



Forum

des Technologies de l'Information et  
de la Communication

# Bibliothèques électroniques : La révolution lente

8 mai 2008

*Paul Thirion*

Université  
de Liège



# Le numérique en bibliothèque universitaire

- Un tournant radical et précoce

En 3 temps

- Numérisation des outils de recherche
- Numérisation des contenus
- Numérisation des services

- Générant de multiples nouveaux rôles

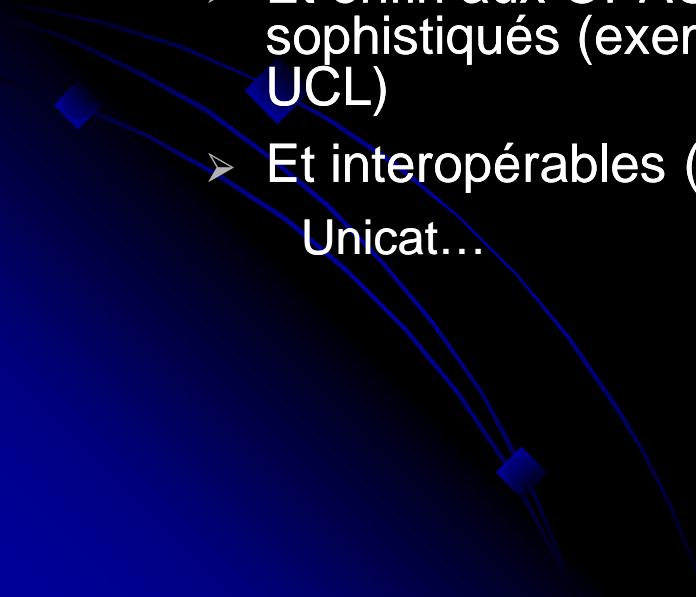
...notamment en lien avec l'Open Access



# Numérisation des outils de recherche

## 1) Catalogues

### En 5 décennies

- Passage des fichiers papiers ...
  - Aux SIGB consultables localement (années 60)
    - Mainframe
    - Solutions locales
  - Et enfin aux OPAC sous WWW de plus en plus conviviaux et sophistiqués (exemple Source à l'ULg, Cible à l'ULB, Libellule à UCL)
  - Et interopérables (formats MARC, XML...)  
Unicat...
- 

# Numérisation des outils de recherche

## 2) Bases de données bibliographiques

En 4 décennies :

- Passage des répertoires bibliographiques papier (revues d'abstracts...)  
Accès très lent et fastidieux
- Aux interrogations « on line » (années 70)  
Accès rapide MAIS coûteux ET via le bibliothécaire
- Aux bases de données sur CD Rom  
Accès direct par l'utilisateur à partir de postes dédiés
- Puis en WWW (années 90)
- Pour aller vers des portails éditeurs (ScienceDirect...), GoogleScholar...

# Numérisation des outils de recherche

## Le côté pile :

- Quantité d'informations disponibles (en centaines de milliers voire millions de notices par outil...)
- Puissance et rapidité
- Précision des recherches (indexation... )
- Convivialité des interfaces
- Accessibilité de plus en plus large :  
accès campus et même campus élargi, voire Internet pour les OPAC

**Amener la documentation là où travaille l'utilisateur**

- De plus en plus de services associés, liens croisés....

# Numérisation des outils de recherche

## Et le côté plus sombre

- Le coût ! Même si :
  - Pour les SIGB : développement de solutions Open Source (PMB, Koha...)
  - Gratuité de certaines base de données (Medline, Eric...)
- Manque d'exhaustivité
  - Tout n'est pas dans les OPAC !
    - Rétrocatalogage
    - Solutions alternatives (ex. Scribe)
  - Ni dans les bases de données
    - Couverture temporelle
    - Géographique et linguistique
    - Choix éditoriaux
- Multiplicité des outils ; effet de « niche »
- Logique booléenne essentiellement (peu de langage naturel, de logique de causalité...)
- Quel usage réel ???
  - Différences subtiles de fonctionnement entre les interfaces
  - Simplisme de la plupart des recherches : « Googlemania »
  - Illusion de compétence

