

Le cas de l'université de Lille 1 : Grisemine

Jean-Bernard MARINO
SCD de l'université de Lille 1

Note : compte-rendu rédigé par l'enssib à partir d'une transcription de la communication orale de Jean-Bernard Marino du 16 novembre 2004.

Résumé :

Cette communication présente Grisemine, la bibliothèque numérique de littérature grise développée par l'université Lille 1 depuis septembre 2001. Elle décrit les modes de traitement dans Grisemine (DTD Marc, langages documentaires) et un bilan de trois ans de mise en ligne, à la fois en terme de contenus et d'utilisations. Les facteurs qui poussent à faire évoluer Grisemine seront également présentés – en particulier tous ceux qui touchent au développement de la FOAD (passage du Marc au Dublin Core et utilisation de Dspace pour traiter toutes les ressources pédagogiques et documentaires numériques). Le développement du partenariat bibliothèque numérique/TICE se dessinera à travers ces grandes lignes d'évolution de Grisemine.

Mots clés :

indexation, ressources pédagogiques numériques, normes et standards pédagogiques, Grisemine, Dublin Core, Dspace, circuit du document

Présentation générale de Grisemine

Grisemine est une bibliothèque numérique de littérature grise universitaire¹, possédant près de 1 097 documents, dont 411 thèses et 96 cours.

Les accès à Grisemine interviennent essentiellement pendant les heures d'ouverture de la bibliothèque universitaire, mais 4 accès sur 10 sont effectués en dehors – ce qui peut constituer un élément de réponse partiel à la question de l'amplitude horaire des bibliothèques universitaires.

Il est intéressant par ailleurs de noter que les documents les plus fréquemment consultés sont les cours, à raison de quinze fois par mois en moyenne. Viennent ensuite les thèses, chacune étant consultée en moyenne huit fois par mois. Tous documents confondus (articles de périodiques,

¹ La présentation en a été faite en décembre 2003 au 5e colloque international sur la littérature grise (GL5), Amsterdam, 4-5 décembre 2003. Texte disponible sur : http://archivesic.ccsd.cnrs.fr/sic_00000892.html.C

rapports de recherche, documents de travail), la moyenne se situe à 5,5 consultations par mois. Il est intéressant de comparer ces chiffres avec les activités de consultation des documents en magasins, où le rapport est de 0,01. Ce qui veut dire qu'un document présent dans la bibliothèque numérique trouve une vie beaucoup plus active que le même document disponible dans un magasin de la bibliothèque.

Evolutions

Grisemine donnera lieu dans le courant de l'année 2005² à une autre application qui prendra le relais, avec d'autres types documents. Trois facteurs rendent nécessaire ce changement.

Les facteurs locaux

Les concepteurs de l'outil utilisé, CinDoc, ne disposent pas du temps nécessaire pour faire évoluer la version actuelle. Par ailleurs, l'université de Lille souhaite que tous les contenus soient désormais validés. Enfin, l'évolution attendue devrait permettre une meilleure adaptation aux applications des TICE.

Les facteurs documentaires

Grisemine est actuellement monolingue français mais une diversité des langues sera réalisée. Ensuite, les contenus seront diversifiés, la seule littérature grise ne suffit pas, et les mécanismes de recherche dans la base seront optimisés afin de se rapprocher des modes de recherche standardisée auxquels les gens sont habitués dans Google, Yahoo, etc.

Les facteurs liés aux contextes de l'information ouverte

Tout d'abord, il est choisi de suivre la tendance générale de l'utilisation des logiciels libres et de mettre en place le protocole OAI-PMH³ permettant le moissonnage des métadonnées pour les archives ouvertes, afin d'entrer dans la danse internationale de l'OAI (*Open Archives Initiative*). Enfin, la qualité de structuration des signalements doit être améliorée, par l'utilisation du Dublin Core notamment.

Le point principal réside dans l'adoption d'applications relevant des TICE et de la FOAD⁴. La mission de Grisemine ne se limite pas à la constitution d'une bibliothèque numérique en accompagnement des TICE, mais plus largement au service des études et de la recherche.

² Le projet n'a pas encore vu le jour en décembre 2005.

