



Bibliothèque numérique de l'enssib

L'indexation des ressources pédagogiques numériques : un partenariat à créer entre les SCD et les services TICE au sein des universités, 16 novembre 2004

Bertin, Gilles,
Bourda, Yolaine,
Cherhal, Elizabeth,
Gómez de Regil, Rosa María,
Joly, Monique,
Pernin, Jean-Philippe

Bertrand, Annie,
Chartron, Ghislaine,
Flory, Laurent,
Hennequin, Xavier,
Marino, Jean-Bernard,

BERTIN, Gilles, BERTRAND, Annie, BOURDA, Yolaine, et al. *L'indexation des ressources pédagogiques numériques : un partenariat à créer entre les SCD et les services TICE au sein des universités, 16 novembre 2004, Lyon-Villeurbanne* [en ligne]. Lyon : école nationale supérieure des sciences de l'information et des bibliothèques, 2004, 87 p. Format PDF.
Disponible sur : <<http://www.enssib.fr/bibliotheque-numerique/notice-1232>>

Ce document est « **tous droits réservés** ». Il est protégé par le droit d'auteur et le code de la propriété intellectuelle. Il est strictement interdit de le reproduire, dans sa forme ou son contenu, totalement ou partiellement, sans un accord écrit de son auteur.

L'ensemble des documents mis en ligne par l'enssib sont accessibles à partir du site :
<http://www.enssib.fr/bibliotheque-numerique/>

Table des matières

Préambule	3
Marianne FOLLET Elisabeth NOËL	
Partie 1 – Les enjeux de l'indexation des ressources pédagogiques numériques	4
Pourquoi indexer les ressources pédagogiques numériques ?	5
Yolaine BOURDA	
Les caractéristiques d'une ressource pédagogique et les besoins d'indexation qui en résultent	9
Laurent FLORY	
Partie 2 – Normes et standards : un processus de normalisation en cours	15
De Marc à Dublin Core	16
Élizabeth CHERHAL	
Le Dublin Core (DC)	22
Élizabeth CHERHAL	
Présentation des standards : (LOM) – Learning Object Metadata	29
Rosa María GÓMEZ DE REGIL	
Sharable Content Object Reference Metadata (SCORM)	37
Gilles BERTIN	
LOM, SCORM et IMS-Learning Design : ressources, activités et scénarios	44
Jean-Philippe PERNIN	
Les évolutions du LOM	52
Yolaine BOURDA	
L'indexation des ressources pédagogiques numériques : questions transversales	58
Ghislaine CHARTRON	
Partie 3 – Rôle et positionnement des acteurs : quels partenariats entre SCD et TICE ?	64
Le cas de l'université de Lille 1 : Grisemine	65
Jean-Bernard MARINO	

Les partenariats autour du portail documentaire de l'université de technologie de Compiègne **70**

Annie BERTRAND

Xavier HENNEQUIN

Synthèse **76**

L'indexation des ressources pédagogiques numériques : un partenariat à créer entre les SCD et les services TICE au sein des universités **77**

Préambule

Marianne FOLLET

Service TICE de l'enssib

Coordination scientifique de la journée d'études

Elisabeth NOËL

Service FORMIST / enssib

Coordination scientifique de la journée d'études

Mots Clés : enssib, normes et standards pédagogiques

Au sein des établissements d'enseignement supérieur, tous les acteurs impliqués dans les TICE et l'enseignement à distance sont en train de prendre conscience de la nécessité d'indexer les ressources pédagogiques numériques pour pouvoir les échanger.

Mais les normes et standards pédagogiques sont difficiles à appréhender, tant par les acteurs des TICE qui ne sont pas des spécialistes de l'indexation que par les bibliothécaires qui ont du mal à se repérer dans l'utilisation de ces différents standards, dont le processus de normalisation est en cours.

L'enssib, qui a pour mission de former les personnels d'encadrement des bibliothèques, se propose donc de participer à la sensibilisation des acteurs concernés par l'indexation des ressources pédagogiques numériques, en particulier au sein des bibliothèques universitaires.

C'est dans cette perspective que cette journée d'études a été organisée en novembre 2004, visant à rassembler les producteurs de ressources pédagogiques et les personnels de bibliothèques qui pourraient être amenés à participer à leur indexation, afin qu'ils examinent ensemble comment nouer des partenariats autour de cette question.

Nous sommes heureux de restituer les interventions et les débats qui ont fait la richesse de cette journée, à travers cette publication en ligne qui marque le lancement d'une nouvelle collection « d'actes en ligne » à l'enssib.

Partie 1 – Les enjeux de l'indexation des ressources pédagogiques numériques

Pourquoi indexer les ressources pédagogiques numériques ?

Yolaine BOURDA
Professeur d'informatique à Supelec, membre de la CN 36¹

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

Résumé :

Cette contribution définit les ressources pédagogiques numériques, en précisant leurs rôles et caractéristiques (interchangeables, évaluables, livrables, réutilisables, abordables, adaptables, durables, fiables, gérables...) et rappelle l'utilité d'indexer ces ressources. Les métadonnées doivent suivre cependant des normes et standards bien précis pour permettre de retrouver facilement ces ressources pédagogiques. Il s'agit aussi de décrire les principales métadonnées (Dublin Core, Dublin Core Education, Learning Object MetaData...) et d'évoquer rapidement les problèmes d'implémentation.

Mots clés :

ressources pédagogiques numériques, indexation, métadonnées, LOM (Learning Object MetaData), normes et standards pédagogiques, interopérabilité.

Nécessité d'une indexation des ressources pédagogiques

Selon la définition, très vaste, du LOM (Learning Object Metadata) d'IEEE², une ressource pédagogique correspond à toute entité (numérique ou non) utilisée dans un processus d'enseignement, de formation ou d'apprentissage et qui est :

- disponible librement (web) ou vendue (consortium, campus virtuel...);
- réutilisable ;
- abordable, adaptable, composable, découvrable, durable, fiable, gérable ;
- interchangeable, évaluable, livrable, réutilisable ;
- décrite par des métadonnées.

¹ Commission de l'AFNOR en charge de la normalisation des technologies de l'information et de la communication pour l'enseignement, la formation et l'apprentissage, responsable du groupe 4 sur les métadonnées.

² IEEE : Institute of Electrical and Electronics Engineers : <http://www.ieee.org/portal/site>

Pourquoi indexer les ressources pédagogiques ? Bien évidemment pour les retrouver. Mais l'enjeu est plus important. En effet, le temps est révolu où un enseignant pouvait se glorifier de mettre son cours en ligne sur un site pour ses propres étudiants. Une ressource pédagogique est longue et coûteuse à produire. C'est pourquoi elle doit être développée en respectant des normes et des standards. La seule façon de la rentabiliser se trouve dans la vente et/ou la mutualisation. Cela nécessite de réfléchir à la fois au fond et à la forme, ce qui demande du temps.

De telles ressources pédagogiques peuvent être disponibles librement sur le web, via des consortiums, des campus virtuels, des universités numériques thématiques. Leur réutilisation est importante, ce qui soulève un certain nombre de questions. Comment retrouver ces ressources ? Est-ce bien la dernière version de la ressource ? Est-ce que ce cours a déjà été produit ? Comment assembler plusieurs ressources automatiquement ? La présentation peut-elle être différente selon les utilisateurs ? D'où l'utilité de la description de la ressource pédagogique, à travers l'indexation, qui se situe à plusieurs niveaux.

Rôle des métadonnées

C'est à cela que servent les métadonnées, données de nature sémantique qui décrivent des ressources. Mais ce qui semble simple à transcrire sur une fiche est beaucoup plus complexe à intégrer dans un logiciel. Un lecteur qui cherche un écrivain précis dans des fiches papier pourra établir une équivalence avec le nom de l'auteur. En revanche un logiciel ne saura pas qu'un auteur peut être un écrivain. Tim Berners Lee, l'inventeur du web, dit que les ressources d'information numérique doivent être compréhensibles par les êtres humains et traitables par des logiciels.

L'indexation doit alors être précise, ce qui implique parfois un très grand nombre de champs à remplir, plus de cinquante dans le LOM. Une partie de l'indexation peut être réalisée automatiquement, comme la reconnaissance automatique du format de fichier par le logiciel. Mais la description de l'objectif pédagogique d'une ressource, par exemple, devra être réalisée manuellement.

Ensuite, l'affichage des champs de métadonnées sera adapté en fonction du type d'utilisateur : un internaute n'aura pas besoin de voir les mêmes champs qu'un indexeur. Sur le web, les moteurs de recherche comme Google permettent en principe de retrouver tout ce qui est accessible. Mais une recherche via Google sur *Notre-Dame de Paris* – l'ouvrage de Victor Hugo – envoie d'abord sur la comédie musicale, les plans de Paris, des critiques de la comédie musicale, etc. Il faut donc indiquer que la recherche porte sur l'œuvre précise, par son titre ou son auteur, et non pas sur des ressources sur cette œuvre. Cet exemple montre que les métadonnées sont vraiment

nécessaires. Elles aident non seulement à retrouver plus facilement les ressources, mais aussi à les évaluer, même si ce ne sont pas les mêmes descripteurs qui sont utilisés dans les deux cas. Enfin, les métadonnées facilitent aussi la maintenance, par la gestion des différentes versions d'un document, et l'assemblage automatique de ressources. Il est alors indispensable d'anticiper et de penser à toutes les utilisations d'une ressource avant de définir l'ensemble des métadonnées associées.

Complexité de l'indexation et normalisation

Dans certains contextes, très centralisés, il est possible d'imposer le système d'indexation, les descripteurs, les logiciels, les formats de fichiers, le système d'exploitation, l'instanciation du LOM, etc. Cependant, un changement du contexte va alors rencontrer certains problèmes et impliquer de transformer des données. Or l'indexation est coûteuse. Dans d'autres cas, le problème est plutôt la manière de dialoguer avec d'autres plates-formes. Le but est donc de ne pas recommencer l'indexation chaque fois qu'intervient un changement de plate-forme ou d'interlocuteurs. L'utilisation de normes ou de standards est alors rendue obligatoire.

Quelques définitions (Educn³) permettent de préciser ces notions :

- Norme : ensemble de règles de conformité, édictées par un organisme de normalisation au niveau national ou international (ISO⁴, AFNOR⁵).
- Standard : ensemble de recommandations émanant d'un groupe représentatif d'utilisateurs réunis autour d'un forum, comme l'IETF (Internet Engineering Task Force), le W3C (World Wide Web Consortium), le Dublin Core.
- Position dominante : domination d'un produit logiciel particulier sur le marché.

Les normes et les standards sont les seuls moyens qui permettent l'interopérabilité et l'évolution des systèmes au cours du temps. Une norme ou un standard ne peut pas forcément répondre aux besoins de chacun, mais doit être extensible. C'est sur un socle de besoins communs que les normes et les standards doivent être bâtis.

Prenons un exemple : pour une communication ou un cours, une présentation PowerPoint va être demandée. Cela correspond à une position dominante de ce logiciel. Mais que deviennent les présentations PowerPoint si ce logiciel disparaît ? Si la version évolue ou si l'utilisateur change de

³ <http://www.educnet.education.fr/tech/normes/0402.htm#definition>

⁴ ISO : International Organization for Standardization, <http://www.iso.ch/iso/fr/ISOOnline.frontpage>

⁵ AFNOR : Agence française de normalisation, <http://www.afnor.fr/portail.asp>

machine, passe de Mac à PC ? Les documents peuvent ne plus être lisibles et se perdre. Si la source des données est gardée, avec une présentation en LaTeX par exemple, le contenu reste accessible même si l'affichage est perdu.

De la même manière, pour que l'indexation survive au logiciel et au matériel, elle doit impérativement reposer sur des standards. Des transcriptions permettant le passage d'un environnement à un autre seront forcément écrites lorsque les normes ou standards évolueront.

Normes et standards permettent donc de maintenir une pérennité des ressources. L'indexation garantit leur accessibilité, et, à travers des échanges, l'interopérabilité entre systèmes. Les ressources peuvent alors être réutilisées et adaptées.

Implémentation

Les ressources pédagogiques peuvent être décrites par des métadonnées comme le Dublin Core ou une version étendue du Dublin Core (Dublin Core éducatif), ou le standard de l'IEEE, le Learning Object Metadata (LOM)⁶. Mais quel que soit le système retenu, leur implémentation n'est pas décrite par le standard. Ainsi, la description de plusieurs chansons pourrait se faire avec une implémentation technique en HTML, et des tags META ; ou bien dans une base de données SQL, dont le schéma et les tables sont à définir ; ou en XML, pour lequel un choix devra être fait entre bases de données XML ou fichiers XML (avec une DTD ou un schéma, et lesquels ?) ; ou encore avec un des langages issus du web sémantique : le RDF (Ressources Description Framework) fait pour les annotations et les métadonnées, ou les OWL (Web Ontology Language), qui ajoutent plus de sémantique et permettent de passer aux ontologies. Aussi, pour implémenter des données dans des systèmes informatiques, les standards doivent être accompagnés d'une transcription, d'un formalisme informatique.

Dans un monde ouvert dans lequel l'interopérabilité est requise, dans lequel la mutualisation des ressources est recherchée, la production de métadonnées est indispensable afin de pouvoir retrouver les ressources et les échanger, voire les vendre.

⁶ Ces normes ou standards sont décrits dans d'autres communications de cette journée d'étude.

Les caractéristiques d'une ressource pédagogique et les besoins d'indexation qui en résultent

La vision du côté TICE (technologies de l'information et de la communication pour l'enseignement)

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

Laurent FLORY
Practice – Université Lyon 1

Résumé :

Appréhender la complexité et l'enjeu de l'indexation des ressources pédagogiques implique en préliminaire de comprendre ce qu'est un objet pédagogique et de définir la notion de grain.

Il est ensuite indispensable de s'interroger sur les besoins des acteurs qui participent à la mise en place de tels objets – en particulier les équipes d'enseignants et les équipes techniques.

Il convient alors de prendre en compte la diversité des projets qui sont mis en œuvre : depuis les « gros projets » aboutissant à un produit fini, volumineux et de très bonne qualité, avec des délais de production longs et des budgets importants ; jusqu'aux « plus petits projets », moins structurés, mais dont la foison se traduit par une production extrêmement riche et diversifiée, avec des mises à jour fréquentes.

En effet, ces deux types de projets soulèvent des questions différentes du point de vue de l'indexation.

Seront aussi abordées les problématiques liées au coût de l'indexation, au niveau de granularité auquel l'indexation doit s'arrêter, à l'indexation automatique, à l'adaptation de l'indexation aux mises à jour et au cycle de production des objets pédagogiques.

Mots clés :

ressources pédagogiques numériques, grain pédagogique, normes et standards pédagogiques, indexation, interopérabilité, circuit du document.

L'objet d'apprentissage : une dimension variable

L'objet d'apprentissage se définit comme une entité, numérique ou non :

- qui peut être utilisée ou citée dans un apprentissage,
- qui possède une unité de sens,
- dont le contenu, la présentation et le traitement peuvent être séparés.

Le grain pédagogique fonctionne comme le lego. Il existe seul et a sa propre entité, mais peut être utilisé dans différents contextes. Un lego rouge, carré, à deux crans existe en soi, mais peut aussi servir à construire tour à tour une station spatiale ou une navette Challenger de milliers de pièces.

