



Date : 21/07/2006

Les bases de données scientifiques en libre accès :  
un survol du paysage

Allocution d'ouverture de la section Asie et Océanie de l'IFLA

Professeur Peter Jacso  
Département des sciences de l'information et de l'informatique  
Université d'Hawaï, Etats-Unis

*Traduction et adaptation :  
Robert Bilodeau, bibliothécaire  
Service du catalogage  
Université du Québec à Montréal, Canada*

<b>Meeting:</b>	<b>157 Asie et Océanie</b>
-----------------	----------------------------

<b>Simultaneous Interpretation:</b>	-
-------------------------------------	---

WORLD LIBRARY AND INFORMATION CONGRESS: 72ND IFLA GENERAL CONFERENCE AND COUNCIL  
20-24 August 2006, Seoul, Korea  
<http://www.ifla.org/IV/ifla72/index.htm>

### Résumé

Le concept et le mouvement du libre accès ne représentent pas un nouveau phénomène, mais ce n'est que tout récemment qu'ils ont acquis un intérêt international tant dans les pays développés qu'en voie de développement. L'occasion est propice pour la section Asie et Océanie de l'Ifla de tenir une session sur le thème « Libre accès : en promouvoir l'implantation en Asie et en Océanie ». L'infrastructure semble être disponible, abordable et réalisable, même pour les pays les moins développés, avec l'aide d'organismes internationaux .

Je vous présenterai ici un survol du paysage formé par les bases de données scientifiques partiellement en libre accès avec une emphase sur les revues et les congrès, lesquels représentent les plus importantes sources dans le domaine des sciences, des techniques, de la médecine et de plus en plus en sciences sociales. La gamme de ces bases s'étend des versions en libre accès de bases bibliographiques commerciales ou non comprenant des résumés aux bases de données en libre accès contenant des millions d'articles en texte intégral de revues scientifiques et de comptes rendus de congrès, offrant de nouvelles fonctionnalités logicielles permettant de maximiser le repérage des ressources.

Entre ces deux extrémités, il existe de nombreuses archives d'éditeurs pour les abonnés aux versions imprimées des revues et des actes de congrès, offrant un libre accès à des dizaines de millions de résumés, éparpillés dans les archives numériques d'éditeurs scientifiques ou de «facilitateurs numériques».

La croissance rapide de réservoirs et de dépôts interdisciplinaires ou spécialisés fournit une autre avenue pour le libre accès en facilitant l'auto-archivage ou le dépôt de versions « *(p)reprint* »<sup>1</sup> d'articles de recherche par les auteurs ou par des groupes et entrepreneurs qui en font l'agrégation et l'intégration à l'intérieur de dépôts disciplinaires et multidisciplinaires. Le libre accès à la documentation scientifique est de première importance pour la recherche, le développement et l'éducation, non seulement dans les pays à faible revenu mais également dans les pays plus développés où l'accès complet aux publications scientifiques se situe aussi au-delà des moyens financiers de plusieurs bibliothèques et lecteurs. L'emphase sera mise sur les bases de données en libre accès qui fonctionnent bien et qui peuvent servir de modèle pour la région de l'Asie et de l'Océanie.

## INTRODUCTION

L'implantation intensive et à grande échelle des bases de données scientifiques d'archives numériques, de dépôts d'articles de revues, de comptes rendus de congrès, de rapports de recherche et de dissertations (thèses) à accès libre sur l'Internet représente la meilleure option qui soit si l'on veut faciliter l'accessibilité à une information de haute qualité, structurée et repérable en totalité.

Le mouvement et le concept de libre accès (*Open access*) ne représentent pas un nouveau phénomène, mais ils n'ont atteint que tout récemment un haut niveau d'intérêt et de faisabilité tant pour les pays développés que ceux en voie de développement. Le moment est propice pour la section de l'Asie et de l'Océanie de l'Ifla de tenir une session ad hoc au présent congrès sur le thème de « Libre accès : promouvoir l'implantation en Asie et en Océanie ». L'infrastructure semble être disponible en vue d'accélérer la mise en œuvre du libre accès à la documentation scientifique. L'implantation pourra mieux se réaliser par l'intermédiaire des bibliothèques universitaires et collégiales, grâce à la formation des bibliothécaires et des usagers sur les tenants et aboutissants du libre accès et grâce au déploiement d'installations à large bande. Subbiah Arunachalam insistait sur de telles actions il y a déjà trois ans, en présentait les initiatives et en proposait d'autres (1).

À titre d'ancien consultant de programmes de l'*Unesco* et de l'*Unido* (*United Nations industrial development organization*), je sais d'expérience qu'il y a eu des projets qui ont très bien réussi dans le domaine, généreusement appuyés par des organisations internationales et régionales avec le déploiement à grande échelle de logiciels en accès libre, notamment la distribution du *CDS/ISIS* de l'*Unesco*, l'un des logiciels les plus performants de gestion bibliographique. À leur tour, les programmes éducationnels et de formation ont rendu possible l'implantation locale de bases de données bibliographiques en texte intégral pertinentes à l'échelle régionale ou locale.

De plus, l'attrait et la libre disponibilité du logiciel ont permis, notamment au Brésil, de produire plusieurs améliorations importantes ainsi que le développement de nouvelles versions pour l'environnement du Web et pour différents systèmes d'exploitation. Ce n'est pas une surprise que le Brésil soit à l'avant-garde de l'implantation de la version numérique en libre accès des revues non seulement brésiliennes mais également des revues scientifiques imprimés au-delà de ses frontières.