³ OAI-PMH : Open Archives Initiative - Protocol for Metadata Harvesting.

⁴ FOAD : Formation Ouverte et A Distance

L'harmonisation des approches au sein de l'université

Pour ce faire, les bibliothécaires doivent s'efforcer de comprendre les schémas mentaux de leurs partenaires des TICE et d'admettre leurs contraintes. Ces métiers ne se recouvrent pas tout à fait. Certaines compétences sont du ressort des TICE ; d'autres, du ressort des bibliothèques, peuvent leur être utiles. Les promoteurs des TICE doivent pouvoir s'appuyer sur le savoir-faire des bibliothèques.

Par exemple, la démarche de numérisation en région avec la charte des TICE est vue comme une activité nouvelle par la communauté universitaire, alors que pour la bibliothèque, il s'agit d'une suite logique : après la bibliothèque physique sur le campus physique, vient la bibliothèque numérique sur le campus numérique.

La mise en ligne de contenus est perçue comme une tâche nouvelle par l'université, ainsi que la nécessité de l'indexation liée à la multiplication des objets. De nouveaux critères d'indexation doivent donc être pris en compte.

Les TICE se préoccupent beaucoup d'interactif alors que la bibliothèque n'a pas cette culture, elle doit donc appréhender ce nouveau type d'information. Ses partenaires sont prêts à apporter des cours, des supports de cours, des documents annexes conçus par un enseignant ou un groupe d'enseignants. Les bibliothécaires ont l'habitude d'ouvrir la bibliothèque et d'y intégrer tout document qui vient en accompagnement de l'enseignement et de la recherche. Une coopération est donc nécessaire, pour apporter un volet documentaire validé élargi par ces différents interlocuteurs, et renoncer à toute tentative d'exclusivité. Il s'agit de mutualiser pour mettre en ligne.

Le dépassement de la seule littérature grise

Les documents didactiques sous forme numérique sont assez mal distribués selon les disciplines : certaines en sont très riches, d'autres pauvres. L'offre doit donc être élargie. Il faut penser à ajouter des textes de base, des documents beaucoup plus fragmentaires, des plans, des lexiques, des tables comme le tableau de Mendeleïev pour les chimistes, des constantes physico-chimiques, des annales d'examens, etc. – quitte à les numériser a posteriori. Enfin, il faut aussi tenir compte des différents niveaux, même si nous disposons de peu de documents aux niveaux master et doctorat.

L'université attend aussi une différenciation des droits d'accès qui doit donc être prévue par l'application : les bénéficiaires (ceux qui ont payé les cours par correspondance), ceux qui font partie de l'annuaire interne et ceux qui n'en font pas partie. Des périmètres concentriques se dessinent alors, qui partent du plus large au plus étroit : internet, intranet universitaire et intranet des bibliothèques universitaires.

La granularité des contenus

La granularité des contenus préoccupe les bibliothécaires depuis longtemps, notamment lorsqu'ils font le choix d'indexer l'unité documentaire la plus fine pour sa pertinence et lorsqu'elle correspond aux habitudes de recherche actuelle, en particulier dans les moteurs de recherche. Par exemple dans les périodiques, c'est l'article qui est indexé ; dans le livre, c'est le chapitre, l'annexe, le lexique, la carte ; pour un colloque, ce n'est pas le compte-rendu de colloque, mais la communication. L'indexation se fait avec un signalement très « sain », avec le titre, le titre traduit s'il est traduit, le résumé, les notes de contenu, c'est-à-dire les parties principales de l'article, les têtes de paragraphes. Ainsi, la recherche pourra s'effectuer non seulement sur les métadonnées réalisées par le bibliothécaire (Rameau, etc.) mais aussi sur les mots du chercheur. La combinaison des deux possibilités multiplie les chances de retrouver un contenu pertinent.

Les lignes d'évolution du projet

Grisemine doit évoluer. Le projet a démarré avec une DTD Marc, il doit maintenant passer dans le monde du Dublin Core, avec ses 15 champs principaux.

