



CNRS – DIRECTION DE L'INFORMATION SCIENTIFIQUE

Un modèle éditorial en transition : l'édition scientifique à l'horizon 2020

01/03/2015



Table des matières

Un modèle éditorial en transition : l'édition scientifique à l'horizon 2020	3
1 - La place de l'édition scientifique dans l'écosystème de la recherche publique : éléments pour une approche macroéconomique	3
<i>Tâches IST : 35% des coûts globaux de main d'œuvre de la recherche publique</i>	3
2 - Les coûts de l'édition scientifique : captation de valeur, réduction des charges, et rétention des marges	5
3 - Taille et structuration de l'édition scientifique au plan mondial	6
<i>Un marché de 12,8 milliards de dollar au plan mondial</i>	6
<i>Une industrie des services numériques</i>	6
<i>Une structuration en « longue traîne » dominée par les 12 plus grands éditeurs</i>	7
<i>Offre mondiale de revues de recherche : la concentration est une réalité</i>	8
4 - Un clivage croissant et irréversible entre les grands éditeurs « for-profit » et tous les autres éditeurs ... 8	
<i>Les grands éditeurs croissent plus vite que le marché en captant les activités d'édition des sociétés savantes</i>	9
<i>Les Big Deals : un avantage compétitif important réservé aux grands éditeurs</i>	9
<i>La taille des corpus gérés est une donnée-clé de la proposition de valeur des grands éditeurs</i>	10
<i>L'obsession de l'évaluation par la mesure de la publication en revues à facteur d'impact fait le jeu des grands éditeurs</i>	10
<i>Le basculement du modèle économique du Gold Open Access plus facile gérable par les grands éditeurs</i>	10
5 - Un secteur de concurrence imparfaite où la concurrence par les prix n'existe pas	11
<i>Une donnée-clé : l'insubstituabilité des revues entre elles</i>	11
<i>Une captation par les grands éditeurs de la rentabilité de l'édition scientifique</i>	11

6 - La question centrale de la titularité des droits de propriété intellectuelle de la publication scientifique	11
<i>D'une économie de la prestation éditoriale à une économie d'exploitation de droits de propriété intellectuelle</i>	12
<i>La captation des droits de propriété intellectuelle par les éditeurs est un obstacle majeur au développement de la science ouverte</i>	12
7 - Dynamiques de croissance et de rentabilité de l'édition scientifique de recherche.....	12
8 - La conversion du modèle économique des éditeurs au Gold Open Access	13
<i>Des stratégies Gold OA articulant systématiquement la création de méga-revues et le développement de catalogues de revues spécialisées.</i>	13
<i>Contrer le phénomène de « Lost Science »</i>	14
<i>Le Gold OA modifie en profondeur les logiques de concurrence entre éditeurs</i>	14
<i>La conversion du modèle économique des éditeurs au Gold Open Access ne garantit pas le maintien des marges</i>	15
9 - La stratégie de plates-formes au cœur des stratégies des éditeurs	15
10 - La concentration inachevée de l'édition scientifique for-profit.....	16
<i>Concentrations : l'effet de taille des éditeurs scientifiques est décisif à l'heure de la « science numérique »</i>	17
<i>Les investissements lourds amortis sur des positions de marché élargies</i>	17
11 - Les fortes positions des éditeurs « for profit » laissent cependant la place à des initiatives innovantes.....	17
En conclusion : les pouvoirs publics doivent promouvoir une Science Ouverte	18
<i>La construction des référentiels directeurs d'identification :</i>	19
<i>La production collaborative d'ontologies ouvertes</i>	19
<i>La redéfinition du droit de la propriété intellectuelle applicable aux objets de science</i>	19
<i>L'expérimentation systématique et l'échange de bonnes pratiques pour mettre au point des outils publics partagés de Text et Data Mining</i>	19
<i>Une réflexion sur les outils de « Search » et la mise au point de moteurs de recherches indépendants ..</i>	20

Un modèle éditorial en transition : l'édition scientifique à l'horizon 2020

Une étude DIST/CNRS

1 - La place de l'édition scientifique dans l'écosystème de la recherche publique : éléments pour une approche macroéconomique

L'édition scientifique est un maillon de la chaîne de la communication scientifique, qui elle-même est un maillon des activités de recherche publique. Mesurer l'incidence en termes de coûts (au niveau des seuls coûts de main d'œuvre) des tâches IST (se documenter, publier...) au sein de la recherche publique ; articuler la place de l'édition scientifique marchande (portée par les éditeurs) dans l'écosystème de la publication scientifique est un préalable indispensable à une analyse stratégique de l'édition scientifique.

Tâches IST : 35% des coûts globaux de main d'œuvre de la recherche publique

Les tâches IST représentent environ 35% des coûts globaux (exprimés en termes monétaires) de main d'œuvre de la recherche publique¹ et des activités d'édition, la « production de la recherche »² en représente 62,5%, les tâches d'édition 2,5% (en partie prises en charge par les éditeurs).

Ces 35% se composent des tâches de recherche documentaire des chercheurs (9,5% des coûts), des tâches de lecture de la documentation scientifique (19,7% des coûts), des tâches d'écriture des articles de recherche (4,3%), enfin des tâches de mise à disposition de la documentation dans les bibliothèques de recherche (1,2%)³. L'incidence importante des tâches IST dans le coût global de la recherche implique *qu'une réflexion sur la productivité de la recherche ne peut faire l'économie d'une réflexion sur la productivité des tâches de documentation et des tâches liées à la publication scientifique.*

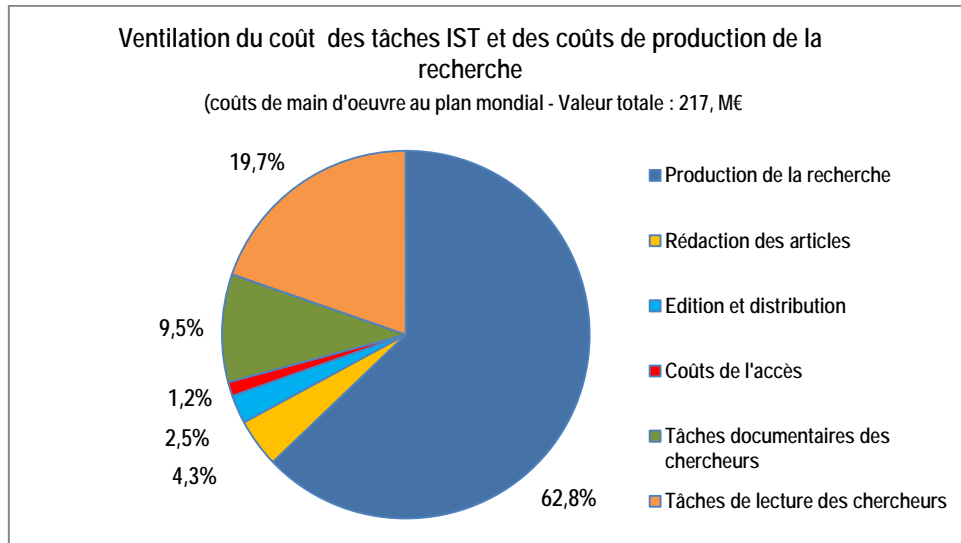
Les grands éditeurs scientifiques « for profit » l'ont bien compris : depuis 2 ans *la mise à disposition d'outils d'aide à la gestion des tâches documentaires et des tâches de rédaction est une composante nouvelle des offres des éditeurs et s'inscrit dans une stratégie de pénétration et de captation du « workflow » des tâches documentaires liées à la recherche* (lire ci-dessous).

¹ Source : "Activities, Costs and Funding Flows in the Scholarly Communications System: Key Findings" ; étude réalisée par le Research Information Network (www.rin.ac.uk) - novembre 2008.

² Effectuation des tâches de recherche stricto sensu hors tâches IST.