De la même manière, un enseignant va constituer son cours en assemblant différents éléments, comme des fragments de texte, des annotations, des images, des sons, des simulations... qui vont constituer des chapitres, puis des cours ou modules et enfin des formations complètes, comme l'illustre le schéma suivant :

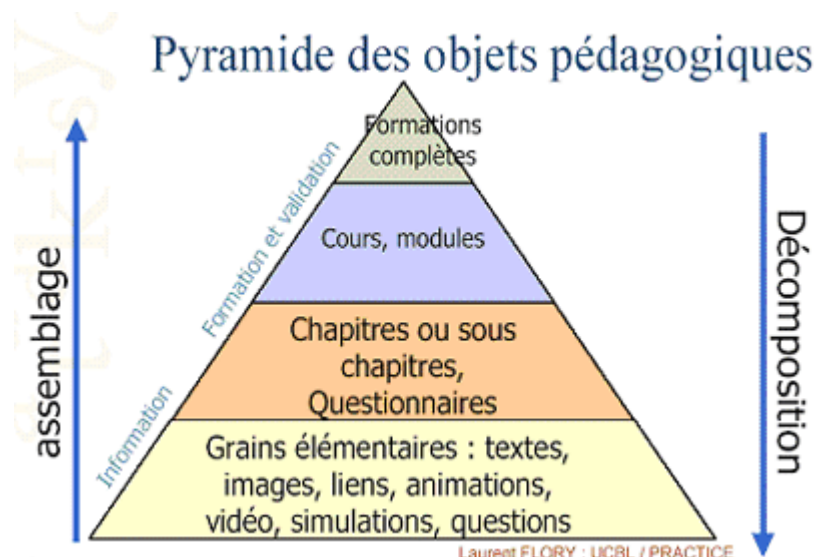


Figure 1 : la pyramide des objets pédagogiques

C'est cette logique d'agrégation des objets pédagogiques qui impose que chaque grain soit « retrouvable, réutilisable, indexable » (pour reprendre trois des dix items proposés par Philippe Parmentier⁷ dans sa description des contraintes à respecter pour un objet pédagogique).

La réutilisation des objets pédagogiques tient avant tout à des questions de coût : des milliers

⁷ « Chaque objet pédagogique doit être durable, adaptable, gérable, fiable, abordable, évaluable, interopérable, retrouvable, réutilisable, indexable ». Parmentier, Philippe. 1999.

de producteurs offrent des sujets de cours communs, dont l'élaboration est dispendieuse⁸. Essentiel, le partage devrait permettre de réaliser des économies importantes en concentrant les moyens pour la création de valeur ajoutée nouvelle. L'objectif final de ces objets est d'offrir aux acteurs (enseignants et élèves) des outils de communication, d'échange, de partage, de validation des savoirs adaptés à leur besoin et à leur rythme. Les nouvelles technologies permettent en effet d'individualiser les apprentissages tout en touchant un plus grand nombre.

Les acteurs de la création

Différents acteurs interviennent dans la création de ces objets pédagogiques :

- l'équipe pédagogique,
- les équipes techniques multimédias,
- les équipes informatiques,
- les conseils en scénarisation,
- les conseils en pédagogie,
- les spécialistes en indexation.

Quels sont les besoins de ces différents acteurs ?

Les enseignants ont besoin de retrouver les objets pédagogiques pour pouvoir les ré-assembler à leur guise. Il y a quelques années, le ministère a financé un projet baptisé PCSM⁹ (Premier Cycle Sur Mesure) comportant des enseignements médiatisés dans toutes les disciplines scientifiques pour le premier cycle (vidéos, simulations, exercices, outils d'auto-évaluation...). Peu d'enseignants l'ont utilisé, car seuls peuvent être sélectionnés des modules complets, ce qui correspond rarement aux besoins. Ce produit est très bien fait, mais il ne peut pas être décomposé en grains élémentaires, si bien que seuls les gens qui l'ont créé l'utilisent...

Côté équipe TICE, le besoin est de ne pas produire un objet pédagogique qui existe déjà, et donc de pouvoir guider les enseignants vers des « grains existants ». C'est une perte de temps et d'argent considérable que de récréer en infographie une image qui existe déjà ou de refaire une simulation, qui une fois créées peuvent être utilisées dans différents cours. Durant le dernier colloque « TICE 2004¹⁰ », qui a eu lieu à Compiègne, le responsable du campus thématique UNIT (Université Numérique Ingénierie et Technologie, dans le domaine des sciences pour l'ingénieur) a

⁸ Downes, Stephen. 2000. *Stephen's Web* : <http://www.downes.ca/>

⁹ <http://www.uel-pcsm.education.fr/consultation/presentation/index.html>

¹⁰ 4e Colloque TICE 2004, organisé par l'université de technologie de Compiègne, 20 au 22 octobre 2004 - Compiègne, France

bien souligné que sa priorité était la réutilisation des objets par les équipes pédagogiques.

Les types de production et les problèmes d'indexation

On distingue deux grands types de productions pédagogiques :

- la grosse production, qui implique des équipes hétérogènes et de nombreuses compétences,
- la production locale réalisée par des équipes souvent plus réduites.

Par exemple, le SCD de Lyon 1 a réalisé un module complet de formation à la recherche d'information, sous Spiral, dans un environnement Flash. Ce programme baptisé Iridoc est destiné à tous les étudiants en licence à Lyon 1 formés à la recherche documentaire : c'est un exemple de grosse production. Mais il existe aussi, à Lyon 1, un grand nombre de « petites » productions réalisées par les enseignants et qui sont diffusées via la plate-forme pédagogique Spiral. Elles peuvent comporter des illustrations, des QCM, voire des arborescences complètes.

Le coût relatif de l'indexation est alors différent : pour les grosses productions qui cumulent plusieurs *années-homme* de travail, ce coût (sur quelques jours ou semaines) reste marginal au regard du coût total. À l'inverse, ce coût devient rédhibitoire pour de toutes petites productions. En revanche, l'indexation fine est plus facile à réaliser pour les petites productions que pour les grosses, où l'encapsulage et la protection des objets font que l'accès aux grains est potentiellement délicat.

Le recours à l'indexation automatique est une possibilité, mais dans le cas d'une grosse production faisant intervenir plusieurs auteurs et plusieurs types de compétences, ce n'est pas forcément simple. Il vaut donc mieux indexer automatiquement les petites productions, en conservant seulement comme mots-clefs la discipline de l'auteur et le domaine d'application. L'indexation par l'usage serait aussi très intéressante, mais elle est encore à l'état de projet de recherche. Il serait pourtant vraiment intéressant de corréler l'indexation à l'utilisation de la ressource dans tel ou tel parcours...

La fréquence de création des objets est aussi un élément à prendre en compte en matière d'indexation. Sur un gros projet, des moyens humains peuvent être affectés à l'indexation des ressources, qui sont quantifiables et limitées. À l'inverse, quand les producteurs d'information sont multiples et qu'ils créent en permanence des contenus pédagogiques, il y a un vrai souci

d'indexation lié au volume.

Par exemple, sur la plate-forme forme Spiral de Lyon 1, en 12 mois calendaires d'ouverture officielle, il y a eu création de plus de 16 000 médias, 490 questionnaires et 845 cours. Cela représente 80 médias, 2 questionnaires et 4 cours à indexer par jour ouvrable, dans des domaines aussi variés que le management du sport, les mathématiques, la santé et la mécanique quantique... Qui est capable d'indexer un tel volume ? Comment répondre à ce besoin ? La question reste posée.

La question de l'actualisation complique encore le problème : outre la production et la création, les mises à jour des cours par les enseignants sont susceptibles d'entraîner des modifications en terme d'indexation.

Cela aboutit donc à un besoin d'indexation presque instantané, en temps réel, qui ne correspond pas à la culture des bibliothèques où la notion de collection rend l'urgence moindre... Or ce qui intéresse les enseignants pour rendre vivants leurs cours, c'est une information à jour : la dernière animation flash, la dernière image qui va s'appuyer sur ce qu'a trouvé Mars express... et pas une ressource datant de dix ans.

Le dilemme, c'est que tous les acteurs ont besoin d'accéder aux objets des autres mais ne souhaitent pas toujours rendre les leurs disponibles et trop souvent l'indexation est considérée comme une perte de temps. On rencontre enfin un problème d'échelle de partage : une indexation locale sur des mots clés (et des thésaurus partagés) est difficile à généraliser au niveau national ou international, les normes et standards doivent alors être utilisés.

Conclusion

En guise de conclusion, un certain nombre de questions subsistent.

Première question : qui va indexer ?

Les enseignants le font peu, seuls 5% renseignent l'objectif pédagogique sous Spiral. Les équipes techniques font ce qu'elles peuvent, mais ce n'est pas leur métier. C'est pourquoi il faut se tourner vers les spécialistes de l'indexation que sont les bibliothécaires. Sinon, la seule réponse possible est la « machine en force brute », en « full text ». Cela permet de retrouver un cours, mais pas encore une simulation, une vidéo, une animation, une bande-son.

Deuxième question : quelle organisation pour cette indexation ?

Faut-il indexer uniquement en automatique ? Ou automatiquement dans un premier temps, avec une intervention humaine ensuite ? Faut-il réserver cette indexation humaine à des cas

spécifiques de ressources labellisées par un comité ? Comment contrôler la qualité de l'indexation ?

Troisième question : comment partager et agréger à un niveau supérieur ?

Par exemple, l'université Lyon 1 et l'ENS ont des cours et des enseignants communs. N'est-il pas possible de réutiliser des objets faits dans les autres établissements, au moins à l'échelon régional dans un premier temps ? Si oui, comment et où les trouver ? Est-ce que ce sont des exports volontaires, des exports manuels ? Faut-il interroger les méta-moteurs des SCD, en utilisant l'OAI¹¹ ? Utiliser le Dublin Core ? Ou le LOM et quel LOM ?

Dernière question : comment monter des équipes mixtes TICE/SCD ?

L'indexation des ressources TICE par des bibliothécaires n'est pas simple. Ce sont deux cultures différentes, deux approches, voire deux communautés différentes, et encore peu habituées à collaborer. Les facteurs clés de succès dans les TICE sont :

- le pragmatisme ;
- la proximité du terrain ;
- la vitesse et la qualité de la réaction ;
- l'adéquation des réponses aux besoins.

Il convient de les garder à l'esprit dans le domaine de l'indexation.

Un seul exemple : l'enseignant préférera avoir 15 images sous forme de vignettes, où il pourra effectuer un contrôle visuel très rapide, plutôt que de renseigner 375 critères de sélection qui permettraient de trouver du premier coup la bonne image. Il est donc important de rester très proche des besoins et des usages des utilisateurs. En effet, la quantité de matière produite et non indexée, le besoin d'indexation en « temps réel », le besoin des acteurs de trouver des objets indexés et leur mauvaise volonté, voire leur refus d'indexer leurs objets, telles sont les contraintes qui font de cette question une urgence à laquelle des réponses adaptées doivent être proposées, avant qu'une solution par défaut (et pas forcément optimale) ne se généralise.

¹¹ OAI : Open Archive Initiative.

Partie 2 – Normes et standards : un processus de normalisation en cours

De Marc à Dublin Core

Élizabeth CHERHAL
Cellule MathDoc, UMS5638, CNRS/Université Joseph Fourier,
Grenoble.

Note : article rédigé par l'auteur à la suite de sa communication orale du 16 novembre 2004.

Résumé :

La question des normes de description bibliographique « classiques » sera abordée d'abord d'un point de vue historique, puis d'un point de vue « utilitaire » selon l'adaptation aux besoins et aux types de documents décrits.

Nous verrons le contexte historique de création des normes MARC¹² et de son corollaire, la norme de catalogage « sur fiches » ISBD¹³, mais aussi l'utilité de cette normalisation « universelle » pour l'échange de métadonnées entre bibliothèques et pour l'accès réparti à des bases de données distantes et différentes.

Nous verrons ensuite le contexte de l'initiative Dublin Core et nous citerons des applications importantes, telles qu'OAI¹⁴.

En conclusion, et au-delà des normes, nous aborderons l'interconnexion souhaitable de deux mondes encore un peu distincts : celui des bibliothèques et celui de la documentation en ligne et des archives ouvertes.

Mots clés :

indexation, ressources pédagogiques numériques, normes et standards pédagogiques, description bibliographique, métadonnées, archives ouvertes, Open Archives Initiative, Marc, Dublin Core.

Marc et Dublin Core, des normes pour la description bibliographique « généraliste »

Marc et Dublin Core ont pour caractéristique commune de se proposer de décrire toute ressource documentaire. En ce sens, ils peuvent être qualifiés de normes (ou standards) généralistes de description bibliographique. Comme toutes les normes (ou standards) de description

¹² Marc : MACHine-Readable Cataloging.

¹³ ISBD : International Standard Bibliographic Description (description bibliographique internationale normalisée).

¹⁴ OAI : Open Archives Initiative.

bibliographique, Marc et Dublin Core ont pour but de décrire l'objet documentaire afin de le retrouver, puis éventuellement de le partager. Ces deux standards ont néanmoins des utilisations et histoires différentes.

La norme Marc est née en 1965 dans le but de procéder à « l'informatisation » des catalogues des bibliothèques. De manière concomitante, les normes ISBD, publiées à partir de 1971, constituent un ensemble de spécifications pour la description bibliographique, ainsi que sa présentation sur fiches. Marc apparaît aujourd'hui comme très représentatif des contraintes imposées à la fois par le catalogue sur fiches et par l'informatique naissante des années 60 et 70. Cependant, il s'agit d'une norme aujourd'hui incontournable et utilisée dans le monde entier par toutes les bibliothèques importantes.

Le standard Dublin Core, qui date de 1995, a pour premier objectif de décrire les ressources en ligne. Le but était de permettre une certaine interopérabilité dans un monde sans description bibliographique, le monde du web. Les acteurs du Dublin Core ont voulu établir un « plus petit dénominateur commun » de la description, quitte à permettre des extensions connues sous le nom de « Dublin Core qualifié ».

Marc aujourd'hui

Dans le passé, il existait pratiquement une version de Marc par pays, ce qui était assez dommageable pour l'interopérabilité. Aujourd'hui deux formats sont reconnus par l'IFLA¹⁵ : les formats Marc21 et Unimarc. Unimarc est le résultat d'une normalisation des différentes variétés de Marc et est surtout utilisé en Europe. Marc21 est le successeur de LC-Marc, utilisé par la Bibliothèque du Congrès et plus généralement en Amérique du Nord.

Marc permet une description bibliographique très détaillée, groupée (logiquement) en blocs, zones et sous-zones, comprenant des informations y compris sur les zones qui doivent être indexées.

¹⁵ IFLA : International Federation of Librarians Association.

```

01235c:arn:2200337_450_(label)
001002100000100031000210200017000520210053000069100004100122101
000800163102000700171105001800178106000600196200009400202205002      répertoire
600296210006300322215002700385320001000112324006200122345001800
484606002900502676001200531686007000543700004700613801003900660
930004200699960004100741960003800782960003900820960003800859-
FRBNF 390005800000008- $a2-87647-210-4$bbr$d56
EUR- $aFR$b00317406- $575999999:39000580001001$aFR$bDLE-20030521-
19871 $a20030521e1979
    k yOfrey0103    ba-0 $afre- $aFR- $ay z 000yy- $ar-1 $aPanorama des      Contenu et
mathématiques pures$bTexte imprimé$ele choix bibliographique$fJean      indicateurs
Dieudonné $a[Reprod. en fac. simi.] $aParis$cJ. Gabay$d2003$e14 Cuidé sur      de traitement
Noireau
$gImpr. Corlet- $axv-IIIIf; p $cIII $d74 cm- $aIndex- $al ac-sim du second tirage de
l'éd. de Paris : Gauthier-Villars, 1979- $b9/020/64/2105- $J119J24J4
$aMathématiques- $a51U$v21- $a51U$2Cadre de classement de la Bibliographie
nationale française- 1$J119UU1U6$aDieudonné$bJean$119U6-
1992$4070- 0$aFR$bBNF$c20030521$gAFNOR$2intermrc-$575999999:39000
500001001$a2003 115J0J- $J119J24J4$aMathématiques
appliquées- $311932131$aMathématiques modernes- $311932131$aMéthodes
mathématiques- $311932131$aSciences mathématiques- 0
    
```

Exemple du format physique d'échange ISO2709

Cet exemple permet de deviner ce que le format hérite de contraintes techniques qui, aujourd'hui, n'ont plus lieu d'être.

Pour résumer, sont à l'actif de Marc :

- ses spécifications ouvertes ;
- une reconnaissance universelle ;
- l'existence du format d'échange qui a permis d'une part le partage de métadonnées entre bibliothèques et la mise en place de catalogues tels que le SUDOC¹⁶ et d'autre part le déploiement de l'outil Z39.50, permettant une interrogation normalisée et à distance de différents catalogues (intégration dans de nombreux outils de catalogage et mise en place d'interfaces d'interrogation de plusieurs catalogues simultanément¹⁷).

Et au passif :

- sa complexité, qui le rend difficilement utilisable en dehors du contexte des bibliothèques classiques : son utilisation serait difficile dans le cadre d'une grande collection de documents numériques ;
- sa relative inadaptation aux conditions technologiques actuelles ;
- le fait que, vu la complexité, la plupart des logiciels qui s'y conforment sont assez coûteux, (mais il faut souligner une exception : le logiciel libre Koha).

