

L'indexation des ressources pédagogiques numériques : questions transversales

Ghislaine CHARTRON
INRP et URSIDOC/enssib, Lyon

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

Résumé :

Le contexte récent du développement de l'enseignement à distance et, de façon plus générale, la mise à disposition de contenus numériques d'enseignement et de recherche renouvellent à la fois les méthodes pédagogiques, les rapports enseignants-apprenants et les processus organisationnels liés à la gestion des documents dans les universités.

À l'appui de quelques observations et de quelques considérations comparatives, nous nous attacherons à interroger le processus d'indexation des ressources pédagogiques à trois niveaux.

Tout d'abord en identifiant les besoins repérés pour cette indexation : il n'est pas sûr que l'on puisse parler d'« une » indexation mais de plusieurs niveaux d'indexation avec des finalités diversifiées, variant de surcroît selon les contextes d'usage de ces documents. Le risque d'une incompréhension mutuelle de ces différents niveaux est latent.

En quoi l'indexation de ressources pédagogiques numériques diffère-t-elle de l'indexation d'un autre document numérique ? Cette interrogation nous amènera à revenir sur la question des normes et standards et de leurs spécificités concernant le document pédagogique.

Enfin, la question des acteurs et de leur rôle respectif sera introduite : quels rôles partagés des auteurs-enseignants, des bibliothèques et des ingénieurs-informaticiens pour cette indexation ? Mais aussi, de façon plus large, l'évolution des services de médiation et le déplacement de leur valeur ajoutée seront évoqués dans le contexte du document numérique.

Mots clés :

ressources pédagogiques numériques, indexation, normes et standards pédagogiques, métadonnées, circuit du document

Introduction

Nous questionnerons tout d'abord la problématique du document pédagogique numérique par rapport à celle du document numérique dans sa généralité. Nous essaierons de montrer en quoi le document pédagogique pose et renouvelle des questions, notamment en ce qui concerne l'indexation. Puis nous tenterons d'identifier les besoins repérés pour cette indexation ou ces indexations, car il n'est pas sûr que l'on puisse parler d'« une » indexation mais de plusieurs niveaux d'indexation avec des finalités diversifiées... Nous analyserons enfin les premiers retours d'usages, afin d'essayer de déterminer si des conclusions se dégagent sur l'utilisation et les utilisateurs du LOM et sur le mouvement réel de mutualisation des ressources pédagogiques, afin de ne pas négliger la dimension sociale de ce processus.

Le document pédagogique numérique

La première question est donc de s'interroger sur la spécificité du document pédagogique numérique, par rapport au document numérique en général qui, depuis longtemps, fait évoluer les missions des bibliothèques et des centres de documentation. L'émergence du document pédagogique peut être rapprochée de la montée en puissance de l'enseignement à distance. Quatre niveaux généraux de standardisation caractérisent le document numérique.

L'identification

Les bibliothèques et les professionnels du document connaissent bien désormais le problème de l'identification du document numérique. Ainsi, le DOI¹ (Digital Object Identifier), mis en place par les éditeurs scientifiques internationaux depuis 2000, est devenu une norme. Il fait office de référence dans l'édition scientifique commerciale et pèse de tout son poids dans les pratiques du document numérique, lors de la recherche et notamment lors de celle d'un article. La finalité de cette norme DOI est le contrôle des droits liés à la propriété intellectuelle dans l'environnement électronique (norme ISO Z39.84). Ainsi, toutes les grandes collections de revues en ligne des éditeurs scientifiques l'utilisent pour l'achat d'articles à l'unité.

¹ <http://www.doi.org>

Le référencement, les métadonnées

En ce qui concerne les métadonnées, la question centrale est de savoir si la version Dublin Core DC Education n'est pas suffisante. Le LOM est-il vraiment nécessaire ? Qu'apporte-t-il de plus que le DC Education ? Le Dublin Core est un standard largement utilisé. Sa finalité est de décrire des documents numériques pour la recherche d'informations. L'idée était d'indexer très rapidement des ressources web.