Il existe maintenant des logiciels dotés d'excellentes caractéristiques qui facilitent le processus d'auto-archivage et la création de bibliothèques numériques multimédias et textuels sans nécessiter un trop fort apprentissage. Les deux applications les plus importantes sont le programme *Eprint* de l'Université Southampton en Grande-Bretagne (3) et le logiciel *Greenstone Digital Library* (4) de l'université Waikato en Nouvelle-Zélande.

Il y a eu de nombreux documents scientifiques gratuits sur le Web depuis plus d'une décennie mais la plupart n'étaient pas facilement repérables ou n'étaient gratuites que pour une catégorie d'usagers. La plupart du temps, ils étaient maintenus grâce aux contributions directes et indirectes des usagers, sous forme de frais d'inscription ou de taxes.

Je vous présenterai ici un survol du paysage formé par les nombreuses bases de données scientifiques en libre accès. Leur registre va des bases bibliographiques contenant des résumés (facilitant le repérage des documents pertinents mais dont la structure tarifaire dépasse les moyens de nombreux établissements de recherche ou d'enseignement des pays en voie de développement) aux bases en texte intégral en libre accès total, contenant

---

<sup>1</sup> N.D.T. Construction difficilement traduisible : *preprint* signifiant publication préliminaire et *reprint* reproduction de l'article.

des millions d'articles de revues scientifiques et de comptes rendus de congrès. Les documents primaires ont été inaccessibles à de nombreux chercheurs même dans les pays développés. Il en alla de même de la livraison de documents. La solution réside dans l'élargissement de l'accès.

Le concept de libre accès est bien défini et expliqué par Peter Suber, l'un des apôtres les plus actifs du mouvement (5). Steve Harnad (6), quant à lui, tel un évangéliste, ne ménage aucun effort pour expliquer de façon intelligente et convaincante, tout en fournissant des faits et des chiffres, les bénéfices du libre accès et les nombreuses options possibles pour l'implantation. Par conséquent, je passerai ce volet sous silence.

Qu'il me suffise de mentionner ici que le libre accès s'avère supérieur à l'accès gratuit. Dans mon esprit, cela signifie que la ressource numérique (ou une des ses parties) bien structurée devient disponible pour tous indépendamment du lieu et le moment, et ce, sans frais. Cela ne signifie pas qu'il ne faille pas s'inscrire et s'enregistrer ou que l'on puisse imprimer ou sauvegarder électroniquement toute l'information souhaitée de la collection. Mais le document peut être repéré ou consulter et certains passages pourront être affichés.

Une grande partie des archives numériques des éditeurs permettent le libre accès à l'information bibliographique et aux résumés des articles publiés, disons, à l'intérieur d'une décennie mais non pour des articles publiés antérieurement ou très récemment. Pour les restrictions fonctionnelles, prenons l'exemple du service de la *Ebrary*<sup>2</sup>. Celui-ci permet l'affichage en libre accès du contenu de 35 000 livres et les frais ne s'appliquent qu'à l'impression et à la sauvegarde des pages sélectionnées.

Quand on observe le paysage des bases de données scientifiques, on s'aperçoit que les sources en libre accès ne cessent d'émerger et de se développer. Elles sont particulièrement présentes dans le domaine des sciences, de la médecine et de la technologie. Mais il en existe également dans le domaine des sciences sociales, des arts et des lettres. La prédominance des revues de différentes disciplines en libre accès constitue une excellente nouvelle si l'on considère les statistiques largement acceptées indiquant qu'à l'intérieur du budget des périodiques d'une bibliothèque collégiale typique, les revues du domaine scientifique, 30% de l'ensemble, accaparent 70% des frais d'abonnement.

### QUI MET LES SOURCES EN LIBRE ACCÈS?

L'une des raisons de la croissance généralisée des bases de données en libre accès tient au fait que les éditeurs, les organisations sans but lucratif et les presses universitaires, directement ou indirectement par leurs «facilitateurs numériques<sup>3</sup>», ont contribué à la croissance continue du corpus scientifiques en libre accès.

Des agences gouvernementales et des organisations non gouvernementales (*ONG*), tels le *Department of Labor* (États-Unis), la Banque mondiale, le Fonds monétaire international (*FM*), figurent parmi les plus importants contributeurs aux bases de données factuelles, statistiques ou de répertoires. Elles sont également, notamment les agences des Nations Unies de même que des fondations et organisations caritatives, derrière des bases de données bibliographiques ou partiellement en texte intégral très connues.

Des éditeurs commerciaux de bases de résumés et d'index ainsi que des agrégateurs de bases de données ont également sauté dans le train du libre accès. Par exemple, EBSCO qui a offert les bases de données *Teacher's Resource Center* et *Library Resource Center* pendant plusieurs années et qui, récemment, a fait de cette dernière une base beaucoup plus considérable en libre accès, la *Library and Information Science & Technology Abstracts* (LISTA). La *National Information Services Corporation* (*NISC*), qu'on pourrait classer parmi les organismes sans but lucratif en raison de son attitude philanthropique, a offert le libre accès et des frais de

---

<sup>2</sup> N.D.T. Ce service devrait révolutionner les méthodes de recherche et d'acquisition d'informations sur le Web en associant un moteur de recherche puissant, une consultation en texte intégral et des outils de référence de pointe constitués de livres, périodiques, cartes et travaux d'archives. *Ebrary* assure pouvoir éviter le piratage et le "photocopillage" grâce à des procédés de cryptage destinés à bloquer l'impression ou le téléchargement de textes à moins d'un paiement préalable. De grands ténors de l'édition pédagogique américaine, tels *Random House*, *Pearson* et *McGraw-Hill*, ont investi dans *Ebrary*. Voir <http://urfiist.univ-lyon1.fr/risi/32-modes.html>

