

Le libre accès à l'information scientifique et technique

L'état de l'art

**Rapport général rédigé pour l' INIST par Jack Franklin (ASFRA BV)
- Edam, Pays-Bas (Novembre 2002).**

Le but du présent rapport est de faire ressortir les principaux problèmes liés aux évolutions des technologies du libre accès et à leur adoption dans le monde de l'information scientifique, technique et médicale. L'auteur de ce rapport est seul responsable de son contenu qui ne saurait en aucun cas refléter un positionnement officiel de la part de l'INIST ou de toutes autres parties organisatrices de ce séminaire.

Sommaire

1. Introduction, portée et objectifs du présent rapport
2. L'objectif de l'Initiative pour les Archives Ouvertes
3. Le logiciel Eprints, l'initiative pour le libre accès et le monde de l'édition scientifique et technique
4. L'acquisition et l'évaluation
5. L'interrogation, la localisation et la création de liens
6. L'archivage
7. Les abonnements, les licences et autres obstacles à un accès libre et gratuit
8. Les coûts
9. Le monde en développement
10. L'avenir et ce qu'il réserve

1. Introduction, portée et objectifs du présent rapport

L'objectif des défenseurs du libre accès (*Open Access*) est d'établir des « normes communes selon lesquelles les archives stockées sur des serveurs conformes à ces normes pourraient constituer une bibliothèque mondiale qui serait interrogeable, interopérable et pérenne » (www.openarchives.org). Il s'agit donc d'une initiative technologique qui permettrait d'ouvrir les portes d'une bibliothèque d'informations qui tout en étant à des endroits différents, seraient reliées entre elles.

De nombreux chercheurs, quant à eux, espèrent bien que le Libre Accès, par définition, sera pour eux l'occasion de changer la manière dont les informations sont stockées et accessibles aujourd'hui. C'est pourquoi l'initiative a pris une telle ampleur dans le secteur de l'information scientifique et technique (IST) qui conteste de plus en plus la manière dont les informations (les articles de revues) sont traitées. Certains chercheurs estiment que le libre accès permettrait de contourner les procédés de l'édition traditionnelle et donc les éditeurs. Les articles pourraient être archivés dans des archives personnelles ou institutionnelles, voire dans des « archives conviviales et gratuites », ce qui permettrait de ne plus avoir recours aux revues scientifiques actuelles.

Le libre accès a donc une forte teneur politique et en fait, ce côté politique s'immisce de plus en plus dans les débats sur le sujet et dans leurs résolutions. Bizarrement, de nombreux éditeurs commerciaux voient dans le libre accès une solution d'avenir. Les leaders du marché de l'édition scientifique et technique, tels que Elsevier Science, John Wiley & Sons et Springer, admettent tous que les objectifs du libre accès sont utiles et réalisables. Certains d'entre eux proposent déjà aux utilisateurs une grande partie de ce que le libre accès veut proposer, à savoir les liens, l'interrogation multibase, l'accès au texte intégral, etc. Par contre, bien souvent la seule différence entre le point de vue des défenseurs du libre accès et celui des éditeurs concerne la manière de financer ces innovations. Compte-tenu des nombreux sujets que ce séminaire pourrait aborder, il est important de rappeler ici que le présent rapport est destiné à servir de canevas aux débats et n'a aucunement la prétention de fournir une description exhaustive ou une évaluation du libre accès. Il essaie plutôt d'esquisser certains des points à débattre et à développer au cours de ce séminaire. Comme l'objectif du séminaire est de débattre du libre accès dans le monde de l'IST, le rapport se concentrera plus particulièrement sur les aspects politiques de la question, à savoir est-ce que les solutions proposées par le libre accès pourraient ou devraient remplacer avantageusement les revues scientifiques traditionnelles. Si tel est le cas, il faudra au minimum que le service proposé soit capable de gérer l'acquisition, la vérification, le contrôle qualité, la production, la diffusion (y-compris la localisation et l'accès) et l'archivage des informations. Pour être viable et concurrencer le système d'édition scientifique actuel, le libre accès devra accompagner les articles de services à valeur ajoutée permettant de mieux trouver et localiser l'information, comme par exemple la fouille de données (*data mining*).

2. L'objectif de l'Initiative pour les archives ouvertes

L'initiative ou mouvement pour les archives ouvertes (*Open Archives Initiative – OAI dans ce qui suit* - www.openarchives.org) couvre une grande partie des activités liées aux technologies du libre accès. Il n'y a pas lieu ici de répéter en détails les excellentes informations disponibles sur ce site. La FAQ en particulier explique bon nombre des points soulevés au cours des discussions qui ont abouti à la rédaction du présent rapport.

Toutefois, il paraît utile de rappeler que l'OAI est le résultat d'initiatives précédentes telles que l'Universal Preprint Service (UPS) et la Convention de Santa Fé et qu'elle a été mise en place aux Etats-Unis en juillet 1999. Son but était, et est toujours, d'une part, de voir comment faire en sorte que les archives de documents, les serveurs de prépublications et les bibliothèques numériques puissent être reliés et accessibles au moyen d'une interface conviviale commune, et d'autre part, de développer des services connexes comme des options de recherche distribuée, des systèmes experts d'information, ou autres.

Il est clair que l'initiative tire en grande partie ses origines du mouvement des prépublications (*preprint*) ou des publications électroniques (*eprint*) en IST ; il est donc naturel qu'elle se soit d'abord focalisée sur le secteur de l'IST, bien qu'elle se soit peu à peu étendue à d'autres aires de l'information. Les travaux se sont surtout intéressés au secteur de la recherche publique mais les normes s'appliqueront aux services gratuits comme aux services commerciaux. On trouvera une bonne synthèse du sujet à <http://www.ariadne.ac.uk/issue31/open-archives-forum/>.

Les partisans du mouvement pour les archives ouvertes estiment que le volume grandissant des informations disponibles sur le web rend cette initiative nécessaire. En plus des articles eux-mêmes, certaines bibliothèques ont mis en ligne gratuitement leurs catalogues, leurs documents en texte intégral et d'autres ressources numériques. Toutefois, une grande partie de ces ressources sont hétérogènes, ce qui empêche de les exploiter sous un seul système.

L'OAI s'est rapidement transformée en un mouvement international auquel participent activement de nombreuses personnes, instituts et organismes. Deux groupes d'utilisateurs sont concernés :

- 1) Les fournisseurs de données qui sont responsables des réservoirs d'information et qui ont adopté le protocole de collecte de métadonnées des archives ouvertes (Open Archive Initiative Metadata Harvesting) afin de mettre en valeur les métadonnées du contenu de leurs réservoirs de données.
- 2) Les fournisseurs de services qui récoltent les métadonnées à partir des réservoirs conformes au protocole des archives ouvertes afin de s'en servir pour créer des services à valeur ajoutée.

Le protocole de collecte des métadonnées des archives ouvertes est au centre du mouvement pour les archives ouvertes. D'autres protocoles ont vu le jour au sein de la communauté des bibliothèques depuis la norme Z 39.50, mais celui des archives ouvertes est plus « simple » et peut donc être plus largement utilisé. L'OAI recommande d'utiliser les métadonnées du Dublin Core comme niveau descriptif de base ; ces métadonnées ont été encodées dans une syntaxe XML spéciale mise au point par l'initiative. En mettant en place le protocole des archives ouvertes, les centres d'archives et les fournisseurs de données rendent leurs contenus numériques accessibles. Lorsque les archives sont interrogées

depuis l'extérieur en utilisant le protocole des archives ouvertes, les réponses se font de manière normalisée.

Les fournisseurs de services récoltent les métadonnées mises à disposition par les fournisseurs de données et les utilisent pour apporter de la valeur ajoutée à des services tels que les services de fourniture de documents ou les moteurs de recherche spécialisés. Ces services peuvent être gratuits ou payants. Selon la politique des archives ouvertes, les fournisseurs de services doivent se conformer aux politiques juridiques des archives auxquelles ils renvoient.

3. Le logiciel Eprints, l'OAI et le monde de l'édition scientifique et technique

Un consultation rapide du site web de l'initiative pour les archives ouvertes, montre que de nombreuses initiatives et expérimentations/solutions techniques sont en cours ouvrant ainsi la voie au libre accès. Certains de ces sites sont conformes au protocole de l'initiative pour les archives ouvertes. Toutefois, il serait difficile de nier le côté passionné du mouvement qui accompagne les avancées technologiques. De nombreux partisans de l'OAI voient clairement dans le libre accès une solution pour mettre un terme à ce qu'ils considèrent comme « l'hégémonie des éditeurs ». Dans son article « For whom the gate tolls », Stevan Harnad présente très bien ce point de vue. <http://www.ecs.soton.ac.uk/~harnad/Tp/resolution>.

En effet, le libre accès pourrait bien s'imposer comme une alternative à l'environnement traditionnel des revues. Cependant, afin de briser la dominance actuelle, il faut que le libre accès propose une solution aux auteurs qui voudraient mettre leurs travaux en ligne. Stevan Harnad a donc avancé non seulement certains arguments concrets mais aussi des solutions, la plus importante en date étant le logiciel Eprints (<http://www.eprints.org/>) écrit par deux collaborateurs de Stevan Harnad, Robert Tansley et Chris Gutteridge. Ce logiciel propose une interface web pour gérer, soumettre, explorer et télécharger des documents. Les réservoirs de publications électroniques utilisant E-prints sont, par définition, conformes au protocole de l'OAI et donc, une fois qu'une collection est enregistrée en tant que fournisseur de données conformes, les services d'information sensibilisés à l'OAI pourront en localiser le contenu et y accéder.

Le logiciel semble efficace (voir par exemple la critique disponible à www.arl.org/sparc/core/index.asp?page=g20#6) et est déjà utilisé par certains instituts pour mettre en place des services expérimentaux (voir l'article de William Nixon « The evolution of an institutional e-prints archive at the University of Glasgow » « *Evolution d'une archive institutionnelle de publications électroniques à l'Université de Glasgow* » <http://www.ariadne.ac.uk/issue32/eprint-archives/intro.html>). Ce logiciel a également fait l'objet d'un long et intéressant débat sur le forum de *American Scientist* consacré aux publications électroniques

<http://www.ecs.soton.ac.uk/%7Eharnad/Hypermail/Amsci/date.html#2212>

Les chercheurs et les utilisateurs non commerciaux peuvent télécharger gratuitement le logiciel Eprints à partir du serveur eprints.org. Les premiers rapports montrent que le logiciel est facile à installer et à utiliser. La société Ingenta (<http://www.ingenta.com/>) travaille également sur une version

commerciale du logiciel Eprints. De plus, la société britannique va proposer aux clients de prendre en charge leurs archives car elle pense, probablement à juste titre, que de nombreuses institutions ne pourront pas ou ne voudront pas investir les ressources nécessaires à la maintenance d'une archive complète avec tous les paramètres de sécurité et de stabilité que cela demande.