La TEI² (Text Encoding Initiative) est un autre type de métadonnées qui concernent plutôt le secteur des sciences humaines, sociales et linguistiques. À son origine, se trouvaient des associations comme l'Association for Computers and the Humanities, l'Association for Computational Linguistics, l'Association for Literary and Linguistic Computing. Cette norme date de 1987.

Le format MARC propose aussi une extension du format Marc de catalogage.

Les protocoles d'échanges

En termes de protocole d'échange, certaines innovations fonctionnent bien, comme l'OAI (Open Archive Initiative). Cette initiative a été créée les 21 et 22 octobre 1999 autour des archives ouvertes, avec Carl Lagoze et Herbert Van de Sompel (Cornell University / Computer Science). Sa finalité est de proposer un ensemble des protocoles communs liés aux métadonnées (avec le formalisme du Dublin Core) et au processus de recherche d'information. À l'origine, seules les bases d'E-prints étaient concernées, mais l'initiative dépasse maintenant le cadre des archives ouvertes et s'est maintenant élargie à d'autres types de contenus numérisés (thèses, catalogues...). Des éditeurs scientifiques comme Elsevier ont même adopté le protocole OAI pour leur moteur Scirus³. Gallica⁴ utilise aussi ce protocole pour l'accès à son contenu numérique. Ces pratiques d'accès numériques se sont largement répandues aujourd'hui et ne sont plus des utopies informatiques.

La structuration du contenu

Le problème de la structuration de contenu est beaucoup plus complexe. Il existe des DTD spécifiques pour les archives, comme l'EAD (Encoded Archives description), qui a des finalités très précises concernant les instruments de recherche de type archivistique : inventaires, répertoires, catalogues de collections. Les Archives de France ont adopté cette norme et développé une

² <http://www.tei-c.org>. Liste de projets TEI : <http://www.tei-c.org/Applications/index.html>, 117 projets.

³ <http://www.scirus.com/>

⁴ Gallica, bibliothèque numérique de la Bibliothèque nationale de France : <http://gallica.bnf.fr>

vingtaine de projets environ, avec l'utilisation d'un éditeur spécifique tel que XMetal.

Utilisation réelle des normes et des standards

Normes et standards sont donc bien utilisés dans ce qui relève des processus d'identification, des interfaces d'accès, des protocoles d'échanges et de description. Par contre, la normalisation atteint ses limites dès qu'il s'agit des contenus (DTD spécifique de structuration) de par les contraintes imposées à l'utilisateur. Des freins sociaux apparaissent dès que s'entame une normalisation des contenus ou des activités sociales comme la pédagogie – très liés à la production de l'individu.

L'indexation du contenu du document pédagogique

En ce qui concerne l'indexation du document pédagogique, les enjeux sont importants. Actuellement, leur description et leur normalisation concernent le contexte pédagogique d'usage, c'est-à-dire le niveau, les méthodes et les évaluations. Cela peut être conflictuel, car très enraciné dans les pratiques culturelles de chacun. Un standard français LOM, ne serait-ce qu'au niveau national, pour l'ensemble des universités, risque de rencontrer beaucoup de difficultés, en raison de pratiques très différentes selon les communautés scientifiques. Il faut donc rester raisonnable dans la rationalisation.

L'autre enjeu de l'indexation des documents pédagogiques se situe dans la description du contexte technique (plate-forme, formats, liens avec d'autres objets) pour la réutilisation dans la conception. Ce point est nouveau et peut engendrer des confusions autour de l'indexation. Quand les bibliothèques parlent d'indexation, elles évoquent des unités sémantiques, du sens – mais pas le plus petit niveau de granularité. Ce changement d'échelle permet de comprendre les conflits latents ou les mécontentements qui peuvent apparaître entre les acteurs autour de l'indexation. De même, la description technique sort également du champ de la description du contenu. Ainsi l'indexation d'un contenu pour sa réutilisation sur différentes plates-formes appartient au monde de l'informatique. D'où l'importance de la répartition des rôles autour de l'indexation et de la définition de celle-ci.

Pourquoi indexer ?

Alors pourquoi indexer aujourd'hui ? Du côté des bibliothèques et des centres de

documentation, l'indexation, vise les usagers qui vont utiliser les contenus. Elle se fait par rapport à des thésaurus, des vocabulaires normalisés ou des plans de classement, mais n'est pas destinée à des concepteurs.