<sup>3</sup> N.D.T. Traduction libre de *digital facilitators*, terme englobant des organisations qui assistent des éditeurs dans la numérisation de leurs publications pour le Web. De plus, des services supplémentaires peuvent s'ajouter, tels l'hébergement des textes numériques, l'identification des abonnés et des visiteurs. Quelquefois, ils offrent un service de résumés des documents primaires. On songe notamment à *MetaPress*, *Ingenta Select*, *HighWire Press*. Voir Peter JACSO, 2003 : *Digital Facilitators*. *Information Today*, 20 (6) <http://www.infotoday.com/it/jun03/jacso.shtml>

redevance réduits aux pays moins bien nantis une décennie avant que l'expression «libre accès» ne soit consacré. La NISC a rendu disponibles et exemptes de tous frais la consultation les bases *AIDSearch*, *Child Abuse*, *Child Welfare* et *Adoption* pour les personnes inscrites.

Un peu partout dans le monde, des universités commencent à créer de vastes dépôts numériques institutionnels chargés d'héberger les pré-publications et les reproductions des publications de leurs départements et de leurs doctorants.

Des groupes de volontaires et personnes entreprenantes ont le grand mérite d'avoir lancé d'importantes revues en libre accès, *First Monday*, *Ariadne* par exemple, ou des archives de pré-publications et de documents électroniques, notamment *e-LIS*, *dLIST* et *Dols* en sciences de l'information.

### CE QUI EST DISPONIBLE EN LIBRE ACCÈS

Il est plus facile de répondre à cette question en indiquant ce qui n'est pas consultable en libre accès. Pratiquement, chaque type d'information scientifique traditionnel est disponible numériquement de la sorte. Le *Registry of Open Access Repositories (ROAR)* de l'université Southampton, l'*Experimental OAI Registry* de l'université d'Illinois, le *Directory of Open Access Repositories (DOAR)* créé par l'université de Lund [Suède] et l'université de Nottingham possèdent l'information à jour sur les nombreux dépôts et ils fournissent des outils pratiques de repérage des bases de données scientifiques en libre accès. Il existe beaucoup de recoupements entre eux. *ROAR* offre une précieuse caractéristique : on y dresse le tableau de la croissance de plusieurs dépôts.

Le *Directory of Open Access Journals (DOAJ)*, publié par l'université de Lund, et *OAIster*, géré par l'université du Michigan, fournissent également un bon aperçu de la variété et de l'importance de plus de 2 200 revues et de près de 650 dépôts d'archives numériques mis à disposition en libre service. Tous deux ont un important atout : ils permettent le repérage par document ou par pièce. Le *DOAJ* aligne 96 000 articles et *OAIster*, 7,5 millions de notices. Ces chiffres devraient toutefois être pris avec prudence, notamment pour *OAIster*, puisque les mêmes documents peuvent apparaître à l'intérieur de plusieurs collections. Chez *DOAJ*, le nombre d'articles est nettement inférieur à ceux offerts par les revues en libre accès, car 600 revues seulement permettent la recherche par article. En conséquence, ceux de 1 600 autres revues ne sont pas recensés.

Il existe plusieurs catégories de collections numériques en libre accès. Outre les bases de données textuelles d'articles de revues et de présentations à des congrès, on a les monographies, les dictionnaires, les encyclopédies, les répertoires de bases de données, les bases de données numériques, les bases de données d'images, audio et vidéo. Du côté des bases de données textuelles, on rencontre des collections bibliographiques numériques (sans descripteurs), des bases de données d'indexation, des bases de données de résumés et d'indexation et des bases en texte intégral. Kirsop et Chan (7) ont présenté récemment un excellent aperçu des bases de données de revues en libre accès dans les pays en développement. Le rythme de croissance des corpus est si rapide qu'une mise à jour de l'information s'imposait pour cette présentation, comme vous le constaterez en dernière partie.

### BASES DE DONNÉES BIBLIOGRAPHIQUES ET DE RÉSUMÉS

De nos jours évidemment, les bases de données en texte intégral sont les plus recherchées. Afin d'apprécier l'importance, par exemple, des bases de données d'indexation en libre accès, il faut se rappeler que, jusqu'à la fin des années 1980, elles représentaient le type le plus courant de base de données de référence rapide pour lesquelles les bibliothèques versaient des dizaines de milliers de dollars en redevances.

La vénérable compagnie H.W. Wilson, par exemple, n'a possédé que des bases de données d'indexation (à l'exception de *Reader's Guide Abstracts*) pendant de nombreuses années. Ce n'est qu'au début des années 1990 que *Agricultural Index*, *Education Index*, *Social Science Index*, *Applied Science & Technology Index*, etc., furent enrichies de résumés et devinrent *Agricultural Abstracts*, *Social Science Abstracts* et *Applied Science & Technology Abstracts*. Elles ne sont pas, et c'est compréhensible, en libre accès puisque des fournisseurs de

contenus, tel H.W. Wilson, tirent leurs profits (et investissent beaucoup d'argent) de la création de ces bases de données.

Toutefois, au-delà des bases de données d'indexation et de résumés bien connues produites par des agences gouvernementales, notamment *Medline* (de même que plusieurs autres bases de données de la *National Institute of Health*), *ERIC*, *AGRICOLA*, *NTIS*, *Energy Science & Technology*, il existe d'autres bases de données en libre accès : la base de données *ASCE Civil Engineering*, la base, au nom curieux mais combien remarquable, *PILOTS* du *National Center for Post-Traumatic Stress Disorder*.