Certaines personnes au sein de la communauté du libre accès contestent le bien-fondé de cette initiative commerciale. Selon eux, un progiciel commercial pourrait rapidement mener à un système à deux vitesses, le logiciel commercial étant plus riche en fonctionnalités que le logiciel gratuit des chercheurs. Il n'y a pas de raison que cela se produise surtout si les deux partenaires sont bien décidés à faire en sorte que les logiciels soient complètement interactifs. Il ne faut oublier que le progiciel SRS, quant à lui, a fort bien survécu à la privatisation [1] .

Mettre en place une archive est une chose mais ce n'est qu'un point de départ. Si de telles initiatives sont appelées à remplacer les processus d'édition traditionnels, il leur faudra proposer des services à valeur ajoutée allant de la vérification du sérieux de l'article, à l'évaluation par un comité de lecture, à l'ajout des métadonnées correspondantes nécessaires à l'archivage, etc. De plus, un des rôles majeurs de l'édition est d'assurer et d'indiquer que les documents sont utiles. L'archivage - c'est à dire la garantie que les documents sont disponibles en permanence - de nombreux articles « préférés des auteurs » pourrait faire plus de tort que de bien à la cause du libre accès.

C'est peut-être pour cette raison que la première préoccupation est de convaincre les auteurs d'archiver leurs articles scientifiques avant qu'ils ne soient repris dans le processus des revues traditionnelles. A l'origine de cette utilisation moderne des prépublications, il y a le succès du service de prépublications en physique mis en place par Paul Ginsparg <http://arXiv.org/> qui fut suivi par d'autres services tels que Cogprints pour les articles en psychologie et en biologie www.cogprints.soton.ac.uk.

Bien que la mise à disposition de versions électroniques des prépublications soit relativement récente, l'idée n'est pas nouvelle car cela fait des décennies, voire des siècles, que les scientifiques cherchent à faire part de leurs idées à un groupe de leur pairs avant de les communiquer au reste du monde. Evidemment, l'Internet a permis aux chercheurs de faire beaucoup plus avec leurs propositions d'articles que de simplement les envoyer à quelques personnes triées sur le volet. Le premier service de prépublications a encouragé les physiciens à soumettre les prépublications de leurs articles à un serveur central afin qu'on puisse y accéder et les critiquer et donc les mettre à disposition plus tôt et de manière plus interactive que le format papier des revues auxquelles ils étaient destinés. Toutefois, un élément essentiel du système était, et est toujours, que les articles seraient plus tard soumis à un comité de lecture et obtiendraient ainsi la valeur ajoutée conférée par l'évaluation par les pairs.

La clé du succès de tout service d'information scientifique est la qualité. Un argument fort et courant des éditeurs est que le manque de qualité fait que les prépublications et tout autre forme d'auto-archivage n'apportent que peu de chose à la chaîne de l'information. Plusieurs partisans du libre accès tel que nous le connaissons aujourd'hui, espèrent que les revues finiront par accepter que la version révisée soit mise à disposition gratuitement parallèlement à la prépublication. Toutefois il faudra bien des changements avant que ce ne soit le

cas, bien que, comme nous le verrons plus loin, les éditeurs soient en général d'accord pour permettre des liens entre les prépublications et les publications à condition que ces dernières restent sur les serveurs des éditeurs.

Il y a pourtant une forte tendance à vouloir faire des prépublications une nouvelle forme de publication scientifique. Un ou deux éditeurs proposent déjà des services de prépublications où les auteurs soumettent les articles à un serveur central qui est dans le domaine public. Les articles peuvent être lus, critiqués et commentés. En général, l'auteur peut décider ce qu'il veut faire de son article suite à cet « examen public ». Ils peuvent décider soit de retirer leur article, soit de le soumettre à une revue traditionnelle. Dans certains cas, ces services font en fait partie des procédures de soumission d'une revue, alors que dans d'autres, les auteurs ont le droit de choisir une autre revue après la période de prépublication. On peut citer à titre d'exemple, Net Prints en médecine (www.clinmed.netprints.org), géré conjointement par le groupe du British Medical Journal et Stanford, et le Chemistry Prints Service (www.preprint.chemweb.com) de ChemWeb d'Elsevier.

Toutefois, il ne faudrait pas croire que le problème soit résolu. De nombreuses revues refusent d'accepter les articles ayant été soumis à un serveur de prépublications. Ce refus est conditionné par la politique de la revue et n'a rien à voir avec le droit d'auteur. Les partisans des services de prépublications prétendent que peu de revues refusent. Pourtant, même si cela n'est pas vraiment le cas actuellement, un nombre important d'éditeurs et de titres de revues sont en train de changer d'avis en ce qui concerne cette forme de (pré) distribution [2]. Il est intéressant de noter que ceux qui refusent ne sont pas nécessairement des éditeurs commerciaux et qu'il y a parmi eux certaines sociétés savantes, comme l'American Chemical Society, qui refusent les articles qu'on avait pu lire sur le serveur de ChemWeb.

En fin de compte, le débat autour des raisons de cet embargo se résume en deux mots : « droit d'auteur et évaluation par les pairs ». La question de la législation sur le droit d'auteur, simple à première vue, est en fait fort complexe et ouverte aux débats. En Europe, la Commission Européenne est d'ailleurs en train de finaliser une directive en vue d'harmoniser cette question dans la Communauté Européenne (<http://www.eurorights.org/eudmca/CopyrightDirective.html>) et qui devrait permettre une meilleure répartition des droits. Dans la plupart des cas, les tierces parties n'ont aucun droit d'auteur sur les pré-publications bien que certains « financeurs » revendiquent les droits de propriété intellectuelle liés aux résultats de leur financement, ce qui veut dire qu'ils peuvent se réserver le droit d'auteur sur tout article qui en résulte. On peut donc auto-archiver les articles ; toutefois, si la prépublication est soumise plus tard à une revue, cette dernière aura tout à fait le droit de décider si elle souhaite ou non traiter cet article.

Le problème de l'évaluation par les pairs est plus simple et concerne surtout le niveau de qualité attendu. Net Prints, dans un aperçu de la manière dont il voit l'évolution du processus d'évaluation, explique clairement la situation :

« Les commentaires des lecteurs ne constituent pas une évaluation par les pairs telle que nous la percevons aujourd'hui. Nous fournissons aux auteurs qui souhaitent soumettre leurs articles à une revue avec comité de lecture, la liste des revues qui acceptent ou non les articles ayant fait l'objet d'une prépublication sous forme électronique. Nous espérons un jour faire en sorte que les auteurs puissent soumettre leurs articles à la revue de leur choix directement à partir de ce site web, ce qui simplifierait le processus d'évaluation par les pairs. Nous aimerions, une fois que les articles ont été publiés par les revues en question, pouvoir pointer les lecteurs vers le texte intégral de la version finale,

où qu'elle se trouve. Cela pourrait se faire par l'intermédiaire de liens hypertexte soit vers la revue, soit vers les sites web personnel ou institutionnel de l'auteur. Les auteurs souhaitant déposer sur clinmed.netprints.org leurs articles ayant été soumis à un comité de lecture, le feront sans renoncer à ce droit dans le cadre des conditions de licence ou de droit d'auteur de la revue. Nous espérons publier en temps voulu les politiques des revues à ce sujet. »

4. L'acquisition et l'évaluation par les pairs

De nos jours, il est rare qu'un chercheur puisse lire en profondeur tous les articles scientifiques concernant son domaine. Dans la majorité des cas, on se contente d'une lecture en diagonale (le titre, le résumé, parfois la section sur les Matériels et Méthodes et quelques références). Pourtant le travail a été fait, il est solide, et il pourrait servir à quelqu'un. Les chercheurs sont toujours contraints de « publier ou périr », c'est pourquoi ils choisissent les revues qui correspondent le mieux à leur sujet et à leur type de recherche. Un article sur un sujet standard et ordinaire aura du mal à se faire accepter par *Science* ou *Nature*, voire par une revue spécialisée de haut niveau comme *Brain Research*. Et pourtant la plupart des articles trouvent quelque part une revue qui les publiera. Les chercheurs sélectionnent donc la meilleure revue possible pour leurs travaux et d'autres chercheurs « pré-sélectionnent » un article sur la base du titre de la revue qui le publie. Dans l'ensemble, le procédé d'évaluation d'une revue est très important ; certains diront même que c'est le comité de lecture à lui seul qui fait la qualité et la notoriété d'une revue.

Nombreux sont ceux au sein de la communauté du libre accès qui cherchent comment ajouter de la valeur à leurs documents, comment les faire valider et les faire évaluer. Certaines revues rendent leurs prépublications conformes à l'initiative pour les archives ouvertes, puis elles ajoutent un lien vers la version finale validée. Toutefois, comme nous l'avons vu plus haut, certains auteurs n'aiment pas que leurs articles soient livrés au public avant d'avoir été soumis à un comité de lecture. La revue *Behavioral and Brain Sciences* (<http://www.bbsonline.org/>) par exemple, utilise le logiciel Eprints pour ses archives et pour l'évaluation par les pairs, mais a créé des fichiers de prépublications différents selon que les auteurs souhaitent ou non divulguer leurs articles en cours d'évaluation par *BBS*.

Il est fort probable que les éditeurs seront favorables à un système où la prépublication serait liée à la version finale sur le site de l'éditeur. Toutefois un problème reste entier : de nombreux utilisateurs ne pourront pas, dans les conditions actuelles, accéder aux articles qui les intéressent car ils ne disposeront ni de la licence ni de l'abonnement nécessaires (voir Chapitre 7). La vente à l'article, telle que la pratique Blackwell Science par exemple, résoudrait en partie le problème pour certains. Les autres devront se contenter de savoir que l'utilisateur est conscient que la prépublication n'est pas la version finale, et cela ne manquera pas d'être une cause de frustration.

On espère également que l'auto-évaluation, l'évaluation par la communauté qui gère le serveur, ou les commentaires du marché, permettront de limiter les abus. Bien que les délibérations quant à la place des prépublications dans le processus d'édition soient toujours en cours, il n'y a pas de doute qu'une grande partie de

L'opinion est favorable à la « libre évaluation » des prépublications. On dispose, dans certains domaines, de techniques informatiques de suivi de l'utilisation des publications en ligne qui montrent que les articles en ligne ont tendance à être plus souvent cités que les articles des revues papier. De telles techniques pourraient servir à suivre l'évolution de ces articles et pourraient donner des indications précises sur l'utilisation future des articles. (Pour une discussion détaillée sur cette question et sur ses aspects connexes, voir les débats sur l'accès électronique de *Nature* à <http://www.nature.com/nature/debates/e-access/Articles/>).