# Numérisation des contenus

## 1) E-journaux

### Le côté pile :

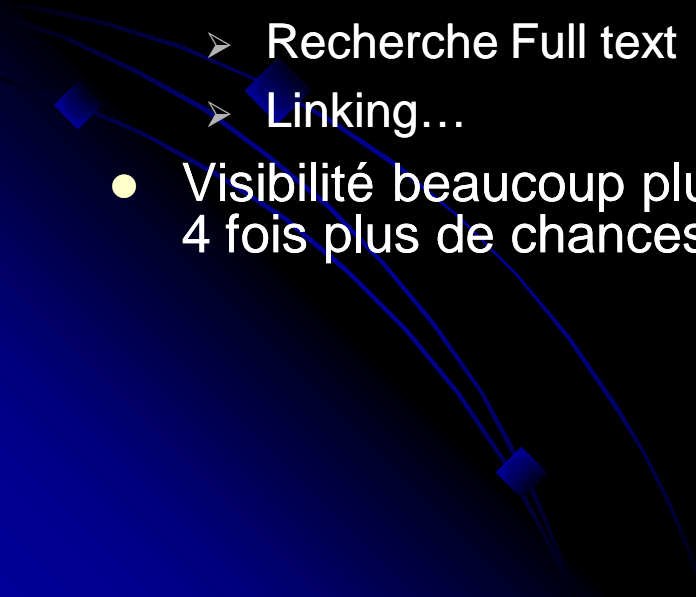
- Développement fulgurant : en moins de 3 décennies !
  - Les tout premiers : dans les années 80 mais surtout depuis les années 90 : une centaine de titres en 1991  
*(Source : Directory of Electronic Journals, Newsletters and Academic Discussion Lists, ARL)*
  - 11600 en 2002  
*(Source : New Jour: Electronic Journals & Newsletters)*
  - Actuellement : au-delà de 100 000 titres électroniques dont des dizaines de milliers de périodiques scientifiques....

Avec développement de « backfiles »

*Exemple ULg : actuellement plus de 13000 e-journaux accessibles*

# Numérisation des contenus

## 1) E-journaux

- Incontournables pour les chercheurs et étudiants
  - Accès campus et même campus élargi (VPN...) :
    - Changement de la relation entre la bibliothèque et ses usagers
    - Illusion de gratuité pour l'utilisateur « Tout est gratuit sur Internet »
  - Possibilités d'exploitation élargies
    - Recherche Full text
    - Linking...
  - Visibilité beaucoup plus large : un article disponible en ligne a près de 4 fois plus de chances d'être cité (S. Lawrence, *Nature*, 2001)
- 

# Numérisation des contenus

## 1) E-journaux

### Et le côté sombre ...

- Modèles commerciaux
  - Abonnement avec surcoût ou e-only
  - « Big deals » (cfr Elsevier)
    - Souvent liés aux abonnements « papier »
    - Titres intéressants mais aussi autres titres
    - Portefeuille de titres fluctuant
  - Package « e-only »
- Licence d'accès et pas vente (pérennité ?)
- Coût se surajoute souvent aux abonnements « papier »
- Commence à apparaître : lien avec « l'usage »
- Statistiques impressionnantes (accès ≠ usage réel )

# Numérisation des contenus

## 1) E-journaux

- Effets pervers de la grande disponibilité : concentration de l'intérêt
- La technique du pêcheur : ferrer le poisson
- Budget fermé => obligation de supprimer des titres d'autres éditeurs
  - => Renforcement de la position dominante des plus puissants et des plus chers
- Différentiel de puissance de développement technologique principaux éditeurs commerciaux <> petites sociétés savantes
  - Qui se vendent au loup
  - Ou risquent de disparaître ....
- Backfiles indépendants des acquisitions antérieures
  - => payer deux fois pour la même chose
- Nécessité de solutions garantissant l'accès à long terme (Portico...) en attendant : obligation de maintenir le papier
- TVA