¹⁶ SUDOC : Système Universitaire de Documentation (catalogue collectif des bibliothèques universitaires) <http://corail.sudoc.abes.fr/>

¹⁷ Le Catalogue Collectif de France (CCF) en est un exemple. <http://www.ccf.bnf.fr/>

Dublin Core, de sa naissance à nos jours

Le Dublin Core (« core » signifie « noyau ») n'est pas une norme et peut être qualifié de « standard de fait ». Il est né en 1995 à Dublin (Ohio), lors d'une réunion entre différents acteurs, dont en particulier l'OCLC¹⁸, lui-même acteur de la communauté des bibliothèques. Aujourd'hui le Dublin Core est maintenu par le DCMI¹⁹ et donne lieu à une conférence annuelle au contenu riche et varié. Le DCMI comporte de nombreux groupes de travail, par exemple le groupe « éducation » qui tient une conférence commune avec le groupe LOM²⁰.

L'objectif de l'initiative, déclaré sur son site web, est :

« The Dublin Core Metadata Initiative (DCMI) is an organization dedicated to promoting the widespread adoption of interoperable metadata standards and developing specialized metadata vocabularies for describing resources that enable more intelligent information discovery systems. »

L'initiative Dublin Core est partie du constat que d'une part beaucoup de documents (et autres ressources) étaient disponibles en ligne sans aucune description, donc difficilement trouvables, et d'autre part qu'il n'était pas possible de « cataloguer » ces ressources en utilisant des moyens traditionnels. Il fallait donc un moyen de décrire simplement des ressources en ligne ou plus généralement les ressources numériques.

Les participants sont arrivés à un consensus sur les points suivants :

- la description devait être assez simple pour être comprise et utilisée par un grand nombre de personnes. Il fallait donc établir un dénominateur commun de base ;
- la description devait être indépendante des formats physiques ;
- la description devait être suffisamment généraliste pour permettre de décrire des ressources autres que les documents classiques (gouvernement, environnement, enseignement, etc.).

Ces principes de base sont toujours respectés²¹.

Une des premières idées des acteurs du Dublin Core était d'améliorer la qualité des résultats de recherche des « moissonneurs » web. Le Dublin Core peut être destiné au moissonnage par un

¹⁸ OCLC : Online Computer Library Center.

¹⁹ DCMI : Dublin Core Metadata Initiative.

²⁰ LOM : Learning Object Metadata, voir d'autres contributions dans ces actes.

²¹ Voir : « Présentation des standards : le Dublin Core » dans ces actes.

moteur de recherche web tel que Google, Yahoo ou encore un moteur plus spécialisé. Dans les faits, c'est cette utilisation spécialisée qui a prévalu dans un premier temps et, dès son lancement, le Dublin Core a servi de base à de nombreux projets et applications dans le domaine des bibliothèques numériques, tels que le projet CISMeF²² (Catalogue et Index des Sites Médicaux Francophones), le « Victorian Education Channel »²³ ou le projet EULER (European Libraries and Electronic Resources in Mathematical Sciences). De fait, ces projets ont souvent aménagé le Dublin Core à leur manière et sa véritable utilité comme dénominateur commun a été démontrée par l'initiative OAI (Open Archives Initiative). OAI a en quelque sorte « établi » le Dublin Core comme standard. Tout serveur OAI a pour obligation de fournir des métadonnées en Dublin Core de base. La préexistence du Dublin Core a grandement facilité le lancement fulgurant de l'initiative OAI et de son protocole OAI-PMH²⁴.

Pour résumer, sont à l'actif du Dublin Core :

- sa simplicité qui le rend facilement compréhensible et utilisable ;
- son aspect généraliste ;
- la grande variété de ressources qu'il peut décrire ;
- son indépendance par rapport aux formats physiques ;
- la vivacité de ses différentes composantes (réunions, groupes de travail)...

Au passif, et corollaire de l'actif :

- son côté généraliste et incomplet, nécessitant souvent des extensions ;
- sa relative jeunesse et le fait qu'il évolue encore (bien que les éléments de base semblent être « gravés dans le marbre »).

Cohabitation et avenir communs ?

Schématiquement, Marc représente le monde des bibliothèques classiques et Dublin Core celui des bibliothèques numériques et du mouvement pour les archives ouvertes ; ils sont apparemment à des lieues l'un de l'autre. En réalité, ces deux mondes se croisent et communiquent relativement souvent. Les bibliothèques classiques se préoccupent de plus en plus de ressources numériques diverses et variées, et les responsables de bibliothèques numériques se rendent compte

²² CISMeF : <http://www.chu-rouen.fr/cismef/>

²³ <http://www.education.vic.gov.au/>

²⁴ OAI-PMH : Open Archives Initiative Protocol for Metadata Harvesting.

que les bibliothèques classiques ont beaucoup à leur apprendre. En témoignent certainement les nombreux « emprunts » à Marc qui se trouvent dans le Dublin Core « qualifié »²⁵.

À l'avenir, sur le plan purement technique, les initiatives en cours, telles que MarcML ou BiblioML, pour exprimer Marc en XML (et s'affranchir des contraintes de l'ISO2709) peuvent également faciliter ce dialogue entre bibliothèques classiques et numériques. Dans le même ordre d'idée, des outils pour la conversion des formats, telles que le « Dublin Core/Marc/GILS Crosswalk » ou le « Marc to Dublin Core Crosswalk » ont pu montrer leur utilité.

Sur un plan un peu plus politique, OAI a indéniablement favorisé la cohabitation, comme le montrent les exemples de la BnF, qui a lancé un serveur OAI devant sa bibliothèque numérique Gallica, ou celui de l'ABES²⁶ qui ouvre un portail où se trouvent de nombreuses notices moissonnées par OAI.

Bibliographie

ISBD

« Présentation des Standards : le Dublin Core ». [http://dublincore.org/
http://www.ifla.org/VI/3/nd1/isbdlist.htm](http://dublincore.org/http://www.ifla.org/VI/3/nd1/isbdlist.htm)

Marc

<http://www.ifla.org/VI/3/p1996-1/sec-uni.htm>

<http://www.loc.gov/marc/marcdocz.html>

LERESCHE, F., « les formats MARC », http://www.rnbm.org/rencontres_2004/leresche-marc.pdf

OAI

<http://www.openarchives.org/>

Outils de conversion

<http://www.loc.gov/marc/dccross.html>

<http://www.loc.gov/marc/marc2dc.html>

Z39.50

<http://www.loc.gov/z3950/agency/>

LAHARY, Dominique, « La norme Z39.50 »,

<http://membres.lycos.fr/vacher/profess/cours/mediadix/z3950/>

²⁵ Voir : « Présentation des standards : le Dublin Core » dans ces actes.

²⁶ ABES : Agence Bibliographique de l'Enseignement Supérieur.

Le Dublin Core (DC)

Élizabeth CHERHAL
Cellule MathDoc, UMS5638, CNRS/Université Joseph Fourier,
Grenoble.

Note : article rédigé par l'auteur à la suite de sa communication orale du 16 novembre 2004.

Résumé :

Nous rappellerons le contexte historique de l'initiative Dublin Core, puis verrons l'ensemble des éléments de base et la notion de Dublin Core qualifié. Nous verrons ensuite que la description bibliographique est indépendante des formats physiques et présenterons quelques-unes des représentations physiques les plus utilisées.

Mots clés :

Dublin Core, métadonnées, indexation, archives ouvertes, Open Archives Initiative, normes et standards pédagogiques, ressources pédagogiques numériques

Rappels

Le Dublin Core n'est pas une norme. Il s'agit d'un ensemble de recommandations qui peut être qualifié de standard de fait. Le Dublin Core est né en 1995, lors d'une conférence tenue à Dublin (Ohio). Il s'agissait de produire une recommandation pour la description bibliographique des documents ou objets se trouvant sur le web, dans une bibliothèque numérique ou même ailleurs. Les acteurs de cette première conférence, qui regroupait des personnes issues de la communauté des bibliothèques, des informaticiens et du web, ont spécifié un ensemble de quinze éléments de base, une sorte de dénominateur commun de la description bibliographique. Le Dublin Core, dans son vocabulaire propre, parle d'éléments et de termes. Éléments, parce qu'il s'agit ici d'un monde issu de SGML, celui de HTML (et plus tard XML). Les termes d'éléments et d'attributs sont donc utilisés là où MARC parle par exemple de zones ou de sous-zones et où d'autres encore pourraient parler de « champs ».

Il est établi que les quinze éléments du Dublin Core sont « gravés dans le marbre » et que tout élément supplémentaire appartient aux extensions.

Les quinze éléments de base

Nom de l'élément en français	Décri	Nom de l'élément en anglais
Titre Sujet Description Source Langue Relation Couverture	contenu	Title Subject Description Source Language Relation Coverage
Créateur Éditeur Contributeur Droits	propriété intellectuelle	Creator Publisher Contributor Rights
Date Type Format Identifiant	instance particulière	Date Type Format Identifier

Le Dublin Core qualifié et les « termes »

Les quinze éléments de base sont considérés comme un dénominateur commun et pour la plupart des cas sont insuffisamment précis. Dès le départ, les éléments de base ont été étendus (ou précisés) par un ensemble d'autres termes, parfois appelés *qualifiers*. Deux classes de *qualifiers* sont reconnues :

- les « raffinements d'éléments » qui rendent plus spécifique le sens d'un élément ;
- et les « schémas d'encodage » ou vocabulaires contrôlés.

Les éléments supplémentaires et « raffinements d'éléments »

Le DCMI²⁷ a défini trois éléments supplémentaires et une trentaine de raffinements d'éléments.

Les éléments supplémentaires : **Audience**, **Provenance** et **RightsHolder** étendent l'ensemble des quinze éléments de base.

Les raffinements d'éléments (« element refinements »), dont voici pour information la liste complète²⁸, précisent le sens d'un élément existant :

²⁷ DCMI : Dublin Core Metadata Initiative, organisme responsable de la définition du Dublin Core.

²⁸ Voir <http://dublincore.org/documents/dcmi-terms/>

Element refinement	Element refined
Abstract	description
Access Rights	rights
Alternative	title
Available	date
Bibliographic Citation	identifiant
Conforms To	relation
Created	date
Date Accepted	date
Date Copyrighted	date
Date Submitted	date
Audience Education Level	audience
Extent	format
Has Format	relation
Has Part	relation
Has Version	relation
Is Format Of	relation
Is Part Of	relation
Is Replaced By	relation
Issued	date
Is Version Of	relation
License	rights
Mediator	audience
Medium	format
Modified	date
References	relation
Replaces	relation
Requires	relation
Spatial	coverage
Table Of Contents	description
Temporal	coverage
Valid	date

Les schémas d'encodage (vocabulaires contrôlés)

Chaque élément ou raffinement d'élément peut disposer d'un ou plusieurs schémas d'encodage. Le DCMI, soucieux de ne pas réinventer la roue, fait référence à des schémas existants et en a inventé un, DCMIType, pour décrire les types (logiques et non physiques) d'objets, d'autres, DCMIBox et DCMIPoint, pour délimiter géographiquement un objet ou lieu et DCMIPeriod pour situer un événement dans le temps.

Nom du schéma (vocabulaire)	« Element refined »
Box (DCMI)	spatial
DCMI Type	type
DDC (Dewey Decimal Classification)	subject
IMT (Internet media type)	format
ISO 3166 (norme iso des noms de pays)	spatial
ISO 639-2 (norme iso pour les codes de langues)	language
LCC (Library of Congress Classification)	subject
LCSH (Library of Congress Subject Headings)	subject
MeSH (Medical Subject Headings)	subject
Period (DCMI)	date / temporal
Point (DCMI)	spatial
RFC 1766 (codes langues- pays)	language
RFC 3066 (codes langues- pays, remplace RFC 1766)	language
TGN (Getty Thesaurus of Geographic Names)	spatial
UDC (Universal Decimal Classification)	subject
URI (Uniform Resource Identifier)	identifiant/ source/ relation
W3C-DTF (dates et heures)	date/temporal

Le vocabulaire DCMIType

Le vocabulaire DCMIType contient une liste de différents types de ressources :

Collection	Groupe de documents (ou ressources)
Dataset	Base de données, liste, tableau
Event	Événement ponctuel (conférence...)
Image	Image, photo, (tous formats, générique)
Interactive Resource	Objet qui demande l'interaction de l'utilisateur
Moving Image	Image animée, vidéo...
Physical Object	Un objet lui-même (et non son image)
Service	
Software	Logiciel disponible pour installation sur une machine
Still Image	
Text	

Les formats physiques

Le Dublin Core ne dépend pas d'une expression physique précise. On peut l'exprimer en HTML ou XML/RDF²⁹ ou encore en XML, selon l'utilisation à laquelle il est destiné. Il existe trois recommandations officielles :

- **Expressing Simple Dublin Core in RDF/XML**
<http://dublincore.org/documents/2002/07/31/dcmes-xml/>
- **Guidelines for implementing Dublin Core in XML**
<http://dublincore.org/documents/2003/04/02/dc-xml-guidelines/>
- **Expressing Qualified Dublin Core in HTML/XHTML meta and link elements**
<http://dublincore.org/documents/dcq-html/>.

Et une recommandation proposée :

- **Expressing Qualified Dublin Core in RDF/XML**
<http://dublincore.org/documents/2002/05/15/dcq-rdf-xml/>

²⁹ RDF : Resource Description Framework, recommandation du W3C.

Voici un exemple simple d'utilisation du Dublin Core dans un fichier HTML :

```
<html>
<head>
  <title> A Dirge </title>
  <link rel = "schema.DC" href = "http://purl.org/DC/elements/1.0/">
  <meta name = "DC.Title" content = "A Dirge">
  <meta name = "DC.Creator" content = "Shelley, Percy Bysshe">
  <meta name = "DC.Type" content = "poem">
  <meta name = "DC.Date" content = "1820">
  <meta name = "DC.Format" content = "text/html">
  <meta name = "DC.Language" content = "en">
</head>
<body>
  <pre>Rough wind, that moanest loudGrief too sad for song ;Wild wind, when sullen
cloudKnells all the night long ;Sad storm, whose tears are vain,Bare woods, whose
branches strain,Deep caves and dreary main, -WailJfor the world's wrong !</pre>
</body>
</html>
```

Voici un exemple d'encodage en HTML, mais en utilisant les raffinements d'éléments (seul l'entête du document HTML est montré).

```
<head profile="http://dublincore.org/documents/dcq-html/">
  <title>Expressing Dublin Core in HTML/XHTML meta and link elements</title>
  <link rel="schema.DC" href="http://purl.org/dc/elements/1.1/" />
  <link rel="schema.DCTERMS" href="http://purl.org/dc/terms/" />
  <meta name="DC.title" lang="en" content="Expressing Dublin Core in HTML/XHTML
meta and link elements" />
  <meta name="DC.creator" content="Andy Powell, UKOLN, University of Bath" />
  <meta name="DCTERMS.issued" scheme="DCTERMS.W3CDTF" content="
2003-11-01" />
  <meta name="DC.identifier" scheme="DCTERMS.URI" content="
http://dublincore.org/documents/dcq-html/" />
  <link rel="DCTERMS.replaces" hreflang="en" href="
http://dublincore.org/documents/2000/08/15/dcq-html/" />
  <meta name="DCTERMS.abstract" content="This document describes how qualified
Dublin Core metadata can be encoded in HTML/XHTML &lt;meta&gt; elements" />
  <meta name="DC.format" scheme="DCTERMS.IMT" content="text/html" />
  <meta name="DC.type" scheme="DCTERMS.DCMIType" content="Text" />
</head>
```

L'exemple le plus courant d'expression du Dublin Core en XML est celui d'un serveur OAI³⁰ (le schéma XML « oai_dc » se trouve sur le site de l'initiative OAI). Voici un exemple extrait d'un serveur OAI.

```
<metadata>
  <oai_dc:dc xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:oai_dc="
http://www.openarchives.org/OAI/2.0/oai_dc" xsi:schemaLocation="
http://www.openarchives.org/OAI/2.0/oai_dc http://www.openarchives.org/OAI/2.0/oai_dc.xsd"
xmlns:dc="http://purl.org/dc/elements/1.1/">
  <dc:creator>Moret, M.</dc:creator>
  <dc:title>Notice nécrologique : Émile Cotton (1872--1950)</dc:title>
  <dc:date>1949</dc:date>
  <dc:identifier>AIF_1949__1__1_0</dc:identifier>
  <dc:identifier>http://www.numdam.org/item?id=AIF_1949__1__1_0</dc:identifier>
  <dc:identifier>oai:numdam.org:AIF_1949__1__1_0</dc:identifier>
  <dc:identifier>citation: Ann. Inst. Fourier 1, 1-4 (1949)</dc:identifier>
  </oai_dc:dc>
</metadata>
```

Conclusion

Précisons pour terminer que :

- Le Dublin Core (ses extensions) évolue assez souvent. Les documents cités en référence peuvent changer.
- La description logique, qui n'est pas liée à un format physique, fait qu'il est possible d'inventer d'autres expressions de celle-ci sans en changer le sens.
- Une des premières idées des acteurs du Dublin Core était d'améliorer la qualité des résultats de recherche des « moissonneurs » web. Le Dublin Core était destiné au moissonnage par un moteur de recherche web tel que Google, Yahoo ou encore un moteur plus spécialisé. Dans un premier temps, c'est cette utilisation spécialisée qui a prévalu, car dans les faits les moteurs web généralistes ont ignoré le Dublin Core (ou plus précisément ont ignoré l'utilisation d'éléments META dans l'entête des documents HTML, à cause du détournement à des fins commerciales de leur utilisation). Aujourd'hui cependant, le projet « Google scholar » préconise l'utilisation du Dublin Core.
- Le Dublin Core est aujourd'hui moissonné efficacement avant tout par le biais d'OAI, à partir de serveurs se conformant au protocole OAIPMH.
- Actuellement, si tous les serveurs OAI utilisent (c'est obligatoire) le schéma oai_dc, bien peu se sont lancés dans l'utilisation de schémas « Dublin Core qualifié ». L'avenir dira si cette tendance s'inversera.