Certains partisans du libre accès pensent que, même si les articles ne sont jamais librement accessibles dans les archives après publication, le libre accès sera quand même utile car les prépublications seront de toute façon dans le domaine public. Une telle opinion peut être dangereuse. Une information inexacte est une désinformation. Un gros inconvénient de l'OAI sera que le lecteur ne saura pas toujours si ce qu'il lit est la version finale de l'article – les ordinateurs se débrouillent très bien avec les liens vers l'amont, mais sont beaucoup moins performants quand il s'agit des liens en aval !

Tout comme le logiciel Eprints a permis de créer des services de prépublications centralisés, il est fort possible que les acteurs du libre accès mettent en place, dans différentes disciplines, des services d'évaluation centralisés qui soient aussi fiables que les procédés d'évaluation actuels de l'édition scientifique traditionnelle. Toutefois, il est plus probable que nous verrons émerger bon nombre de services d'évaluation et de validation indépendants. Certaines universités font déjà preuve d'une certaine réticence à laisser leurs enseignants/chercheurs mettre à disposition du public des articles « non-évalués » qui pourraient nuire à la réputation de l'université. Les entreprises industrielles ont des règles très strictes sur ce que leur personnel a le droit de charger sur des serveurs internes ou externes, ce qui permet d'assurer un certain contrôle de la qualité. De toute façon, le problème de l'évaluation est tellement important que la survie du libre accès en dépend.

A ce jour, probablement à cause de la jeunesse du libre accès, les procédures de contrôle de la qualité des sites en libre accès ne sont encore qu'à l'état d'ébauche. Cependant tout le monde est conscient du problème, et on verra bientôt émerger certaines formes de contrôle d'accès et de validation ainsi que d'évaluation des contenus. Les sociétés savantes sont certainement les mieux placées pour initier ce type d'action. Ce ne serait pas la première fois, car au cours de l'histoire de l'édition scientifique, la contribution de ces mêmes sociétés aux premières revues fut non négligeable, et elles seraient donc à même de savoir lesquels de leurs membres pourraient contribuer aux services de prépublications ou autres dans un domaine spécialisé. Il faudra probablement faire appel à des initiatives parallèles, telles que celles qui concernent les bibliothèques électroniques, pour trouver les compétences nécessaires à la conception de ces services.

Une autre tendance pourrait ouvrir de nouveaux horizons à l'IST. Les bases de données de séquences d'acide nucléique « contrôlent » les données au fur et à mesure de leur traitement. D'autres domaines devraient faire de même et déposer leurs données dans des systèmes électroniques permettant un contrôle extérieur. Avec le libre accès, les chercheurs peuvent mettre leurs données à

disposition afin qu'elles soient revues, voire remaniées. Il faudra mettre en place, et observer, des règles sur le droit d'auteur et le « piratage », mais il est fort probable que, dans un avenir proche, les articles de revues vont de plus en plus s'appuyer sur des données déposées à d'autres endroits.

5. L'interrogation, la localisation et la création de liens

Certains ne manqueront pas de me reprocher de n'avoir pas inclus certaines informations dans mon rapport. Il faut dire je dois beaucoup au World Wide Web sans lequel je n'aurais pu rédiger ce rapport dans les délais impartis et avec les ressources dont je disposais. Le web est une mine d'informations. Les moteurs de recherche comme *Google* permettent de fouiller des millions de pages de données, mais il ne faut pas tomber dans le travers qui consiste à se dire : « J'ai tellement d'informations, donc j'ai tout ce qu'il me faut sur le sujet ».

En réalité, même aujourd'hui, bon nombre d'informations échappent aux interrogations les plus sophistiquées. Ainsi, les universitaires travaillant dans le domaine biomédical, ont l'impression qu'en interrogeant MEDLINE, ils trouvent toutes les informations dont ils ont besoin. La base de données MEDLINE, aussi bonne soit-elle, ne couvre qu'un pourcentage de la littérature pertinente, et une bonne recherche devrait être complétée par l'interrogation d'autres bases de données bibliographiques, telles que EMBASE, CAB Abstracts et PASCAL, pour n'en citer que trois. Si l'on ajoute les nombreuses nouvelles bases de données commentées (*curated*) qui indexent la littérature manuellement ou par d'autres techniques, il est évident qu'il va falloir utiliser de nombreux outils pour retrouver l'information dont on a besoin.

Contrairement à ce que les éditeurs pourraient penser, les utilisateurs ne se dirigent pas spontanément vers un site. Plusieurs revues de grande qualité comme *Nature*, *Science* et *Cell*, ont toutes affirmées par le passé qu'elles pensaient que les chercheurs viendraient naturellement sur leur site pour trouver les informations pertinentes. On avait déjà vu une telle attitude chez les grands éditeurs qui pensaient que, du fait de leur notoriété, les chercheurs se tourneraient tout naturellement vers leur site, alors qu'en réalité, la plupart des chercheurs ne savent pas qui édite telle ou telle revue. En fait, le web exige des services de recherche plus performants et avec la technologie du libre accès, il faudra disposer d'outils capables d'aller chercher l'information en profondeur pour retrouver le détail demandé (les chercheurs de terrain quant à eux, sont clairs : ce qu'il leur faut c'est un guichet unique qui puisse les orienter vers les titres pertinents afin de retrouver facilement les articles dont ils ont besoin).

Une innovation récente consiste à relier les références de la bibliographie d'un article à d'autres articles qui pourraient être intéressants. CrossRef est le principal service qui offre cette fonctionnalité (www.crossref.org). CrossRef est un réseau à but non lucratif qui repose sur des partenariats avec les éditeurs, et dont le but est d'assurer l'efficacité et la fiabilité des liens dans toute la littérature scientifique en ligne. CrossRef, qui est en croissance constante, est à ce jour le principal utilisateur du système d'identificateur d'objet numérique (*Digital Object Identifier* ou DOI - www.doi.org et 153 éditeurs utilisent déjà ce service (en septembre 2002, la base CrossRef comptait environ 6 400 revues et 5 millions d'articles). Dans le système CrossRef, chaque DOI est associé à un ensemble de métadonnées de base et à une URL qui pointe vers le texte intégral, afin d'identifier spécifiquement un élément de contenu et d'assurer un lien

permanent vers sa localisation sur l'Internet (bien que le DOI ne fasse pas partie du protocole des archives ouvertes, les principes sont les mêmes).

Les éditeurs aiment bien le système CrossRef, car le service central ne contient pas le texte intégral mais assure les liens au moyen des DOI qui sont reliés aux métadonnées des articles fournis par les éditeurs participant à CrossRef. En cliquant sur un lien (dont l'éditeur détermine le format, par exemple un bouton CrossRef ou « Article » en html), le chercheur se connectera à une page du site de l'éditeur où il verra la référence bibliographique complète de l'article et, dans la plupart des cas, le résumé. Il utilisera ensuite les mécanismes appropriées pour accéder au texte intégral de l'article ; en général, les abonnés iront directement au texte alors que les autres seront informés des différents moyens d'obtenir le texte : abonnement, fourniture de document, ou paiement à la carte. CrossRef ne coûte rien au chercheur ; les coûts sont couverts par la somme symbolique demandée aux éditeurs pour déposer leurs métadonnées, les cotisations annuelles, et les frais d'accession à la banque de DOI de CrossRef demandés aux éditeurs secondaires voulant créer des liens vers le texte intégral des articles.

Les DOI ne s'appliquent pas qu'aux articles scientifiques. La Fondation encourage le développement de différents profils d'application qui fixent les normes de métadonnées à utiliser dans cette application (le DOI en prend en charge plusieurs dont le Dublin Core et le OA Harvest). Le DOI permet également au groupe de définir les règles commerciales et donc permet à des groupes distincts de définir des DOI dans leur domaine, ce qui en fait un candidat idéal au soutien d'un régime international de libre accès.

La technologie derrière le service CrossRef est semblable à celle que l'on est en train de mettre en place pour certains aspects du libre accès. L'éditeur doit fournir, dans un format défini, un ensemble minimum de métadonnées pour chaque article (il peut fournir des métadonnées supplémentaires s'il le souhaite). Le DOI/URL sera enregistré auprès de la Fondation Internationale DOI dans le Registre des DOI et deviendra ainsi une référence fixe pointant vers cet élément d'information.

CrossRef est un réel progrès pour la communauté de l'IST. Evidemment, l'utilisateur se sentira frustré si l'accès à un site lui est refusé parce que lui ou son institution n'est pas abonné à la revue dans laquelle se trouve l'article. Afin de remédier à cette situation, de nombreux éditeurs autorisent le paiement à la carte (*Pay-Per-View* ou *ppv*) et bien que le prix par article puisse paraître énorme à l'utilisateur moyen, il s'agit là d'un problème plus économique que technique (bien que de nombreux bibliothécaires déplorent qu'il soit plus facile aux utilisateurs d'aller sur le site d'un éditeur plutôt que d'aller voir si le document recherché se trouve dans les collections de la bibliothèque, ce qui pourrait conduire à des achats en double. Pour éviter cela, il faudrait que les services d'accès soient bien contrôlés et qu'ils signalent à l'utilisateur que sa bibliothèque est abonnée et qu'il peut s'adresser à elle pour se procurer le document).

La communauté du libre accès lance des projets similaires. Un projet déjà bien avancé est celui de l'European Mathematical Information Service (EMIS). Ce service, mis en place en 1994, est devenu un réseau de serveurs d'informations de l'European Mathematical Society [3]. Il bénéficie du soutien bénévole de plusieurs partenaires dans le monde entier (voir l'article de Bernd Wegner et

Michael Jost, "EMIS 2001 – A portal to Mathematics in Progress" High Energy Physics Libraries Webzine, numéro 6, mars 2002, <http://library.cern.ch/HEPLW/6/papers/4/>

Citebase Search (<http://citebase.eprints.org/>) permet aux chercheurs de faire des recherches croisées dans des archives électroniques de prépublications en texte intégral, avec la possibilité de trier les résultats selon certains critères comme le facteur d'impact (*citation impact*), et ensuite de naviguer vers cette littérature au moyen des liens des références et de l'analyse. On commence à trouver dans ce service des exemples issus de la physique, des mathématiques, de l'informatique, de la psychologie, des neurosciences et des sciences biomédicales.

CrossRef et les services de liens de ce type, permettent de suivre l'historique d'un article de bout en bout. Traditionnellement, les utilisateurs d'IST se tournaient tout d'abord vers les services bibliographiques. De nos jours, de plus en plus de chercheurs utilisent des services tels que le moteur de recherche *Google* pour trouver l'information.

Les systèmes de recherche de ce type ont leurs bons côtés et Elsevier a lancé Scirus (www.scirus.com) qui, selon l'éditeur, a plusieurs avantages sur *Google*, dont le filtrage des sites non scientifiques, la recherche d'articles évalués par un comité de lecture à partir de bases de données à accès contrôlé. Par exemple, si vous faites une recherche sur la maladie de la vache folle, Scirus vous trouvera des articles de ScienceDirect alors que *Google* ne vous trouvera que des pages web ; Scirus descend au delà des deux premiers niveaux d'un site web et retrouvera donc bien plus d'informations pertinentes. Scirus ne recherche pas que les contenus d'Elsevier, et prend en compte de nombreux projets que la communauté du libre accès défend actuellement comme E-Print ArXiv, CogPrints, BioMedCentral, Mathematics Preprint Server et Chemistry Preprint Server.