# Numérisation des contenus

## 1) E-journaux

### Développement de consortiums (BICfB...)

- Avantageux au départ
- Mais objectifs divergents :
  - Les bibliothèques : avoir plus ou au moins autant dans une enveloppe fermée ou en régression
  - Les éditeurs commerciaux : développer leur chiffre d'affaire
- Contraintes et lourdeur de mise en œuvre
  - Coût de la solidarité
  - Nouvelle loi sur les marchés publics...

# Numérisation des contenus

## 2) E-books

- Numériser le contenu des bibliothèques ?
  - Pas si simple
    - Coûts gigantesques
      - En matériel (rapide, non destructif, espaces de stockage...)
      - Et surtout en ressources humaines  
ex pour ULg si 10 sec/page : 1 ETP pendant 1000 ans !
    - Droits d'auteurs ! => livres anciens
    - Définir les priorités
    - Collaborations nationales et internationales indispensables pour éviter la redondance
    - Normes et standards
    - Paysage institutionnel belge (enseignement supérieur vs culture)

# Numérisation des contenus

## 2) E-books

### ➤ Nombreuses initiatives

- De numérisation (Gutenberg, Gallica, Biographie nationale de Belgique, Digithèque ULB...)
- De collaboration ou de définition de standards (MICHAËL, EDL-Net, PEP'S, DONUM...)

Mais le chemin est encore long...

### ➤ Le cas particulier : « Google books »

- Numérisation de 300 000 livres exempts de droits de la bibliothèque universitaire de Gand
- Perspectives éthiques et patrimoniales...

# Numérisation des contenus

## 2) E-books

- Au niveau commercial

- Développement tardif après des expérimentations avortées
- Accélération récente : initiatives d'éditeurs (Luc Pire, Springer...) ou de diffuseurs (Dawson...)

*Exemple ULG : actuellement environ 14000 titres*

- **Difficultés :**


- Modèles commerciaux en cours d'invention
  - Accueil réservé des lecteurs
    - Bourse délier ?
    - Lecture à l'écran
    - Modèle technique de download par page ou groupes de pages
    - L'ouvrage ne se réduit pas au texte
    - ...
  - Peu d'exploitation WWW : simple copie de l'ouvrage « papier »
- Exceptions : dictionnaires et encyclopédies

# Numérisation des services

Transformation de services existants ou création de nouveaux services

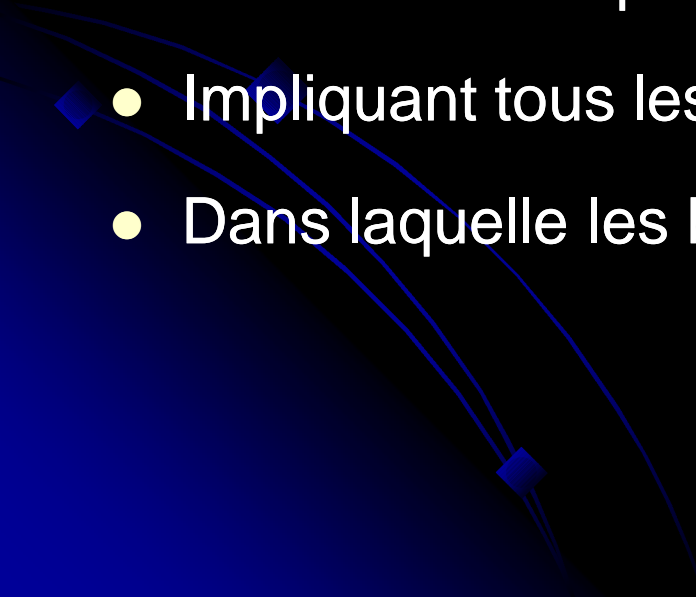
- Portails WWW
- Du prêt inter-bibliothèques traditionnel à la fourniture électronique de documents
- Didacticiels
- My Library
  - Réservations, prolongations à distance
  - Espace virtuel personnel (paniers de référence avec parfois espace électronique de travail)
  - Portefeuille électronique
- Multiplier les liens
  - FRBR (Functional Requirements for Bibliographic Records)
  - Resolveurs de liens
- «Ask a librarian » Développement de services de références virtuels intelligents