Plus remarquable encore demeure le fait que les plus importants éditeurs commerciaux et d'associations de revues scientifiques, de comptes rendus de congrès et de monographies, tels l'*American Chemical Society*, *AMA* [*American Medical Association*], *ACM* [*Association for Computing Machinery*], *BCS* [*British Computing Society*], *IEEE* [*Institute of Electrical and Electronics Engineers*], *IEE* [*Institution of Electrical Engineers*], *IoP* [*Institute of Physics*], *APS* [*American Physical Society*], Elsevier, Springer (incluant Kluwer), *Taylor & Francis*, *John Wiley & Sons*, *Blackwell*, *SAGE*, *Oxford University Press*, *Cambridge University Press*, *Emerald Press* ainsi que plusieurs petits éditeurs offrent de concert un libre accès à des dizaines de millions de notices bibliographiques et de résumés d'articles de revues scientifiques et de présentations à des congrès. Certains le font directement tandis que d'autres le font par l'intermédiaire de leurs «facilitateurs numériques», comme *SAGE*, *Oxford University Press* avec *HighWire Press*, *Springer*, *Taylor & Francis* avec *MetaPress*, *Palgrave Macmillan* et d'autres avec *IngentaConnect*. Des éditeurs passent d'un «facilitateur» à un autre ou confient certaines revues à l'un d'eux et le reste à un autre. La *HighWire Press* à elle seule offre plus de 3,5 millions de notices en libre accès, dont environ deux millions contiennent des résumés et 1,3 millions d'articles en texte intégral de revues ayant un rayonnement des plus élevés.

Le service *Scirus* d'*Elsevier* offre le libre accès aux notices bibliographiques et aux résumés de plus de 6,5 millions d'articles publiés dans les 1 700 revues d'*Elsevier*. On y trouve également près de 20 millions de notices d'indexation et de résumés provenant de *PubMed* ainsi que des archives de revues d'associations telles celles de *IoP* et d'*API*. La recherche y est possible et le service comprend des notices complètes des revues en libre accès de *PubMed Central* et de *BioMed Central* de même que de quelques parapublications et de publications électroniques. J'ai abordé le sujet dans un article (8).

Bien que le projet *African Journals Online* (*AJOL*) se situe à une plus petite échelle que les bases de données mentionnées jusqu'ici, il possède, dans la perspective d'un soutien à l'implantation de bases de données en libre accès dans des pays et des régions à faible revenu, la possibilité de devenir un modèle. (Curieusement, il existe une autre base portant le même nom : la *Australian Journals Online*, un répertoire traditionnel des revues australiennes). *AJOL* contient de l'information bibliographique et des résumés d'articles publiés dans 230 revues africaines, de l'Algérie (2 revues) au Zimbabwe (6). Le Nigéria domine avec près de 100 revues couvertes par *AJOL*, suivi ensuite de l'Afrique du Sud (47 revues).

L'aspect le plus intéressant d'*AJOL* demeure le fait que les pays à bas revenus, Burkina Faso, Lesotho, Malawi et Soudan entre autres, y participent et profitent de la sorte d'un important élan de départ. Il existe des notices d'environ 20 000 articles dans *AJOL*. On peut commander la livraison de documents à la fin de la recherche par le logiciel. Les articles sont acheminés sans frais dans plus de 100 pays. Considérant les coûts et les limites de la livraison d'un document imprimé, il devient évident que la prochaine étape devrait être la numérisation et la reconnaissance optique de caractères<sup>4</sup> des articles les plus souvent demandés. La livraison commerciale de documents imprimés peut s'avérer financièrement prohibitive, peu importe le bailleur de fonds. De plus, la numérisation d'articles offre l'énorme avantage des recherches dans le texte intégral. Elle représenterait la plus importante phase à venir d'*AJOL*.

L'*AJOL* pourrait servir de modèle pour plusieurs pays de l'Asie et de l'Océanie. La création de cette base de données a été conçue et gérée par l'*International Network for the Availability of Scientific Publications* (*INASP*), une organisation caritative basée à Oxford et financée par plusieurs organisations internationales. On ne se

---

<sup>4</sup> NDT Traduction libre de *OCR-ing* : *OCR* est l'abréviation de *Optical character recognition* et est rendu par *reconnaissance optique de caractères*.

surprendra pas que le système soit actuellement administré par le subsidiaire africain de l'*IVASP*, lequel possède une grande expérience de l'automatisation des services d'information dans les pays à faible revenu.

L'un des programmes les plus importants de l'*IVASP* à l'intention des pays les moins développés de la planète est le *PERI (Programme for the Enhancement of Research Information)*. Ce programme a commandité et géré plusieurs entreprises similaires à l'*AJOL* à l'échelle nationale en Asie centrale et en Amérique latine. Les projets visaient l'amélioration de l'accès à la recherche par le soutien à l'édition de revues scientifiques locales et la création de bases de données d'indexation et de résumés. Les efforts ont permis d'améliorer la visibilité des résultats de recherches locales et de faciliter le libre accès aux archives en texte intégral des principaux éditeurs de travaux de recherche.