6. L'archivage

Il faut que les publications restent accessibles aux générations futures. Il est donc essentiel que les documents scientifiques dont il est question dans ce séminaire soient « archivés » de manière à ce qu'ils soient toujours disponibles. Il faut que ces archives soient financièrement et techniquement stables, et qu'elles soient gérées de manière à conserver l'intégrité des documents ; il faut aussi qu'elles soient accessibles à tout ceux qui en ont besoin même si cela nécessite une forme de paiement.

L'OAI utilise assez librement le mot « archive », ce qui pourrait laisser penser que les documents conservés « quelque part dans un ordinateur conforme à l'OAI » seraient suffisamment protégés, mais ce n'est pas le cas. De plus, on ne s'est pas encore vraiment penché sur la définition du type d'archive dont on aura besoin. Le terme d'auto-archivage est souvent utilisé dans les documents de l'OAI pour désigner le processus selon lequel les auteurs peuvent soumettre eux-mêmes leurs travaux au serveur ou à l'archive de leur choix. Il y a les « archives institutionnelles » où les auteurs peuvent déposer leurs articles sur un serveur administré par une organisation ou une société savante, par leur université ou leur institut de recherche, et il y a aussi des archives centrales par discipline ou autre spécialité telles celles que développent les nouveaux acteurs du domaine comme BioMedCentral (<http://www.biomedcentral.com/>).

Le terme d'archive est également utilisé pour décrire des collections d'information servant à un but commun. Eprints et les autres technologies d'archivage permettent de collecter, sous réserve de l'autorisation de l'auteur, des ensembles d'informations afin de les rendre accessibles et de les distribuer. L'objectif de certaines de ces collections n'est pas toujours clair, mais l'Electronic Research Archive in International Health (ERA) mise en place par *The Lancet* en est un bon exemple. Cette archive permet aux chercheurs en médecine de déposer des articles particulièrement pertinents pour les problèmes de santé rencontrés dans de nombreux pays en développement. Les articles soumis sont évalués avant d'être acceptés et après cela, ils sont archivés et disponibles en ligne gratuitement.

Les initiatives actuelles concernant les liens entre les contenus ont aussi donné un nouvel élan à l'archivage. Ainsi, Elsevier et d'autres éditeurs tel que Springer, se lancent dans de grands travaux de numérisation de l'antériorité de revues et de collections entières. De nombreux éditeurs numérisent également des manuels et autres ouvrages de référence afin de les ajouter à leurs contenus électroniques.

Mais il y a archivage et archivage. Il est évident qu'un éditeur a intérêt à garder une collection de fascicules afin de vendre l'antériorité d'une revue si nécessaire. Toutefois, tout le monde est d'accord pour dire que les bibliothèques spécialisées, et surtout les bibliothèques nationales, devraient être investies de la mission d'archivage pour les générations futures (dans de nombreux pays, dont la France, cette mission fait déjà l'objet d'une loi au niveau national). Aux Pays-Bas, la bibliothèque royale néerlandaise a récemment signé, à l'initiative de l'éditeur Elsevier, un accord d'archivage pour toutes les publications d'Elsevier Science. La bibliothèque met ces publications à la disposition des personnes autorisées à accéder à ses collections et s'engage à ouvrir un accès à distance pour tout le monde au cas où Elsevier viendrait à ne plus commercialiser ses titres.

L'archivage est coûteux et on craint que, en laissant les publications chez leurs éditeurs commerciaux, on pourrait perdre les archives si l'éditeur déposait son bilan. Il en va de même des services d'archivage non-commerciaux, en particulier ceux n'ayant pas une raison essentielle de les maintenir – après tout, même la collection de Paul Ginsparg a changé de domicile et qui peut prédire si son nouvel hôte acceptera de l'héberger gratuitement « ad vitam eternam ». Bon nombre de nouveaux projets d'archives ont perdu leur contenu quand les archives ont dû fermer faute de subvention. Bien plus d'universités que ne le pensent les chercheurs ont des trésoriers peu disposés à libérer les fonds nécessaires au stockage des rapports et articles des enseignants/chercheurs ne faisant plus partie de l'université. Leurs nouveaux établissements pourraient quant à eux, hésiter à prendre la responsabilité d'archiver les résultats de travaux réalisés sous une autre tutelle. Et certains rapports pourraient bien « disparaître » dans une archive destinée à d'autres activités ou à d'autres domaines. Il serait donc bien naïf de penser qu'une institution puisse permettre l'auto-archivage sans un but et des règles bien définis.

Même les bibliothèques nationales ne disposent pas de fonds illimités, mais elles sont les mieux placées pour montrer la voie en ce qui concerne les projets d'archivage. Mais avant de se lancer, elles voudront sans doute s'assurer que les documents archivés sont essentiels et de bonne qualité. Il faudra éviter les

doublons ainsi que l'archivage de documents non finalisés alors qu'une version finale est déjà stockée ailleurs. Le volume de documents impliqués est tellement énorme qu'il faudra absolument procéder de manière rationnelle.

L'ICSTI, le conseil international pour l'information scientifique et technique (www.icsti.org) a récemment commencé à œuvrer auprès de certaines communautés du monde de l'IST (chercheurs, éditeurs, bibliothèques, archives, producteurs de données) afin de les sensibiliser au besoin de préserver le patrimoine électronique de la science.

L'ICSTI pense que la création d'une archive numérique et exhaustive des sciences mettra en jeu un réseau complexe d'initiatives tant au niveau des disciplines et des institutions, qu'au niveau national et international. Il faut donc décider des normes à suivre et s'y tenir (il semble que le modèle de référence de l'Open Archives Information System (OAIS) qui va devenir une norme ISO, soit un bon modèle à suivre <http://ssdoo.gsfc.nasa.gov/nost/isoas/overview.html>). L'ICSTI estime qu'il y a encore beaucoup à faire pour définir clairement les politiques d'archivage, les responsabilités de chacun et faire en sorte que l'infrastructure mise en place soit entretenue, financée et durable et puisse résister à l'épreuve du temps.

Pourtant les préconisations de l'ICSTI (*Advocacy Statement*) à ce sujet laissent certaines questions sans réponse :

- ?? Faut-il aussi archiver la forme et la présentation du contenu ou seulement le contenu ?
- ?? Comment pourra-t-on en garantir l'authenticité ?
- ?? Quelle(s) version(s) faudra-t-il conserver ?
- ?? Faudra-t-il conserver les liens--qui sont très importants dans de nombreux services d'information électronique-- et si c'est le cas, comment le faire ?
- ?? Comment aborder les questions de droit d'auteur et de propriété intellectuelle afin que les réservoirs de données puissent conserver des archives perpétuelles tout en permettant l'accès ?

La déclaration de l'ICSTI mentionne d'autres problèmes comme la collecte des données dans des archives dont les auteurs détiennent le droit d'auteur. Ou bien, les lois sur la protection des données et l'harmonisation des bases de données vont-elles nous empêcher de tirer pleinement parti des technologies dont nous disposons ? On s'intéresse beaucoup aux aspects pratiques et juridiques des droits de propriété intellectuelle depuis l'avènement de l'Internet. Ces sujets font actuellement l'objet de directives, de séminaires et de groupes de travail au sein de la Commission européenne. Deux rapports pouvant intéresser la communauté de l'IST, l'un sur les droits de propriété intellectuelle et la bioinformatique et l'autre sur la propriété intellectuelle et l'Internet, sont disponibles à http://europa.eu.int/comm/research/era/ipr_en.html.

L'article de Jeffrey Young intitulé « 'Superarchives' Could Hold All Scholarly Output » [*Des superarchives pourraient accueillir toute la production scientifique*] <http://chronicle.com/free/v48/i43/43a02901.htm> décrit une des manières dont la mise en réseau mène aux archives. L'auteur fait remarquer que de nombreux instituts sont en train d'explorer la manière dont ils pourraient regrouper leurs propres documents afin de pouvoir les échanger et les travailler dans un « environnement libre ». L'article ne mentionne pas les dispositifs de contrôle qualité que ces archives pourraient mettre en place, bien qu'il fasse

allusion à l'éventuelle nomination de « rédacteurs » et limite le contenu au type de documents qui de toute façon sont destinés à être publiés.

Bon nombre d'instituts réputés considèrent en fait les archives comme une « carte de visite » qui reflète leur excellence. Une fois de plus on risque de créer un monde à deux vitesses avec d'un côté les centres « riches » proposant des archives interactives avec toute une gamme d'outils et de contenus, et de l'autre les centres « pauvres » devant se débrouiller avec les moyens du bord.

Un autre facteur important est aussi à prendre en considération. Les gens ne voient souvent pas de grandes différences entre les anciennes et les nouvelles versions des services. De nos jours, le texte de l'article scientifique à l'écran est souvent perçu comme l'équivalent du texte papier. Cependant la technologie moderne permet d'aller beaucoup plus loin, et les articles scientifiques disposeront bientôt de documents ou de services connexes tels que les liens vers des fichiers vidéo ou audio, des illustrations, des tableaux ou des logiciels permettant de manipuler et de visualiser les données. Ces évolutions font l'objet de nombreuses études au sein du nouveau domaine des bibliothèques numériques (voir par exemple le site du groupe européen DELOS à www.delos-noe.iei.pi.cnr.it/). Ces évolutions ont un impact sur l'archivage, notamment la disponibilité à perpétuité des outils permettant de lire ces fichiers, ce qui pourrait représenter un obstacle pour ceux qui ne disposent pas de l'équipement informatique ou des bandes passantes nécessaires.

7. Abonnements, licences et autres obstacles à un accès libre et gratuit

L'objectif des chercheurs quand ils interrogent la littérature scientifique est de trouver le texte intégral de l'article dont ils ont besoin et, comme nous l'avons mentionné plus haut, quand on interroge les chercheurs en biomédecine, la plupart d'entre eux pensent qu'une recherche dans MEDLINE couvre toute la littérature de leur domaine.

Un tel a priori est dangereux. Les estimations diffèrent, mais il y a probablement plus de 20 000 revues ayant une certaine pertinence pour les chercheurs alors que MEDLINE n'indexe qu'environ 4000 d'entre elles. Même ensemble, tous les services bibliographiques n'arrivent pas à couvrir la totalité de la littérature scientifique. De nouveaux services apparaissent sur le marché mais eux aussi semblent se concentrer sur les « meilleurs » articles comme par exemple, le projet « Faculty of 1000 » (www.facultyof1000.com).