# Changement de rôles

- Passeurs de savoirs ...
    - Dans un contexte où le lien avec l'utilisateur se virtualise
      - ⇒ **Maintenir le contact pour rester en adéquation avec les besoins**
    - Aux rôles traditionnels (acquisition, conservation, consultation sur place...)
    - S'ajoutent de multiples nouveaux rôles
      - Développements informatiques de plus en plus complexes
      - Négociation
      - Communication et marketing
      - Formation
      - Évaluation qualitative
      - Participation à la production de contenus (**en particulier Open Access**)
- 

# Open Access

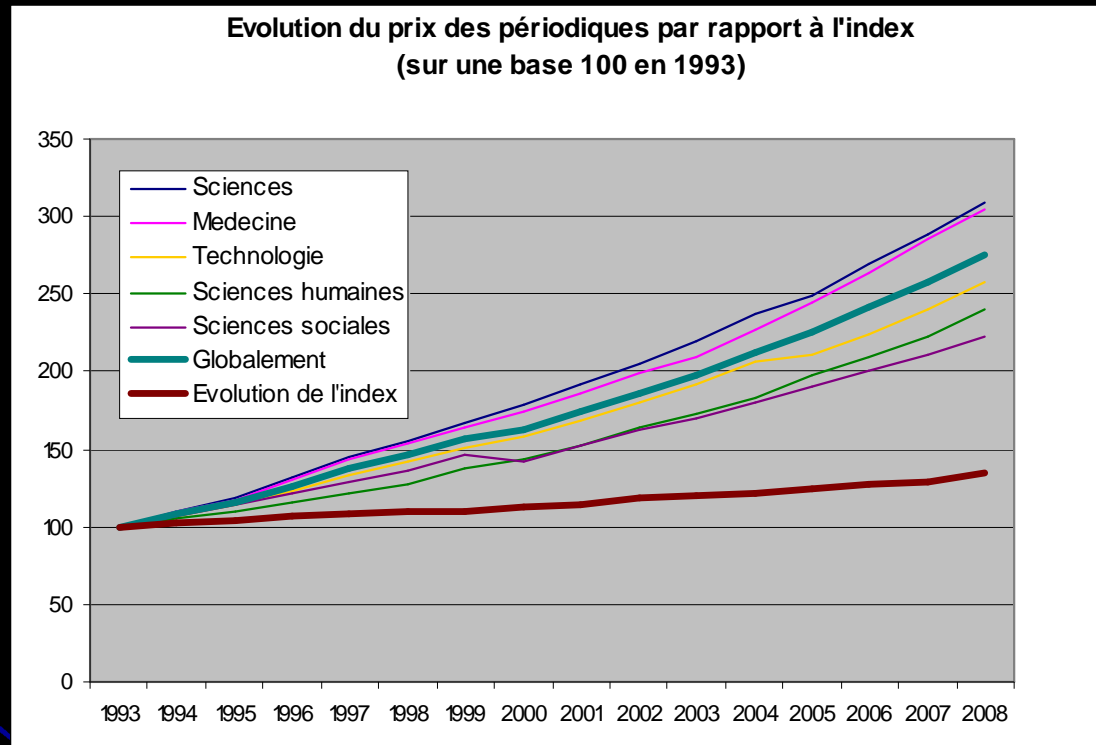
## Mutation profonde

- Une véritable remise en cause du modèle traditionnel de communication scientifique
  - Débutant dans les années 90, avec une forte accélération depuis les années 2000
  - Impliquant tous les acteurs
  - Dans laquelle les bibliothécaires sont très impliqués
- 