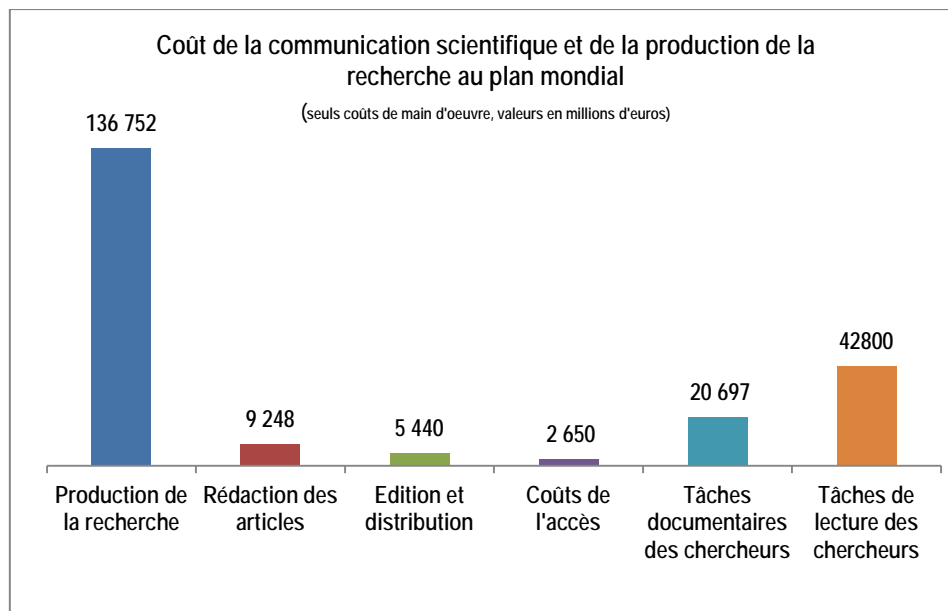
³ les 2,5% restants représentant comme on l'a dit les tâches éditoriales dont seul le *Peer Reviewing* est assuré par les chercheurs issus de la recherche publique.

Graphique 1



Source : "Activities, Costs and Funding Flows in the Scholarly Communications System: Key Findings" ; étude réalisée par le Research Information Network (www.rin.ac.uk) - novembre 2008.

Graphique 2



Source : idem graphique 1

On notera que les tâches liées à la publication scientifique ne représentent que 2,5% du total des coûts globaux de production de la recherche (en y incluant cette fois les tâches IST).

Si l'on ne tient pas compte des coûts de *Peer Reviewing* pour s'en tenir aux seuls coûts et marge des éditeurs, ce pourcentage est alors de 1,8%.

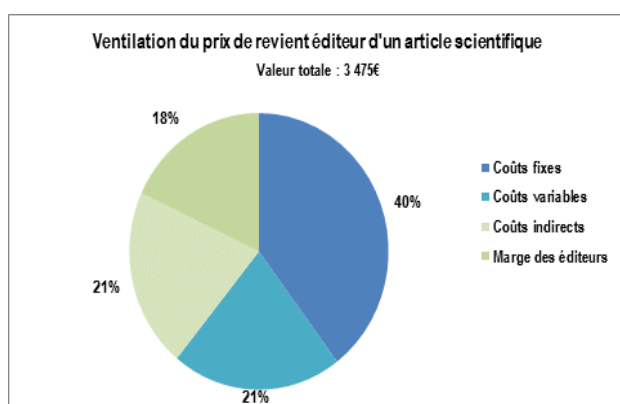
Quel que soit celui de ces deux pourcentages que l'on retienne, un constat se dégage : *alors que la publication scientifique joue au sein de la recherche une fonction pivot de transmission des connaissances elle n'a qu'une incidence très modérée sur les coûts globaux de la recherche.*

On notera que les 1, 8% représentant les « coûts éditeur s » au sein des coûts globaux de la recherche publique donnent une mesure approchée de ce que coûterait aux budgets de recherche publique le passage généralisé au Gold Open Access⁴.

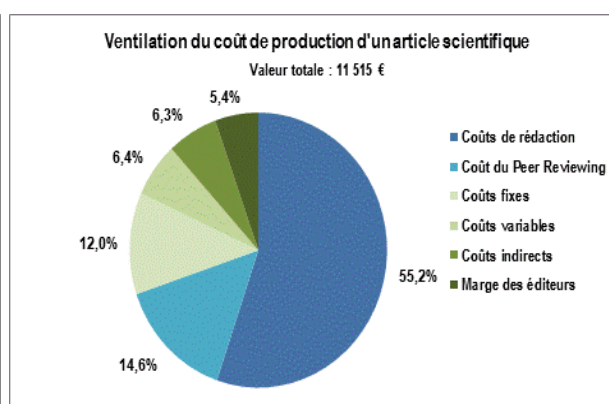
Les agences de financement de la recherche qui imposent une logique de publication *Open Access* autorisent d'ailleurs généralement les organismes de recherche à inclure dans les budgets de recherche un poste d'environ 2% pour couvrir les APC (*Article Processing Charges*), les frais d'édition imputés aux chercheurs publiant.

2 - Les coûts de l'édition scientifique : captation de valeur, réduction des charges, et rétention des marges

Graphique 3 :



Graphique 4



Source : « *Economic implications of alternative scholarly publishing models: Exploring the costs and benefits* » : étude réalisée pour le compte du JISC (Joint Information Systems Committee) ; www.jisc.ac.uk - janvier 2009.

Si l'on analyse les coûts de l'édition scientifique⁵, appréhendés au niveau de la structure de coûts d'un article de recherche « type »⁶, on constate (graphique 4) que **70% des coûts de production de la publication scientifique sont supportés par les budgets de la recherche publique**⁷. Ces 70% se ventilent en 55% de coût de rédaction des articles et 15% environ de coûts de *Peer Reviewing*., deux tâches assurées gratuitement par les chercheurs.

En se faisant céder par les contrats d'édition qu'ils imposent au chercheur désireux de publier l'ensemble des droits de propriété intellectuelle attachés aux articles, les éditeurs procèdent à une véritable captation de valeur qui est ici très directement mesurée.

⁴ Voie dorée pour le libre accès aux publications scientifiques

⁵ Source : « *Economic implications of alternative scholarly publishing models: Exploring the costs and benefits* » : étude réalisée pour le compte du JISC (Joint Information Systems Committee) ; www.jisc.ac.uk - janvier 2009.

⁶ Les coûts éditoriaux par article peuvent varier selon la discipline, le nombre de pages, l'existence ou non d'une version imprimée du même article : les études qui sont à la source de ces données évaluent les coûts d'un « article moyen » après avoir tenu compte de la variabilité de ces coûts. Les données fournies ici sont déjà anciennes (données 2007) : le coût relativement élevé affiché ici (3 475 € par article) reflètent le fait qu'alors les coûts éditoriaux comprenaient les coûts liés à la production et à la distribution d'une version imprimée de l'article. Pour les articles publiés aujourd'hui dans des revues qui n'existent que sous forme numérique, le coût moyen par article est abaissé, se situant dans une fourchette de 2 500 à 3 000 €. Là aussi il existe une grande variabilité selon le type de revue, la discipline, qui fait que les APC peuvent s'élever de 600 à 3 000 €.

⁷ Dans cette structure de coûts, les coûts du peer reviewing et de la rédaction des articles, assurés gratuitement par les chercheurs (donc financés sur budget de la recherche publique) ont été monétisés sur la base des temps évalués d'accomplissement de ces tâches.

La ventilation des seuls coûts par article pris en charge par les éditeurs (30% du coût total, 3475 € par article) est celle illustrée par le graphique 3 ci-dessus. Les coûts fixes représentent 40% des coûts supportés par l'éditeur, les coûts variables 21% (dans le cas d'une revue maintenant une version imprimée), les coûts indirects (coûts commerciaux, frais généraux) 21% et la marge des éditeurs 18%.

On notera que pour une revue qui n'existe que sous forme numérique diffusée en ligne les coûts variables sont fortement réduits : cette réduction des coûts n'a en général pas été répercutée par les éditeurs sur les tarifs d'abonnement. C'est la raison pour laquelle dans la dernière décennie la marge des éditeurs (lire ci-dessous), qui se situait déjà pour les revues imprimées à un niveau élevé chez les grands éditeurs « For-profit » (*Elsevier, Springer, Wiley, Nature publishing Group, Wolters Kluwer, Taylor & Francis*), ont continué de progresser de façon très régulière de 1 à 2% par an.

3 - Taille et structuration de l'édition scientifique au plan mondial

Un marché de 12,8 milliards de dollar au plan mondial

La quantification du marché mondial de l'information scientifique, technique et médicale pose de redoutables problèmes méthodologiques⁸. D'après le cabinet d'études américain *Simba Information*⁹, *le marché mondial de l'information scientifique, technique et médicale (ISTM) peut être estimé (données 2012) à 21 milliards d'euros*¹⁰. Ce marché est traditionnellement segmenté en trois compartiments : l'édition scientifique de recherche (8,6 Md. €); l'édition médicale généraliste (8,2 Md. €) et l'édition pour les sciences humaines et sociales (4,2 Md. €).