Le libre accès permettrait donc d'accéder à la majorité de « ces données scientifiques cachées ou moins accessibles ». Certaines de ces dernières se trouveront sur les serveurs des éditeurs car la plupart des organisations proposent maintenant en ligne leurs contenus récents, et parfois l'antériorité, soit sur leur propre site web, soit sur le site d'aggrégateurs comme Ingenta (www.ingenta.com) ou Highwire (www.highwire.org). Nous pourrions donc accéder à l'article à condition que les métadonnées qui l'accompagnent permettent de le localiser, ou qu'on puisse trouver la bonne adresse ou le bon lien (et un service tel que CrossRef devrait faciliter les choses).

Cependant, un problème majeur est que ces contenus sont protégés et que les utilisateurs ne peuvent y accéder librement. Les éditeurs ont lourdement investi dans la constitution de leurs listes de contenu et tiennent à les commercialiser,

tout du moins en ce qui concerne les revues, sous forme d'abonnements ou de licences, plutôt que de vendre à l'article. Certains affirment que ces licences seront un jour assez « flexibles » pour permettre à « n'importe qui » d'accéder à « n'importe quoi », mais ce n'est pas encore le cas.

Personne ne met en doute que le succès commercial des éditeurs scientifiques commerciaux repose sur le modèle des abonnements aux revues : même les sociétés savantes ont lancé de nouvelles revues afin d'accroître leurs revenus. Il est également vrai que de nombreux clients des éditeurs se considèrent victimes de ce succès, et l'augmentation incessante des prix s'est traduite depuis plus de dix ans par une diminution correspondante des abonnements.

Les ventes de la plupart des éditeurs du secteur de l'IST sont maintenant concentrées sur un petit noyau de bibliothèques orientées vers la recherche et le développement, Les petits instituts, les bibliothèques de départements universitaires et les particuliers ont dû résilier une grande partie de leurs abonnements, et de nombreuses bibliothèques ne sont plus en mesure de proposer les services que leurs bibliothécaires souhaiteraient fournir. Les revenus des éditeurs sont de plus en plus vulnérables car ils dépendent maintenant d'un nombre réduit de centres payeurs.

La "révolution électronique" ouvre de nouveaux débouchés aux éditeurs. Les coûts de l'édition électronique se situent en amont : une fois que les éditeurs ont mis en place l'infrastructure, les contrôles d'accès et les contenus sous forme électronique, l'ajout de nouveaux utilisateurs leur demande peu d'investissement. Les grands éditeurs se tournent donc de plus en plus vers l'octroi de licences, de site ou de consortiums par exemple. En passant d'un système d' « abonnements individuels » à leurs principales revues à un système de licences de « collections de contenus », ils ont réussi à conserver leurs revenus « pré-payés » tout en offrant de la valeur ajoutée aux utilisateurs. Une autre évolution concerne les achats groupés ou de consortium qui permettent aux petites institutions de se regrouper afin d'acheter une licence de consortium. Cela a permis à de nombreuses petites institutions qui avaient résilié des abonnements de revenir sur le marché.

A première vue, l'opération semble avoir réussi. Elsevier Science signale que, en ce qui concerne les revues, la part de revenus des licences de site et autres licences d'accès aux revues est passée de 0 à 65 % en quatre ans. John Wiley & Sons font part d'un succès similaire. Ces éditeurs ainsi que d'autres semblent ravis de ce succès et déclarent (ou espèrent) que le marché est également satisfait. L'utilisation est en hausse et ceux qui disposent d'une licence téléchargent un grand nombre d'articles (on peut toutefois se demander ce qu'on entend par utilisation ; lors que c'est gratuit, la plupart des chercheurs téléchargent les articles pour les lire plus tard, mais il n'y a aucune garantie qu'ils le feront) et un plus grand nombre d'utilisateurs peuvent accéder aux articles dont ils ont besoin. Alors, pourquoi changer le système quand il suffit de s'affilier à quelqu'un qui dispose d'une licence de site pour obtenir l'article dont on a besoin.

Pourtant, bien que de nombreux chercheurs soient affiliés à un établissement qui fait partie d'un consortium ou dispose d'une licence de site, qu'en est-il des nombreux chercheurs qui ne le sont pas ? On est en train d'essayer de trouver des solutions. Comme les coûts liés à l'établissement de nouveaux accès clients

sont minimales, les grands éditeurs étudient de nouvelles formes de licence et même d'éventuelles tarifications selon les pays. Il semble également que le système de licence de site puisse être étendu aux bibliothèques et aux bibliothèques municipales.

A première vue, cela semblerait pouvoir satisfaire tout le monde. Toutefois, certains se demandent encore si les licences répondent vraiment aux besoins du marché. Par exemple, les éditeurs ont tendance à proposer des collections de plus en plus vastes alors que le marché demande des collections d'articles spécialisées selon des critères bien définis. De plus, le fait de proposer des collections « fourre tout » sous couvert de la qualité (autrement dit toutes nos revues sont bonnes) ne sera que source de mécontentement de la part des utilisateurs car les revues d'un éditeur donné sont souvent de qualité variable et le marché le sait. Leo Waaijeers de la bibliothèque de l'Université Technologique de Delft résume la situation en ces termes : « Les consortiums c'est bien pour les éditeurs mais pas pour les bibliothèques. Maintenant, nous devons acheter 1 100 titres d'Elsevier au lieu des 68 dont nous avons besoin. (voir, « A New Role for the Scientific Publisher in the Electronic Age » [*Le nouveau rôle de l'éditeur scientifique à l'ère de l'électronique*] www.iospress.nl). De plus, de nombreux bibliothécaires aimeraient avoir quelque chose de plus tangible en fin de contrat car souvent il n'existe pas de version papier et donc l'archivage et l'antériorité posent problème.

Le volume des données soulève un problème de fond, à savoir celui du coût de l'information. De plus en plus, les organismes constatent qu'ils n'ont plus les moyens nécessaires pour acquérir l'information dont ils ont besoin. Certains déplorent que l'obligation d'acheter une licence grève leur budget au détriment d'autres acquisitions. On est en droit de se demander quelles politiques d'acquisition pourraient justifier l'acquisition d'une licence -- couvrant une partie importante, mais non la totalité des informations -- demandant un tel investissement qu'il serait impossible d'acquérir les documents d'autres éditeurs. Les licences actuellement proposées par les éditeurs s'appliquent plus aux bibliothèques multidisciplinaires qu'aux bibliothèques spécialisées qui, quant à elles, réclament un autre modèle de licence. En fin de compte, l'inclusion dans une licence de revues dont personne n'a besoin, est de plus en plus contestable quand cela empêche d'acquérir des revues essentielles auprès d'autres éditeurs. Les résultats d'une étude récente effectuée par le Ingenta Institute et présentés au cours de conférences organisées par l'Institut (les actes sont disponibles sur le site www.ingenta.com) remettent en question la durabilité des modèles de licence actuels. L'offre globale (*Big Deal*) qui regroupe au sein d'un même contrat avec une bibliothèque ou, de plus en plus avec un consortium, toutes les revues d'un éditeur semblait judicieuse il y a quelques années lorsque les bibliothèques pouvaient trouver les financements nécessaires pour une telle offre groupée. Cependant, ces financements ne concernaient que l'offre de départ, et 3 ou 5 ans plus tard, les consortiums doivent renégocier le contrat de départ et sont obligés de constater qu'ils ne disposent plus des fonds nécessaires au renouvellement du contrat de licence.

On pourrait penser qu'il s'agit une fois de plus des sempiternelles lamentations des bibliothécaires, à savoir un budget d'acquisitions stable face à des augmentations annuelles de 5 à 8 %, sans tenir compte de l'inflation. Aujourd'hui, le problème est pourtant bien réel : le coût et le volume des

informations disponibles augmentent alors que les budgets stagnent quand ils ne diminuent pas. De plus, les subventions publiques ayant permis l'acquisition d'une licence groupée n'étaient pas renouvelables et à l'heure fatidique du renouvellement de la licence, les bibliothèques et consortiums sont obligés d'en assumer le coût sur leurs propres fonds. Les résultats de l'étude de l'Ingenta Institute montrent que l'avenir de telles offres groupées est limité. De plus, on est en droit de penser que les gros investissements ne sont plus de mise sur le marché de l'IST. De toute façon, la solution licence n'est pas une panacée et ne saurait contenter tout le monde.

8. Les coûts

Rien n'est jamais gratuit et il faut toujours que quelqu'un paie, qu'il s'agisse d'une revue ou d'une « archive ouverte ». Même le fameux système de prêt entre bibliothèques universitaires n'est pas aussi gratuit qu'il ne le semble car les coûts indirects de gestion et de personnel des bibliothèques participantes sont conséquents.

Les estimations du coût d'édition d'un article de revue varient. BioMedCentral applique des règles semi-commerciales à ses politiques de libre accès et facture un forfait de traitement de 500 dollars par article. Nombreux sont ceux qui estiment que ce forfait est trop faible pour couvrir les coûts de traitement et l'avenir nous dira s'ils ont ou non raison. Toutefois, cela nous donne une idée de ce que coûte le traitement et la publication d'un article évalué par les pairs dans le cadre des recommandations de l'initiative pour les archives ouvertes.

L'archivage sera également coûteux et il suffit de regarder les coûts globaux de l'édition scientifique pour avoir une idée du type de budget dont il faudra disposer. De plus, avec une augmentation annuelle des flux de production d'articles d'environ 3 %, les coûts ne pourront qu'augmenter et on peut se demander qui les prendra en charge.

Aujourd'hui, le consensus est que c'est « l'Etat » qui paie en subventionnant les bibliothèques. Pourtant, nombreux sont ceux dans le monde universitaire qui pensent que cette situation est injuste. En effet, non seulement « l'Etat » finance la recherche mais en plus il doit payer pour accéder aux résultats de ses propres recherches. Il n'en demeure pas moins que publier coûte de l'argent. Il est rare que les subventions de recherche et développement comportent une ligne budgétaire consacrée à la publication car, bien que les travaux doivent donner lieu à des publications, on estime que les coûts de publication, ainsi que les risques financiers encourus, seront à la charge de quelqu'un d'autre, à savoir les éditeurs.

Il est clair que les budgets des bibliothèques ne peuvent pas tout assumer. Mais si on se tourne vers le modèle du « paiement à la page », il faudra aussi trouver les fonds pour le financer. En théorie, le transfert des subventions des bibliothèques vers les départements scientifiques afin qu'ils puissent publier leurs articles au lieu de les acheter, serait la solution. Mais, on peut se demander si, d'une part le monde universitaire sera d'accord et d'autre part les bibliothécaires accepteront ces transferts de fonds qui signifieront aussi un transfert de pouvoir et d'influence.