### COLLECTIONS PLEIN TEXTE EN LIBRE ACCÈS

Quelques-uns des éditeurs mentionnés précédemment offrent également le libre accès au texte intégral des archives de leurs revues les plus influentes, dont *Clinical Diabets* de l'*American Diabets Association*, *Nucleic Acid Research* de *Oxford University Press*, le *Journal of the American Osteopathic Association*, le *Journal of Bone Joint & Surgery, CA: A Cancer Journal for Clinicians of the American Cancer Society*, *Canadian Medical Association Journal*, pour n'en nommer que quelques-unes qui sont hébergées par l'excellent «facilitateur numérique» *HighWire Press*. Ces revues ne sont pas importantes en elles-mêmes, mais lorsqu'elles sont regroupées, leur importance augmente rapidement. *HighWire Press* représente le plus grand (ou le deuxième) dépôt d'archives de textes intégraux à impact élevé. Plus de 1,3 millions des 3,5 millions articles sont en libre accès chez *HighWire Press*, lequel possède de nouveaux et superbes mécanismes informatiques. Deux autres «facilitateurs» d'importance offrent le libre accès à des livraisons de quelques revues mais, en quantité, ils n'appartiennent pas à la même catégorie.

Nous devons noter ici que des éditeurs commerciaux ont montré qu'ils pouvaient se montrer généreux lorsque plusieurs d'entre eux ont adhéré à *Health InterNetwork Access to Research Initiative (HINARI)* et à *Access to Global Online Research in Agriculture (AGORA)* (9), deux entreprises qui ont bien réussi pendant de nombreuses années. Mis sur pied par l'Organisation mondiale de la santé (*OMS*) et l'Organisation des Nations Unies pour l'alimentation et l'agriculture (*FAO*), avec une aide efficace des universités Yale et Cornell en gestion de ressources numériques, les deux entreprises ont fourni un accès sans précédent pendant plusieurs années aux archives des meilleures revues en médecine et en agriculture des éditeurs les plus connus. En mai 2006, *HINARI* offrait aux chercheurs, aux scientifiques, aux éducateurs et aux praticiens de plus de 2 000 établissements répartis dans les 68 pays aux plus faibles revenus et un PNB par tête de moins de 1 000 \$US un libre accès aux archives en texte intégral de 3 290 revues du domaine de la médecine, de la biologie et de l'agriculture [[http://extranet.who.int/hinari/en/journalList\\_print.php](http://extranet.who.int/hinari/en/journalList_print.php)] provenant de plus de 60 éditeurs [<http://www.who.int/hinari/partners/en/>]. Quarante-cinq autres pays dont le PNB se situait entre 1 000 et 3 000 \$US paient des frais nominaux de 1 000 \$US par établissement pour obtenir le même privilège. *AGORA* regroupe 35 éditeurs partenaires [<http://www.who.int/hinari/partners/en/>] et offre le libre accès à 849 revues [<http://www.aginternetwork.org/en/journals.php>] aux 68 pays (une douzaine de pays admissibles ne se sont toutefois pas encore inscrits pour bénéficier de l'offre exceptionnelle). Évidemment, cela ne constitue pas un libre accès universel mais, néanmoins, les chercheurs du tiers des pays peuvent en profiter.

La croissance rapide des dépôts numériques constitue une conjoncture favorable au libre accès, à l'auto-archivage, au versement des textes de rodage et des tirés-à-part par les auteurs eux-mêmes. Le regroupement et l'intégration de ces textes dans des réservoirs spécialisés ou interdisciplinaires par des groupes de volontaires ou d'entrepreneurs multiplie l'effet boule de neige.

De façon naturelle, les universités se retrouvent à l'avant-garde du mouvement de libre accès. Les professeurs compétents appliquent ce qu'ils enseignent, ce qu'ils prêchent et ce qu'ils écrivent. Ils lancent des initiatives, des logiciels et des services de soutien (avec l'aide de doctorants et de webmestres) à l'auto-archivage par leurs pairs, à l'identification des dépôts de publications de recherche à libre disposition de façon à en arriver au guichet unique de recherche et de repérage.

Le *Harvard-Smithsonian Center for Astrophysics* (avec d'autres composantes de l'université de Harvard) a mis sur pied et maintenu *ADS (Astrophysics Data System)*, la plus grande base de données d'indexation, de résumés et de textes intégraux. Le système contient environ 4,5 millions de notices dont près de la moitié sont disponibles en texte intégral, en format numérique rapidement repérables et en libre accès. L'université Cornell, depuis longtemps défenseur du mouvement de libre accès et particulièrement dans le domaine de l'agriculture, est présentement engagée dans l'implantation d'une autre version du dépôt en libre accès *arXiv* contenant près de 400 000 publications électroniques en physique, en informatique, en mathématiques et en biologie quantitative.

Des universités aux États-Unis, au Royaume-Uni, en Allemagne, au Japon, en Nouvelle-Zélande, en Australie et aux Pays-Bas créent les plus grands dépôts institutionnels en libre accès de textes intégraux et de matériel multimédias de recherche. Quelques-unes d'entre elles participent au développement d'outils de repérage de sources en libre accès dont nous avons déjà parlé.

Les agences gouvernementales et les ONG au niveau national et international représentent un groupe de première importance parmi les contributeurs de contenu en libre accès. Des organisations et des agences aux États-Unis, au Royaume Uni, au Japon et au Canada offrent les plus grandes collections en texte intégral en libre accès. *PubMed Central*, par sa taille et de sa couverture, est d'une importance primordiale pour les chercheurs des pays à la fois développés et de ceux en voie de développement. Cette base de données comprend près de 250 000 articles en texte intégral en libre accès provenant essentiellement de revues en sciences biologiques. Ce ne sont pas tous les articles qui sont immédiatement en libre accès, quelques uns possèdent un moratoire avant d'atteindre cet état, mais *PubMed Central* constitue néanmoins une collection impressionnante, le meilleur exemple de l'argent bien dépensé des contribuables.