# Open Access : facteurs déclenchants

- Développement des NTI : appropriation aisée
- Besoin d'accélérer la diffusion de l'information
- Souci éthique : logique de partage (open source...)
  - Vers les pays émergents
  - Ce qui est financé par des fonds publics doit être librement accessible à tous
- Volonté de réagir au dépouillement de l'auteur de ses droits
- Mais surtout réaction à la « Serial crisis »
  - Lecteurs captifs : absence de concurrence
  - Utilisateurs ≠ payeurs
  - Tyrannie du facteur d'impact (IF)
  - Du mécénat à la poule aux œufs d'or : concentration massive des éditeurs

# Serial price crisis



- Moyenne basée sur plus de 70.000 titres (aussi bien éditeurs commerciaux que sociétés savantes)
- Prix moyen : presque triplé alors que l'index belge des prix à la consommation n'augmentait que de 34,04% sur la même période.
- Touche plus dramatiquement les sciences et les sciences médicales.
- Mais les sciences sociales et humaines sont loin d'être épargnées...

# Le poids des mots ...

Exemple : coût des périodiques pour l'Ulg :

La documentation :  
un luxe ...  
... mais aussi un minimum vital !

2 milliard

... sans mêm

L'équivalent d'une nouvel



# Un équilibre ... presque parfait

## Les chercheurs

- Pensent
- Cherchent
- Trouvent
- Écrivent des articles
- Les corrigent
- Paient pour être publié
- Coordonnent les revues
- Relisent et valident les articles
- Paient les revues
- Lisent les articles
- ...

## Les éditeurs commerciaux

- Impriment
- Diffusent
- ... et reçoivent l'argent



# OA, au fond, c'est quoi ?

## Selon le BOAI, 2002

« Par "accès libre" à cette littérature, nous entendons sa mise à disposition **gratuite** sur l'Internet public, permettant à tout un chacun de lire, télécharger, copier, transmettre, imprimer, chercher ou faire un lien vers le texte intégral de ces articles, les disséquer pour les indexer, s'en servir de données pour un logiciel, ou s'en servir à toute autre fin légale, **sans barrière financière, légale ou technique** autre que celles indissociables de l'accès et l'utilisation d'Internet.

La seule contrainte sur la reproduction et la distribution, et le seul rôle du copyright dans ce domaine devrait être de garantir aux auteurs un contrôle sur **l'intégrité** de leurs travaux et le droit à être correctement **reconnus et cités**. »

**ATTENTION !**

**Libre ≠ gratuit**

# Deux fronts complémentaires

## Publication de documents scientifiques directement « Open Access » (la voie d'or)

- Les « e-prints » (documents non publiés par ailleurs). Critères de qualité ?
- Les revues scientifiques (même critères de qualité)
  - Soit anciennes revues changeant de modèle commercial ex. *Nucleic Acids Research* (IF: 6,3)  
Soit nouvelles revues ex. *Genome Biology* (IF : 7,17), *PLOS Biology* (IF : 14,10)...
  - Voir **DOAJ** (<http://www.doaj.org/>) ou **Open J-Gate** (<http://www.openj-gate.com/>)
    - 65,9% des revues OA peer reviewed (39,1% pour l'ensemble des revues scientifiques)
    - Malgré leur récence 10,8% des revues OA ont déjà un IF (11% pour l'ensemble des revues)
  - Financement par des institutions ou organismes de financement de la recherche ou modèle économique inversé (BioMed Central).  
**Attention au « modèle Hybride »**



## Auto-archivage (la voie verte)

- Répertoires (archives) institutionnels ou thématiques (ArXiv, CogPrints, Pubmed...)
- Plus de 1100 dans le monde. **Interopérabilité** (voir <http://www.opendoar.org/>)
- 67% des 378 éditeurs les plus importants en taille acceptent déjà l'auto-archivage sous une ou plusieurs formes (voir <http://www.sherpa.ac.uk/romeo.php?stats=yes>)



Actuellement, on estime que **20%** de la littérature scientifique mondiale est déjà en OA !