L'édition médicale dans son immense majorité n'est pas une édition de recherche, mais une édition destinée aux praticiens et aux étudiants. De plus ce compartiment inclus des services en ligne d'aide au diagnostic et à la prescription¹¹.

Si l'on s'en tient à l'édition de recherche proprement dite, le marché mondial est donc estimé à 12,8 Md. €. Cette somme se décompose de la façon illustrée par le graphe de la page suivante (*graphique 5*).

Une industrie des services numériques

On voit que les services numériques¹² représentent 60% des chiffres d'affaires. Chez les grands éditeurs, qui ont investi de façon plus précoce dans le numérique et dans une logique de plates-formes ce ratio est de l'ordre de 75%. Chez ces grands éditeurs, au-delà d'une phase d'investissements lourds sur ces plates-formes numériques, cette montée en puissance du numérique a eu un impact très positif sur la rentabilité des éditeurs.

⁸ En raison de la rareté des chiffres disponibles et des incertitudes sur le périmètre exact du secteur ISTM retenu par les différentes études. Deux cabinets américains, *Outsell, Inc.* (www.outsellinc.com) et *Simba Information* (www.simbainformation.com) fournissent des estimations dont les ordres de grandeur comparables. Les chiffres retenus ici sont ceux fournis par *Simba Information*, cités dans le blog *Scholarly Kitchen* (scholarlykitchen.sspnet.org)

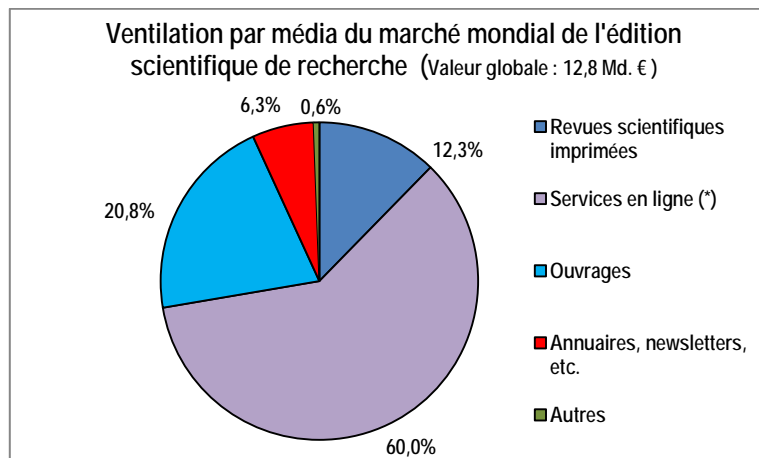
⁹ *Outsell, Inc.* et *Simba Information* donnent des estimations convergentes sur la taille globale du marché de l'ISTM évalué par *Outsell* (données 2013) à 24,5 Md. \$ et par *Simba Information* à 26 Md. \$.

¹⁰ Dans les études auxquelles on se réfère et dans les données financières publiées par les différents éditeurs, les valeurs sont le plus souvent exprimées en \$. On a choisi dans la suite d'exprimer ces valeurs en euros sur la base de la parité 1\$ = 0,80 €.

¹¹ Ainsi le leader mondial de l'édition Scientifique affiche un CA global 2014 de 2,54 Md. € qui se répartit en 1,3 Md. € pour l'édition de recherche (y compris de recherche biomédicale) et 1,24 Md. € d'édition médicale destinée aux praticiens et aux étudiants.

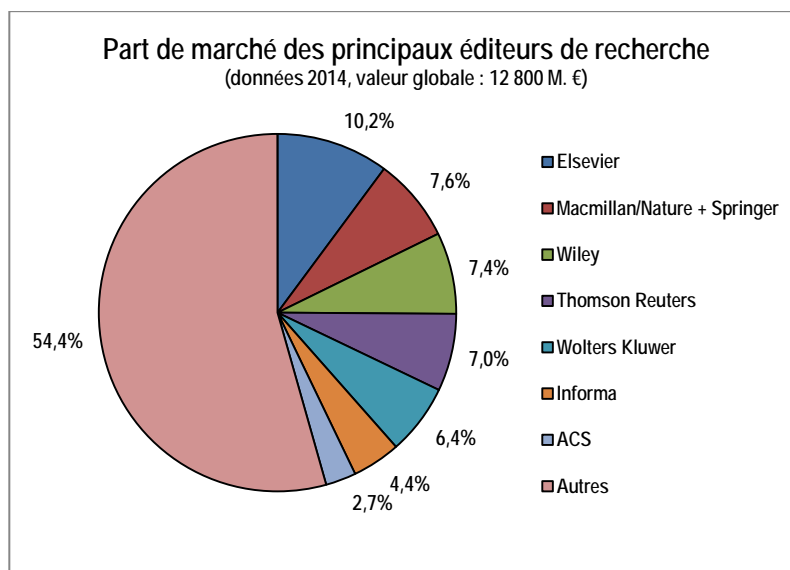
¹² incluant les bibliothèques de périodiques en ligne telles *ScienceDirect* d'*Elsevier* et *SpringerLink* de *Springer* mais aussi les plates-formes de recherche et d'analyse telles *Scopus* et *SciVal* d'*Elsevier* et *Web of Science* de *Thomson-Reuters*

Graphique 5 :



Source : Simba Information

Graphique 6 :



Source : DIST sur la base des résultats 2014 des sociétés cotées et de la presse économique pour les sociétés non cotées

Une structuration en « longue traîne » dominée par les 12 plus grands éditeurs

Les chiffres d'affaires 2014¹³ des 7 premiers éditeurs¹⁴ (les 6 premiers sont des sociétés « for-profit », le 7^{ème} est l'American Chemical Society – ACS -, la plus importante société savante mondiale) sont ceux reportés dans le tableau 1

¹³ 5 sur 6 des éditeurs occupant le haut du classement sont des sociétés cotées disposant d'un portefeuille d'activités diversifiées dans l'industrie de l'information. Elles ont présenté leurs résultats 2014 entre le 11 et le 26 février 2015

¹⁴ On a anticipé ici la fusion annoncée le 15 janvier 2015 du groupe Springer Science+Business Media et du groupe Macmillan Science et Education (dont le fleuron est le groupe Nature).

ci-dessous. Sur la base de ces chiffres et des données *Simba Information* sur la taille globale du marché de l'édition de recherche on calcule les parts de marchés des principaux éditeurs.

Elsevier, N°1 mondial, affiche une part de marché de 10,2%, ses quatre premiers poursuivants affichent des parts de marché de l'ordre de 7%. En analyse des positions concurrentielles ces parts de marchés ne signent pas une situation concentrée de l'offre.

Tableau 1

Données 2014 Valeurs en millions d'euros	CA Total 2014 du groupe d'appartenance	CA IST 2014	Croissance IST (organique) /2013	Taux de marge opération. IST	Contribution de l'IST au CA du groupe	Contribution de l'IST à la marge opér. Groupe
Elsevier	7 159,0	2540,0	2,0%	37,2%	35,5%	43,8%
Macmillan/Nature + Springer	1 255,0	972,0	4,0%	39,0%	77,5%	85%
Wiley	1 617,0	941,5	2,0%	28,6%	58,2%	76,6%
Thomson Reuters	11 140,0	896,5	3,0%	32,4%	8,0%	10,0%
Wolters Kluwer	3 660,0	816,0	5,0%	24,1%	22,3%	25,7%
Informa (Taylor & Francis)	1 557,0	560,0	3,0%	36,7%	36,0%	44,9%
American Chemical Society	402,1	351,0	1,8%	14,5%	87,3%	92%

Source : rapports financiers 2014 des éditeurs cotés et estimation DIST sur la base de données issues de la presse économique.

Offre mondiale de revues de recherche : la concentration est une réalité

La notion de concentration est cependant pertinente si l'on pousse plus loin l'analyse. Les 5 premiers éditeurs contrôlent 40% du marché (en valeur) des revues scientifiques. *Les quatre premiers éditeurs, s'ils éditent au plan mondial 24,8% des titres, publient 50,1% des revues à plus fort facteur d'impact* indexées par l'*ISI Journals Citation Report (JCR)*¹⁵.

Les quatre premiers éditeurs affichent un nombre moyen de plus de 1500 revues à leur catalogue (2200 pour le seul *Elsevier* ; même ordre de grandeur chez Springer/Macmillan-Nature) alors que pour l'ensemble des autres éditeurs, le nombre moyen de revues inscrit au catalogue est à peine supérieur à 2. Seuls 2% des éditeurs publient 100 titres ou plus.