³⁰ Voir « État de l'art en matière de normes et approche critique : du Marc au Dublin Core » dans ces actes.

- Les métadonnées « Dublin Core » sont très souvent exportées d'une base de données ou application autre, dont les métadonnées sont plus précises ou plus spécialisées.

Bibliographie

HILLMAN, Diane, *Guide d'utilisation du Dublin Core*

<http://www.bibl.ulaval.ca/DublinCore/usageguide-20000716fr.htm>

Dublin Core Metadata Initiative Documents : <http://dublincore.org/documents/>

DCMI Metadata Terms: <http://dublincore.org/documents/dcmi-terms/>

DCMI Encoding Guidelines <http://dublincore.org/resources/expressions/>

DCMI Schemas <http://dublincore.org/schemas/>

Présentation des standards : (LOM) – Learning Object Metadata

Rosa María GÓMEZ DE REGIL

Doc'Insa, Lyon

Membre du groupe d'experts sur les métadonnées pédagogiques à l'AFNOR (au sein de la CN36)

Note : article rédigé par l'auteur à la suite à sa communication orale du 16 novembre 2004.

Résumé :

Le LOM (*Standard for Learning Object Metadata* ou recommandation de métadonnées concernant les objets pédagogiques) trouve sa place dans un contexte de production massive d'objets pédagogiques numériques.

Le développement de la FOAD (formation ouverte et à distance) et l'utilisation des nouvelles technologies pour la recherche et la diffusion des ressources rendent nécessaire le partage, l'échange et la réutilisation des ressources pédagogiques électroniques. Le LOM se manifeste comme un outil fédérateur et cohérent, en proposant en neuf catégories des éléments de description communs à tous ceux qui l'adoptent. L'utilisation d'un schéma de métadonnées standard facilite la recherche, l'évaluation, l'acquisition et l'utilisation des objets pédagogiques par les apprenants, par les enseignants et par les producteurs de ressources.

Cette intervention présente les spécificités et les enjeux actuels de l'indexation des ressources pédagogiques autour du LOM dans le but de familiariser les acteurs des bibliothèques, ainsi que ceux des TICE, à de nouvelles structures normalisées de description de ressources.

Mots clés :

LOM (Learning Object MetaData), normalisation, indexation, ressources pédagogiques numériques, métadonnées, AFNOR, circuit du document, normes et standards pédagogiques

Introduction

L'accès aux ressources électroniques via Internet et le développement de la formation ouverte et à distance sont des facteurs qui révèlent le besoin de trouver, évaluer, partager, échanger et utiliser des ressources pédagogiques numériques. La production de ce genre de ressources est coûteuse car elle demande des développements informatiques conséquents et l'investissement de

plusieurs personnes. Une fois que ces ressources sont produites, il est fondamental de pouvoir les exploiter (plus d'une fois). Dans ce contexte, la structuration de l'information pédagogique devient cruciale et les métadonnées s'avèrent utiles.

Il existe plusieurs modèles de métadonnées pour décrire des ressources numériques. Le Dublin Core est un modèle qui répond, au travers de ses quinze éléments, à une description de base d'une ressource numérique. Cependant, il existe des besoins beaucoup plus précis qui demandent une description plus fine. Le *Learning Object Metadata* trouve sa place en tant que modèle de métadonnées pour la description de ressources à but pédagogique.

Le LOM définit un *objet pédagogique* comme « toute entité numérique ou non, utilisée dans un processus d'enseignement, de formation ou d'apprentissage³¹ ». En fait, il est utilisé en général pour décrire des ressources numériques. Le terme d'objet se situant plutôt dans un contexte informatique, c'est le terme de *ressource pédagogique* qui sera utilisé tout au long de cet article.

Les métadonnées (données qui décrivent ou définissent d'autres données) facilitent le partage de l'information, contribuent à la minimalisation de pertes de données, permettent de rechercher des documents dans un entrepôt de données, facilitent le stockage de données.

Les métadonnées des ressources pédagogiques remplissent différentes fonctions³² :

- décrire et identifier ;
- retrouver et évaluer ;
- gérer ;
- gérer les droits ;
- permettre l'interopérabilité entre les plates-formes de télé-formation ou plates-formes pédagogiques³³.

Afin de codifier les pratiques des communautés et d'envisager des approches communes, des normes et standards sont développés. En effet, l'approche normative est indispensable pour que les systèmes de métadonnées fonctionnent correctement.

Il est important d'établir la distinction entre les normes et les standards³⁴ :

- Une *norme* est un ensemble de règles de conformité, édictées par un organisme de normalisation au niveau national ou international.

³¹ Learning Technology Standards Committee of the IEEE. *Draft Standard for Learning Object Metadata*. [en ligne]. New York : IEEE, 2002. Disponible sur : http://ltsc.ieee.org/wg12/files/LOM_1484_12_1_v1_Final_Draft.pdf. (Consulté le 10 juin 2005).

³² Haynes, David. *Metadata for information management and retrieval*. London : Facet, 2004, 186 p.

³³ Aussi appelées *Learning Management Systems* (LMS).

³⁴ Educnet. *Dossier sur les métadonnées, normes et standards*. [en ligne]. 2004. Disponible sur : <http://www.educnet.education.fr/dossier/metadata/default.htm> (consulté le 10.06.2005).

- Un *standard* est un ensemble de recommandations émanant d'un groupe représentatif d'utilisateurs réunis autour d'un forum, comme l'IETF³⁵, le W3C³⁶, le LTSC³⁷ de l'IEEE³⁸.

Présentation du Learning Object Metadata

Le *Learning Object Metadata* (LOM) est un standard publié en 2002 par le *Learning Technology Standards Committee* (LTSC) de l'IEEE (*Institute of Electrical and Electronics Engineers*). Le standard est composé de quatre parties :

- IEEE 1484.12.1 – Modèle conceptuel de métadonnées ;
- IEEE 1484.12.2 – Implémentation de la norme ISO/IEC 11404 au modèle de métadonnées LOM ;
- IEEE 1484.12.3 – Définition et application du schéma XML pour le LOM ;
- IEEE 1484.12.4 – Définition du cadre d'application RDF (*Resource Description Framework*) pour le LOM.

Cet article présente exclusivement le modèle conceptuel des métadonnées (IEEE 1484.12.1), qui détermine les éléments de données du modèle ainsi qu'une série de vocabulaires de référence.

Le modèle spécifie également combien de fois il est possible de répéter les éléments dans une même instance (notice) de métadonnées. Tous les éléments du LOM sont facultatifs, c'est-à-dire que le modèle peut fonctionner sans que tous les champs soient remplis. Néanmoins, il est souhaitable de fournir la plupart des informations afin que les ressources puissent être exploitées au maximum. Dans le modèle se trouve aussi le type de donnée informatique de chaque élément. Les types de données que le LOM propose sont la chaîne de caractères (*CharacterString*), la chaîne de langue (*LangString*), le vocabulaire (*vocabulary*), la date et le temps (*DateTime*) et la durée (*Duration*). Ces types de données indiquent les contraintes informatiques pour remplir les champs.

Structure du LOM

Le LOM est structuré en neuf catégories accomplissant une fonction différente. Les éléments contenus dans chaque catégorie peuvent être observés dans la figure 1. La sémantique des éléments est déterminée par le contexte de la catégorie qui la contient. Par exemple, l'élément

³⁵ IETF : Internet Engineering Task Force.

³⁶ W3C : World Wide Web Consortium

³⁷ LTSC : Learning Technology Standards Committee

³⁸ IEEE : Institute of Electrical and Electronics Engineers, <http://www.ieee.org/portal/site>

nommé *contribution* apparaît en 2.3 et en 3.2 : chacun s'applique au contexte de sa catégorie, la contribution en 2.3 sera celle intégrée au cycle de vie de la ressource et en 3.2 celle de la personne qui a travaillé sur le remplissage ou la validation des métadonnées.

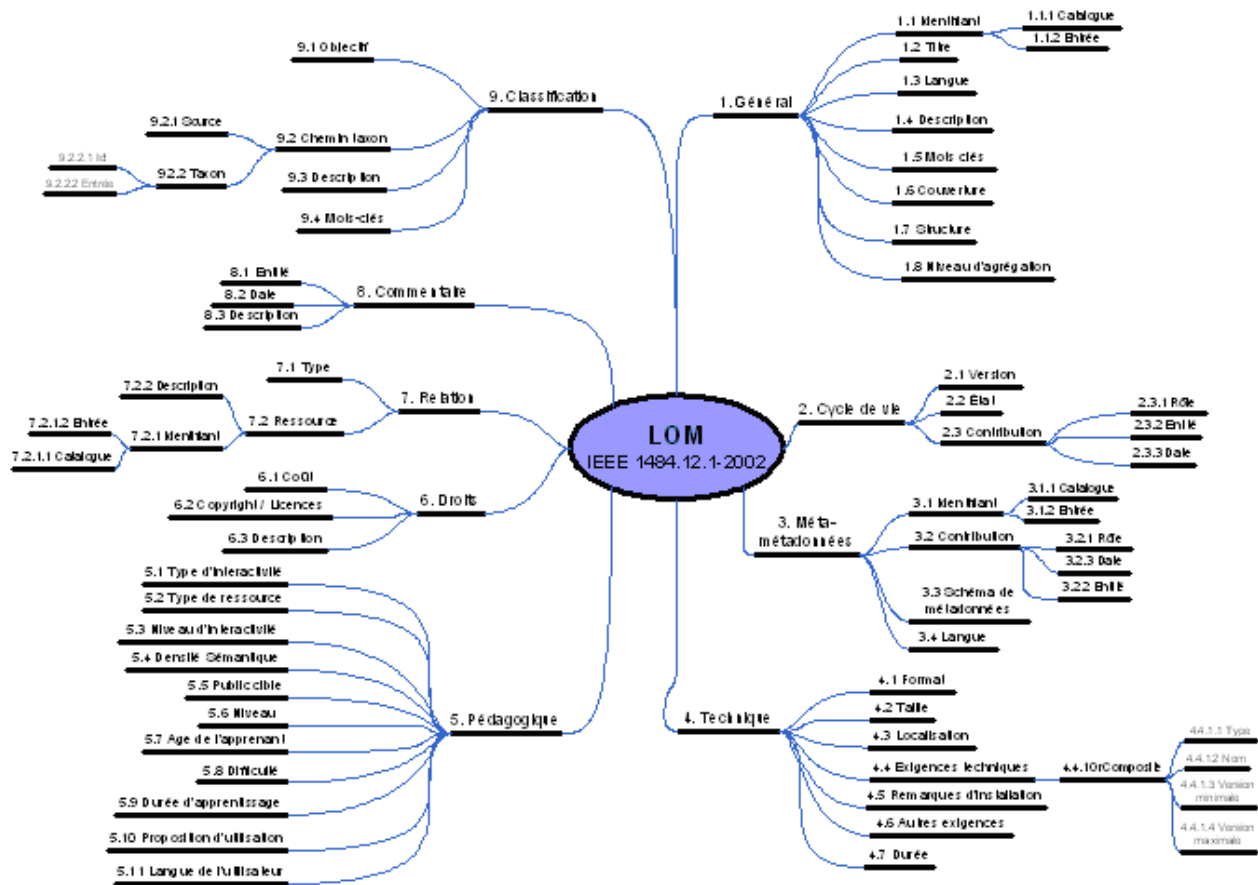


Figure 1 – Aperçu général du LOM

Général – La fonction de cette catégorie est la description et l'identification de la ressource. Les informations contenues permettent de se référer aux autres catégories et d'établir en général les interactions avec les systèmes d'information.

Cycle de vie – Le cycle de vie permet d'avoir des informations sur l'histoire de la ressource et sur ses contributeurs. Les informations collectées ici serviront à la gestion de la ressource.

Meta-métadonnées – Les méta-métadonnées permettent de garder une trace (assez détaillée) du schéma de métadonnées utilisé pour décrire une ressource. Ces informations seront utiles pour envisager l'interopérabilité des différents systèmes et pour l'échange de données.

Technique – Cette catégorie présente les caractéristiques techniques qui permettront l'exécution de la ressource sur un système informatique. Les contraintes techniques d'utilisation peuvent être exprimées ici. Cette catégorie sert à décrire, évaluer et gérer les ressources.

Pédagogique – La pédagogie est le cœur du LOM. C'est ce qui le rend spécifique par rapport à d'autres modèles. Cette catégorie accomplit une fonction de description pédagogique de la ressource. S'y retrouvent, pour ne citer que certains des éléments, le contexte d'utilisation (scolaire, secondaire, supérieur, formation continue, etc.), le type de public cible (enseignant, auteur, apprenant, gestionnaire), la difficulté, l'âge des apprenants auxquels la ressource est destinée, le temps d'apprentissage type, la langue des apprenants, le niveau et le type d'interactivité.

Droits – Ici figurent les conditions légales d'utilisation de la ressource : indications de coût, de licences, de droits d'auteur.

Relation – Cette catégorie permet de lier deux ressources physiques entre elles, en donnant une valeur à cette relation, un sens (est partie de, est version de, est format de, est référencée par, etc.).

Commentaire – L'annotation ou commentaire permet de faire des remarques à propos de la ressource, tout en gardant en mémoire la description de la personne qui commente et la date de ce commentaire. Cette catégorie apporte des informations complémentaires utiles à l'évaluation de la ressource.

Classification – Il s'agit de la catégorie qui permet d'assigner des classifications et des vocabulaires contrôlés à une ressource. Des classifications peuvent être attribuées pour la discipline, les pré-requis pédagogiques, les objectifs pédagogiques, l'accessibilité, les restrictions, les compétences.

Une normalisation en cours

Le LOM a été proposé à l'ISO³⁹ pour devenir une norme internationale, mais il a été refusé. Plusieurs pays, dont la France et la Chine, se sont opposés à sa publication en tant que norme ISO parce que le LOM véhicule une idée américaine de l'enseignement (plutôt behavioriste) et ne permet pas d'exprimer d'autres visions de la pédagogie.

Le LOM est toutefois le standard que les acteurs de l'enseignement et la formation sont sur le point d'adopter. Soit en l'appliquant tel quel, soit en l'adaptant à leurs besoins au travers des profils d'application. Dans la région Rhône-Alpes, l'Insa de Lyon⁴⁰, l'ENS Lyon⁴¹, l'université Claude-Bernard Lyon 1 et l'enssib intègrent le LOM dans leurs projets de gestion de ressources pédagogiques. Au niveau européen, plusieurs institutions se sont approprié le LOM, les pays les plus actifs dans le développement des profils d'application et de la réflexion sur les ressources

³⁹ ISO : International Organization for Standardization.

