jou,1bel,1don,1prêt	0	0
4 Thèses ;	Pas de liens	-1%, 1URL :1art -2%,3URL3pret - 2%,4URL : 1rap,1jou,2out -0%	-1%, 1 URL:1 sip -0% -1%, 5 URL : 4 sit, 1out -0%	0 dans les 4 thèses
Actes de colloque francophone : Actes du XXIIème du congrès AFM, Nantes, Association Française de Marketing 11/12 mai 2006 Analyse de 15 textes/90	0%	0,5% 4 URL : 1 pret, 3 art	0%	11% 4 URL: 4 sit

sic_00359488, version 1 - 9 Feb 2009

Actes de colloque <u>anglophone</u> Annual Meeting Allied Social Science Association, Nouvelle Orléans, American Economic Ass., 4/6 janv 2008, Analyse de 15 textes/200	0%	0,5% <u>2 URL</u> : 2 sit	0%	0,4% <u>2 URL</u> : 1sit, 1 out

Impact des préconisations d'éditeurs sur les pratiques de citations ?

Pour comprendre le nombre très faible de citations électroniques sous forme d'URL dans l'échantillon traité, au moins dans les revues, il était intéressant de savoir si les prescriptions des éditeurs pour la rédaction des articles dans ces revues pouvaient avoir un impact sur les pratiques des auteurs. Certes, la seule analyse de ces préconisations ne peut suffire, il faudrait aussi interroger les auteurs et comprendre leur mode de travail. Quoiqu'il en soit cet éclairage fut déjà révélateur des freins qui pouvaient être engendrés.
Nous ne reproduisons ici que quelques extraits de préconisations d'éditeurs transversales à l'échantillon traité.

<u>Tableau 2 : Exemples de préconisations d'éditeurs concernant les ressources électroniques</u>

Revues	**Recommandations des éditeurs concernant les ressources électroniques**	**Extrait de ces recommandations sur les sites d'éditeurs**
Genome Research, (CSHLP)	Demande d'intégration des citations d'URL indiquant où a été obtenu le matériel pour faire le travail, dans le texte de l'article	*"If the paper includes Web site URLs as references to places where material was obtained for doing the work, the URL should be included next to that information in the text"*
Plant Cell (American society of plant biologists)	- Demande d'intégration des citations de sites web dans le texte, non dans la bibliographie - Demande de tests des URLs et liens -Demande d'intégration du numéro PMID si l'article cité est publié en ligne et pas encore sur papier.	*"Citations for web sites (other than for primary literature) should be handled parenthetically in the text and not included in the reference list. Authors should test all URLs and links"* *"If you are citing an article that only exists as an early online version, include the PubMed ID (PMID) number to allow reviewers to link directly to the article."*
Applied Physics A: Materials Science & Processing, (Springer)	Demande d'intégration du DOI pour les articles publiés en ligne et pas encore sur papier.	*"A paper published online but not (yet) in print can be cited using the Digital Object Identifier (DOI). The DOI should be added at the end of the reference in question. Example:* *J. Ward, P. Robinson: Eur. Radiol., DOI: 10.1007/s00330-004-1450-y (2004)*

sic_00359488, version 1 - 9 Feb 2009

Quarterly Journal of Economics (MIT press)	Préconisation de citer les textes non publiés, les sites web personnels dans le texte ou les notes de pages	*"Papers submitted to journals but not forthcoming... and personal Web sites (where there is no commitment to archiving data) are better not cited as references. Depending on the situation: incorporate these into text or set them as page footnotes.* *« News/Web-type references : Factiva (http://factiva.com, Dow Jones and Reuters)."*

Les éditeurs des revues de cette étude ont manifestement le souci de donner la possibilité aux auteurs d'intégrer des ressources électroniques dans leurs textes, en particulier les revues en biologie moléculaire, les revues anglophones en économie-gestion de l'échantillon.
Une des raisons essentielles concerne la rapidité de publication: les versions en ligne des articles sont désormais diffusées dans bien des cas avant la version papier. L'éditeur demande alors de notifier les identificateurs pérennes de ces versions : numéro dans Medline (PMID) ou DOI par exemple.