Les 12 plus grands éditeurs publient 74% des 27 000 revues scientifiques actives. Au-delà aux 5 opérateurs « for profit » déjà identifiés il faut ajouter dans ce top 12 les grandes sociétés savantes, *ACS, RSC, IEEE, IET*¹⁶ et les grandes presses universitaires anglo-saxonnes, *Cambridge University press, Oxford University Press, California University Press.*

La structuration du marché de l'IST répond parfaitement au modèle de la « longue traîne ». Un nombre limité d'éditeurs occupent le haut du classement avec des parts de marché qui en apparence n'en font pas des acteurs dominants. Derrière eux figure un peloton très nombreux de très petites structures.

4 - Un clivage croissant et irréversible entre les grands éditeurs « for-profit » et tous les autres éditeurs

Jusqu'au début des années 2000 petits et grands éditeurs faisaient fondamentalement le même métier, les grands éditeurs bénéficiant seulement d'économies d'échelle et de plus fortes capacités d'investissement. Le clivage de l'édition scientifique entre grands éditeurs « for-profit » et éditeurs de petite taille s'est fortement

¹⁵ Source : Sally Morris, "Mapping the Journal publishing landscape", *Learned Publishing*, vol. 20, N° 4, octobre 2007.

¹⁶ American Chemical Society, Royal Society of Chemistry, Institute of Electrical & Electronics Engineers, Institution of Engineering & Technology

accentué dans les 15 dernières années et est aujourd'hui irréversible. Ce clivage s'explique par la conjonction de plusieurs facteurs

Les grands éditeurs croissent plus vite que le marché en captant les activités d'édition des sociétés savantes

Le premier facteur est la captation par les grands éditeurs de l'activité d'édition des sociétés savantes. Dans une logique d'« affermage » des revues éditées par les sociétés savantes en échange d'un montant fixe annuel de royalties supérieur aux revenus que dégagent les sociétés savantes en gérant elles-mêmes ces activités d'édition, les grands éditeurs ont fait croître rapidement leur catalogue. *Ce mouvement de satellisation des sociétés savantes autour des grands éditeurs « for-profit » a été continu et régulier et se poursuit aujourd'hui avec de bonnes réserves de croissance.*

Tableau 2 : Ventilation de la production éditoriale des 4 plus grands éditeurs scientifiques mondiaux¹⁷

	Elsevier	Springer	Wiley Blackwell	Taylor & Francis	A : Total "Big four"	B : Nbre. de titres au plan mondial	% A/B
C : Nombre de revues éditées	1 934	1 612	1 199	1 025	5 770	23 277	24,8%
D : Nombre de revues éditées pour le compte de sociétés savantes	390	328	549	324	1 591	11 578	13,7%
% D/C	20,2%	20,3%	45,8%	31,6%	27,6%		

Cette logique de captation des activités d'édition des sociétés savantes a permis aux grands éditeurs de faire croître leur catalogue beaucoup plus rapidement que celui des éditeurs de plus faible taille et surtout beaucoup plus rapidement que s'ils avaient mené ce développement de leur catalogue en compte propre.

Outre l'impact en termes de croissance du chiffre d'affaires, cette captation des activités d'édition des sociétés savantes est importante dans le cadre d'une stratégie numérique : les très grands éditeurs sont devenus aussi de grands agrégateurs qui proposent de vastes bibliothèques de périodiques en ligne (*ScienceDirect* pour *Elsevier*, *SpringerLink* pour *Springer*).

Ces bibliothèques numériques, qui peuvent être segmentées en fonction du besoin de populations de chercheurs, sont aujourd'hui au cœur de la proposition de valeur des grands éditeurs.

Les Big Deals : un avantage compétitif important réservé aux grands éditeurs

Le second facteur tient aux logiques de mise à disposition des bibliothèques numériques dans le cadre de « Big Deals ». Par « Big Deal » on désigne des contrats de licences d'usage fédérant de nombreux utilisateurs¹⁸ (qui ont accès à l'ensemble des fonds numérisés d'un éditeur).

Ces « Big Deals » accentuent le clivage entre les grands éditeurs et les autres puisque les catalogues limités des petits éditeurs ne leur permettent pas de proposer une valeur ajoutée d'agrégateur.

¹⁷ Source : Sally Morris, "Mapping the Journal publishing landscape", *opus citus*, note 14 ci-dessus.

¹⁸ Ces contrats peuvent même correspondre à des licences nationales ouvrant un droit d'usage à tous les organismes relevant de l'enseignement supérieur et de la recherche publique.

La taille des corpus gérés est une donnée-clé de la proposition de valeur des grands éditeurs

Les licences d'usage de ces grandes bibliothèques numériques d'articles scientifiques (et plus généralement des plates-formes numériques associant contenus et services) sont aujourd'hui au cœur de la proposition de valeur des grands éditeurs. Au cœur de ces plates-formes figurent de très vastes corpus de documents scientifiques : articles de recherche, ouvrages, actes de colloques, etc.

Cet avantage compétitif lié à la taille du corpus d'articles et de documents mis à disposition en ligne par les grands éditeurs sera demain amplifié par la généralisation des usages du Text et du Data Mining (TDM) : la taille des corpus exploitables est un paramètre important de l'intérêt de telles technologies. Pour consentir les investissements se chiffrant en plusieurs centaines de millions d'euros sur des plates-formes multiservices, les capacités financières des grands éditeurs « for profit » sont bien sûr un avantage compétitif majeur en raison des investissements très élevés que nécessitent ces grandes plates-formes numériques.

L'obsession de l'évaluation par la mesure de la publication en revues à facteur d'impact fait le jeu des grands éditeurs

Mal conçues (c'est-à-dire prenant en compte essentiellement les revues à haut facteur d'impact), les logiques d'évaluation à base d'indicateurs bibliométriques renforcent l'attractivité des grands éditeurs auprès des chercheurs. Les quatre plus grands éditeurs « for-profit » contrôlent 25% des 28 000 revues de recherche actives recensées au plan mondial mais 50 % des revues à facteur d'impact. Parce qu'ils savent que cela améliorera leur évaluation, les chercheurs vont en première intention vouloir publier dans les revues à facteur d'impact. *Le contrôle majoritaire des revues à facteur d'impact est donc pour les grands éditeurs un atout pour capter dans leurs publications les meilleurs articles des meilleurs chercheurs*, s'adjudicant par là un avantage compétitif décisif.

Le basculement du modèle économique du Gold Open Access plus facile gérable par les grands éditeurs

La conversion généralisée des éditeurs d'un modèle économique de « Toll Access » (fondé sur l'abonnement payant aux revues) à un modèle de « Gold Open Access » (fondé sur le paiement par les chercheurs publiant d'APC¹⁹) accentue également le clivage au sein de l'édition scientifique car la transition d'un modèle économique à l'autre est d'autant plus facile à conduire si cette évolution peut être « lissée » dans le temps grâce à un vaste catalogue de revues.

Même si l'activité d'édition scientifique consiste encore majoritairement pour tous les éditeurs à capter, évaluer, valider, mettre en forme des publications scientifiques émanant des chercheurs, même si toutes les maisons d'édition grandes ou petites partagent souvent une même culture éditoriale, il n'est pas exagéré de dire que *les très grands éditeurs « for profit » ne font plus aujourd'hui le même métier que les milliers d'éditeurs qui constituent la « longue traîne » déjà évoquée.*

¹⁹ Article Processing Charges

5 - Un secteur de concurrence imparfaite où la concurrence par les prix n'existe pas

Une donnée-clé : l'insubstituabilité des revues entre elles

Mais la réflexion sur la concentration au sein de l'édition scientifique amène à constater que *l'édition de recherche est un secteur obéissant à une logique de « concurrence imparfaite » qui explique les taux de marge très élevés des grands éditeurs « for-profit », souvent supérieurs à 30%.*

Cela s'explique par *un phénomène d'insubstituabilité des revues l'une à l'autre* : si la publication scientifique dans une discipline à pour vecteurs principaux N revues²⁰ le chercheur ou son institution s'abonneront à l'ensemble de ces revues. La concurrence entre éditeurs scientifiques est une concurrence pour capter les meilleurs chercheurs publiant ou les meilleurs *Peer Reviewers* ; pour capter les réserves de croissance liées à l'ouverture de nouveaux champs scientifiques donnant lieu au lancement de nouvelles revues ; enfin pour satelliser dans le cadre d'un « mercato » très animé les publications des sociétés savantes. *Elle n'est jamais une concurrence par les prix.*

Cette « concurrence imparfaite » explique les niveaux de rentabilité exceptionnellement élevés des grands éditeurs scientifiques « for profit » qui les distinguent des autres éditeurs de plus petite taille. Les sociétés savantes (à l'exception des très grandes déjà citées), les éditeurs n'ayant que quelques dizaines ou moins de revues à leur catalogue ne dépassent pas par en général un taux de marge opérationnelle de l'ordre de 15%.