Cela peut paraître simpliste, mais c'est en fait le modèle de financement adopté pour de grands projets tels que le projet sur le génome humain (Human Genome Project) et autres banques de données génétiques. En substance, les agences de financement subventionnent des institutions aux Etats-Unis, en Europe et au Japon, pour que les utilisateurs puissent accéder gratuitement à leurs données. Autrement dit, soit l'institution est payée pour rendre sa base ou ses bases de données accessibles gratuitement, soit le chercheur est payé pour acheter les données à l'institution.

L'OAI est financée à partir de fonds publics centralisés et de nombreuses nouvelles initiatives du secteur public adoptent les recommandations de l'initiative. En Europe, le projet E-BioSci (www.E-biosci.org) qui propose toute une palette de nouveaux services de stockage et d'interrogation dans le domaine biomédical, sera conforme à ces recommandations. Une fois opérationnel, le projet permettra aux utilisateurs de naviguer entre des bases de données bibliographiques, factuelles ou de séquences, et le texte intégral des articles pertinents. Il mettra également en place un nouveau service d'édition électronique permettant aux auteurs de déposer leurs articles sur le web, ce qui, sur le long terme, devrait permettre de créer une infrastructure de recherche européenne.

Une partie de E-BioSci va donc remettre en question le statu quo de l'édition scientifique. FIGARO (www.figaro-europe.net), un groupe d'universités et d'éditeurs européens ayant pour objectif de créer un réseau d'institutions qui apporterait à la communauté scientifique une aide à l'édition électronique, devrait avoir un rôle similaire. FIGARO va étudier de nouveaux modèles économiques d'édition scientifique et encouragera le libre accès aux publications produites et distribuées au sein de son infrastructure, rendant ainsi l'édition scientifique plus rapide, plus simple et moins coûteuse.

La prise en compte d'un modèle économique dans le projet FIGARO est un point important et on peut se demander si les autres projets de ce type ne devront pas aussi planifier au-delà de leurs subventions initiales. Le projet d'E-BioSci de créer des liens entre les bases de données de séquences nucléotides ne présente pas de grandes difficultés, et a déjà fait l'objet du projet ADLIB -- Advanced Database Linking in Biotechnology -- projet financé par la Commission européenne. Ce projet pilote (*Proof-of-Concept* -- POC), une coopération du secteur public et du secteur privé, a avait montré dès 1998 que un tel projet était techniquement faisable. Cependant, les participants au projet n'ont pas réussi à rentabiliser le service soit et ont abandonné le projet malgré son intérêt évident. ADLIB n'a pas demandé d'autres financements publics, mais le projet E-BioSci devra-t-il compter sur un financement continu de la part de la Commission européenne ? Un financement continu et illimité, aurait tôt fait de consumer une grande partie des fonds de recherche et développement de l'Union européenne, à moins que de nouvelles lignes budgétaires ne soient créés pour maintenir en vie de tels services centralisés (puisque c'est ce qu'ils deviendront). En effet, il faut que le libre accès soit durable pour être à même d'effectuer de vrais changements dans la manière de traiter l'information. Les projets initiaux de l'OAI et les archives initiales sont financés par des organismes « riches », mais ces fonds ne seront pas suffisants pour gérer le volume de documents générés si le rêve du libre accès venait à se réaliser.

Il faut des financements à long terme, indépendants de toute influence politique. Le palmarès du système des subventions n'est pas brillant quand il s'agit de lancer des industries novatrices. L'étude des dossiers de financement des projets novateurs de la Commission européenne montre que bien peu de ces projets se sont révélés économiquement viables. La Commission européenne, pour sa part, fait de plus en plus savoir que les projets qu'elle finance devront à terme voler de leurs propres ailes. Il ne s'agit pas d'une simple déclaration politique pour faire plaisir aux partisans du libre marché, car la Commission ne peut garantir les financements que pour la période d'un programme cadre. Il n'y a donc aucune garantie que les projets financés dans le cadre d'un programme donné pourront demander le renouvellement des financements dans le cadre du programme suivant, car les priorités et les règlements peuvent changer du tout au tout. [4]

Cela veut dire que dans le contexte de l'Union européenne, un projet ne peut compter que sur un maximum de quatre ans de financement. De nombreux éditeurs commerciaux seraient ravis de pouvoir atteindre en aussi peu de temps, le seuil de rentabilité, voire les bénéfices, qui leur permettraient de survivre de manière autonome. L'information a besoin de continuité, cette continuité repose sur les financements et le système des financements publics n'est pas en mesure de garantir des financements à vie.

Toutefois, les financements publics initiaux sont une aide précieuse pour des projets novateurs tel que le projet sur le libre accès. Ensuite, de tels projets devront faire preuve d'un sens des affaires, ce qui veut dire que les options dont ils disposeront ne seront pas vraiment différentes de celles du monde actuel de l'édition.

William Y. Arms dans son article sur les modèles économiques possibles de l'édition en libre accès (« Economic Models for Open Access Publishing » [*Modèles économiques pour l'édition en libre accès*] (http://www.cisp.org/imp/march 2000/03_00arms.htm) établit un parallèle entre les modèles économiques de la télévision, de l'information en ligne avant l'Internet et le web, et de l'édition traditionnelle, et les éventuels modèles économiques de l'édition électronique en libre accès, qui en fait ne sont pas nouveaux, à savoir :

- ?? Accès réservé : paiement en fonction de l'utilisation,
- ?? Accès réservé : abonnements,
- ?? Accès libre : publicité,
- ?? Accès libre : financement extérieur.

Pour le moment, seul le dernier modèle semble répondre aux besoins politiques et économiques de la communauté du libre accès. La solution résidera probablement dans une combinaison d'initiatives publiques et commerciales, telle qu'elle existe aujourd'hui. D'un côté, il y a la publicité qui selon certains membres récents de la communauté du libre accès, pourrait représenter 20 % des revenus et qui, avec la facturation du traitement de l'article, générerait des revenus suffisants pour couvrir les frais d'édition des revues en libre accès. D'autres sont plus sceptiques et rappellent que souvent dans l'édition électronique, la publicité n'a pas été en mesure d'accomplir ce qu'on attendait d'elle. Le modèle de l'abonnement, aussi impopulaire soit-il, risque donc de refaire son apparition.

Une plus grande transparence en ce qui concerne les coûts d'exploitation de services d'édition et d'archivage électroniques en accès libre devrait aider à trouver une solution à ce problème des coûts. Dans son article, William Arms, fait remarquer que, à l'université Cornell, le montant total des fonds investis par chaque département dans son site web dépasse celui du budget d'acquisition de la bibliothèque. Les coûts de fonctionnement d'un nœud de réseau Internet dans une université coûte beaucoup plus cher que les utilisateurs pourraient le croire. On observe déjà d'importants déplacements de crédits au sein des institutions, et le libre accès pourrait être considéré comme une infrastructure et faire partie des coûts de fonctionnement d'une institution, au même titre que les autres infrastructures de cette dernière. Toutefois, nous n'en sommes aujourd'hui qu'au stade des opinions et des conjectures et on ne peut pas dire qu'un plan concret de financement à long terme du libre accès soit près de voir le jour.

9. Le monde en développement

On a souvent dit que l'évolution de l'édition électronique allait permettre aux pays en développement de rattraper leur retard en ce qui concerne l'accès et l'utilisation de l'information scientifique ou autre. Les pays pauvres n'auront plus à acheter des bibliothèques entières -- ce dont d'ailleurs ils n'ont pas les moyens -- mais pourront accéder comme ils veulent à des collections centralisées. Cela permettrait de surmonter les obstacles qui entravent les flux d'information entre le Nord et le Sud, le Sud et le Sud ou le Sud et le Nord, et en particulier dans le cas Sud/Sud et Sud/Nord où les nouvelles technologies permettent aux chercheurs de diffuser à peu de coût leurs travaux auprès d'un large public (voir www.ariadne.ac.uk/issue30/oai-chan/).

Le libre accès permet aux travaux de recherche des chercheurs du monde en développement d'être accessibles et utilisés par tout le monde. En théorie, les universitaires des pays défavorisés pourraient se servir de tous les serveurs de par le monde proposant des services conformes à l'initiative pour le libre accès sans avoir pour autant à mettre en place leurs propres serveurs. Pour le moment, un des principaux obstacles est une méconnaissance des différents mécanismes de distribution et d'accès des travaux de recherche actuellement disponibles ; le fait que cette disponibilité passe par l'Internet ne fait d'exacerber le problème car le monde en développement ne dispose pas d'infrastructures de télécommunications adéquates.

Toutefois, de grandes opérations d'investissements dans les infrastructures sont en cours, tant au niveau local qu'au niveau international, et on pense de plus en plus avoir trouvé une solution au problème de la « fracture numérique ».

L'initiative de Budapest sur le libre accès (Budapest Open Archive Initiative ou BOAI <http://www.soros.org/openaccess/read.shtml>) fait partie des initiatives les plus connues qui font état d'une volonté d'aider spécifiquement le monde en développement. L'initiative de Budapest, lancée fin 2001, repose sur le postulat que la science est « publiée sans rétribution aux auteurs pour le bien de la recherche et de la connaissance » (auquel il faudrait ajouter la visibilité). Si on pousse ce postulat à sa conclusion logique, il s'en suit que « (...) si l'information était librement accessible, le monde entier en bénéficierait encore plus (...). La suppression des obstacles qui entravent l'accès à cette littérature va accélérer la recherche, enrichir l'enseignement, favoriser le partage des connaissances entre

les riches et les pauvres et vice versa, rendre cette littérature aussi utile que possible, et jeter les fondations de l'unification de l'humanité à travers un dialogue intellectuel et une quête du savoir communs. »

L'initiative de Budapest conteste, non seulement les prix pratiqués dans le monde de l'édition commerciale, mais également les marges réalisées par les éditeurs, tout en reconnaissant que la production d'une littérature scientifique évaluée par les pairs a un coût. Les signataires de l'initiative sont en faveur de toute tentative visant à briser le carcan de l'édition actuelle et à le remplacer par une nouvelle infrastructure. Afin de mettre en place le libre accès à la littérature scientifique, les signataires recommandent les deux stratégies complémentaires suivantes :

1. L'auto-archivage – Pour les chercheurs, l'auto-archivage est une pratique selon laquelle ils peuvent déposer eux-mêmes dans des archives électroniques ouvertes leurs articles publiés dans des revues à comité de lecture. Il faut donc mettre à disposition des auteurs, les outils et les aides nécessaires à l'auto-archivage. Lorsque les archives sont conformes aux normes de l'initiative pour les archives ouvertes, les moteurs de recherche et autres outils de recherche peuvent interroger des archives distinctes comme une seule et même archive. L'utilisateur peut donc accéder au contenu des archives sans avoir à se soucier de savoir si telle ou telle archive existe et où elle se trouve.
2. Les revues en libre accès – Il faut que les chercheurs aient les moyens non seulement de lancer une nouvelle génération de revues fidèle au concept du libre accès, mais aussi d'aider les revues existantes qui souhaitent migrer vers le libre accès. Ces revues ne restreindront plus l'accès et l'utilisation des documents qu'elles publient au nom du droit d'auteur, afin d'assurer la plus grande diffusion possible. Elles utiliseront ces mêmes droits d'auteur pour assurer le libre accès permanent à tous les articles qu'elles publient. L'accès étant bloqué par les coûts, ces nouvelles revues ne feront payer ni abonnement ni droit d'accès, mais se tourneront vers d'autres moyens pour couvrir leurs frais. De nombreuses sources de financement existent comme les fondations et organismes publics qui financent la recherche, les universités ou les laboratoires qui emploient les chercheurs, les fonds de dotation de sociétés savantes ou d'institutions, les amis de la cause du libre accès, les bénéfices dégagés par la vente d'informations supplémentaires, les fonds libérés par l'annulation des abonnements aux revues traditionnelles faisant payer des abonnements et des droits d'accès, ou même par la disparition de ces revues, et enfin les contributions des chercheurs eux-mêmes. Il n'y a pas lieu de préconiser une solution plutôt qu'une autre selon les disciplines ou les pays ou d'arrêter de chercher des solutions de remplacement nouvelles et novatrices.