- 15% par la voie verte
- 5% par la voie dorée

# Exemples d'initiatives OA

## 1) PoPuPs



- Portail de Publication de Périodiques Scientifiques de l'Académie Wallonie-Europe (ULg & FUSAGx)
- Diffusion Open Access au format XML et/ou pdf
  - Revues existantes (7 actuellement)
  - Revues nouvelles entièrement WWW (1 + 1 en préparation)
- Interopérabilité



- Globalement actuellement : plus de 550 articles accessibles
- Taux de consultation près de 600 articles consultés par jour (hors robots...)
- Origine : 68% Google; 19% navigation dans PoPuPs

# Exemples d'initiatives OA

## 2) BICTEL/e



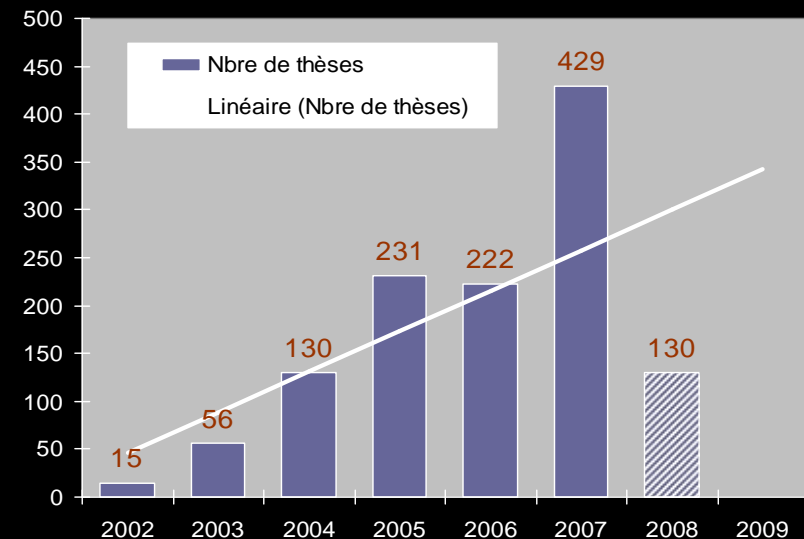
- Projet initié en 2003 par la BICfB
- Dépôt des thèses de doctorat sous forme électronique
- Une instance par université
- Interopérabilité :
  - Intégration dans ND LTD (Network Digital Library of Theses and Dissertation)
  - et dans DART-Europe E-theses Portal (plus de 80000 thèses européennes électroniques)
  - ...
- Actuellement : plus de 1200 thèses accessibles en tout ou en partie

(exemple pour ULg : 63% des thèses déposées : accessibles en totalité)

- ...Une consultation massive

Exemple ULg pour les thèses accessibles :

- Plus de 1500 fichiers téléchargés/semaine (= ± 80000/an)
- Par plus de 1100 machines différents
- Chaque fichier, en moyenne 11 fois/semaine
- Origine : 70,6% Google



# Exemples d'initiatives OA

## 3) Archives/répertoires institutionnels



- Initiative BICfB
- Phase pilote : 2005-2008
- Cadre commun (spécifications communes en vue d'une interopérabilité...)
- Avec un certain nombre de degrés de liberté...

En particulier : politiques institutionnelles différentes

- Quels contenus ?
- Obligatoire/optionnel ?

La question du mandat

# Mandat RI

- **Facultatif**

- Les « mordus »
- Croissance lente et anecdotique (max 10-12% de la production scientifique)
- Peu d'impact réel sur la communication scientifique

- **Obligatoire**

- Possible au niveau d'un organisme important de financement de la recherche (NIH, Wellcome Trust...)
- Difficile pour une institution seule (réaction des chercheurs...)