La base de données de *BioMed Central* est plus petite, avec une emphase sur les séries des revues populaires du *BMC* en libre accès, mais mérite notre appréciation puisqu'elle offre à l'utilisateur la même interface à la fois pour *PubMed*, *PMC* et *BMC*. Le Conseil de recherches du Canada offre les versions numériques de 30 000 articles de 16 revues publiés par son *Research Press* au cours des dix dernières années. Il y a de l'information supplémentaire sur les bases de données en texte intégral en libre accès à la référence figurant en (10).

## **LES MEILLEURS MODÈLES DE BASES DE DONNÉES SCIENTIFIQUES EN LIBRE ACCÈS POUR LA RÉGION DE L'ASIE ET DE L'OCÉANIE**

Pour le développement de bases de données d'indexation et de résumés concernant des articles de revues scientifiques et des comptes rendus de congrès publiés par les auteurs d'un pays ou d'une région et/ou dans des revues et comptes rendus publiés dans un pays ou une région, le projet *African Journals Online (AJOL)*, abordé précédemment, peut s'avérer un bon modèle pour les pays à faible revenu. Toutefois, avec le parrainage d'agences, de fondations ou d'organisations caritatives internationales, il peut valoir la peine de considérer d'ignorer cette étape et de se diriger résolument vers la numérisation en texte intégral des articles et des comptes rendus de congrès, étant donné qu'une portion significative des documents primaires sont susceptibles d'être déjà disponible en format numérique, même si ce n'est seulement qu'en format de traitement de texte.

Une approche mixte pour la création de telles bases de données est également praticable. Le meilleur modèle est celui de la *New Zealand Digital Library (NZDL)* Bibliothèque numérique de Nouvelle Zélande) qui a été créée par les développeurs du très populaire logiciel *Greenstone*<sup>5</sup> avec le support financier de l'*UNESCO* et de l'organisation caritative non gouvernementale *Human Info*. L'expérience de ces organisations dans les pays en développement est d'une importance particulière pour démarrer avec succès des ressources numériques. L'utilisation de la *NZDL* comme modèle semble bien fonctionner pour la *African Digital Library Center*. Sa première collection en libre accès a été créée par l'université de Cape Town.

Le Brésil et l'Inde pourraient revêtir un intérêt particulier pour de nombreux pays de la région de l'Asie et de l'Océanie à titre de contributeurs importants pour des archives de revues en texte intégral en libre accès. Ces deux pays se situent respectivement au 71<sup>e</sup> et 129<sup>e</sup> rang en termes de PNB par habitant, mais, relativement au

---

<sup>5</sup> NDT : *Greenstone* est un ensemble logiciel multilingue en *open-source* (logiciel libre) de création et de distribution de collections numériques de bibliothèques. Il fournit une nouvelle façon d'organiser l'information et de l'édition pour Internet ou sur un cédérom.

phénomène du libre accès, ils se situent devant des nations plus riches. Ils peuvent montrer les meilleurs modèles pour la promotion du mouvement du libre accès en Asie et en Océanie par l'intermédiaire d'une coopération régionale et nationale entre les pays possédant les revenus les plus élevés et ceux possédant les revenus les plus bas.

Au Brésil, la majorité des projets de numérisation de revues scientifiques est fortement centralisée à la *FAPESP*<sup>6</sup> et à la *BIREME*<sup>7</sup>. *SciELO*, la *Scientific Electronic Library Online*, en a pris la tête en favorisant le libre accès en mettant gratuitement à disposition, il y a près de 10 ans, une douzaine de revues brésiliennes. Ce fut le prolongement naturel des premières contributions significatives de la *BIREME*, à partir des années 70 : création de plusieurs bases de données d'indexation et de résumés des publications des chercheurs de l'Amérique latine et des Caraïbes, adaptations et améliorations substantielles du très populaire et très puissant logiciel *open source CDS/ISIS* de l'*Unesco*.



**FIGURE 1: Résultats de recherche sur SciELO**  
(N.D.T. L'image diffère légèrement de celle du texte original)

En date de mai 2006, mes vérifications indiquaient qu'il y avait près de 57 000 articles, tirés de 150 revues brésiliennes, en texte intégral dans *SciELO*. La *BIREME* a étendu sa portée en hébergeant sur *SciELO* des revues publiés au Chili (53 revues), en Espagne (28), à Cuba (18) et au Venezuela (6). Des revues éditées au Pérou, au Mexique, en Argentine et en Uruguay pourront être bientôt disponibles dans *SciELO*, lequel utilise le logiciel en libre accès *WWW\_-SIS*, conçu par la *BIREME*. Ce projet a mérité au Brésil une reconnaissance régionale pour sa performance au titre de la technologie de l'information de même qu'à celui de l'intégration et de la gestion d'un projet de base de données de revues en libre accès hautement efficace et substantiel. Évidemment, la présence de l'Espagne permet à ce programme de dépasser le cadre régional.