S'il est clair que, dans l'idéal ces notions sont largement partagées dans le monde du libre accès, il n'en demeure pas moins que la recherche de solutions reste vague et trop générale. Dresser simplement la liste des sources de financement possibles sans indiquer la manière dont elles sont impliquées serait faire preuve d'ignorance. Il faudra plus que quelques « déclarations » pour organiser une infrastructure capable de remplacer l'infrastructure de l'IST

actuelle, avec les changements de financement à grande échelle que cela implique ainsi que les changements de comportement, voire les changements dans l'évaluation scientifique, les facteurs d'impact et les citations étant aujourd'hui étroitement liés aux des titres de revues.

Les défenseurs de ces évolutions pensent tout naturellement que l'initiative de Budapest sera couronnée de succès grâce au financement de l'Open Society Institute, le réseau de fondations fondé par le philanthrope George Soros. L'Institut est d'avis que, malgré les fonds importants dont son réseau dispose, l'initiative aura besoin pour aboutir des ressources d'autres organisations ainsi que d'une administration et d'une planification de haut niveau.

Pour réussir, le libre accès pourrait suivre la voie tracée par l'Internet : chaque acteur apportera sa contribution à la nouvelle infrastructure et de la somme de ces contributions naîtra un nouvel environnement.

10. L'avenir et ce qu'il réserve

Il est fort probable que le libre accès, ainsi que l'initiative pour les archives ouvertes, donneront lieu à de nouvelles méthodes de stockage et de recherche de l'information scientifique. La technologie va bientôt permettre aux utilisateurs d'interroger en une fois des centres et des archives dispersés. Des instituts ou des groupes d'instituts vont mettre leurs fichiers en ligne de sorte que leurs contenus puissent être agrégés, exploités ou reliés à des services à valeur ajoutée. Les bibliothèques numériques vont mettre en place des collections d'informations validées et agrémentées de documents connexes et les articles scientifiques vont évoluer au gré des critiques et des commentaires venus de l'environnement des prépublications.

Est-il envisageable que les nouvelles infrastructures d'édition électronique puissent remplacer le système d'IST actuel et en particulier les revues scientifiques ? Après tout, l'industrie de l'édition est bien implantée, bien financée et, de plus, elle est à la pointe des nouvelles technologies. Les éditeurs disposent également d'une infrastructure dont l'efficacité n'est plus à démontrer, et dont les millions de pages de contenu sont la base de nouveaux services et de nouvelles avancées.

Pour le moment, il n'en est pas question. Le libre accès n'est encore qu'un ensemble de technologies en pleine évolution. Il ne dispose pas (encore) d'une infrastructure « prête à l'emploi ». Certains doutent que la collecte des métadonnées préconisée par l'OAI soit aussi efficace qu'on voudrait le croire, les initiatives actuelles sont encore trop modestes et avancent trop lentement pour porter sérieusement atteinte au système actuel, et le manque de bases financières solides fait qu'il est difficile de compter sur la confiance et l'adhésion à long terme de la communauté scientifique. L'OAI se limite encore à un « club » de personnes ayant les mêmes convictions. Bien que l'importance des problèmes soulevés soit plus qu'évidente, les débats restent confinés au sein d'un groupe minoritaire de chercheurs et la question fondamentale du financement commence tout juste à être débattue.

Parallèlement, le coût actuel des revues risque de remettre en cause le statu quo de l'édition telle que nous la connaissons aujourd'hui. Autrement dit, le marché estime désormais que le prix des revues est prohibitif. Le monde de l'édition commerciale est critiqué pour son contrôle absolu du marché de l'IST. L'augmentation du prix des revues est bien supérieure à l'inflation, des volumes ou des suppléments sont arbitrairement ajoutés aux offres actuelles sans que le client puisse ni protester ni refuser, et des nouvelles revues sont créées dans des domaines du marché déjà saturés. Le prix des revues ne fait qu'augmenter et nombreux sont ceux qui pensent que le point de rupture est atteint car les budgets ne peuvent plus faire face aux prix demandés par l'industrie de l'édition. Sans vouloir entrer dans un débat sur les torts ou les raisons de l'édition commerciale, il faut cependant souligner que ce ne sont pas les éditeurs qui génèrent les articles qu'ils publient. Les chercheurs, quant à eux, ont besoin de publier. L'article scientifique représente l'aboutissement d'un projet de recherche et le succès de la recherche dépendra pour beaucoup de la notoriété de l'article. Le coût d'une revue dépend en grande partie du nombre d'articles qu'elle contient. Malheureusement, la plupart des articles ne vont intéresser qu'un petit nombre de chercheurs et donc, le pourcentage d'articles intéressants susceptibles d'attirer un lectorat important est faible comparé aux articles plus ésotériques.

Cependant les coûts d'édition des articles ésotériques et des articles intéressants sont les mêmes. Le marché adopte donc le modèle de l'abonnement qui permet de vendre des articles dont seul un petit nombre intéressera le chercheur. Ce modèle génère d'importants bénéfices pour les éditeurs qui l'utilisent pour continuer à publier. Sans ces rentrées d'argent, moins d'articles seraient publiés, mais on pourrait penser que le prix de l'article baisserait en conséquence. Le débat fait rage depuis des décennies et n'est pas prêt de s'arrêter. Le libre accès pourrait peut-être trouver sa place dans une solution globale en permettant, par exemple, que certaines formes d'information soient disponibles « gratuitement ». Ainsi pourrait-on voir évoluer le rôle de la revue.

Pour le moment, les partisans du changement ne font que reproduire, certes à moindre coût le modèle de la revue traditionnelle.. Aux Etats-Unis, l'initiative de la Public Library of Science (PloS www.plos.org) tire ses origines d'un mouvement qui s'élevait contre les prix des revues traditionnelles et contre le fait que peu de ces revues étaient disponibles gratuitement sur le serveur PubMedCentral aux Etats-Unis. Les initiateurs du mouvement ont tout d'abord appelé les auteurs à boycotter ces revues et plus de 30 000 chercheurs ont signé la pétition. Toutefois, le mouvement n'a pas vraiment eu l'effet escompté. Le nombre d'articles soumis à ces revues est plus ou moins le même qu'avant la pétition de la Public Library of Science. Le groupe de la Public Library of Science reconnaît qu'il ne dispose pas d'éléments concrets permettant de dire si son initiative a ou non fait son effet, mais il pense que pour « gagner », il leur faudra lancer des services concurrents. Dans une lettre adressée à l'auteur du présent rapport en septembre 2002, on peut lire :

« [nous sommes] sur le point d'annoncer que nous disposons d'un soutien financier solide pour lancer nos propres revues. Nous pensons commencer à accepter des articles pour ces revues au cours du premier semestre de l'année prochaine et commencer à les publier au cours de l'été prochain. »

La Public Library of Science souligne également que les articles soumis à BioMedCentral, un éditeur en libre accès, sont en constante augmentation, ce qui les conforte quant au succès de leur campagne.

BioMedCentral finance ses activités d'édition en faisant payer des frais de traitement de l'article (ou facturation à la page comme on dit couramment). BioMedCentral espère que la somme de 500 dollars par article suffira à financer ses services, mais il reconnaît qu'il faudra le vérifier dans la pratique. BioMedCentral a rédigé une charte bien définie qui répond en grande partie aux attentes de la Public Library of Science et devrait donc répondre aux questions éthiques soulevées par cette dernière. Selon cette charte, les auteurs conservent le droit d'auteur de leurs articles et les articles sont également chargés sur PubMedCentral et d'autres services d'archivage. BioMedCentral est particulièrement décidé à s'assurer que les articles stockés chez lui soient conservés sur le long terme, ce qui prouve son engagement dans les nouvelles technologies et aussi dans les services expérimentaux nécessaires pour garantir que ce qu'il publie est accessible à l'avenir.

Pour le moment, BioMedCentral n'est encore qu'un acteur bien modeste. L'avenir nous dira si le groupe de la Public Library of Science sera capable de lancer son offensive, et même s'il y arrive, nombreux sont ceux qui se demandent s'il arrivera vraiment à modifier notablement l'habitude qui consiste à soumettre les articles à des revues prestigieuses. Le nombre d'articles pour publication augmente tous les ans et donc on verra probablement des auteurs se tourner vers les nouveaux projets ; il n'en demeure pas moins qu'il faudra bien du courage à un chercheur pour renoncer à la possibilité de publier son article dans une revue prestigieuse ayant un fort facteur d'impact. De plus, nonobstant le problème du prix des revues, les administrateurs des universités, ainsi que les bailleurs de fonds, préfèrent voir les résultats des travaux de recherche qu'ils financent publiés dans les revues les plus réputées. Un article qui n'est pas cité ou qui n'est pas publié dans une revue indexée dans MEDLINE ou dans une autre grande base de données bibliographiques telle que PASCAL ou EMBASE, peut être considéré comme « mort » ; cela peut être également une cause de résiliation de subvention, car certaines agences de financement utilisent le facteur d'impact comme barème pour évaluer la qualité des résultats des travaux d'un chercheur.

Pour que le monde de l'édition scientifique traditionnelle se mette à évoluer, il va falloir que BioMedCentral se mette à grandir de manière exponentielle et que la Public Library of Science réussisse à lancer son nouveau service d'édition. Et il faudra de nombreuses années avant que l'infrastructure actuelle ne soit vraiment remise en question.