Mais par ailleurs, toute référence mentionnant des ressources identifiables par des adresses URL sont majoritairement rejetées de la bibliographie, il est demandé à l'auteur d'intégrer ces références dans les notes, dans le texte ou dans une liste séparée à la fin de la bibliographie. L'argument de cette distinction semble être (il faudrait le confirmer par des entretiens) la stabilité d'adressage de ces ressources, la bibliographie devant assurer au lecteur une certaine garantie d'accès aux textes cités. Il n'y a pas de différence entre les exemples des revues des sciences de la nature et des sciences sociales anglophones ici, en ce qui concerne ces ressources électroniques de type URL.

Par contre, les revues francophones en économie-gestion n'ont encore pas spécifié de recommandations aux auteurs, risquant alors de laisser s'installer une diversité de pratiques. L'absence de préconisations pourrait aussi avoir des effets d'autocensure chez les auteurs : si rien n'est dit sur l'intégration d'URL, le raisonnement pourrait conduire à masquer ces ressources.

Les préconisations aux auteurs ne sont donc certainement pas neutres sur les pratiques de citations des auteurs. Une enquête qualitative serait nécessaire pour affiner en tout cas la compréhension de ces pratiques.

Mécanismes de liens à l'œuvre

L'analyse de bibliographies d'articles de revues conduit également à apprécier le fort développement des processus d'automatisation des liens avec le texte intégral, comme l'ont déjà mis en évidence d'autres travaux (Chang, 2007), (Culling, 2007).

Une exploration élargie des revues a permis de proposer et de vérifier ici une typologie des liens de façon transversale. Pour la biologie moléculaire et la physique appliquée, respectivement 24 et 17 revues classées en tête du *Journal Citation Report* ont été consultées. Concernant les sciences sociales (principalement la psychologie) : analyse de 11 revues d'éditeurs et de diffuseurs numériques anglophones et francophones.

sic_00359488, version 1 - 9 Feb 2009

sic_00359488, version 1 - 9 Feb 2009

Tableau 3 : Exemple de liens à partir des références d'articles

Revues	**Commentaires**	**Exemples de référence avec boutons de liens**
Laser and particle beams (Cambridge University Press)	- Lien vers Google Scholar, PubMed, Cross Ref - Lien OpenURL (fenêtre de requête OpenURL, possibilité résolveur public ou local) - lien "CJO Abstract" vers le site de l'éditeur ("journals.cambridge.org")	Wagner, T., Eberl, E., Frank, K., Hartmann, W., Hoffmann, D.H.H. & Tkotz, R. (1996). XUV amplification in recombining z-pinch plasma. Phys. Rev. Lett. 76, 3124–3127. OpenURL PubMed CrossRef Google Scholar Fiedorowicz, H. (2005). Generation of soft X-rays and extreme ultraviolet (EUV) using a laser-irradiated gas puff target. Laser Part. Beams 23, 365–373. OpenURL CrossRef [CJO Abstract] Google Scholar
-Nature Medicine - Nature Structural and molecular Biology	- Lien sur le site de l'éditeur (Article avec utilisation de doi) - Lien vers PubMed (résumé), ISI (résumé), Chemport (résumé)	Urano, T. et al. Efp targets 14-3-3 sigma for proteolysis and promotes breast tumour growth. Nature 417, 871–875 (2002). | Article | PubMed | ISI | ChemPort |

Bilan général

Biologie moléculaire et Physique appliquée :
Dans les deux disciplines de sciences de la nature, la mise en place de services de liens se vérifie chez tous les éditeurs, qu'ils soient commerciaux (Elsevier, Springer, Wiley, Nature) ou plus proches des sociétés savantes.
Le mécanisme Crossref est fortement présent, mais aussi les liens vers des services spécialisés par domaine (PubMed, Chemport…), des services spécifiques aux éditeurs (Scopus et Science Direct d'Elsevier, SpringerLink…).

Sciences sociales :
Les mécanismes de liens automatiques associés aux citations sont repérables pour les revues de certains éditeurs anglophones (Taylor &Francis, Cambridge University Press) mais pas tous (Muse, SAGE) à la date de cette étude (fin 2007). Nous n'en observons pas non plus à cette date pour les revues des éditeurs ou agrégateurs francophones (CAIRN, Erudit).