Grisemine propose un prototype à 48 champs, qui fonctionne très bien ; le catalogage est facile et rapide. Des champs peuvent être ajoutés car Dspace permet de répondre à toutes ces contraintes. Son choix, effectué par la bibliothèque, a permis de se mettre dans des conditions favorables à un dialogue avec les interlocuteurs des TICE.

L'analyse des besoins a fait ressortir les points suivants :

- le système doit pouvoir être alimenté par n'importe quel type de documents ;
- il doit permettre la différenciation des droits d'accès et le traitement des documents au niveau le plus fin ;
- il doit inclure la fonction de validation par un responsable de comité de lecture ;
- il doit être fondé sur le Dublin Core ;
- il doit répondre aussi bien au système de dépôt des universités (repository) qu'aux besoins des bibliothèques numériques des bibliothèques universitaires ou SCD.

Le choix s'est donc porté sur DSpace⁵, application libre conçue au MIT⁶ pouvant être francisée. L'application est composée de deux modules, Ariane et Iris. Le module Ariane est destiné aux chercheurs : il s'agit d'une « boîte vide » mise à la disposition des chercheurs pour y déposer leurs documents validés. Le module Iris est destiné à la bibliothèque numérique du SCD, alimentée

⁵ <http://www.dspace.org>

⁶ MIT : Massachusetts Institute of Technology, <http://ocw.mit.edu/index.html>

régulièrement par des métadonnées produites par les bibliothécaires et intégrant aussi des éléments d'Ariane, selon le plan de développement de la bibliothèque.

DSpace est disponible tout à fait librement ; il est utilisé au MIT, à Cambridge et Cornell University, ainsi que dans certaines universités européennes, notamment au Pays-Bas et en Italie. Enfin, il est compatible avec la classification Dewey, puisque des classes principales et des sous-classes Dewey peuvent être créées pour remplacer les catégories et les sous-catégories visibles en bas de l'écran (figure 1).

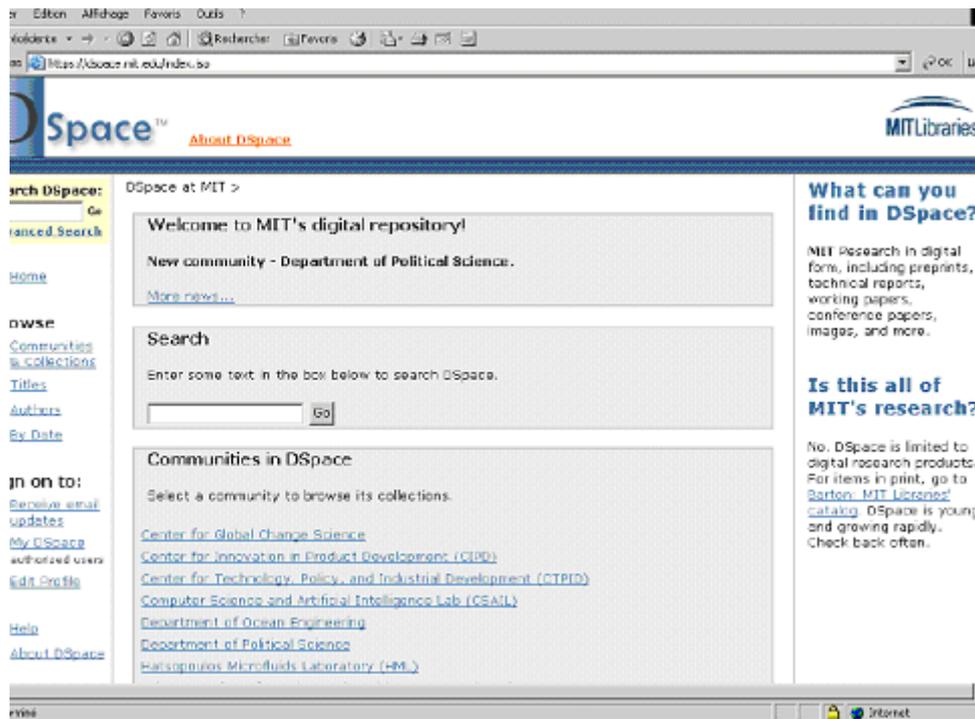


Figure 1 : Site internet de DSpace