Une captation par les grands éditeurs de la rentabilité de l'édition scientifique

Les très grands éditeurs captent à leur profit, au niveau de leur marge, une fraction importante de la valeur dégagée globalement par les activités d'édition de recherche, ce qui leur permet à la fois de dégager des dividendes importants pour leurs actionnaires, mais aussi d'avoir une capacité d'investissement dans les technologies numériques qui creuse encore l'écart de compétitivité avec les éditeurs de faible taille.

Cette non concurrence par les prix se traduit par une donnée simple : *le prix des revues scientifiques publiées par les grands éditeurs « for profit » sont en moyenne deux fois plus élevés que le prix de revues du même type éditées par les éditeurs « not-for-profit », au premier rang desquels les sociétés savantes²¹.*

6 - La question centrale de la titularité des droits de propriété intellectuelle de la publication scientifique

²⁰ ce dernier chiffre est relatif à la publication en physique des hautes énergies comme l'a montré l'initiative Scoap3 conduite par le CERN

²¹ Des travaux économiques démontrent largement ce point. Les deux études séminales sur ce sujet sont : « *Pricing of scientific Journals and Market Power* » ; Mathias Dewatripont, Victor Ginsburgh, Patrick Legros and Alexis Walckiers ; *Journal of the European Economic Association*, Volume 5, Issue 2-3, pages 400-410, April-May 2007 et « *Study of the economic and technical evolution of scientific publication markets in Europe* » ; Commission européenne, Direction générale de la Recherche, janvier 2006.

D'une économie de la prestation éditoriale à une économie d'exploitation de droits de propriété intellectuelle

Les éditeurs peuvent maintenir ces niveaux de prix élevés pour une autre raison : *en se faisant transférer par les chercheurs-auteurs les droits de propriété intellectuelle sur les articles qu'ils publient et en faisant des titres de revues des marques protégées, les éditeurs scientifiques²² opèrent dans un modèle économique d'exploitation de droits de propriété intellectuelle*. Or dans ce modèle d'exploitation de « droits d'auteurs » les prix finaux n'ont pour seule limite que la capacité des clientèles à supporter ces prix²³, et sont détachés des structures de coûts. Ce qui est paradoxal puisque *l'édition scientifique est fondamentalement une activité de prestation*. Cette prestation consiste à recueillir, évaluer et valider (par le Peer Reviewing) enfin à mettre en forme et de diffuser des publications scientifiques, en un mot de gérer un « Workflow » éditorial.

Le modèle économique du Gold Open Access (lire ci-dessous) offre l'occasion de mener un débat non pas sur la valeur ajoutée des éditeurs, qui est incontestable, mais sur le juste prix de la prestation éditoriale²⁴.

La captation des droits de propriété intellectuelle par les éditeurs est un obstacle majeur au développement de la science ouverte

Pour ce faire il faut rompre avec la cession par les chercheurs des droits de propriété intellectuelle aux éditeurs. Cette question de la publication ouverte (le plus souvent dans le cadre de licences Creative Commons) est essentielle au devenir de l'écosystème de la publication scientifique. *Dans cette perspective la refonte des droits de propriété intellectuelle attachés aux objets de science, et avant tout à la publication scientifique est indispensable et urgente*. L'objectif doit être de rattacher ce droit aux logiques de Creative Commons tarissant ainsi à sa source la captation par les éditeurs des droits de propriété intellectuelle.

Une articulation plus équitable est à trouver entre acteurs de la recherche publique et éditeurs scientifiques. C'est d'autant plus urgent que même si les tâches éditoriales constituent encore l'essentiel de l'activité (mesurée au niveau des flux éditoriaux gérés par les éditeurs²⁵) des grands éditeurs, ceux-ci sont déjà sur un autre positionnement : celui de fournisseurs de plates-formes agrégeant de vastes corpus scientifiques, des outils d'exploitation de ces contenus (outils de « Search » ou de TDM), des outils logiciels d'appui à la productivité des chercheurs et enfin des services analytiques aidant au pilotage et à la gouvernance de la recherche (lire ci-dessous).

7 - Dynamiques de croissance et de rentabilité de l'édition scientifique de recherche

²² Ici la distinction entre petits et grands éditeurs n'est pas pertinente.

²³ Le mouvement en faveur de l'Open Access est né en 2001, date de la publication de la pétition PLoS (Public Library of Science), lancée par des chercheurs américains). Il s'est nourri de la lassitude des bibliothèques de recherche – et quelquefois des chercheurs – face à l'inflation croissante des prix d'abonnement aux revues scientifiques.

²⁴ L'initiative Scoap3 conduite par le CERN dans le contexte de la publication en physique des hautes énergies est de ce point de vue exemplaire.

²⁵ En 2014 plus d'un million d'articles ont été soumis à Elsevier pour publication, environ un tiers ont été publiés en revues, deux tiers ont été rejetés.

Les éditeurs scientifiques, indépendamment de leur taille, enregistrent, depuis deux décennies des taux de croissance organique modérée mais continue qui, selon les éditeurs varient de 2% à 4%, les « petits » éditeurs se situant en général dans le bas de cette fourchette. La raison de cette croissance linéaire s'explique par la croissance régulière des populations de chercheurs (en particulier dans les grands pays émergents : Brésil, Russie, Inde, Chine), la taille de ces populations de chercheur étant historiquement le principal facteur auquel est corrélée la croissance de l'édition scientifique²⁶. On relèvera que la croissance, exponentielle et non linéaire, du nombre d'articles publiés mondialement ne se répercute pas intégralement sur la croissance du chiffre d'affaires des éditeurs. Le passage au numérique et l'externalisation du gros des tâches éditoriales à des prestataires du sous-continent indien²⁷ dans les quinze dernières années a permis aux éditeurs d'absorber dans leurs coûts la croissance des flux éditoriaux qu'ils ont à gérer .

Entre 2006 et 2010, la croissance organique des grands éditeurs a même été de l'ordre de 5% par an : les logiques de « *Big Deal* » prévoient souvent une progression régulière des facturations aux bibliothèques de recherche dans le cadre de contrats pluriannuels. La justification de ces augmentations tarifaires contractualisées sur la base du budget constaté l'année précédant la signature d'une licence était que les chercheurs pourraient désormais, dans le cadre des licences de « *Big Deal* », accéder à l'ensemble de la production d'un éditeur (logiques de « *Eat-all-you-can* »)

Cette dynamique de croissance est plus vive pour les plus grands éditeurs qui proposent des plates-formes numériques plus complexes que de simples bibliothèques de périodiques. L'impact de ces offres de plates-formes multifonctionnelles sur la croissance des éditeurs ne s'est sans doute pas encore complètement déployé, car la pénétration du marché de la recherche publique par ces plates-formes n'en est qu'à ses débuts. *On peut raisonnablement pronostiquer que dans les années à venir on va assister à un découplage des taux de croissance organique des grands éditeurs de ceux constatés pour les autres éditeurs.*

8 - La conversion du modèle économique des éditeurs au Gold Open Access²⁸

Des stratégies Gold OA articulant systématiquement la création de méga-revues et le développement de catalogues de revues spécialisées.

Dans les deux dernières années tous les grands éditeurs (successivement *Springer*, *l'American Society of Chemistry*, *Sage*, *Nature*, *Taylor & Francis*, *Elsevier*, *la Royal Society of Chemistry* et *Wiley*) ont annoncé le lancement de « méga-revues »²⁹. L'articulation d'une nouvelle offre de « méga revues » multidisciplinaires et du développement d'un catalogue de revues spécialisées en Gold OA est clairement au cœur d'une mutation de la

²⁶ Source : *Growth dynamics of scholarly and scientific journals*, Michael Mabe, M Amin - *Scientometrics*, 2001 - *akademai.com*

²⁷ En 2011, dans une présentation aux investisseurs²⁷, *Elsevier* indiquait que sur un effectif de 15 000 ETP travaillant pour le groupe, 4 500 étaient directement employés par l'éditeur dans des fonctions marketing-vente et 10 500 dans des tâches de gestion des flux éditoriaux. Sur ces 10 500 emplois de « *Back Office* » seuls 2 500 relèvent de l'emploi de salariés directement par *Elsevier*, les 10 500 autres ETP sont externalisés, le plus souvent auprès de prestataires du sous-continent indien. Chez *Springer* ce sont 40% des tâches éditoriales qui sont ainsi externalisées.