Typologie des liens identifiés

Pour tous les domaines, différents types de liens sont proposés à l'usager, sans qu'ils en comprennent très certainement la signification, d'où peut-être l'importance d'être le premier bouton cliquable …

Ces liens peuvent être classés sur les catégories suivantes :

Modèles généraux de liens :

sic_00359488, version 1 - 9 Feb 2009

Ces modèles dominent et sont transversaux à tous les domaines. Ils s'appuient sur des standardisations internationales, il s'agit principalement de :

- Crossref, Cross Reference , mécanisme initié par les principaux éditeurs STM et s'appuyant sur le DOI, http://www.crossref.org
Editeurs intégrant ici ce mécanisme : Annual Reviews, Nature, Elsevier, Wiley, Springer, Blackwell Publishing, Taylor et Francis, Cambridge University Press, Oxford University Press (OUP), PloS, petits éditeurs
- OpenURL, la version 1.0 est devenue un standard ANSI/NISO» (Z39.88-2004), ces liens sont contextuels, ils s'appuient sur les métadonnées du contenu (Machefert, 2007), (Van de Sompel, 1999a, 1999b)
Editeurs intégrant ici ce mécanisme : Oxford University Press mais développement en cours pour beaucoup d'offreurs de contenus.

Modèles de liens associés à des banques de données propres à chaque domaine:
Dans ce cas, le lien trouvé dans la citation pointe généralement vers une notice bibliographique permettant à son tour de rebondir si possible vers le texte intégral de l'article.

Par exemple :
- Liens vers PubMed (pour les éditeurs : Annual Reviews, Nature, Wiley, Springer, Taylor et Francis, OUP, petits éditeurs)
- Liens vers ISI (éditeurs : Annual Reviews, Nature, OUP)
- Liens vers Chemport (éditeurs : Nature, Springer, Wiley, Blackwell Publishing)
- Lien vers Astrophysics Data System (Springer)
- Lien vers Searchable Physics Information Notices (American Institute of Physics)
- Lien vers Inspec (IOP, AIP)
- Lien vers Google Scholar (Cambridge University Press).

Mécanismes internes propres aux éditeurs :
Il s'agit de liens internes pointant sur la plate-forme de l'éditeur. Ces liens peuvent s'appuyer sur différents identifieurs ou sur une codification interne propre à l'éditeur comme dans l'exemple suivant pour l'éditeur Biomedcentral : http://www.biomedcentral.com/1471-2164/7/25

Exemples relevés :
- Elsevier: liens vers Scopus et Science Direct (produits d'Elsevier)
- Springer : liens vers SpringerLink (banque de données multidisciplinaire de Springer)
- Wiley: liens vers la plate-forme Wiley Interscience qui oriente vers PubMed et ChemPort
- Taylor et Francis : liens vers Informaworld (plate-forme de publications en ligne développée par Routledge, Psychology Press et Informa Healthcare)
- IOP: technologie de liens « Hypercite », STACKS (système pour intégrer des liens vers les revues IOP).

Une exploration et une quantification détaillée de ces liens sur l'échantillon du corpus de l'étude menée pourra être consultée dans le rapport.

Tendances dominantes pour ces mécanismes de liens

sic_00359488, version 1 - 9 Feb 2009

Les mécanismes de génération de liens automatiques sont bien à l'œuvre dans le domaine des revues scientifiques, autant en sciences de la nature qu'en sciences humaines et sociales dont il faut toutefois souligner l'absence d'initiatives visibles des éditeurs et agrégateurs francophones, à la date de cette étude. Ce service a un coût, l'économie des liens doit être clairement identifiée et ce sont logiquement les grands éditeurs internationaux anglophones qui ont mis en œuvre cette valeur ajoutée en premier.

En sciences de la nature, l'échantillon nous montre la prédominance du mécanisme Crossref, en particulier en Physique appliquée, constat prévisible dans la mesure où les fondateurs de Crossref sont les principaux éditeurs internationaux en STM.

L'autre modèle générique de liens à savoir l'OpenUrl apparait peu développé dans ce corpus (toujours à la date de cette étude) alors qu'il devrait répondre à une plus grande intégration dans la navigation des ressources électroniques. L'une des raisons en est certainement la difficulté d'obtenir des métadonnées correctes sur les plateformes éditoriales sauf pour des acteurs tels qu'Ebsco qui maintiennent aussi de grandes bases de données bibliographiques structurées.

Le lien entre le texte intégral des articles et les différentes bases de données sectorielles (factuelles, bibliographiques, archives) est aussi en développement, assurant une continuité de travail pour le lecteur. Ces liens sont particulièrement développés en biologie moléculaire.

Enfin, l'échantillon montre qu'au-delà de la dimension technique que recouvrent ces mécanismes de liens viennent se greffer de réels enjeux stratégiques de captation de l'attention de l'usager : une même revue diffusée par deux acteurs différents peut proposer une circulation différente dans les services en ligne, privilégiant les services qui lui sont associés (c'est le cas des liens Scopus proposés par Elsevier et des liens internes proposés par Ebsco par exemple). Le terrain n'est donc pas tout à fait neutre.