Si l'indexation doit varier en fonction du public, il faut donc savoir à quels usagers elle est destinée. Ceux-ci peuvent être des enseignants ou des étudiants. Un récent mémoire de DEA⁵ montre que les données, les cours, les ressources pédagogiques qui sont aujourd'hui réutilisées le sont essentiellement par des étudiants et non par des enseignants.

L'indexation pour des concepteurs⁶ a un objectif différent : la réutilisation de cours ou d'éléments de cours (image, vidéo, exercice...).

Enfin, un troisième type d'indexation concerne la gestion des documents dans les organisations, leur cycle de vie, qui va de l'acquisition à la gestion des droits d'accès, en passant par les différentes versions et exemplaires, l'archivage, la mise à disposition, etc.

Indexations : plusieurs acteurs ?

Il y a donc plusieurs indexations des ressources pédagogiques.

L'accès, c'est-à-dire le repérage d'un contenu de cours, relève des métadonnées, à peu près l'équivalent d'une notice bibliographique. Les bibliothèques et les centres de documentation sont les plus compétents pour ce type d'indexation.

L'indexation qui relève de la conception pédagogique incombe à l'auteur, l'enseignant, qui connaît son objectif pédagogique et le contexte d'utilisation de son cours. Difficile de se substituer à cet auteur...

L'indexation qui rend compte de la conception technique, notamment en vue de la réutilisation sur d'autres plates-formes, est vraiment une compétence des cellules TICE.

Enfin, la question du cycle de vie du document, des versions, des droits relatifs au document relève aussi des bibliothèques et des centres de documentation.

Réalité de la mutualisation ?

En ce qui concerne la question de la mutualisation et de l'interopérabilité, il ne faut pas négliger les freins sociaux. En effet, un enseignant réutilise rarement le cours d'un autre enseignant. Une enquête menée par une étudiante de DEA sur 17 campus français à dominante SHS au début de

⁵ Kreczanik, Thomas, « Vers une rationalisation de l'indexation des ressources pédagogiques électroniques », Diplôme d'études approfondies, Université Lyon 3, 2004. [en ligne] http://memsic.ccsd.cnrs.fr/mem_00000163.html

⁶ Se reporter à l'intervention de Laurent Flory.

l'été dernier⁷ donne les résultats suivants : les cours en ligne sont produits localement, sans partenariat extérieur ; il n'y a pas de mutualisation avec d'autres universités et peu de synergies avec les ressources numériques (sous licence et libres d'accès) mises à disposition par les bibliothèques. En conclusion, il s'agit donc d'une situation d'autoproduction des ressources pédagogiques, et non de partage – ce qui correspond peut-être à une phase préliminaire.

LOM, quelles utilisations ?

Une enquête de OCLC⁸, acteur majeur du côté des bibliothèques pour pister les innovations, a mené une étude comparative sur l'usage du LOM.

Cette étude, qui prend en compte 35 pays anglophones utilisant le standard LOM, montre que les éléments les plus utilisés du LOM sont ceux communs avec le Dublin Core, tandis que les moins utilisés sont les éléments de description pédagogiques, les éléments techniques, sophistiqués, les éléments relationnels entre objets. L'interopérabilité constatée dans cette enquête (convergence des profils) a un périmètre extrêmement limité, souvent autour d'un même pays, voire autour d'une même institution.

Conclusion

L'indexation des ressources pédagogiques doit exister, mais il convient de rester raisonnable et de ne pas négliger les facteurs de coût et de temps. Il n'est pas certain d'arriver à un dénominateur commun national, en raison des spécificités d'usages propres à chaque discipline. Les résistances doivent être prises en compte pour éviter de s'épuiser et de dépenser beaucoup d'argent sur ces programmes, sans grands résultats au final.

⁷ Baczkowski, B., « Campus numériques: ressources et contenus », Diplôme d'études approfondies DISIC, École doctorale Informatique et information pour la société, 2004.

⁸ Enquête de C. Jean Godby/OCLC, publiée dans Ariadne : <http://www.ariadne.ac.uk/issue41/godby>, octobre 2004.