²⁸ Modèle économique de mise à disposition en libre accès, en ligne, après une période d'embargo de l'ordre de 6 mois à un an, des publications scientifiques, les coûts éditoriaux étant couverts par des « *Article Processing Charges* » ou APC payés par les chercheurs ou leur institution de rattachement .

²⁹ Ces revues scientifiques à spectre disciplinaire large ont toutes des caractéristiques communes : elles publient un grand nombre d'articles pour un niveau d'APC de bas de fourchette compris entre 600 et 1400 \$ (ce dernier est le niveau moyen des APC exigés par *PLoS One* qui a été pionnier de ce type de service éditorial).

stratégie des éditeurs « for profit » qui ont entrepris de s'approprier le modèle économique de la « voie dorée » de l'édition en libre accès. Cette mutation s'accélère nettement depuis juin 2014. *La création de nouvelles revues en Open Access représente désormais près de 80% des nouveaux titres lancés sur le marché*³⁰.

Première composante de ces stratégies des grands éditeurs, *la création de méga-revues a pour objectif de permettre aux éditeurs d'absorber la croissance exponentielle du nombre d'articles publiés mondialement, qui ne peut plus être contenue dans le périmètre des revues sur abonnement préexistantes*. En effet pour atteindre son point mort une nouvelle revue répondant au modèle économique de l'abonnement (« Toll Access ») suppose la construction patiente d'une audience payante, sur un délai de l'ordre de 5 à 8 ans. Cette temporalité économique de la création de nouvelles revues sur abonnement n'est plus compatible avec l'accélération du rythme de la publication scientifique.

Contre le phénomène de « Lost Science »

*Le second objectif auquel visent à satisfaire cette création de méga-revues est de trouver une réponse au problème de « Lost Science »*³¹. Pour maximiser leur facteur d'impact, *les revues spécialisées sont très sélectives, laissant sans débouché de nombreuses publications de qualité*. Les éditeurs sont les mieux placés pour mesurer l'ampleur de ce phénomène de « Lost Science ». Ils sont aussi conscients, même si cet argument a jusqu'ici été peu utilisé dans les débats sur la publication scientifique, que la question de la « Lost Science » est l'une des critiques majeures que l'on peut adresser au système de la publication scientifique malthusien qui va de pair avec le Toll Access et la recherche de la maximisation du facteur d'impact des revues.

Le Gold OA modifie en profondeur les logiques de concurrence entre éditeurs

La création ou la conversion de revues scientifiques en *Gold Open Access* modifie en profondeur le scénario concurrentiel de l'édition scientifique. L'éditeur allemand *Springer*, qui s'apprête à fusionner avec *Macmillan Science & Education* (dont le groupe *Nature* est une filiale) publie plus de 400 revues en OA (y compris celles de sa filiale *Biomed Central*). *Alors que sa part de marché dans l'édition de revues scientifiques de recherche en « Toll Access » (accessibles sur abonnement) est de 17% sa part de marché dans le Gold Open Access est de 33%*³².

Cette montée en puissance des méga-revues des éditeurs « for profit » déstabilise les pionniers de l'édition scientifique « not-for-profit » en Open Access : PLoS One, la plus ancienne et la plus importante (par le nombre d'articles publiés) des méga-revues, créée en décembre 2006, a vu chuter de 25% les articles parus dans son titre phare depuis décembre 2013 (date qui marque un pic pour le nombre d'articles publiés)³³.

La montée en puissance du Gold OA ravive la concurrence entre grands éditeurs, jusque-là relativement figée. Si un grand éditeur n'est que marginalement présent dans une spécialité scientifique, entre autres parce qu'une ou plusieurs revues sur abonnement occupent déjà le terrain, il lui sera loisible de créer de toute pièce, sans consentir d'investissement lourd, une revue en Gold OA sur ce même créneau. *Dans le modèle économique du Gold OA, la concurrence entre éditeurs se déplace sur un nouveau terrain : celui de la captation des*

³⁰ Etude DIST/CNRS sur la base des communiqués de presse des éditeurs, cf. <http://www.cnrs.fr/dist/distint.html>. Dist Info N°8

³¹ On désigne par-là les connaissances contenues dans des publications scientifiques qui n'ont pas été retenues pour publication dans les revues sur abonnement.

³² Ces deux segments de l'édition scientifique ne sont pas de même taille : le marché de l'abonnement aux revues scientifiques est évalué à 10,5 Md. \$ (9,2 Md. €) contre 300 M\$ (263,4 M€) pour le marché des APC (estimations 2014). Mais ce dernier marché croît au rythme de 25% par an, quand les revues sur abonnement voient leur chiffre d'affaires globalement stagner.

³³ Source : *blog Scholarly Kitchen* : <http://scholarlykitchen.sspnet.org/2014/12/02/peak-plos-planning-for-a-future-of-declining-revenue/>

meilleurs chercheurs publiant et des meilleurs peer-reviewers. La sélectivité obligée des revues sur abonnement à fort facteur d'impact crée des opportunités pour développer de nouvelles revues en Gold OA qui accueilleront les chercheurs (et en particulier les jeunes chercheurs) dont les contributions n'ont pas été publiées (sans que cela soit la conséquence d'un manque de qualité) dans les revues prestigieuses préexistantes, accessibles en « Toll Access ».

La conversion du modèle économique des éditeurs au Gold Open Access ne garantit pas le maintien des marges

Le « portage » des marges élevés et des revenus dans le nouveau modèle économique du Gold OA ne va pas de soi : *le Gold Open Access fait apparaître clairement que l'activité d'édition scientifique est une activité de prestation* dans le cadre de laquelle il sera plus difficile de justifier des marges opérationnelles de 32% comme c'est le cas aujourd'hui chez Springer, Elsevier, Wiley ou Taylor & Francis. *Le Gold OA est intrinsèquement porteur d'une exigence de transparence sur la structure des coûts de l'édition scientifique. Cette transparence devrait pouvoir profiter aux communautés de chercheurs publiant.*

9 - La stratégie de plates-formes au cœur des stratégies des éditeurs

Les grands éditeurs devraient cependant préserver leurs marges et leur taux de croissance. Dès aujourd'hui leur proposition de valeur est moins centrée sur l'édition de revues en tant que telle³⁴ que sur la mise à disposition de vastes plates formes³⁵ pour lesquelles la question du libre accès ne se pose tout simplement pas.

Comme c'était le cas pour les revues en « toll access », ces plates-formes sont insubstituables l'une à l'autre : une mobilisation exhaustive des ressources documentaires scientifiques suppose un abonnement institutionnel pour toutes. *La concurrence par les prix entre ces plates-formes « must-have » n'existe pas. Ce qui devrait permettre aux grands éditeurs « for-profit » de préserver leurs taux de marge élevés.* Cela se vérifiera d'autant plus que pour les clients issus de la recherche publique, la notion de « juste prix » de ces plates-formes est difficile à appréhender autrement que par les limites de leurs budgets.

Ces plates-formes, qui en raison des coûts d'investissement qu'elles requièrent ne peuvent être proposées que par de très grands-éditeurs, partagent des traits communs :

- *elles agrègent de vastes corpus scientifiques*, formés de millions d'articles et d'images, de dizaines de milliers d'ouvrages, des centaines de milliers d'actes de conférences,
- *elles proposent des outils d'exploitation de ces contenus, au travers d'outils de « Search » ou de TDM,*
- *elles se complètent d'une offre des outils logiciels d'appui à la productivité du workflow documentaire des chercheurs* : Elsevier a racheté en 2013 la jeune-pousse Mendeley, un outil de partage collaboratif des ressources documentaires des chercheurs affichant une large base d'utilisateurs (on est ici dans une logique de réseau social dédié à la recherche) ; Springer avec Papers et Nature avec Read Cube explorent aussi ce créneau. Cette gamme d'outils inclut aussi des utilitaires d'aide à la rédaction des articles scientifiques (EndNotes de Thomson Reuters pour gérer les logiques de citations, FigShare de Digital

³⁴ Et une méga revue est-elle d'ailleurs encore une revue ou un gisement documentaire sur lequel est apposée une marque commerciale?