Avant tout, il va falloir convaincre les chercheurs que leurs travaux ne seront pas perdus car, pour le moment, ils croient plus en la pérennité des revues actuelles qu'en celle d'un serveur hébergé dans un coin de campus. La visibilité pourrait être un argument majeur du libre accès dans sa remise en cause de l'état actuel des choses. Steve Lawrence, dans l'article intitulé « Free Online Availability Substantially Increases a Paper's Impact » [*La libre disponibilité en ligne augmente de beaucoup l'impact d'un article*] qu'il a envoyé au forum de *Nature* sur l'accès électronique (<http://www.nature.com/nature/debates/e-access/Articles/>), démontre que les auteurs en informatique sont plus souvent cités quand leurs articles sont librement disponibles en ligne, et il y a tout lieu de

penser qu'on va retrouver le même phénomène dans d'autres disciplines. Cette meilleure visibilité va attirer les auteurs et il faudra que les éditeurs réagissent en rendant leurs articles plus accessibles afin d'en accroître la visibilité. Une fois de plus, cela ne veut pas dire que les auteurs doivent abandonner les revues, mais plutôt que les revues doivent faire en sorte que leurs articles soient plus visibles et accessibles.

Il est important de rappeler que, en général, les éditeurs ne sont pas opposés dans l'ensemble au concept du libre accès. On peut même dire que c'est tout le contraire, car de nombreux éditeurs sont totalement en faveur des normes, des liens et des moteurs de recherche communs. Il faut dire que les éditeurs commerciaux, bien qu'assez réticents au départ, investissent maintenant des sommes énormes dans les services sur l'Internet et il est dans leur intérêt que ces services réussissent. Commercialement parlant, il n'est pas étonnant que les éditeurs des revues actuelles essayent de tirer parti de leur portefeuille d'abonnement pour dégager des revenus du nouvel environnement électronique et donc on est en droit de penser que l'édition commerciale va se servir des initiatives sur le libre accès au lieu de s'y opposer.

Les partisans du logiciel Eprints et de l'auto-archivage pensent que les archives contiendront à la fois les prépublications et les postpublications (*post-prints*) et que ces versions publiées seront éventuellement disponibles gratuitement. Mais il faut être réaliste : l'édition coûte de l'argent et, dans une économie de marché, il faut récupérer cet argent. Cependant, l'environnement du libre accès et celui de l'édition commerciale sont plus complémentaires qu'on ne le pense. En effet, il va de soi que le secteur de la recherche et le secteur commercial ont intérêt à collaborer sur la durabilité du monde de l'IST plutôt que d'essayer de se faire concurrence dans un contexte de restrictions budgétaires.

Quant au libre accès, la meilleure façon d'avancer, serait d'utiliser l'OAI pour trouver de nouvelles méthodes de diffusion de l'information plutôt que de vouloir imiter les revues actuelles. De nos jours, il n'est plus nécessaire d'aller dans une bibliothèque pour trouver l'information. Avec l'Internet, on peut consulter plusieurs sites à la fois. La revue n'est plus confinée aux rayonnages de la bibliothèque, mais peut-être utilisée en tant que source d'information dynamique. De plus, le format papier ne correspond plus aux nouvelles exigences de présentation et de manipulation des données modernes. De nombreuses données doivent être analysées (et donc évaluées) au moyen de programmes informatiques spécialisés. En général, ce sont souvent les ultimes résultats d'une recherche qui font l'objet d'une publication traditionnelle et toutes les différentes données que le chercheur et son équipe ont pu analyser avant d'arriver à leurs conclusions ne sont souvent pas prises en compte. Les chercheurs aimeraient que dans certains domaines, comme par exemple la cristallographie et les essais cliniques, leurs articles puissent être accompagnés de données, d'illustrations, ou autres données complémentaires permettant de mieux comprendre comment ils sont arrivés à leurs conclusions.

Jusqu'à maintenant, la publication d'un article dans une revue fournissait la preuve qu'un projet de recherche avait été réalisé. De nos jours, les informations déposées par les auteurs dans des banques de données servent aussi à mesurer la qualité des travaux effectués. Alors, l'article publié ne représentera plus qu'une partie d'un ensemble, et les éditeurs vont s'en rendre compte. Il se

pourrait que le nombre de données devant être publiées dans les revues diminue, ce qui veut dire que le marché dans son ensemble devra évoluer. Les éditeurs se concentreront peut-être sur les articles de premier plan nécessitant une évaluation approfondie ou proposeront des synthèses regroupant un domaine, ses disciplines et sous disciplines. Ils pourront aussi développer et proposer des chaînes complètes d'information payantes, montrant ainsi que l'information scientifique n'est pas qu'un « rapport final » mais joue un rôle tout au long du processus de recherche. Pour sa part, le libre accès pourrait permettre de trouver les informations avant qu'elles ne soient figées. Certaines informations n'atteindront donc pas le stade de la publication et resteront sur les serveurs de prépublications et dans les banques de données.

En conclusion, il faut se rendre à l'évidence que rien n'est gratuit dans ce bas monde. La gratuité a toujours un coût et il faut bien que quelqu'un règle la facture. De petits services peuvent se permettre d'être gratuits car eux-mêmes ne coûtent pas cher, mais une fois qu'un service atteint la notoriété et se développe, il ne peut plus se le permettre. C'est là la rançon du succès ! Donc, il ne faut pas se faire d'illusions et tout service d'information scientifique entièrement basé sur le libre accès devra un jour ou l'autre passer à un modèle payant. Le *Journal of High Energy Physics* (<http://www.iop.org/journals/jhep>) est un cas d'école. Cette revue électronique, publiée par l'Institute of Physics, est diffusée gratuitement depuis ses débuts en 1997. Elle est conforme à l'OAI et la soumission et l'évaluation des articles sont prises en charge par l'International School of Advanced Studies (SISSA) à Trieste en Italie. Jusqu'à maintenant, la revue était gratuite grâce aux contributions des deux organismes qui la parrainent, mais au fur et à mesure qu'elle a pris de l'importance, ce type de financement n'était plus viable. Et donc, à partir de janvier 2003, alors que l'antériorité des archives reste accessible gratuitement, les institutions devront payer un abonnement annuel.

[1] Le logiciel d'interrogation SRS, très répandu dans le monde de la bioinformatique, a été développé par les chercheurs d'EMBL, avec l'aide du groupe EMBnet. Il a été commercialisé plus tard par LION Bioscience (www.Lion.com) mais, à l'époque de sa privatisation, le logiciel était un pilier de l'infrastructure de la bioinformatique académique. Cette tournure des événements a donc mécontenté beaucoup de monde, mais l'entreprise a fait comprendre qu'elle continuerait à offrir des licences gratuites aux universités et autres centres de recherche affiliés. Cela a bien marché et, au moment où ce rapport a été rédigé, la version 7 du logiciel reste la même afin d'assurer la compatibilité de tous les sites SRS. Les clients des licences commerciales bénéficient d'une importante assistance et d'autres expertises.

[2] Certains éditeurs scientifiques envisagent l'auto-archivage de la manière suivante :

Nature : "Nature ne souhaite pas entraver la communication entre les chercheurs.... Ni les serveurs d'actes de congrès, ni les serveurs de

prépublications ne sauraient être considérés comme constituant une publication antérieure."

Elsevier : "En tant qu'auteur [publié par Elsevier] vous avez le droit de conserver une version de la prépublication d'un article sur un serveur public tel que le World Wide Web."

American Physical Society : Dans sa déclaration sur le droit d'auteur la société de physique américaine encourage l'auto-archivage : " Le(s) auteur(s) disposeront des droits suivants... Le droit de déposer et de mettre à jour l'Article sur un serveur de publications électroniques à condition que les fichiers préparés et/ou formatés par l'APS ou ses prestataires de services ne soient pas utilisés à cette fin. Tout affichage ou mise à jour une fois que l'Article a été accepté pour publication devra être accompagné d'un lien vers le résumé en ligne présent dans la revue de l'APS ou vers la page d'accueil de cette même revue."

Le contrat-type de licence de l'Association of Learned and Professional Society Publishers (ALPSP)

précise : "Vous...conservez le droit d'utiliser votre propre article (sous réserve que vous mentionniez la référence bibliographique de l'article original dans un format standard) de la manière suivante à condition que ne le vendiez [ou ne le donniez] pas dans des conditions qui iraient directement à l'encontre de nos intérêts commerciaux. Vous êtes libre d'utiliser votre article... chargé sur votre propre site web ou sur celui de votre institution; [déposé sur des serveurs publics gratuits de prépublication et/ou d'articles dans votre discipline]..."

[3] Un élément clé du service est la bibliothèque électronique de EMIS (ELibM) qui propose une collection de publications électroniques en accès gratuit. L'objectif est d'être aussi exhaustif que possible bien qu'il y ait encore du travail à faire. Le service a développé son propre service d'évaluation, à partir des critères qu'utilisent les revues traditionnelles, afin d'assurer une qualité optimale. EMIS reconnaît que certains éditeurs ne voudront pas investir pour rien et donc la plupart des revues proposées sont publiées à petit budget, ce qui fait que le risque de désabonnements à la revue papier en raison de la disponibilité d'une version électronique gratuite, est faible. Certains éditeurs fournissent la version électronique à EMIS avec un certain retard, ce qui avantage la version papier. EMIS propose également des liens vers des bases de données dont MATH, la version en ligne de Zentralblatt MATH ; MATHDI, la version en ligne d'un service semblable pour l'enseignement des mathématiques ; MPRESS, un index de prépublications mondiales ; et une base d'objets géométriques.

A l'évidence, les commentaires des membres ont contribué au développement du service. Le service a bénéficié de subventions, mais il dépend beaucoup du volontariat. De plus, il faut s'abonner au service complet.

Les auteurs soulignent que EMIS ne pourrait se maintenir au niveau actuel sans le grand nombre de ses adeptes, qui assurent de manière bénévole la maintenance et l'installation des revues électroniques, s'occupent des soumissions et des transferts de contenu et font fonctionner les sites miroirs. EMIS ne pourrait exister sans eux. Cela prouve qu'un service viable peut fonctionner avec quelques bénévoles, mais, contrairement à ce que dit EMIS, cela ne prouve pas que le service puisse se développer de manière à couvrir exhaustivement les mathématiques dans leur ensemble. Cela ne veut pas dire non plus qu'un tel service doit être purement commercial mais, comme ils font d'ailleurs payer des frais d'abonnement, on peut se demander si un aspect mercantile est en jeu.

[4] SwissProt, la base de données de structures de protéines, reconnue comme étant la pierre angulaire de la bioinformatique moderne, a fortement bénéficié d'une série de subventions de la Communauté européenne. Suite à l'adoption d'un nouveau Programme Cadre, le financement de la CE n'a pas été renouvelé. Les producteurs ont dû de toute urgence faire payer les utilisateurs commerciaux de la base de données, transformant ainsi un produit gratuit en un produit commercial. Cette démarche a fortement mécontenté de nombreux utilisateurs et sérieusement inquiété les utilisateurs en général. SwissProt est maintenant géré par l'institut de biotechnologie suisse (Swiss Biotechnology Institute) et est financé par un mélange de redevances commerciales et de subventions.