⁴⁰ INSA : Institut National des Sciences Appliquées.

⁴¹ ENS Lyon : École Normale Supérieure de Lyon.

pédagogiques sont le Royaume-Uni et la Finlande. Ailleurs, ce sont les États-Unis et le Canada qui ont utilisé le LOM et qui réalisent plusieurs guides d'implémentation du modèle et des entrepôts de ressources pédagogiques. Certains pays d'Asie deviennent des acteurs importants dans le monde de la normalisation des métadonnées pédagogiques et dans le développement des systèmes de formation ouverte et à distance : Singapour, Corée, Chine et Japon.

Les enjeux

Les communautés de l'enseignement et de la formation vont disposer de plus en plus de ressources électroniques, qu'il sera nécessaire de structurer afin de ne pas se noyer dans la masse de l'information, d'un côté, et aussi pour avoir une bonne visibilité sur Internet. La structuration de cette information pédagogique passe par l'indexation, définie comme la caractérisation d'un document ou du contenu du document au moyen d'un mot ou d'un ensemble de mots (qui peut être issu d'un vocabulaire contrôlé) qui servent à le désigner⁴². Afin de réussir l'indexation des ressources pédagogiques, il est nécessaire d'identifier l'objectif de l'indexation. Elle peut servir aux apprenants à repérer des ressources, aux enseignants et formateurs à construire et réinterpréter les ressources existantes, aux bibliothèques – dans le sens large (numériques et physiques) – à garantir l'accessibilité du document.

Partenariats

L'objectif de cette journée d'étude était de réunir les acteurs des bibliothèques et des TICE⁴³ afin de montrer les limites des métiers et de les sensibiliser à l'importance des métadonnées. Il est sans doute possible de construire un partenariat entre les enseignants, les producteurs et les bibliothèques, qui viserait un travail en commun, chacun effectuant son métier, pour arriver à obtenir des métadonnées de qualité. Les acteurs des TICE pourront s'appuyer sur les bibliothèques pour compléter l'indexation. Les bibliothèques renforceront leurs compétences, dans le domaine du droit, par exemple, afin de pouvoir participer à la mise en place de contrats entre les enseignants, les auteurs et l'établissement, pour suivre les autorisations de publication de contenus ; mais aussi dans le domaine de la gestion des droits d'accès électroniques, qui permettront un bon suivi du workflow de publication de la ressource, et veilleront à avoir des personnes pro-actives qui puissent interagir avec les enseignants pour avoir des métadonnées complètes.

À l'Insa de Lyon, il est envisagé de mettre en place le schéma d'indexation suivant :

⁴² Office québécois de la langue française. *Le grand dictionnaire terminologique*. [en ligne]. Québec : OQLF, 2002. Disponible sur : <http://www.granddictionnaire.com> . (Consulté le 10 juin 2005).

⁴³ TICE : Technologies de l'Information et de la Communication pour l'Enseignement

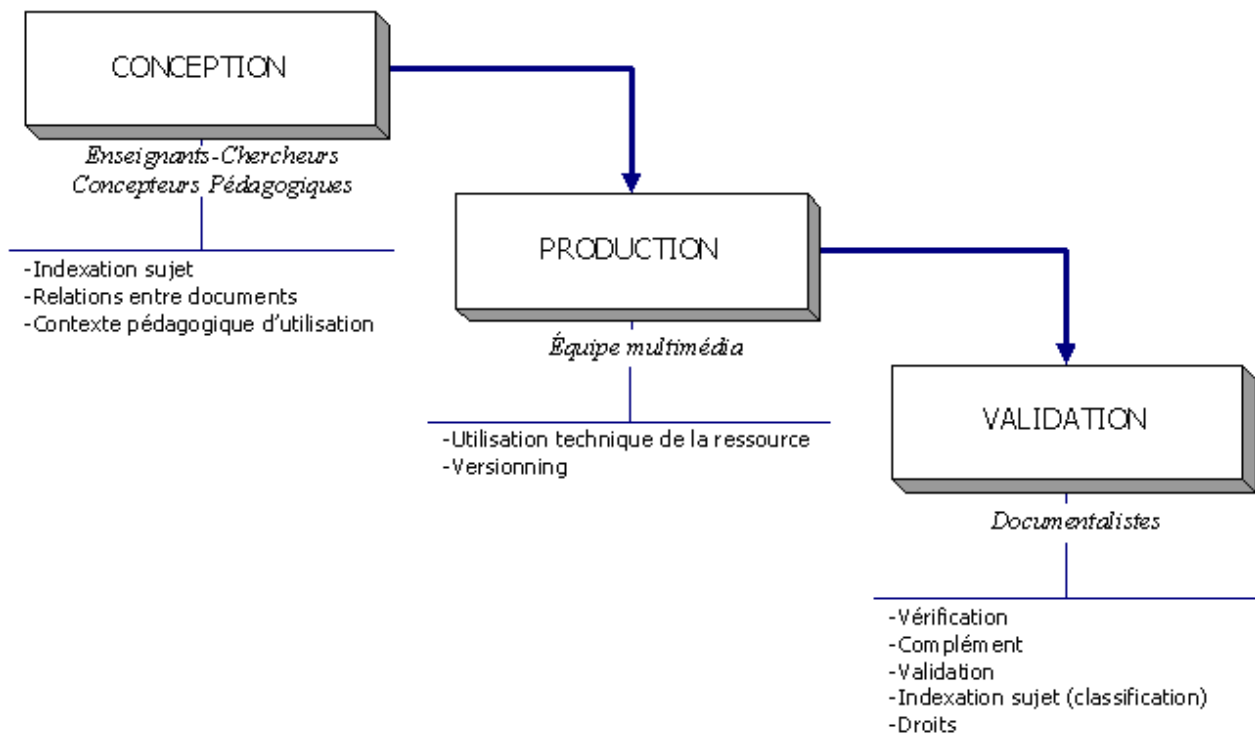


Figure 2 – Procédure d'indexation des ressources pédagogiques à l'Insa de Lyon

Conclusion

Le modèle de métadonnées LOM est très complet. Il est possible d'y repérer plusieurs types de métadonnées : descriptives, administratives, gestion de droits, préservation et interopérabilité.

Les travaux de normalisation dans le domaine des technologies éducatives et de la formation sont en cours de développement. En France, la CN36 (commission nationale de normalisation) s'occupe des travaux sur les technologies de l'enseignement et l'apprentissage. Un de ses projets est la publication d'un profil d'application français du LOM. Au niveau international, l'organisme ISO qui s'occupe des standards de l'éducation (ISO/IEC JTC1 SC36 WG4) travaille sur la norme internationale des métadonnées pour décrire les ressources pédagogiques, « Metadata for Learning Resources » (MLR).

Par ailleurs, la mise en œuvre de ce type de modèle nécessite un grand accompagnement de la part des décideurs et une intégration dans le système d'information dans lequel la gestion des ressources pédagogiques est inscrite. Il serait utopique de penser qu'une seule personne pourrait renseigner tous les champs de données du LOM. Il est indispensable d'avoir des bases de données de ressources pédagogiques cohérentes et de bonne qualité. Les utilisateurs (apprenants, enseignants et producteurs) auront tout à y gagner.

Bibliographie

Dublin Core Metadata Initiative. [en ligne]. Disponible sur : <http://dublincore.org>. (Consulté le 10.06.2005).

Educnet. *Dossier sur les métadonnées, normes et standards*. [en ligne]. 2004. Disponible sur : <http://www.educnet.education.fr/dossier/metadata/default.htm> (consulté le 10.06.2005).

HAYNES, David. *Metadata for information management and retrieval*. London : Facet, 2004, 186 p.

JEAN GODBY, Carol. *What do application profiles reveal about the Learning Object Metadata Standard?* [en ligne]. 30 octobre 2004. Disponible sur : <http://www.ariadne.ac.uk/issue41/godby/> (consulté le 10.06.2005).

Learning Technology Standards Committee of the IEEE. *Draft Standard for Learning Object Metadata*. [en ligne]. New York : IEEE, 2002. Disponible sur : http://ltsc.ieee.org/wg12/files/LOM_1484_12_1_v1_Final_Draft.pdf . (Consulté le 10 juin 2005).

Office québécois de la langue française. *Le grand dictionnaire terminologique*. [en ligne]. Québec : OQLF, 2002. Disponible sur : <http://www.granddictionnaire.com>. (Consulté le 10 juin 2005).

Sharable Content Object Reference Metadata (SCORM)

Gilles BERTIN
Chef de projet à Ancoly, Université de Lyon

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

Résumé :

SCORM est une norme qui prend de plus en plus d'importance, tant du côté des plates-formes pédagogiques – pour ses capacités d'interopérabilité – que du côté des contenus pédagogiques, pour les mêmes raisons ainsi que pour ses capacités d'indexation.

Après avoir présenté les besoins auxquels SCORM tente de répondre, nous décrirons ses deux volets : son modèle d'agrégation puis son environnement d'exécution. Nous terminerons en évoquant sa mise en œuvre et les problématiques liées.

Mots clés :

ressources pédagogiques numériques, SCORM (Sharable Content Object Reference Metadata), plate-forme pédagogique, LCMS (Learning Content Management System), indexation, normes et standards pédagogiques

Introduction

ANCOLY⁴⁴ est une émanation du pôle universitaire de Lyon (PUL) qui accompagne le développement des TICE dans les principaux établissements lyonnais : les trois universités, les deux écoles normales supérieures, les écoles... soit en tout quinze établissements regroupant 98 000 étudiants.

Cette communication s'appuie sur une expérience antérieure au sein de la société icampus : il s'agissait de déployer le campus virtuel de la société Auchan dans quatorze pays pour 26 000 employés ; c'est dans ce contexte que SCORM avait été utilisé – l'objectif étant l'interopérabilité et l'actualisation des contenus pédagogiques.

⁴⁴ ANCOLY : ANimation du Campus Ouvert de Lyon, <http://www.ancoly.net/>

L'origine de SCORM

SCORM est issu d'une démarche de standardisation entre acteurs américains gouvernementaux, universitaires et industriels. Le projet ADL⁴⁵ du DoD (Département de la Défense américaine) regroupait une fédération d'acteurs hétérogènes comprenant Ariadne⁴⁶, AICC⁴⁷, IEEE⁴⁸ et IMS⁴⁹. Cette initiative avait été lancée par l'AICC au départ, mais plutôt pour le off-line – notamment pour l'industrie aéronautique : il s'agissait de fédérer le travail fait dans les universités et les entreprises privées.

La norme la plus couramment utilisée actuellement est la version 1.2 de SCORM, qui traite du modèle d'agrégation et de l'environnement d'exécution. La nouvelle version de SCORM sortie en 2004 clarifie certains points de la version 1.2 en précisant le séquençement et la navigation. L'exposé qui suit a été conçu à partir de la version 1.2.

Pourquoi utiliser SCORM ?

SCORM répond à plusieurs besoins :

- L'interopérabilité entre les contenus pédagogiques et les applications informatiques. Un cours interopérable peut fonctionner sur différentes applications informatiques. Concrètement, il s'agit de pouvoir utiliser un même cours dans des plates-formes différentes ou inversement intégrer à sa propre plate-forme des cours de provenances diverses.
- La pérennité et la gestion du risque. Les plates-formes sont des outils très volatiles : les coûts d'exploitation d'une plate-forme peuvent évoluer très rapidement ; quant aux sociétés éditrices, elles disparaissent régulièrement. Par conséquent, il faut pouvoir changer facilement de plate-forme, sans avoir à réadapter les contenus. Lier ses contenus pédagogiques à une plate-forme particulière serait un risque trop important : il faut pouvoir pérenniser les contenus face aux évolutions technologiques.
- L'accessibilité, la réutilisabilité, l'adaptabilité. Il faut pouvoir chercher et trouver des ressources selon des mots-clés, assembler plusieurs ressources pour constituer un nouveau cours, maintenir et adapter les contenus.

⁴⁵ ADL : Advanced Distributed Learning, <http://www.adlnet.org>

⁴⁶ ARIADNE : Alliance of Remote Instructional Authoring and Distribution Networks for Europe

⁴⁷ AICC : Aviation Industry CBT Committee

⁴⁸ IEEE : Institute of Electrical and Electronics Engineers, <http://www.ieee.org/portal/site>

⁴⁹ IMS : Instructional Management System, <http://www.imsproject.org/metadata/index.html>

L'usage typique de SCORM est de gérer les interactions entre un entrepôt de cours et de ressources pédagogiques (ce que l'on appelle en informatique un « *repository* ») et un LMS⁵⁰ c'est-à-dire une plate-forme pédagogique chargée de diffuser ces ressources auprès des apprenants et des enseignants.

Les principales caractéristiques de SCORM

SCORM définit un modèle d'agrégation du contenu mais aussi un environnement d'exécution. Ce sont deux aspects complètement séparés – ce qui peut être un peu déroutant au départ.

Le modèle d'agrégation du contenu

Ce modèle d'agrégation du contenu vise à décrire et échanger des éléments de formation, au sein d'un entrepôt de documents pédagogiques. Il comporte trois niveaux de métadonnées, qui établissent pour chaque composante du modèle SCORM la correspondance avec les éléments LOM de l'IEEE :

- Le premier niveau est la composante « **Asset** » ou « Ressources multimédia ».
- Le deuxième niveau est la composante « **SCO** » ou « grains de formation ».
- Le dernier niveau est la composante « **Agrégation de contenus** » ou « Cours, parcours ».

Ces métadonnées sont regroupables dans des catalogues ou conditionnées avec les ressources qu'elles décrivent.

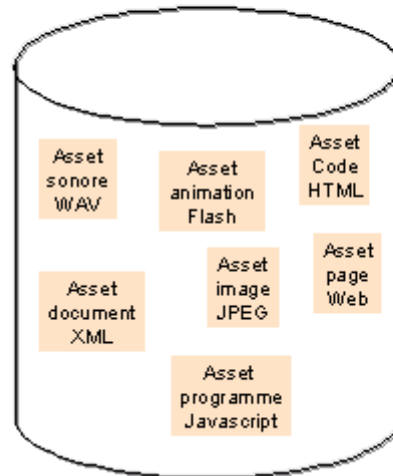
⁵⁰ LMS : Learning Management System

Modèle d'agrégation niveau 1 : les assets

Assets : forme la plus élémentaire du contenu pédagogique

Représentations électroniques de médias :

- texte,
- images,
- séquences sonores,
- pages web,
- objets d'évaluation,
- autres éléments pouvant être envoyés à un navigateur Web.



Un asset peut être décrit avec des métadonnées pour l'indexer dans des dépôts de données.

Ce premier niveau est celui du producteur de ressources.

Modèle d'agrégation niveau 2 : les SCO

SCO : Sharable Content Object (Objet de contenu partageable)

Ensemble comprenant :

- un ou plusieurs Assets,
- un actif lançable qui utilise l'environnement d'exécution du SCORM pour communiquer avec le LMS.

Le SCO est le niveau de granularité le plus fin des ressources d'apprentissage pouvant être suivi par un LMS.

Tout LMS supportant SCORM peut :

- lancer des SCO, peu importe qui les a générés,
- assurer le suivi d'un SCO,
- lancer n'importe quel SCO de la même façon.

Un SCO peut être décrit avec des métadonnées pour l'indexer dans des dépôts de données.