³⁵ Scopus pour Elsevier, Springer Link, Thomson Web of Science.

Science pour l'édition et le partage des données de la recherche et *OverLeaf*, toujours chez *Digital Science* un utilitaire d'assistance à la rédaction d'articles de recherche).

On l'a déjà dit : *les outils de productivité proposés par les éditeurs sont des outils « propriétaires », payants, non interopérables, basés sur des logiciels non « ouverts qui rendent incompatibles avec une logique d'« Open Process »* qui serait pourtant nécessaire pour rendre le workflow documentaire de la recherche publique moins dépendant des offres des éditeurs,

- *Pour Elsevier (services Scopus et SciVal) et Thomson Reuters (services Web of Sciences et InCites) ces offres de plates-formes s'enrichissent de services analytiques qui visent le marché nouveau de l'aide au pilotage et à la gouvernance de la recherche* (cf. encadré SciVal trends ci-dessous).

Ce qu'est SciVal Trends

Elsevier a annoncé³⁶ le 25 février 2015 le lancement d'un nouveau module au sein de son outil analytique *SciVal* baptisé *SciVal Trends*. Ce module permet pour la première fois d'utiliser au sein de la plate-forme d'analyse d'Elsevier les « traces » laissées par les usagers des services *ScienceDirect* (bibliothèque de périodiques en ligne) et de *Scopus* (moteur de recherche intégrant bases bibliographiques et bases de citations). Jusqu'ici les analyses de *SciVal* portaient exclusivement sur les publications scientifiques. Or comme le précise le communiqué de presse d'Elsevier un tiers des chercheurs ne sont pas publiants : les centres d'intérêts de ces chercheurs étaient jusqu'ici une « zone aveugle » inaccessible aux analyses de *SciVal*. En combinant les données d'usage (très volumineuses : on est là dans la logique du « Big Data ») de ses services en ligne les plus utilisés et la puissance analytique des technologies propriétaires de *RELX Group* (nouveau nom du groupe *Reed Elsevier*), *SciVal Trends* va permettre à ses usagers de visualiser sur la base de mots-clés définissant un profil thématique les « traces » laissées par ces utilisateurs anonymes et non publiant - mais aussi les traces des chercheurs-publiant utilisant les services en ligne d'Elsevier dans leurs recherches documentaires. Cette nouvelle fonctionnalité s'appuie sur la technologie *High Performance Computing Cluster* (HPCC), une technologie propriétaire du Groupe RELX (connu jusqu'ici sous la raison sociale *Reed Elsevier*). Cet outil est l'une des « dorsales technologiques » du groupe sur laquelle s'appuie aussi bien la division *LexisNexis* pour analyser des fonds de jurisprudence ou des fonds d'information brevets que la division *Elsevier* au travers de *SciVal*. Ce dernier service permet désormais d'analyser 75 billions (75 000 milliards) de données métriques d'usages actualisées chaque semaine. L'économie numérique est une économie des données et de la trace : les technologies analytiques traitant des données d'usage des services en ligne étaient jusqu'ici réservées à des acteurs comme *Google* ou *Facebook*. Cette valorisation par des technologies de Big Data, dans le cadre de *SciVal Trends*, des « traces » (anonymisées) laissées par les usagers des services en ligne est une première dans le champ de l'IST.

10 - La concentration inachevée de l'édition scientifique for-profit

Malgré une croissance modérée mais régulière des chiffres d'affaires et les niveaux de marge opérationnelle très confortables des grands éditeurs scientifiques, *d'autres opérations de concentration sont à attendre dans les*

³⁶ <http://www.elsevier.com/about/press-releases/science-and-technology/elsevier-launches-the-new-generation-sci-val,-providing-dynamic-real-time-analytics-and-insights-into-the-global-research-landscape#sthash.M6VFPF5s.dpuf>

prochains mois et les prochaines années, au-delà du rapprochement annoncé de *Springer* et de *Macmillan Science & Education/groupe Nature*.

Ce mouvement de concentration n'a d'ailleurs jamais vraiment cessé. Reed Elsevier (RELX Group depuis le 27 février) est né en 1993 du rapprochement du britannique *Reed* et du néerlandais *Elsevier*, soucieux de bâtir un portefeuille d'activités contra-cycliques. En 1998 une opération très avancée de fusion entre *Reed Elsevier* et *Wolters Kluwer*, qui aurait donné naissance au N°1 mondial de l'information professionnelle, échouait en raison des garanties demandées par les autorités de la concurrence européennes. En 2004 Springer absorbait *Kluwer Academic*. En 2006 Wiley mettait la main sur le britannique *Blackwell Publishing*. Très régulièrement les analystes financiers (en particulier ceux d'*Exane Paribas*³⁷ pronostiquent la fusion probable de *Thomson Reuters* et de *Wolters Kluwer*, un « perfect fit » qui rapprocherait les activités d'édition juridique de ces deux groupes, mais aussi qui permettrait à *Thomson Reuters* de muscler son pôle *Intellectual Property & Science* qui ne représente aujourd'hui que 8% de ses revenus consolidés et 10% de sa marge opérationnelle. Le groupe britannique *Informa*, qui détient une pépite, le groupe d'édition scientifique *Taylor & Francis*, est parfaitement opérable et pourrait intéresser aussi bien *Reed Elsevier* que *Wiley*.

Concentrations : l'effet de taille des éditeurs scientifiques est décisif à l'heure de la « science numérique »

Si la motivation d'une opération de fusion-acquisition majeure est toujours in-fine financière, *ce sont bien des impératifs stratégiques qui vont motiver les prochaines opérations de fusion-acquisition*. Les éditeurs scientifiques sont en effets confrontés à deux impératifs et le facteur « taille » est important pour y répondre. D'une part, on l'a dit ci-dessus, les éditeurs sont en train d'effectuer la conversion de leur modèle économique au modèle de « Gold Open Access ». Etaler le calendrier de ces mutations sur une base éditoriale large permet de piloter finement l'impact de cette mutation sur le compte d'exploitation des éditeurs.

L'autre motivation sous-jacente à ces futures concentrations est explicitée dans le communiqué de *Reed Elsevier* : ces grands groupes éditoriaux vont muter vers un profil de sociétés associant technologie, contenus et services analytiques. L'une des raisons du rapprochement de *Springer* et de *Macmillan Science & Education/Nature Publishing Group* tient au fait que ces deux enseignes à forte notoriété n'ont pas pour des raisons différentes investi dans des plates-formes multiservices comparables à *Scopus d'Elsevier* et au *Web of Science* de *Thomson Reuters*.

Les investissements lourds amortis sur des positions de marché élargies

Or ces plates-formes seront à la base de la nouvelle proposition de valeur de ce que l'on désigne encore du terme d'« éditeurs ». *Les investissements importants qui seront nécessaires à cette mutation afficheront une meilleure rentabilité s'ils sont amortis sur des positions de marché élargies*. Or l'écosystème de la recherche a pour ces grands groupes un avantage certain : celui d'être d'ores et déjà largement globalisé, et donc de permettre la rentabilisation des investissements à l'échelle planétaire. La concentration à venir permettra aux quelques groupes capables d'acquérir une taille critique suffisante (supérieure à 1 Md. € de revenus) non seulement de consentir ces investissements importants sans éroder leur rentabilité élevée, mais surtout de se détacher définitivement de la « longue traîne » d'éditeurs de faible taille avec qui ils partageaient pourtant jusqu'à hier un même métier.

11 - Les fortes positions des éditeurs « for profit » laissent cependant la place à des initiatives innovantes

Tout ce qui précède ne doit pas conduire à une vision déterministe du scénario de l'édition scientifique. Les problématiques de la communication scientifique sont suffisamment riches pour qu'il existe un espace de

³⁷ <http://deweybstrategic.blogspot.fr/2011/12/thomson-reuterswolters-kluwer-merger.html>

développement en dehors de l'emprise croissante des grands éditeurs : il existe encore de nombreux éditeurs de faible taille (on peut citer en France) qui vivent et se développent de façon satisfaisante même (ou parce que) ils se contentent de taux de marge opérationnelle inférieurs ou égaux à 15%.

De jeunes maisons d'édition en Open Access (*Hindawi, Frontiers, EJPress*, etc) se sont créées dans les dix dernières années et ont prospéré (avant souvent d'être rachetés par les grands éditeurs).

Des plates-formes de plus faible taille peuvent être développées et rencontrer la faveur des usagers : c'est le cas en France, dans le champ des SHS, avec les plates-formes *CAIRN* et *Revue.org*.