Conclusion : la citation dans un paradigme renouvelé

Le constat majeur de cette étude concerne la très faible mention de ressources électroniques sous forme d'URL dans les textes et le développement important de mécanismes de liens dans les publications scientifiques. Ce résultat est à nuancer pour la biologie moléculaire.
Ceci peut paraitre paradoxal dans le contexte d'une très grande diversité et richesse des ressources en ligne, en dehors des vecteurs traditionnels de publication, par exemple les sites personnels de chercheurs, les archives ouvertes, les nombreux rapports sur différents sites officiels…

Pour les revues, un des facteurs explicatifs non négligeables semble être les préconisations des éditeurs sur la localisation des ressources électroniques « non conventionnelles » . Mais pour les thèses et les actes de colloques, la liberté de l'auteur est plus grande et donc les explications restent à trouver. La taille de l'échantillon étudié ne peut conduire à généraliser les propos mais les spécialités scientifiques sondées ici sont tout de même révélatrices de pratiques qui tendent à converger.

sic_00359488, version 1 - 9 Feb 2009

L'écart (présumé) qui semble exister entre la consultation réelle de ressources électroniques « non conventionnelles » par les auteurs et l'inscription effective dans leurs textes pourrait être en partie levé par :

- La stabilisation des liens de ces ressources « non conventionnelles » avec des identifieurs stables (de type doi ou autres, comme le font désormais les archives ouvertes),
- L'ouverture des préconisations des éditeurs : la possibilité d'intégrer ces ressources stabilisées dans les bibliographies et non uniquement dans le texte ou les notes dans bien des cas,
- La formalisation précise de recommandations concernant les ressources électroniques, de nombreux éditeurs n'en formulant aucune, ceci peut conduire à une autocensure de la part des auteurs.

L'étude met aussi en évidence l'automatisation en marche des liens par les éditeurs dans les revues. Cette automatisation progresse rapidement, des différences ont pu être relevées entre le début et la fin de l'étude. L'interopérabilité va très certainement privilégier de plus en plus des modèles génériques et il n'est pas étonnant de constater la domination grandissante de Crossref, même en sciences humaines et sociales.
L'OpenUrl constitue un cadre élargi pour l'automatisation de ces liens, ne privilégiant pas un identificateur particulier (comme le doi par exemple), ni un type de liens particuliers comme peuvent le faire les éditeurs. Le résolveur de liens est paramétré par les bibliothèques et non par les offreurs des contenus. Mais sa mise en place au niveau des sources est encore complexe, elle nécessite un gain de qualité sur les métadonnées des contenus et une coopération de l'ensemble des éditeurs pour alimenter les bases de connaissances des résolveurs de liens, les conflits d'intérêts peuvent exister.

Plusieurs recommandations peuvent au final être formulées pour un développement harmonieux et efficient de l'économie des liens dans les publications scientifiques :

- une régulation des pouvoirs publics serait appréciable concernant l'inscription coordonnée au niveau national dans le registre mondial DOI-Crossref, concernant également la mise en place d'un registre national des résolveurs de liens des organismes français afin de faciliter le paramétrage des offreurs de contenus, et enfin concernant des modalités d'accompagnement pour mettre en oeuvre ces technologies un peu complexes, à l'instar du groupe KBART[3] au Royaune-Uni, dans la suite du rapport de J. Culling (Culling, 2007),
- une révision des préconisations des éditeurs diffusées à leurs auteurs pour les citations des ressources électroniques, et une gestion optimisée des métadonnées dans leur chaîne de production éditoriale,
- un encouragement aux auteurs à faire connaitre leurs besoins concernant les matériaux qu'ils souhaitent citer et à les faire respecter.

[3] KBART/JISC : http://www.uksg.org/kbart

Ce nouveau paradigme de la citation des publications scientifiques concerne donc le renouvellement des contenus, mais aussi la complexité des accès associés aux textes sources. Une coordination efficace de tous les acteurs impliqués est aujourd'hui un objectif à atteindre pour offrir à l'usager une circulation riche et facilité dans ces espaces numériques de travail (Meyer, 2008).

