

Journée RERO

Mardi 23 mai 2006- Lausanne

Alexandre Moatti

Secrétaire Général du Comité de Pilotage français
en vue de la création de la « Bibliothèque
Numérique Européenne »



A spiral-bound notebook with a light beige, textured cover and a silver metal spiral binding on the left side. The notebook is open to a blank page with faint horizontal lines. The text "1. Documents et références" is centered on the page in a dark brown, serif font.

1. Documents et références

Le Comité de pilotage BNUE

créé au Journal Officiel du 12 juillet 2005.

Réunions le 13 juillet, 30 août, 17 octobre 2005 et 11 janvier 2006.

composition: Editeurs (SNE, Editis, Documentation française), Industriels (Thomson), Administration (Ministères de la Culture, de l'Education et de la Recherche, de l'Industrie)

présidé par le Ministre de la Culture et vice-présidé par le président de la BnF.

Le Comité de pilotage BNUE (2)

Secrétariat général:

- Alexandre Moatti (Ingénieur en chef des Mines) SG^{al} depuis le 30 août 2005.
- Valérie Tesnière (BnF) SG^{al} adjointe depuis le 30 août 2005.
- Jean-Manuel Bourgois (SNE Syndicat National de l'Édition/ Albin Michel) rapporteur depuis le 20 février 2006.

Documents de référence

- 📄 Rapport du comité de pilotage remis au Ministre de la Culture et au Président de la République (31 janvier 2006):
 - www.bibnum.net
- 📄 Décisions Conseil des Ministres du 8 février 2006:
 - [\(lien\)](#)
- 📄 la présentation qui suit porte sur le travail effectué de septembre 2005 à mars 2006.

Documents de référence (2)

📄 Le 2 mars la Commission Européenne a rendus publics les résultats de sa consultation en ligne (réponses du 20 janvier):

- [Communiqué de presse](#)
- [Synthèse des réponses](#)

📄 Le 27 mars première réunion du « groupe d'experts de haut niveau » conseillant la Commission:

- [lien](#)

A spiral-bound notebook with a light beige, textured cover and a silver metal spiral binding on the left side. The notebook is open to a blank page with faint horizontal lines. The text is centered on the page.

2. Éléments de contexte

Un environnement en effervescence depuis septembre 2005

📄 Annonce Google 13 décembre 2004.

📄 Une course à l'indexation des contenus de qualité est lancée entre acteurs majeurs du Web depuis septembre 2005:

- Yahoo, consortium Open Content Alliance
www.opencontentalliance.org septembre 2005
- Accord MSN/ British Library nov. 2005 (100 000 livres sur 2006)
- Accord Google/ Library of Congress nov. 2005

Un environnement en effervescence depuis septembre 2005 (2)

Projet privé allemand Libraires + Editeurs:

- sous la houlette de l'éditeur M. Ullmer
- « Volltextsuche online » (recherche plein texte)

Au bout d'un an, fin 2005, relativiser les résultats effectifs de Google Book Search:

- 20 000 livres en un an, dont 14 000 provenant des éditeurs (contenus électroniques natifs, pas de numérisation).
- Éditeurs francophones: De Boeck (Belgique) notamment.
- Et donc 6 000 provenant d'une numérisation spécifique Google avec les bibliothèques partenaires (période 1800-1870)

Autres éléments de contexte

📄 BNUE est différent du projet franco-allemand Quaero:

- Quaero est un projet à moyen terme de moteur de recherche multimedia.
- La BNUE a besoin d'un moteur de recherche (le cas échéant avec sémantique et linguistique améliorées), mais pas d'un nouveau moteur.
- Bien distinguer la fonction technique « moteur » des fonctions « page d'accueil drainant le public » et « indexation ».


Autres éléments de contexte (2)

☞ Une coopération BNUE avec les grands moteurs (Google, MSN, Yahoo,...) est nécessaire:

- Ce sont eux qui drainent le public et indexent.
- Appétence de leur part pour les contenus francophones (public canadien notamment)
- Réponse commune des BN européennes aux sollicitations de ces entités?

☞ La France est en avance avec ses 80 000 documents Gallica en ligne:

- 1 500 000 documents consultés par mois
- mais numérisés en mode image, non indexables

A spiral-bound notebook with a light beige, textured cover. The spiral binding is on the left side. The text is centered on the page.

3. Définition possible d'une BNUE (suivant travaux du Comité de pilotage français)

Que pourrait être la BNUE?

- ☞ Une « *bibliothèque des savoirs* » tournée vers le plus large public possible, encore à conquérir.
- ☞ Une dynamique permettant l'association de partenaires publics et privés: éditeurs notamment.
- ☞ Une interface de recherche commune à des contenus sous droits et hors droits:
 - Contenus hors droits (avant 1935): numérisation par les BN, *a priori*.
 - Contenus sous droits après 1990: ils sont déjà sous format électronique.
 - Contenus sous droits 1935-1990: qui numérise?