Exemple de SCO

```
Api = getAPI();  
Var result = api.LMSInitialize(" ");  
Var val =  
api.LMSGetValue("exemple.doc");
```

Asset XML doc Asset HTML object Asset Flash object

Asset JPEG Asset JPEG Asset Javascript

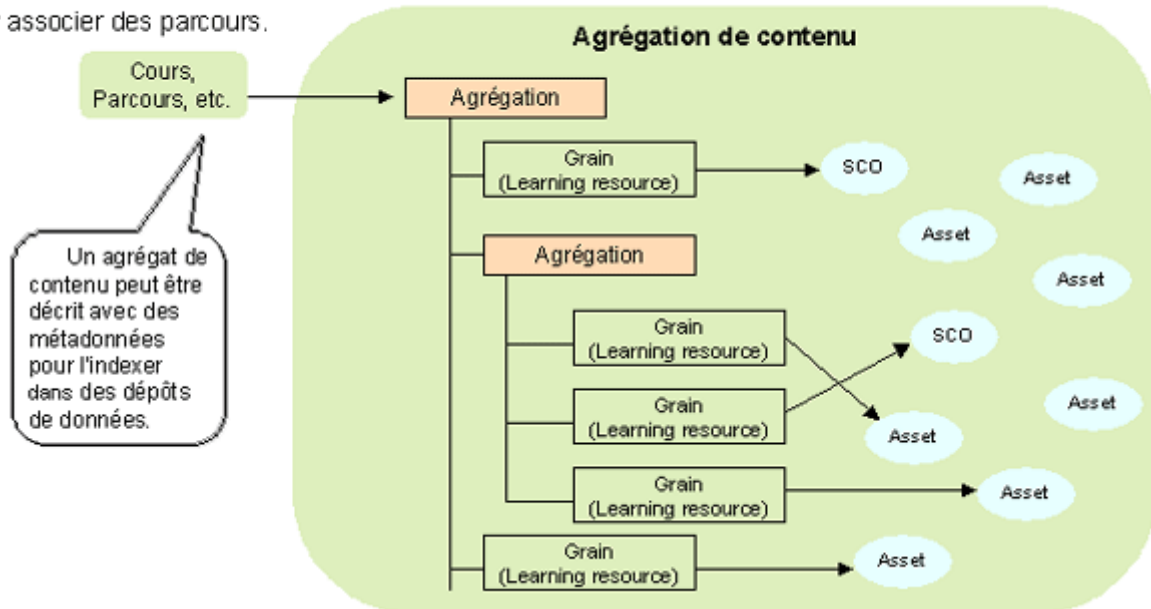
```
Api.LMSFinish(" ");
```

Ce deuxième niveau est celui de l'objet proprement pédagogique : un grain exécutable par un navigateur et qui peut être documenté de la même façon que la ressource multimédia.

Modèle d'agrégation niveau 3 : les agrégations de contenus

Correspondance utilisée pour :

- regrouper les grains en cours, chapitre, module, etc.
- y associer des parcours.



Ce troisième niveau est celui où l'enseignant assemble tous les grains pour constituer un cours et organiser des activités sur un trimestre ou sur une année.

L'environnement d'exécution

La partie environnement de SCORM est ce qui permet d'utiliser un cours sur n'importe quelle plate-forme. L'analogie avec les prises de courant peut être faite : elles ont été normalisées au niveau européen, de ce fait, un appareil électrique peut être branché dans n'importe quel pays d'Europe.

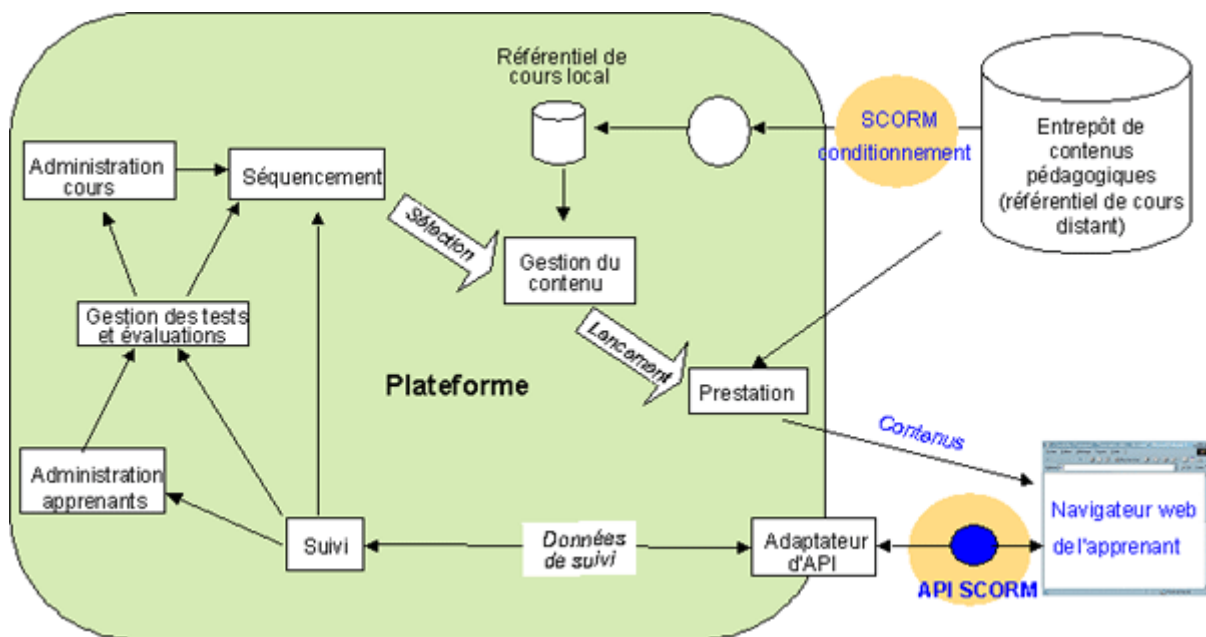
Comment cela fonctionne ?

Cet environnement d'exécution fonctionne à partir d'un API (*Application Programming Interface* – interface avec les applications informatiques) qui est un modèle de données permettant la standardisation des échanges entre le LMS et le navigateur web de l'apprenant – quel qu'il soit. C'est une boîte à outils dans laquelle l'informaticien va piocher pour faire remonter des informations vers la plate-forme sur l'activité de l'apprenant. Par exemple, s'il s'est arrêté à la moitié du cours, l'enseignant et le tuteur le sauront et il pourra reprendre au bon endroit la fois suivante.

L'environnement d'exécution fournit aux grains (SCO) une façon normalisée de communiquer avec les plates-formes et autres applications. Il « masque » aux techniciens créateurs de contenus les détails de la réalisation de l'interface de communication entre le grain et l'application qui le « jouera ».

Il est basé sur 3 catégories de fonctions :

- État d'exécution : LMSInitialize(""), LMSFinish("")
- Gestion de l'état (erreurs) : LMSGetErrorString()
- Transfert des données apprenant : LMSGetValue(), LMSSetValue()



Les outils pour « scormer »

Les outils permettant de générer des métadonnées SCORM sont de trois types :

- les outils de base XML, comme « Reload Scorm », utilisé à Lyon,
- les outils de production, dont certains intègrent déjà la possibilité de faire du SCORM (par exemple les logiciels appartenant à la suite Macromédia)
- enfin les outils auteurs, comme les plates-formes pédagogiques de type LCMS, qui permettent également de charger des fichiers SCORM ou de « scormer » des ressources.

Conclusion

Si SCORM est complexe à cause de ses trois niveaux, et orienté contenu plus qu'activité, ce standard présente néanmoins un certain nombre d'avantages : il intègre et complète très bien le LOM (ce qui a été précisé dans SCORM 2004) ; il n'a pas de concurrents et il est de mieux en mieux supporté par les plates-formes. SCORM constitue donc le choix pragmatique d'une interopérabilité fonctionnelle, réellement opératoire dès aujourd'hui...

LOM, SCORM et IMS-Learning Design : ressources, activités et scénarios

Jean-Philippe PERNIN
Institut National de Recherche Pédagogique, ERTé e-Praxis, Lyon,
Laboratoire CLIPS-IMAG, Grenoble

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

Résumé :

Cette présentation se concentre sur le concept d'objet d'apprentissage (Learning Object) qui focalise aujourd'hui un grand nombre d'attentes, que ce soit au niveau pédagogique, économique, politique ou culturel. Un des problèmes majeurs réside dans le manque de cohérence des différentes définitions, provenant principalement de la variété des points de vue adoptés pour aborder la question.

À travers différentes propositions (LOM, SCORM et IMS-LD), nous analysons trois des principales approches, apparues successivement : les langages d'indexation de données, les modèles de mise en œuvre informatique et enfin les langages de modélisation pédagogique. Ces derniers, qui font à l'heure actuelle l'objet d'études approfondies, présentent de réelles avancées en remplaçant le point de vue pédagogique au centre du dispositif et en proposant de décrire l'ensemble des relations sémantiques liant les activités et les ressources de manipulation de connaissance. En particulier, cette nouvelle approche nécessite de s'intéresser aux usages effectifs des ressources d'apprentissage au sein de communautés de pratique (enseignants, formateurs et documentalistes), tant en termes de mutualisation, de réutilisation que de stratégies d'indexation.

Mots clés :

indexation, LOM (Learning Object MetaData), SCORM (Sharable Content Object Reference Metadata), IMS-LD (IMS-Learning Design, scénarios pédagogiques, ressources pédagogiques numériques, normes et standards pédagogiques.

Évolutions des représentations du rôle de l'enseignant/formateur

Le rôle de l'enseignant utilisant les TICE a considérablement évolué depuis dix ou quinze ans. Au départ, dans une approche très artisanale, l'enseignant était auteur et il était formé à des

outils lui permettant de créer lui-même ses cours.

La deuxième approche était plutôt éditoriale, l'enseignant devenait concepteur et prescripteur de contenu : il donnait son cours à réaliser à d'autres – les services multimédias, en l'occurrence.

La logique qui a prévalu ensuite repose sur une approche « documentaliste », où l'enseignant est prospecteur, référenceur et organisateur de ressources. Il compose son cours avec des objets d'apprentissage, qu'il crée de toutes pièces ou réutilise. Pour cela, il doit être capable de chercher et de trouver ces objets d'apprentissage, de les référencer, de les indexer et bien évidemment de les organiser dans des unités de plus haut niveau.

L'enseignant est ensuite devenu plutôt un scénariste au travers de ce que l'on nomme les unités d'apprentissage ou les scénarios : l'approche est alors centrée sur l'activité.

Des objets d'apprentissage aux langages de modélisation pédagogique

Un objet d'apprentissage est « toute entité numérique ou non, qui peut-être utilisée, réutilisée ou référencée lors d'une formation dispensée à partir d'un support technologique⁵¹. »

Objet d'apprentissage : un concept au centre des tensions

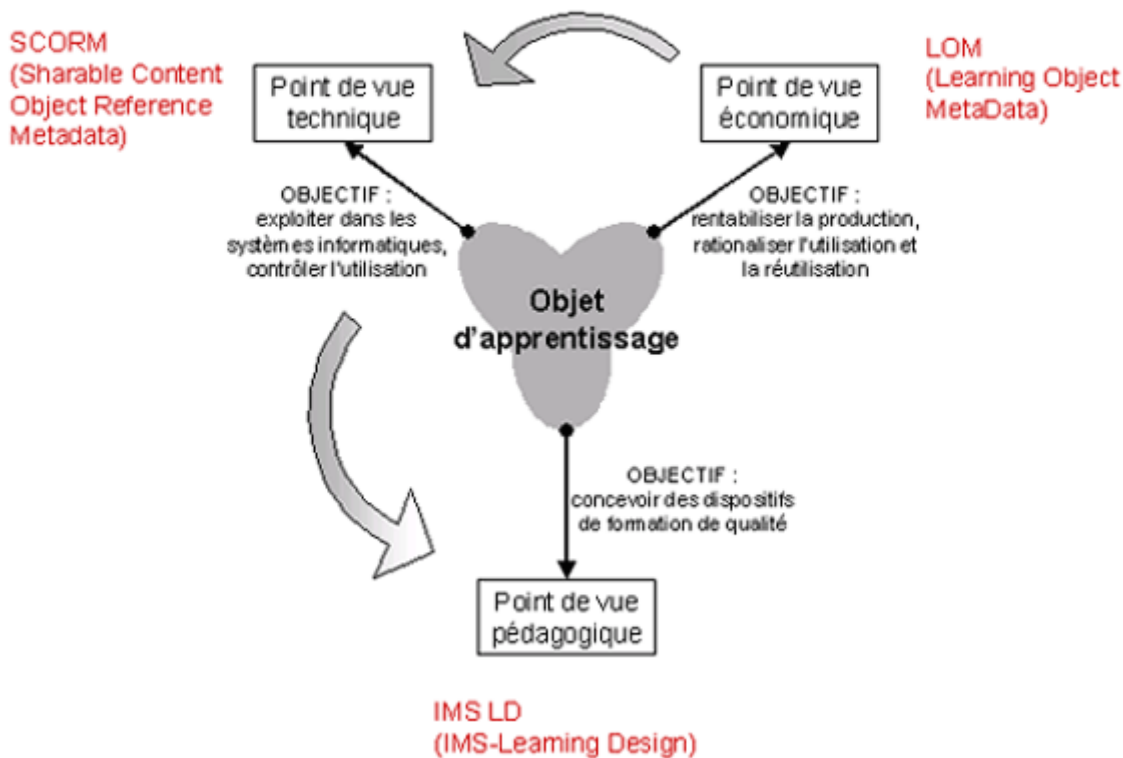


Figure 1 : Objet d'apprentissage, un concept au centre de tensions

Trois approches sont apparues successivement et ont débouché chacune sur des normes ou propositions de normes : LOM, SCORM et IMS-Learning Design.

Le terme « objet d'apprentissage » émerge au milieu des années 1990 au sein de consortiums internationaux comme IMS⁵² ou Ariadne⁵³ – ce qui amènera à proposer un standard, le LOM (Learning Object Metadata), au début des années 2000. L'objectif est alors de rentabiliser la production et de développer la réutilisation (point de vue économique).

⁵¹ Définition de l'objet d'apprentissage selon l'IEEE, Institute of Electrical and Electronics Engineers : <http://www.ieee.org/portal/site>

⁵² IMS : Instructional Management System, <http://www.imsproject.org/metadata/index.html>

⁵³ Ariadne : Alliance of Remote Instructional Authoring and Distribution Networks for Europe, <http://www.ariadne-eu.org/>

À la fin des années 1990, l'apparition des LMS et LCMS⁵⁴ (plates-formes de formation ouverte et à distance) conduit à se poser la question de l'exploitation technique des objets d'apprentissage avec internet, d'où l'apparition de SCORM⁵⁵ au sein des consortiums « corporatistes » américains – comme AICC⁵⁶. L'objectif est de pouvoir exploiter la ressource dans un système informatique et d'en contrôler l'utilisation (point de vue technique).

Dix ans plus tard, la question qui prévaut est l'intégration des objets dans la conception de situations d'apprentissage : les activités – et non les ressources – deviennent centrales⁵⁷. D'où la création d'IMS-LD Design, qui vise la conception de dispositifs de formation de qualité (point de vue pédagogique).

LOM et le modèle d'agrégation

Objectif du LOM (Learning Object Metadata)

Le LOM ne provient pas du monde documentaire : il a été créé par des informaticiens. L'objectif du LOM est d'indexer des objets d'apprentissage pour les réutiliser dans des *curricula*⁵⁸. Il est fondé sur le principe du « share and reuse ».

À l'origine, il s'agissait de réutiliser les objets pédagogiques dans des unités de plus haut niveau, comme des hypertextes, des simulations, des ressources d'apprentissages plus ou moins interactives. Ce principe de réutilisation pose cependant un certain nombre de questions : quels sont les usages effectifs de réutilisation ? Pour quels besoins ? Qui veut réutiliser ces grains et pourquoi ? Comment construire une solution de formation par agrégation ?

Un modèle de structuration sous-jacent

Le modèle de structuration du LOM comprend :

- des unités de structuration (curriculum, cours, leçon),
- 4 niveaux de ressources de granularités diverses.

⁵⁴ LMS : Learning Management System ; LCMS : Learning Content Management System.

⁵⁵ SCORM : Sharable Content Object Reference Metadata, <http://www.scorm.com/>

⁵⁶ AICC : Aviation Industry CBT Committee

⁵⁷ Koper, Rob. *Modeling units of study from a pedagogical perspective: the pedagogical meta-model behind EML*. Open Universiteit Nederland, 2001, [en ligne] <http://dspace.ou.nl/handle/1820/36>.

⁵⁸ Curriculum : « Enoncé d'intentions de formation comprenant : la définition du public cible, les finalités, les objectifs, les contenus, la description du système d'évaluation, la planification des activités, les effets attendus quant à la modification des attitudes et des comportements des individus en formation. » Raynal Françoise, Rieunier Alain. *Pédagogie : dictionnaire des concepts-clés*, Paris, ESF éditeur, 1997.