Bibliographie

Apps, A., Mac Intyre, R. "*Emerging uses for the OpenURL Framework"*, 2005, <http://eprints.rclis.org/archive/00005220/>

Boukacem C., Shöpfel, J., "Statistiques d'utilisation des ressources électroniques : le projet COUNTER", *Bulletin des Bibliothèques de France*, 2005, t.50, n°4, pp. 62-64

Chang S.H., Full-Text Article Linking: "Where Are We Now?", *Chinese Librarianship*, 2007, n° 23, <http://www.white-clouds.com/iclc/cliej/cl23.htm>

Boulogne A., "Comment rédiger une bibliographie", Paris : Nathan, Collection 128. Information, documentation, 2005.

Chang S.H., "Full-Text Article Linking: Where Are We Now?" , *Chinese Librarianship*, 2007, n° 23, <http://www.white-clouds.com/iclc/cliej/cl23.htm>

Clarke M.E., Oppenheim C., "Citation behaviour of information science students II : postgraduate students", *Education for Information,* 2006, vol.24 n°1, p.1-30

Culling J., "Link Resolvers and the Serials Supply Chain ", *rapport UKSG*, 2007, <http://www.uksg.org/projects/linkfinal >

Dalbello M., Lopatovska I., Mahony P., Ron N. ,"Electronic Texts and the Citation System of Scholarly Journals in the Humanities: Case Studies of Citation Practices in the Fields of Classical Studies and English Literature", *Proceedings of Libraries in the Digital Age*, 2006, <http://dlist.sir.arizona.edu/1638/01/Dalbello_posterrev.pdf>

Jacso P., "Savvy searching, link-enabled cited references", *Online Information Review*, 2004a, vol. 28 · n°4,p. 306-311

Jacso P., "Citation-enhanced indexing/abstracting databases", *Online Information Review*, 2004b, vol. 28 n°3, p. 235-238

Machefert S., *" L'OpenURL dans les institutions françaises, une chance pour la valorisation des ressources électroniques ? "*, mémoire INTD, Titre 1, 2007. <http://intd.cnam.fr/servlet/com.univ.utils.LectureFichierJoint?CODE=1212418434613&LANGUE=0>

Mahé A., « La communication scientifique en (r)évolution. L'intégration des revues électroniques dans les pratiques informationnelles des chercheurs en sciences de la nature comme révélateur des mutations du modèle traditionnel de la communication scientifique »,

sic_00359488, version 1 - 9 Feb 2009

Thèse de doctorat en sciences de l'information et de la communication, 2002, Lyon 1, <http://www.enssib.fr/bibliotheque/documents/theses/mahe/mahe.pdf>

Meyer C.A., "Reference Accuracy: Best Practices for Making the Links", *Journal of Electronic Publishing*, 2008, vol. 11, n°. 2, http://dx.doi.org/10.3998/3336451.0011.206

NISO. The OpenURL Framework for context-sensitive services, z 39.88, 2004, 15 avril 2005.

Van de Sompel H. and Hochstenbach P., 1999a, "Reference Linking in a Hybrid Library Environment. Part 1: Frameworks for Linking", *D-Lib Magazine*, 5(4).
< http://www.dlib.org/dlib/april99/van_de_sompel/04van_de_sompel-pt1.html>

Van de Sompel, H. and Hochstenbach, P.. 1999b, "Reference Linking in a Hybrid Library Environment. Part 2:SFX, a Generic Linking Solution". *D-Lib Magazine*, 5(4),
<http://www.dlib.org/dlib/april99/van_de_sompel/04van_de_sompel-pt2.html>

sic_00359488, version 1 - 9 Feb 2009

Annexe 1

Les différentes URL relevées ont été codifiées de la façon suivante :

- acl : actes de colloques
- art : articles
- bdd (propre à la biol. moléculaire) : collection (bactéries, virus…), cartographie de gènes
- bdd (nac) (propre à la biol. moléculaire) : numéro d'accès d'une séquence de gène
- bel : brèves en ligne
- don : données, statistiques
- dtc : document technique
- ebo : livre électronique
- fou (propre à la biol. moléculaire) : fournisseur de produits chimiques, de souches , d'appareils de mesure
- jou : article de journal
- log : logiciel, programme
- met : méthode, protocole
- out : outil d'aide à la recherche (en biol. moléculaire : serveur prédicteur de séquences de gènes, planche anatomique, analyse d'images)
- pho : photos
- pret : prétirage électronique
- rap rapport
- sit : site d'éditeur et autres
- sip : site personnel
- tex : textes de loi (directives…)

sic_00359488, version 1 - 9 Feb 2009