Le programme, grâce au libre accès des textes archivés de *e-publications* de *Bioline International* hébergées par le *CRIA*<sup>8</sup> a contribué à la réputation internationale du Brésil. *SciELO* a développé le logiciel de la collection en libre accès de 56 revues de recherche éditées au Bangladesh, au Brésil, en Chine, en Colombie, en Croatie, en Égypte, en Inde, au Kenya, en Malaisie, au Nigeria, au Sénégal, en Turquie, en Ouganda et au Venezuela [ <http://www.bioline.org.br/journals> ]. L'université de Toronto apparaît parmi les partenaires de l'entreprise : elle offre le service de *Bioline* en libre accès avec un logiciel différent. Le nombre d'articles en texte intégral de la

<sup>6</sup> NDT FAPESP : Fundação de Amparo à Pesquisa do Estado de São Paulo (Fondation de la recherche de l'État de Sao Paulo)

<sup>7</sup> NDT BIREME : Biblioteca regional de medicina (Bibliothèque régionale de médecine, actuellement Centre latino-américain et caraïbe d'information scientifique médicale)

<sup>8</sup> NDT CRIA Centro de Referência em Informação Ambiental (Centre de référence sur l'information environnementale du Brésil )



base est de moins de 3 000 mais elle représente une solution pratique et efficace de coopération entre pays à revenu élevé et pays moins bien dotés qui désirent implanter des archives en libre accès.

En Inde, la centralisation de la numérisation des revues scientifiques sur le Web n'est pas aussi centralisée qu'au Brésil, mais la pratique demeure un bon modèle pour le mouvement du libre accès en Asie et en Océanie. *MedKnow* met en libre accès 20 000 articles provenant de 30 revues publiées en Inde. Ces articles sont disponibles en format HTML ou en PDF et demeurent facilement repérables grâce à un logiciel de recherche intuitif et convivial. Quatorze des trente revues sont également consultables dans les services brésilien ou canadien à *Bioline*. Ces revues sont indexées et résumées par plusieurs services spécialisés tels *MEDLINE*, *EMBASE*, *PsycINFO*, *CAB* et *CINAHL*.

The screenshot shows the MedKnow Publications website interface. At the top, there is a navigation menu with links for 'About Us', 'Services', 'Journals', 'For Librarians', 'For Authors', 'Advertise', and 'Contact Us'. Below the menu, a search bar indicates the search term '[WORD]toxoplasmosis' and shows 'Showing 1 to 10 of 77 results'. A list of search results is displayed, including titles like 'REVIEW ARTICLE: Mother-to-child transmission and diagnosis of toxoplasma gondii infection during pregnancy' and 'Central nervous system toxoplasmosis in acquired immunodeficiency syndrome: An emerging disease...'. To the right of the search results, there is a search form with fields for 'Author' and 'Keyword(s)', a 'Search' button, and promotional banners for 'Subscribe Online Now' and 'Register for Free Newsletter/ToC'.

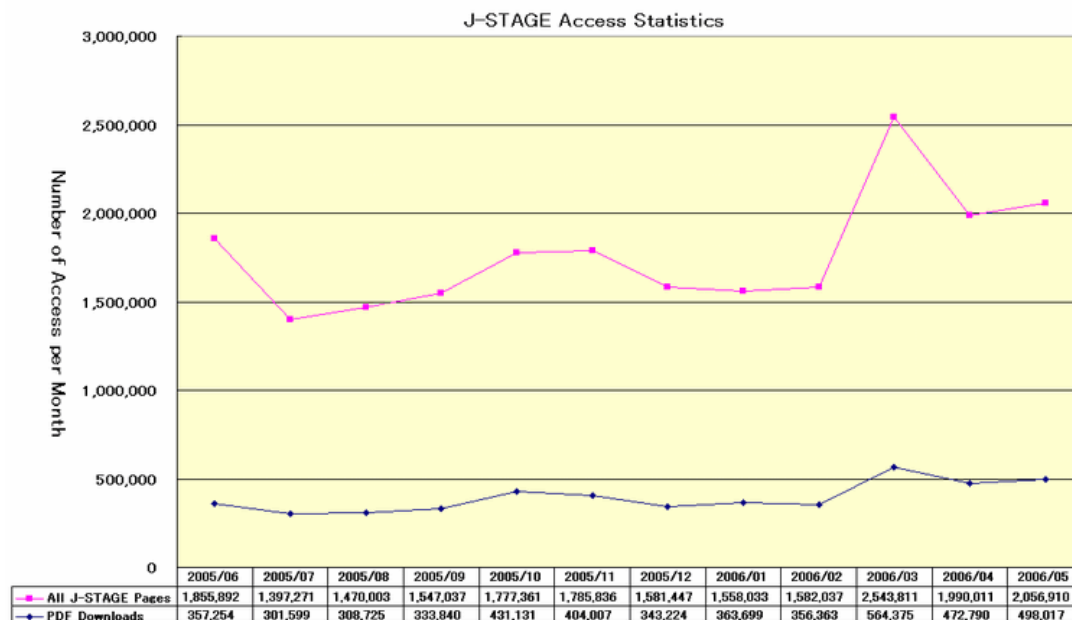
**FIGURE 2 : Résultats de recherche dans la base de données MedKnow**  
(N.D.T. L'image diffère légèrement de celle du texte original)

*MedKnow* est le système le plus perfectionné et contenant le plus grand nombre de revues en libre accès en Inde, mais il existe au moins trois autres dépôts importants de revues en libre accès. La base de données en texte intégral *medIND*, au nom plutôt ambigu et créée à partir des textes de 38 revues biomédicales, est une extension de la base de données originale *indMED* issue du *MEDLARS Center of India*, lequel offrait des notices d'indexation et de résumés de 33 500 articles publiés dans 90 à 95 revues biomédicales indiennes (total variant en raison des changements de titres). Présentement, *medIND* possède une petite collection d'environ 3 200 articles, mais il est probable que les praticiens désirent l'essentiel sur le *toxoplasmosis* préféreront les huit articles en texte intégral de la base plutôt que d'avoir à compulser les 29 références et résumés repérés sur le sujet dans *IndMed*. L'interface de recherche allège toutefois les dilemmes en présentant une case à cocher à ceux qui voudraient limiter la recherche aux articles disponibles en texte intégral.