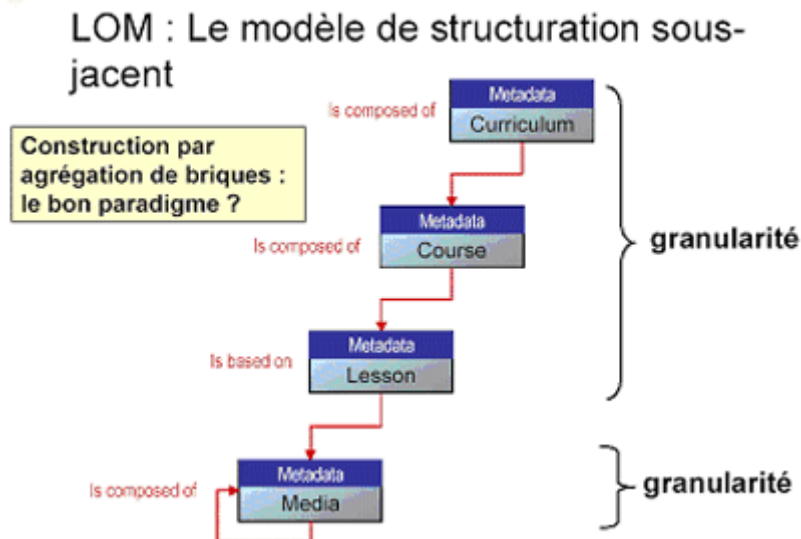


Figure 2 : LOM, le modèle de structuration sous-jacent

C'est un modèle de structuration basé sur des grains de quatre niveaux. Des médias peuvent être constitués d'autres médias qui peuvent eux-mêmes être décomposés. Cette notion de granularité est aujourd'hui parfois remise en cause. Les structures de haut niveau – leçon, cours et curriculum – se prêtent bien à certains enseignements et très mal à d'autres. Ce qui conduit à s'interroger sur la pertinence d'une construction par agrégation de briques.

L'approche technique : SCORM

SCORM : Sharable Content Object Reference Metadata est le second langage, fondé sur le concept « d'objet de contenu partageable ». Il vient des États-Unis et plus précisément de la formation dans l'industrie aéronautique. Cette initiative a été développée par l'AICC puis reprise par le consortium ADL59.

Ces propositions concernent principalement les contenus en ligne. Il s'agissait de faire de la formation totalement libre à travers le web. SCORM enrichit le standard LOM avec notamment un modèle d'agrégation un peu différent et un environnement d'exécution qui permet de surveiller l'activité d'un apprenant dans un LMS.

L'organisation sur trois niveaux, est la suivante :

- le niveau de base appelé *asset* ;
- un niveau intermédiaire où se trouvent des objets de contenus partageables. Il s'agit de la plus petite unité pouvant donner des informations sur l'utilisation, par exemple : tel document a-t-il été parcouru ? telle simulation a-t-elle été effectuée ?

⁵⁹ ADL : Advanced Distributed Learning, <http://www.adlnet.org>

- un niveau supérieur, concernant l'agrégation de contenus (des cours par exemple).

C'est donc au niveau intermédiaire que la notion d'activité sur la ressource est prise en compte.

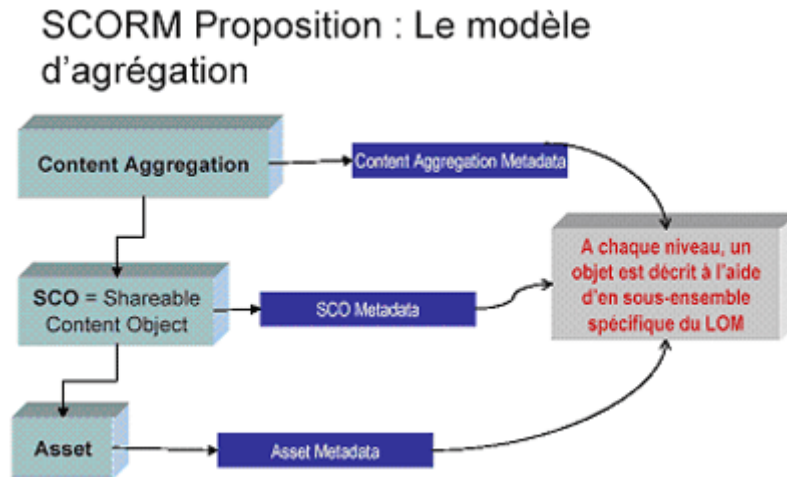


Figure 3 : SCORM, le modèle d'agrégation

La deuxième composante de SCORM concerne l'environnement d'exécution. Le principe consiste à renvoyer vers une plate-forme de formation des informations sur l'état d'utilisation d'un objet (a-t-il été parcouru ? lu ? effectué avec succès ? etc.). Ces informations peuvent être stockées dans une base de données et exploitées par la suite sur le plan pédagogique.

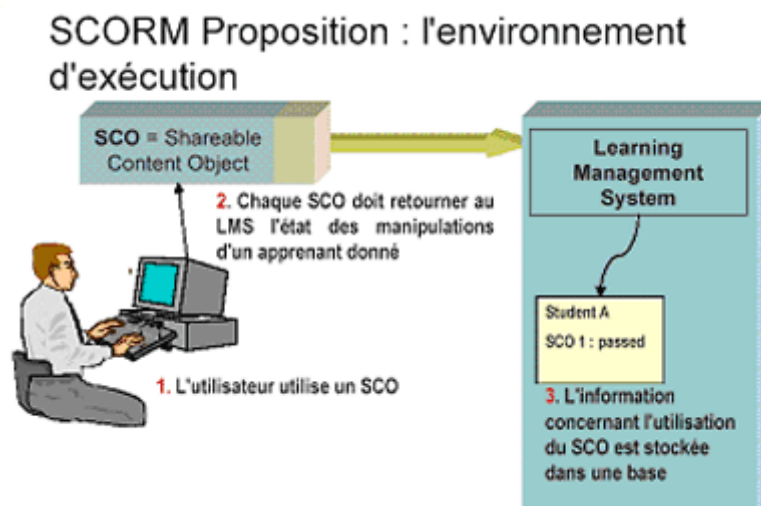


Figure 4 : SCORM, l'environnement d'exécution

L'approche pédagogique : les langages de modélisation pédagogique

Constats

Ces deux premiers points de vue ont été fortement critiqués au travers du courant qui peut être nommé « langage de modélisation pédagogique », et qui met l'accent sur la mise en scène des ressources dans un processus de formation, plus que sur l'agrégation de ressources. Ce courant, issu d'EML60 en 2001, promu par Rob Koper de l'« Open University of Netherlands » a été adopté en 2003 par IMS comme « Learning Design ».

Le constat est le suivant : ce ne sont pas les objets pédagogiques qui sont centraux dans le processus d'apprentissage mais les activités qui leur sont associées. Par conséquent, cela remet en cause certains aspects du LOM, qui est peu adapté à cette représentation, puisque les objets ne peuvent être a priori définis en dehors de leur usage dans des « unités d'apprentissage ».

Objectifs

L'objectif est de fournir des modèles adaptés à la conception pédagogique de situations d'apprentissage diversifiées. L'activité est mise au centre du processus ; elle se définit comme une tâche avec un objectif pédagogique précis ; elle est effectuée par un certain nombre de personnes qui tiennent différents rôles (enseignants, tuteurs, etc.) dans un certain environnement et en s'appuyant sur un certain nombre de ressources, numériques ou non, parmi lesquelles les objets pédagogiques, des ressources documentaires, des simulations...

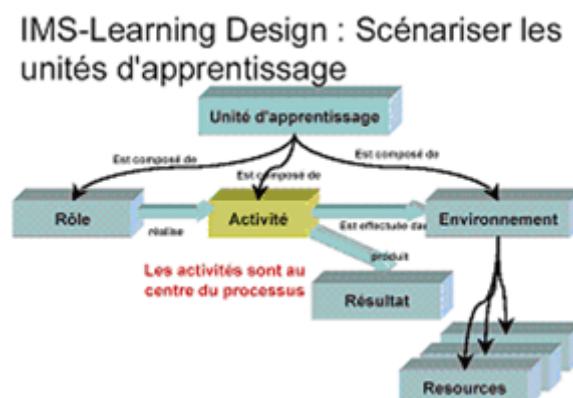


Figure 5 : IMS-Learning Design, scénariser les unités d'apprentissage

⁶⁰ EML : Educational Modelling Language.

Ce modèle différencie le catalogage et l'indexation de la ressource de celui des activités liées à ces ressources de base.

Synthèse : LOM, SCORM, IMS-LD

Trois niveaux doivent être différenciés dans une comparaison entre ces trois modèles, :

- un niveau de base qui traite des médias bruts (texte, image ou son) ;
- un niveau intermédiaire qui essaie d'introduire la notion d'activité : cette ressource doit être consommée, consultée, exécutée ;
- enfin un niveau supérieur de structuration des activités avec des données de plus haut niveau comme en particulier la façon de les exécuter dans le temps : leçon, cours, curriculum (les noms diffèrent suivant les modèles).

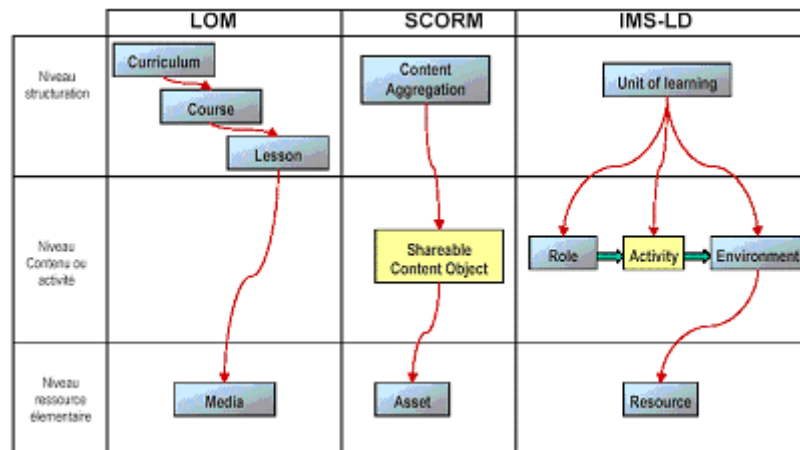


Figure 6 : Synthèse entre LOM, SCORM et IMS-LD

L'idée est de différencier les champs qui vont qualifier chacun de ces niveaux. Un champ traite uniquement des données documentaires brutes que l'on peut combiner entre elles ; un champ intermédiaire va fournir des informations sur l'activité effectuée sur ces ressources ; et un troisième niveau est l'agrégation de ces activités dans des structures de plus haut niveau.

Les évolutions du LOM

Yolaine BOURDA
Professeur d'informatique à Supelec, membre de la CN 36⁶¹

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

Résumé :

Cette communication présente les principales évolutions du LOM (Learning Object MetaData) et notamment ses différents profils d'application. Nous nous intéresserons plus spécialement à l'objet d'apprentissage ou « learning resource type » décrit dans le LOM. Nous parlerons aussi du « Metadata Learning Ressource » ou MLR, développé au sein de l'ISO et du Working Group 4, et censé être compatible LOM. Nous ferons enfin le point sur les métadonnées françaises et notamment sur la publication prochaine par l'AFNOR d'une norme expérimentale (EXP).

Mots clés :

Indexation, ressources pédagogiques numériques, normes et standards pédagogiques, LOM (Learning Object Metadata), AFNOR, normalisation, MLR (Metadata Learning Ressource).

L'utilisation du LOM⁶² (Learning Object Metadata) est rendue difficile par le manque de précisions autour des définitions des éléments et de leurs valeurs. Les exemples ne suffisent pas à lever les ambiguïtés. Ainsi, lors d'une procédure rapide de normalisation auprès de l'ISO⁶³, au printemps 2002, un certain nombre de commentaires ont été apportés, relatifs à des définitions, des champs à scinder ou des éléments à ajouter. Mais les IEEE ayant refusé de discuter, le processus de normalisation, qui nécessite que tout commentaire soit discuté, n'a pu aboutir.

⁶¹ Commission de l'AFNOR en charge de la normalisation des technologies de l'information et de la communication pour l'enseignement, la formation et l'apprentissage, responsable du groupe 4 sur les métadonnées.

⁶² Le LOM est un produit des IEEE (<http://www.ieee.com>), qui maîtrisent seuls les évolutions autour de la norme 2 du LOM. Pour pouvoir participer à la réflexion autour de ces évolutions, il faut cotiser auprès des IEEE.

⁶³ ISO : International Organization for Standardization, <http://www.iso.org/iso/fr/ISOOnline.frontpage>

Les profils d'application du LOM

En fait le LOM est toujours utilisé dans le cadre d'un **profil d'application**, c'est-à-dire un choix parmi les éléments du LOM. Un profil d'application permet aussi d'affiner, de spécialiser des définitions (mais il n'est pas possible de généraliser). Il peut aussi contenir des éléments qui ne sont pas dans le LOM, intégrer des éléments de métadonnées qui viennent d'ailleurs. Des valeurs peuvent être spécifiées pour certains éléments. Les éléments facultatifs peuvent être rendus obligatoires. En fait l'objectif d'un profil d'application est de préserver l'interopérabilité du schéma de base.

Au sein de l'ISO, le groupe de travail chargé des métadonnées est le Working Group 4 (WG 4). Le WG 4 a fait une étude comparative⁶⁴ des différents profils d'application utilisés. La conclusion de cette étude est que l'interopérabilité ne fonctionne pas. Il n'est pas possible de prendre des descriptions venant d'un profil chinois, d'un profil canadien et d'un profil français : certains vocabulaires et certains éléments sont incompatibles. Autre point intéressant, les éléments réellement utilisés sont principalement ceux provenant du Dublin Core...

Pour mieux comprendre ce qu'est un profil d'application, on peut prendre pour exemple le célèbre champ 5.2, le « learning resource type » dont la définition anglaise est la suivante :

« Specific kind of learning object. The most dominant kind shall be first; Values: exercise, simulation, questionnaire, diagram, figure, graph, index, slide, table, narrative text, exam, experiment, problem statement, self-assessment, lecture⁶⁵. Conform to the DCMI type element (Collection, Dataset, Event, Image, InteractiveResource, Sound, Service, Software, Text. »

Qu'est-ce exactement que cet « objet d'apprentissage » ? Les valeurs sont trop vagues pour le comprendre, avec des mélanges de données de natures différentes : une « figure » se retrouve sur le même plan qu'un « exercice » ; certains éléments correspondent à des notions techniques (slides, texts, graphs), d'autres à des notions pédagogiques. En outre, la conformité énoncée avec Dublin Core est fautive, puisque toutes les valeurs de Dublin Core ne se retrouvent pas ici.

L'objectif d'un profil est de pouvoir retrouver une ressource – par exemple un exercice de mathématiques de niveau licence – à l'intérieur d'entrepôts de données différents, basés sur des profils d'applications différents.

⁶⁴ Document numéro 109.

⁶⁵ À noter : aucune définition de ces éléments n'est proposée.

Mettre en place un processus de standardisation et de normalisation, c'est déjà choisir des éléments, bien les définir et s'intéresser aux valeurs qu'ils peuvent prendre. Se contenter de chaînes de caractères reviendrait à une recherche en plein texte.

Le « Metadata Learning Ressource » ou MLR

L'ISO travaille actuellement sur le « Metadata Learning Ressource » ou MLR, au sein du Working Group 4 (WG 4). La France est coéditrice de la première partie de cette norme (dont le nombre de parties n'est pas encore défini). L'enjeu consiste à faire en sorte que le MLR soit compatible avec le LOM – afin de préserver ce qui est déjà acquis – tout en prenant en compte certains points qui ne sont pas abordés dans le LOM.

Dans le LOM, le nom d'un champ est confondu avec sa signification, or celle-ci n'est pas neutre : elle dépend du contexte, de la culture, de la langue. L'idée est donc de s'intéresser aux définitions à un niveau abstrait, afin que tous, quels que soient leur culture et leur langue, puissent s'accorder sur la signification de chaque élément, sans se focaliser sur son nom. Une fois que la signification d'une valeur est bien définie, il est possible de lui donner plusieurs noms, ce qui résoudrait le problème du multilinguisme. Des intersections ou des unions de concepts peuvent aussi permettre d'avoir différents niveaux d'implémentation, avec le même niveau de définition conceptuel.