Quels contenus pour une BNUE?

☰ Des axes thématiques définissant l'identité européenne:

- le mouvement de la pensée et des sciences en Europe.
- le développement des démocraties en Europe (politique, droit, histoire...)
- le patrimoine culturel et linguistique : identités européenne/ identités nationales.
- les territoires, sources de l'enracinement des individus et des sociétés.


☰ Des axes « européens » permettant de rassembler les contenus nationaux:

- sans perte des contenus francophones essentiels.

Ce qui n'est pas dans la BNUE

Articles de recherche récents:

- le débat éditeurs spécialisés/ « Open access » pour la recherche est hors sujet.
- Contenus non adaptés aux publics visés (voir cas particulier des revues SHS).
- La Commission européenne traitera séparément en 2006 ce sujet « information scientifique et technique»

 Une grande part de la littérature générale récente.

A spiral-bound notebook with a light beige, textured cover and a silver metal spiral binding on the left side. The notebook is open to a blank page with faint horizontal lines. The text is centered on the page.

4. Quelques choix importants

Portail BNUE ou site BNUE?

☞ Portail ouvrant vers les sites des BN et bibliothèques de chaque pays:

- la consultation du document se fait dans l'environnement du site propriétaire du document.

☞ Site et consultation propres à la BNUE:

- la BNUE est en effet une bibliothèque de consultation.
- Format unifié de visualisation des documents: mode texte et image combinés (p.e. format XML Alto), surbrillance jaune des résultats de recherche.
- Interface BNUE analogue à interface Google Book.

Numériser et/ou rendre accessible?

📄 En France l'enjeu prioritaire en 2006 est l'« OCRisation » des contenus Gallica.

📄 Au-delà se pose la question de la taille idéale d'une bibliothèque numérique:

- le site BNUE doit être connu du grand public, priorité à donner à l'accès, à la notoriété du site, à l'indexation des contenus.
- Équilibre des dépenses à trouver entre notoriété du site vs. Numérisation intensive.

📄 Test de numérisation de masse mené par la BnF septembre 2006: 0,1 à 0,2 € la page?

Associer les éditeurs en amont

- ☞ L'originalité d'une BNUE pour le grand public est la possibilité de chercher *aussi* dans des contenus sous droits (éditeurs):
 - Ce ne doit pas être qu'un projet patrimonial
 - seul moyen de rivaliser avec GBS, MSN Book.
- ☞ Modèle économique à définir pour la partie payante de la BNUE.
- ☞ Mutualisation d'un service de bibliothèque numérique pour des éditeurs indépendants.

Options actuelles Commission

La Commission a annoncé qu'elle s'appuierait sur www.theeuropeanlibrary.org:

- portail ouvrant vers les BN, pour l'instant pas un site propre avec format de visualisation unifié.
- Site TEL à mieux faire connaître du grand public.
- La communication de la Commission (2 mars) ne mentionne pas pour l'instant les contenus sous droits

Plan de route - France

(Conseil des Ministres 8 février)

Fonds numérique Gallica (BnF):

- OCRisation de 60 000 documents sur 80 000.
- Numérisation de masse en 2006, 40 000 nouveaux documents.

Travail commun avec les éditeurs: intégration de leurs contenus, modèle économique.

Fédération d 'une ou deux BN européennes intéressées à aller rapidement.

Une priorité- deux points connexes

☞ La mise en ligne et la bonne indexation des contenus francophones est une priorité absolue.

☞ Deux points connexes importants:

- la préservation à long terme des contenus numériques sur support informatique.
- Le dépôt légal *via* Internet.

A spiral-bound notebook with a light beige, textured cover and a silver metal spiral binding on the left side. The notebook is open to a blank page.

5. Démonstration rapide & indicative (si accès internet)

www.bibnum.org

- 📄 OCRisation à partir de numérisation directe.
- 📄 OCRisation à partir de numérisation microfiche ou microfilm.
- 📄 Editeur (Odile Jacob), contenu électronique natif.
- 📄 « Mise en page » d'un document en mode texte.
- 📄 Moteur de recherche et surbrillance des résultats recherchés.

www.bnue.org

(si en fonctionnement)

📄 recherches croisées dans des documents d'origine diverses:

– Gallica, Documentation française (discours Ministres), portail SHS Cairn, 1 livre éditeur, Numdam

📄 y.c. documents mode image dans la même visualisation (Champollion MOM)

6. Coordonnées

alexandre.moatti@mines.org

www.bibnum.net (rapport BNUE)

www.indispensables.net (blog sciences)