La *Indian Academy of Sciences* possède des archives en texte intégral de ses onze revues. Les livraisons peuvent être parcourues en ligne, mais l'interface n'est pas unifiée et, plus important encore, il n'y a pas de moteur de recherche. (Naturellement, on peut toujours utiliser *Yahoo* et *Google* pour limiter une requête en utilisant le paramètre de recherche : *ias.ac.in*). L'Académie héberge ses propres publications. Elle offre un libre accès pour une demi-douzaine de revues et de comptes rendus.

Le Japon possède un excellent modèle pour la conversion et la numérisation haut de gamme à grande échelle vers une possibilité partielle de libre accès. La *Japan Science and Technology Agency (JST)* vient d'améliorer sa bibliothèque numérique, laquelle comprend *J-EAST*, une énorme base de données d'indexation et de résumés de textes en langue anglaise de chercheurs japonais publiés dans 3 000 revues, d'actes de congrès et de rapports techniques; *READ*, une base de données (répertoire) des établissements et des scientifiques liés à des activités de recherche et développement au Japon; et *BIRD*, une collection d'outils de recherche et de

compilations de données numériques relatives au développement et à la recherche en bio-informatique. L'Agence a ajouté une soixantaine de revues à sa base de données *J-STAGE* consacrée essentiellement aux revues en libre accès et a lancé la base de données *Journals@archive* formée avant tout des numéros anciens en libre accès de plusieurs revues, le tout totalisant environ 490 publications et 735 000 articles, de même que d'autres documents provenant d'annuaires, d'actes de congrès et de rapports techniques.



**FIGURE 3 : Statistiques de fréquentation pour la base de données J-STAGE**

(NDT : Les données s'échelonnent de juin 2005 à mai 2006, alors que celles du texte original couvraient avril 2005 à mars 2006)

Ce dernier chapitre de *JST* venait de se terminer lorsque j'entrepris la rédaction de mon allocution. Les statistiques de fréquentation diffusées sur le site constituent un message sans équivoque à ceux qui envisagent de promouvoir la visibilité des activités savantes d'un pays ou d'une région par l'implantation de bases de données libre service. Le nombre d'accès, qui se situait à 1,5 million par mois, a grimpé à 2,5 millions en mars après que la base *J-STAGE* ait été enrichie de 75 000 articles, dont la plupart en libre accès.

Ces statistiques démontrent le pouvoir d'attraction du libre accès pour les documents de qualité, elles illustrent admirablement le message de ma conférence : allez et faites-le!

## Références

[1] Information for Research in Developing Countries – Information Technology, a Friend or Foe? Arunachalam, Subbiah (2003) Information for Research in Developing Countries – Information Technology, a Friend or Foe?. Bulletin of the American Society for Information Science and Technology 29(5):pp. 16-21. disponible au <http://dlist.sir.arizona.edu/832/01/Arun%5F01.pdf>

[2] Packer, Abel (2000). SciELO — a Model for Cooperative Electronic Publishing in Developing Countries. D-Lib Magazine,, 6 (10) disponible au . <<http://www.dlib.org/dlib/october00/10inbrief.html#PACKER>>.

[3] Brody, Tim (2001). Mining the Social Life of an ePrint Archive, OpCit Blog disponible au <http://opcit.eprints.org/tdb198/opcit/>

[4] Ian H. Witten, Ian, H., Bainbridge, David and Boddie, Stefan J. Greenstone: Open-Source Digital Library Software (2001). D-Lib Magazine, 7(10) disponible au <http://dx.doi.org/10.1045/october2001-witten>

- [5] Suber, Peter (2005). Open Access Overview. The SPARC Open Access Newsletter, March 10, 2006 disponible au <http://www.earlham.edu/peters/fos/overview.htm>
- [6] Harnad, Steven (2006). Open Access Archivengalism [weblog] disponible au <http://openaccess.eprints.org/>
- [7] Kirsop, Barbara and Leslie Chan (2005). Transforming Access to Research Literature for Developing Countries. *Serials Review* 31(4): 246-255. disponible au <https://tspace.library.utoronto.ca/handle/1807/4416>
- [8] Jacso, Peter (2006). Open Access Scholarly Indexing and Abstracting Databases (2006). *Online Information Review*, 30(4) disponible au <http://www2.hawaii.edu/jacso/savvy-mcb.htm>
- [9] Ochs, Mary, Aronson, Barbara, and Jane Wu (2004), HINARI and AGORA: revolutionizing access to scientific information in the developing world *Serials* 17(2):175-182 disponible au: <http://uksg.metapress.com/openurl.asp?genre=article&id=doi:10.1629/17175>
- [10]. Jacso, Peter (2006). Open Access Scholarly Full Text Databases and Digital Archives (2006). *Online Information Review* 30(5) disponible au <http://www2.hawaii.edu/jacso/savvy-mcb.htm>

### Note biographique

Peter Jacso est professeur au département des sciences de l'information et de l'informatique de l'université d'Hawaï. Il a reçu le *Outstanding Information Science Teacher* de *ASIS&T* et de l'*Institute for Scientific Information*, le *Pratt-Severn Faculty Innovation Award in Library and Information Studies of the Association of Library and Information Science Educators*, et le *Louis-Shores-Oryx Press Award* pour la qualité des comptes rendus des bases de données fournis à la *ALA Reference & User Services Association*. Il est un participant assidu des congrès internationaux et il a été chroniqueur pour de nombreuses revues.