Des services innovants comme *Faculty of Thousands* (offrant une valeur ajoutée de sélection et de hiérarchisation du meilleur de la publication scientifique) sont nés et se sont développés par leurs propres moyens.

Outre-Atlantique des initiatives comme *Hathitrust* ou *Highwire Press* démontrent qu'il est possible de construire de vastes corpus scientifiques en dehors du périmètre des grands éditeurs.

Les innovations de process dans la gestion des tâches documentaires ne sont pas nées au sein des grands groupes (même si *Mendeley, Papers, ReadCubed, FigShare*, etc. ont rejoint le giron de grands éditeurs. Une attention plus grande à l'innovation dans les tâches IST et un soutien financier public au meilleur de ces innovations peut empêcher la captation de celle-ci par les très grands éditeurs pour qu'elles puissent se développer dans une logique d'Open Science.

Sur un autre registre le développement de grands agrégateurs tels qu'*EBSCO* ou *ProQuest* rend quelquefois une mise en concurrence partielle des grands éditeurs.

Développées dans le cadre de politiques publiques de longues haleines et bien financées, des initiatives ouvertes peuvent prospérer et devenir des infrastructures de science ouverte. Nées aux USA, *Pubmed* et *PubChem* développés par les *National Institutes of Health (NIH)* sont devenus des pièces majeures du système global de la recherche biomédicale. Le portail hispano-lusitanien *Scielo* irrigue très largement le tissu de l'ESR au Brésil et en Amérique latine et a redonné à la publication scientifique de ces pays une large visibilité dans le cadre d'un modèle d'Open Science.

Rappelons aussi (pour contrer tout pessimisme) qu'Internet et le World Wide Web qui de facto sont des infrastructures essentielles d'un hypothétique « *World Wide Web of Open Science* » se sont développés dans les milieux de la recherche grâce à des logiques collaboratives non marchandes.

En conclusion : les pouvoirs publics doivent promouvoir une Science Ouverte

La montée en puissance des logiques de plates-formes multifonctionnelles, l'opacité des structures de coûts de ces dernières, la capacité d'innovation des très grands éditeurs, leur puissance financière, leur contrôle d'une part croissante des flux de publications scientifiques : *tous ces facteurs risquent à terme de déséquilibrer l'équilibre des forces entre ces fournisseurs et les acteurs de la recherche publique et de l'enseignement supérieur au détriment de ces derniers*. La mutualisation des achats de ressources et de services au sein de l'ESR ne suffira pas à corriger ce déséquilibre. *Il est nécessaire que les pouvoirs publics prennent conscience de la nécessité d'une action de longue haleine visant à promouvoir une « science ouverte » se définissant par l' « Open Access » aux ressources documentaires scientifiques et par l' « Open Process » c'est-à-dire la prise en charge du workflow des tâches IST par des outils non propriétaires.*

Dans cette perspective plusieurs chantiers semblent prioritaires :

La construction des référentiels directeurs d'identification :

La production de métadonnées d'identification (des publications, des chercheurs, des entités de recherche, des sources de financement, des projets et programmes) au sein de référentiels partagés est une question centrale pour la problématique d' « *Open Science* ». Des métadonnées « propriétaires » (sur lesquelles leurs producteurs détiennent des droits de propriété intellectuelle) sont incompatibles avec le développement de l'Open Science. Ces référentiels relèvent de *la notion d'infrastructure immatérielle de l'Open Science*. On peut regrouper tous ces référentiels sous le vocable de « données pivot ». En effet la disponibilité, la qualité de ces données conditionnent l'exploitation des contenus IST des plates-formes *Open Science*, leur *interopérabilité* et la « valeur » *dégagée pour les chercheurs et l'ESR* par ces outils. L'Open Science étant par construction internationale, se pose aussi la question du champ géographique de ces référentiels. Les logiques de coopération internationale sur ces chantiers s'imposent, même si elles ne conditionnent pas le développement d'une réflexion nationale sur ce thème.

La production collaborative d'ontologies ouvertes

Ce thème est assez similaire (en ce qu'il relève de la notion de données-pivot) à celle des référentiels d'identification évoqués au point précédent. L'enjeu fonctionnel est cependant différent. *Les ontologies seront au cœur du web sémantique et des technologies de Text Mining. C'est du langage commun de la science qu'il s'agit*. Dans une perspective de science ouverte, la multiplication d'ontologies « propriétaires » serait hautement problématique. La mise à disposition d'ontologies ouvertes doit prendre en compte la spécificité des champs disciplinaires (alors que la problématique des données d'identification est complètement transversale). Le champ potentiel est immense : des priorités thématiques argumentées devraient être établies.

La redéfinition du droit de la propriété intellectuelle applicable aux objets de science

La question des droits de propriété intellectuelle sur les objets de science (publications, métadonnées, jeux de données, API, algorithmes) est centrale dans une perspective d'Open Science : elle sous-tend les modèles économiques, elle impacte les usages, elle régule la collaboration scientifique. Les tensions entre la sphère ESR publique et les opérateurs privés s'enracinent souvent dans la question du régime juridique des objets de science. La délégation de « copyright » aux éditeurs est par nature antinomique avec la notion d'Open Science. Ce problème est transposable à la question centrale des métadonnées (métadonnées « propriétaires » vs. métadonnées ouvertes). *L'élaboration de contrats types relevant des Creative Commons encadrant les rapports entre éditeurs et les chercheurs, la définition de politiques normatives interdisant le transfert gratuit des droits de propriété intellectuelle aux opérateurs for-profit semblent des priorités*.

L'expérimentation systématique et l'échange de bonnes pratiques pour mettre au point des outils publics partagés de Text et Data Mining

Les technologies TDM (fouille des données et des textes) seront centrales à moyen terme dans l'exploitation des plates-formes d'Open Science et de la connaissance scientifique accumulée. Alors que les technologies de « l'Information Retrieval » dont sont issues les techniques du « Search » ont pris naissance dans les années 70 dans le secteur public (action de la Nasa, du DoE, du CEA), les technologies du Text & Data Mining dans leur dimension applicative sont aujourd'hui assez fermées car les développements les plus aboutis touchent soit au monde du renseignement (NSA...) soit à des projets internes de grandes entreprises (dont les grands éditeurs). Il est nécessaire que des outils ouverts soient mis à la disposition de la communauté de recherche, que des projets expérimentaux soient soutenus que soient mises en place des échanges de bonnes pratiques au sein de la recherche publique.

Une réflexion sur les outils de « Search » et la mise au point de moteurs de recherches indépendants

L'émergence des outils de TDM ne doit pas faire oublier que les technologies de type « moteur de recherche », même s'ils nous sont assez familiers restent une brique technologique essentielle de l'exploitation des ressources de science, dans une logique de recherche fédérée, ignorant le cloisonnement des « silos » d'articles scientifiques mis en place par les grands éditeurs. Le développement rapide des archives ouvertes rend cruciale cette question de l'existence de moteurs de recherche fédérée. Il ne serait pas acceptable que *Google Scholar* reste le seul moteur de recherche fédérée sur les contenus de science issus de la recherche publique.

Conclusion :

Le scénario de l'édition scientifique à l'horizon 2020, tout proche, sera à coup sûr marqué par le prolongement des tendances qui laissent entrevoir la constitution de positions dominantes par les très grands éditeurs internationaux « For-Profit » qui ont vocation à devenir des acteurs majeurs de la science numérique.

Mais il sera aussi façonné par la plus ou moins grande vigueur des initiatives publiques visant à promouvoir des logiques de science ouverte. Ceci suppose une réflexion stratégique accrue et partagée au sein de toutes les composantes de l'ESR. Cela passe par des solutions organisationnelles nouvelles et par une meilleure coordination de tous les organismes de recherche publique.

In fine la possibilité d'action des pouvoirs publics, outre l'élaboration de législations pertinentes, est liée au fait que le financement global de la communication scientifique repose entièrement sur les budgets publics. Cet argument doit être utilisé dans le rééquilibrage des positions respectives des organismes publics et des opérateurs privés.

Ce jeu ne se fait cependant pas à armes égales : les grands éditeurs disposent de capacités d'élaboration stratégique importantes, de possibilités d'exploration de l'innovation et des moyens de financiers de sa captation. Ils disposent aussi de savoir-faire marketing, de services commerciaux étoffés, d'une expérience globale de la négociation. Ils profitent de leurs positions globales pour aborder des situations et des marchés nationaux fractionnés entre lesquels la coopération internationale est insuffisante et doit être renforcée.