- **« Immediat deposit - Optional access » (Mandat ULg)**

- Archivage obligatoire sur un serveur institutionnel **dès acceptation de publication**
- Accès ouvert (OA) chaque fois que possible/autorisé
- Accès fermé dans les autres cas (avec système de demande de tirés-à-part)
- Concerne **tous** les articles de périodiques

*« Le seul moyen de contrer le Cheval de Troie de l'Open Choice, c'est de rendre obligatoire le dépôt institutionnel, et optionnelle la mise en accès libre »*

Bernard Rentier,  
mars 2007 (blog)



# Exemples d'initiatives OA

## 3) Répertoires institutionnels



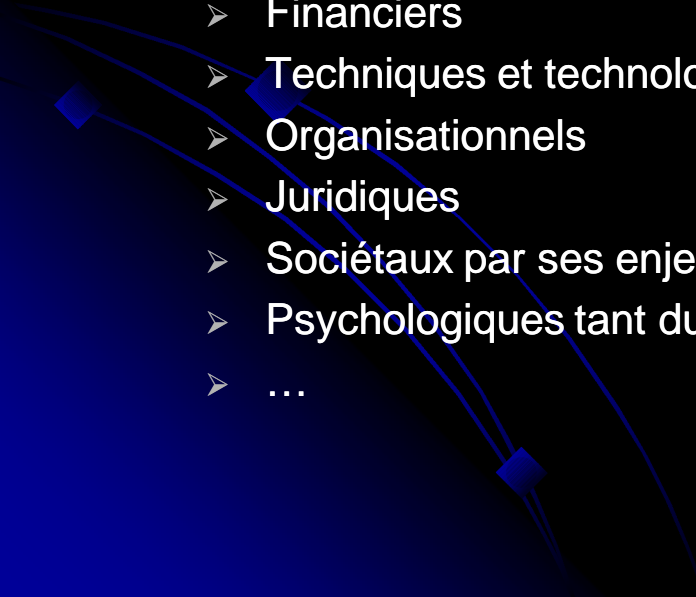
ULg : **ORBi**



Développé par le Réseau des Bibliothèques

- A la fois
  - Bibliographie académique complète (environ 3000 références/an)
    - Tout type de document publié par au moins un membre ULg
    - Au moins depuis 2002
  - Répertoire institutionnel OA
    - Obligatoire pour TOUS les articles de périodiques depuis 2002 (+ Thèses BICTEL/e)
    - Optionnel pour le reste
- Dépôt par les auteurs eux-mêmes
- Incitant : seule source de référence pour les évaluations de CV
- Mise en production : mai 2008. Objectif : 1500 full text en 2008
- Développements de nombreuses aides à l'encodage (intégrations diverses) et de services à valeur ajoutée

# Pour conclure...

- Les axes de développement de la bibliothèque numérique sont multiples
  - Répondant à une aspiration profonde des différents acteurs (multiplier, faciliter et accélérer les accès à l'information)
  - Impliquant de nombreuses et profondes transformations des rôles et des fonctions de la bibliothèque et des bibliothécaires
  - Mais aussi des freins puissants
    - Financiers
    - Techniques et technologiques
    - Organisationnels
    - Juridiques
    - Sociétaux par ses enjeux (pérennité d'accès, liberté d'accès...)
    - Psychologiques tant du côté des usagers, des auteurs que des bibliothécaires
    - ...
- 

# Pour conclure...

- Pendant longtemps encore, la bibliothèque traditionnelle ne sera pas remplacée par la bibliothèque électronique. Elle la complétera
- Pour la suite, de nouveaux modèles sont à inventer
  - Des espaces probablement très centralisés de conservation
    - Du patrimoine historique « papier »
    - De l'information électronique
  - Des outils de sélection et de diffusion de l'information scientifique, de développement de services ET de formation
  - Mais aussi de participation et de soutien à la production de l'information scientifique
- Le chemin à parcourir reste considérable et les objectifs, ambitieux
  - Il nécessitera l'acquisition de nombreuses nouvelles compétences
  - Et le développement de multiples collaborations et synergies
- **La révolution est en route, elle court ! Soyons patients**