La France préconise pour le MLR l'utilisation du schéma ISO 11179, norme utilisée pour les registres de métadonnées, dont voici les définitions :

“Data Element Concept (DEC) : a concept that can be represented in the form of a Data Element, described independently of any particular representation.

Conceptual Domain (CD) : a set of valid Value Meanings. The Value Meanings may either be enumerated or expressed via a description.

Value Domain (VD) : a set of Permissible Values. The Value Domain provides representation, but has no implication as to what Data Element Concept the Values may be associated with nor what the Values mean.

Data Element (DE) : a unit of data for which the definition, identification, representation and Permissible Values are specified by means of a set of attributes

Value Meaning (VM) : the meaning or semantic content of a Value. The representation of Value Meanings in a registry shall be independent of (and shall not constrain) their representation in any corresponding Value Domain.

Permissible Value (PV) : an expression of a Value Meaning allowed in a specific Value Domain.”

ISO/IEC 11179

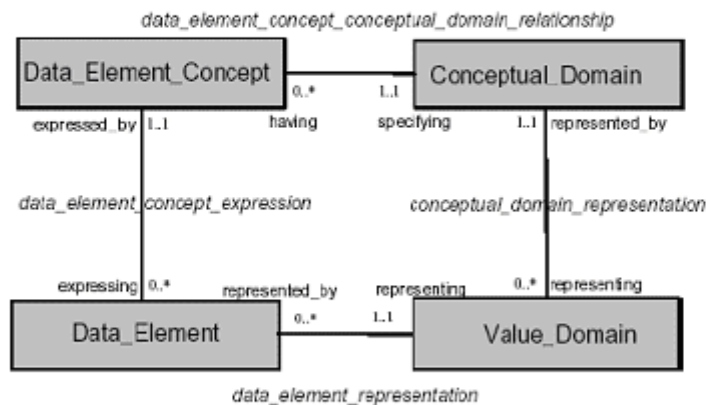


Figure 3: High-level metamodel

Le « *data element concept* » permet de définir toutes les données.

Le « *conceptual domain* » est un ensemble de significations (et non de valeurs), sur lequel il est possible de faire des relations, des opérations, des unions.

Le « *value domain* » est le niveau des valeurs : à une définition est associée une valeur. La cardinalité indique que pour un domaine conceptuel, il peut y avoir autant de « *value domain* » que nécessaire.

Pour faire une analogie avec les profils d'application du LOM,

- le niveau du « *conceptual domain* » correspond à celui des définitions,
- le niveau des « *value domain* » correspond à tous les profils d'application.

Comme les profils d'application seront reliés au même domaine conceptuel, le passage d'une valeur à l'autre sera automatique. Les logiciels sauront faire cela. Pour l'instant, n profils d'application nécessitent 2^n traducteurs ; avec le MLR, n profils d'application nécessiteront n traducteurs.

Voici un exemple concret :

- **Au niveau Conceptuel (MLR)**

(définition de la signification de « interactivity type » du LOM et de la signification de ses valeurs)

DEC	name InteractivityType DEC
DEC def	predominant mode of learning supported by this learning object
CD name	InteractivityType CD
CD def	Enumeration of all the modes of learning
VM	<V1, The learner produces some actions> <V2, The learner is passive> <V3, For some parts of the resource, the learner produces actions and for other parts, the learner is passive>

- **Au niveau de l'instanciation LOM**

(valeurs anglaises de « interactivity type » du LOM)

VD	name InteractivityType VD en
VD	definition Enumeration in english
PV	<V1, active> <V2, expositive> <V3, mixed>

- **Au niveau de l'instanciation MPE**

(valeurs françaises de « interactivity type » du LOM)

VD	name InteractivityType VD fr
VD	definition Enumeration en français
PV	<V1, actif> <V2, passif> <V3, mixte>

Les métadonnées françaises

À la demande de la direction de la technologie du ministère de l'Éducation nationale, l'AFNOR⁶⁶ doit produire une norme française de métadonnées pour les ressources numériques

⁶⁶ AFNOR : Agence française de normalisation, <http://www.afnor.fr/portail.asp>

éducatives. Ce profil d'application du LOM doit répondre aux besoins de l'Éducation nationale et du plus grand nombre d'acteurs français de la formation initiale et continue.

Ce profil d'application français sera une « norme expérimentale »⁶⁷, c'est-à-dire une norme avec une durée de vie assez courte, car le domaine est mouvant. La base en est ManUeL⁶⁸, un profil d'application du LOM réalisé pour C@mpusScience. Ce profil a été diffusé au sein du groupe de travail G4 de l'AFNOR, puis à un certain nombre d'autres acteurs, pour commentaires. Des discussions ont ensuite eu lieu (novembre 2004), avant une dernière enquête élargie (janvier 2005) ; la publication de la norme expérimentale est prévue en mars 2005 par l'AFNORR⁶⁹.

Chacun peut participer à la construction de cette norme expérimentale. Plus grande sera la diversité des acteurs impliqués dans l'élaboration de cette norme, plus le résultat produit sera à même de répondre aux besoins de tous les acteurs de l'enseignement français.

⁶⁷ Norme expérimentale (EXP) : document publié alors que les domaines en question sont en forte évolution. Valable pour une durée limitée (3 ans maximum) et ensuite soit révisée, soit retirée, soit transformée en norme homologuée.

⁶⁸ Brigitte de La Passardière, Pierre Jarraud, ManUeL, un profil d'application du LOM pour C@mpuSciences®, Revue Sticef, Volume 11, 2004, ISSN : 1764-7223, mis en ligne le 30/12/2004, <http://sticef.org>

⁶⁹ XP Z76-040, Août 2005. Technologies de l'information pour l'éducation, la formation et l'apprentissage - Profil français d'application du LOM (LOMFR) - Métadonnées pour l'enseignement.

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 DOI⁷⁰ (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.

⁷⁰ <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 TEI⁷¹ (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 Scirus⁷². Gallica⁷³ 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

⁷¹ <http://www.tei-c.org>. Liste de projets TEI : <http://www.tei-c.org/Applications/index.html>, 117 projets.

⁷² <http://www.scirus.com/>

⁷³ 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 DEA⁷⁴ 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 concepteurs⁷⁵ 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

⁷⁴ 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

⁷⁵ Se reporter à l'intervention de Laurent Flory.

l'été dernier⁷⁶ 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 OCLC⁷⁷, 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.

⁷⁶ Baczkowski, B., « Campus numériques: ressources et contenus », Diplôme d'études approfondies DISIC, École doctorale Informatique et information pour la société, 2004.

⁷⁷ Enquête de C. Jean Godby/OCLC, publiée dans Ariadne : <http://www.ariadne.ac.uk/issue41/godby>, octobre 2004.

Partie 3 – Rôle et positionnement des acteurs : quels partenariats entre SCD et TICE ?

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 universitaire⁷⁸, 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,

⁷⁸ 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 2005⁷⁹ 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-PMH⁸⁰ 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 FOAD⁸¹. 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.

⁷⁹ Le projet n'a pas encore vu le jour en décembre 2005.

⁸⁰ OAI-PMH : Open Archives Initiative - Protocol for Metadata Harvesting.

⁸¹ 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 DSpace⁸², application libre conçue au MIT⁸³ 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

⁸² <http://www.dspace.org>

⁸³ 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).

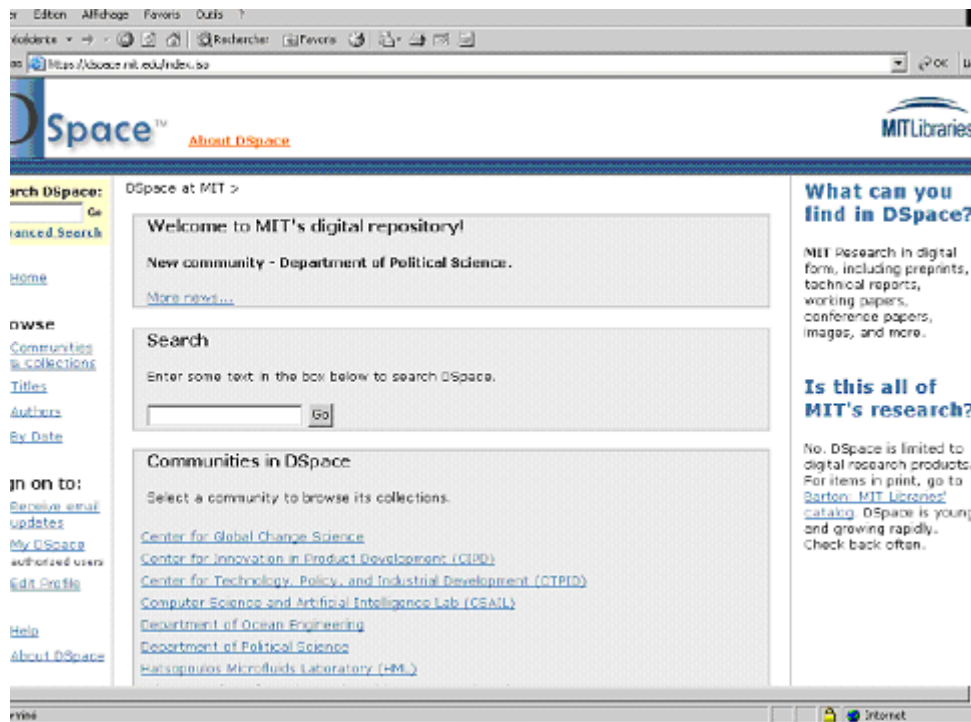


Figure 1 : Site internet de DSpace

Les partenariats autour du portail documentaire de l'université de technologie de Compiègne

Annie BERTRAND
Service Commun de la Documentation (BUTC),
Université de Technologie de Compiègne,
Directrice

Xavier HENNEQUIN
Unité d'Innovation « Ingénierie des Contenus et des Savoirs »,
Université de Technologie de Compiègne,
Ingénieur – Équipe projet

Note : article rédigé par les auteurs à la suite de leur communication orale du 16 novembre 2004.

Résumé :

L'université de technologie de Compiègne (UTC) a, en juin 2004, ouvert l'accès à son système d'information documentaire. Porteur de la stratégie de l'établissement en matière de TIC, il signale, indexe et met à disposition de tout utilisateur autorisé non seulement des ressources électroniques acquises ou sélectionnées par le SCD mais également des ressources produites en interne (thèses, cours, rapports...). L'enrichissement du portail et la production des référencements de ces ressources ont été imaginés dans une approche totalement collaborative : un auteur potentiel (thésard, enseignant-chercheur) ou un service autorisé (école doctorale...) peut déposer un document et participer à son signalement. Le respect de normes de description (Dublin Core, LOM...) permet la récupération automatique de métadonnées, par exemple celles qui proviennent de la chaîne éditoriale Scenari, conçue par l'unité d'innovation ingénierie des contenus et des savoirs (ICS) de l'UTC.

Mots clés :

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

Introduction

L'université de technologie de Compiègne (UTC) a ouvert, en juin 2004, l'accès à son système d'information documentaire. Porteur de la stratégie de l'établissement en matière de TIC, il signale, indexe et met à la disposition de tout utilisateur autorisé non seulement des ressources électroniques acquises ou sélectionnées par la bibliothèque (BUTC) mais également des ressources produites en interne (thèses, cours, publications de chercheurs, rapports d'étudiants et mémoires divers...). L'enrichissement du portail documentaire et la production des référencements de ces dernières ressources ont été imaginés dans une approche collaborative : un auteur potentiel (thésard, enseignant-chercheur...) ou un service autorisé (école doctorale...) peut y déposer un document et participer à son signalement.

Le respect des normes de description de ressources électroniques (Dublin Core, LOM...) doit permettre la récupération automatique de métadonnées. Cette récupération est déjà opérationnelle pour les coordonnées des auteurs (noms) et les informations sur les enseignements (intitulés et sigles de l'unité de valeur [UV]). Elle a été étendue à d'autres métadonnées documentaires du modèle pédagogique « académique » de la chaîne éditoriale Scenari⁸⁴ appliquée aux cours et conçue par l'unité d'innovation ingénierie des contenus et des savoirs (ICS) de l'UTC.

L'enrichissement du portail documentaire de l'UTC⁸⁵

Le portail, point central de diffusion de la production de l'UTC

Au-delà d'un outil fédérateur de ressources permettant une recherche facilitée sur l'ensemble des sources acquises ou sélectionnées et décrites par la BUTC, il a été de la volonté officiellement affirmée de la direction de l'UTC d'en faire le point central de gestion, de diffusion et de valorisation de la production interne de l'établissement.

C'est ainsi que, depuis l'automne 2004, une quarantaine de cours numériques sont référencés dans le portail. À ces références est associé soit le texte intégral du document si celui-ci a été déposé sur le serveur du portail, soit une adresse active si le document est stocké sur un autre serveur de l'université.

⁸⁴ Bachimont B., Cailleau I., Crozat S., Majada M., Spinelli S. *Le procédé SCENARI : Une chaîne éditoriale pour la production de supports numériques de formation TICE 2002*, Lyon 13-15 novembre, 2002], p. 183-192. [en ligne] <http://docinsa.insa-lyon.fr/tice/2002/cs/cs036.html> [30/09/2003].

⁸⁵ <http://eliot.utc.fr>

Lors de la phase de réalisation de ce portail documentaire, dans la double perspective à la fois d'optimiser les processus internes et de mobiliser les personnels de l'UTC autour de ce projet, il a été choisi de mettre en place des technologies de workflow qui devaient avoir pour objectif de permettre à tout auteur potentiel et autorisé de déposer lui-même son document et donc, en particulier, son cours.

De cette collaboration, chacun devient ainsi un partenaire actif dans l'enrichissement du portail.

Lors du dépôt de son cours, chaque enseignant-chercheur propose, de manière totalement libre, des mots ou expressions sujets décrivant le contenu de son œuvre, et décide des autorisations de consultation (restrictions éventuelles sur la population des lecteurs et/ou sur la période de consultation). Les personnels de la BUTC habilités, informés par mail automatique du dépôt d'un document, contrôlent et enrichissent la partie descriptive par, entre autres, la saisie de mots Rameau.

La représentation des éléments du Dublin Core dans le modèle conceptuel de données (MCD) du portail

Le Dublin Core a été retenu pour décrire les ressources électroniques du portail. Une sélection des éléments de cette norme qui sont le plus souvent utilisés et qui ont été adaptés à nos besoins spécifiques est présentée ci-après.

Éléments Dublin Core	Modèle conceptuel de données du portail	Commentaires
Table « NOTICE GÉNÉRALE »		
• Identifier	UTIL_CREATE DATE_CREATE UTIL_MAJ DATE_MAJ	Identifiant créateur notice
• Type	TYPE_DOC UNITÉ_VALEUR	Exemple : cours Exemple : GE15
• Title	TITRE SOUS_TITRE	
• Contributor	AUTRES_AUTEURS RESPONSABLE	Responsable de l'UV
• Publisher	NOM_ÉDITEUR	
• Date	ANNÉE_PUBLICATION	
• Subject	SUJETS MOTS_CLÉS DOMAINE_DISCIPLINE SEMESTRE	Rameau Automne ou Printemps
• Description	DESCRIPTION_1 DESCRIPTION_2	Résumé en français Résumé en anglais
• Language	LANGUE	
Table « DOCUMENT »		
• Rights	UTILS_AUTO GROUPS_AUTO DATE_DEBUT DATE_FIN	
• Identifier	URL DOCNAME	
• Title	TITRE_DOC	
• Format	FORMAT_DIFFUSION FORMAT_TAILLEDOC	

Un exemple de partenariat : la chaîne éditoriale Scenari⁸⁶ et le portail documentaire de l'UTC

La chaîne éditoriale SCENARIchain permet la publication multisupports (version écran, version imprimable, etc.) de contenus de cours rédigés en XML. Depuis maintenant deux ans, certains enseignants-auteurs de l'UTC ont pu utiliser cette chaîne éditoriale. Une des exploitations possibles de ces contenus de cours est la mise à disposition à travers le portail documentaire de l'UTC. Un travail commun a donc été réalisé pour faire le lien entre les documents issus de la chaîne éditoriale et le portail documentaire.

Nature du lien existant entre les documents Scenari et le portail documentaire de l'UTC

Le premier lien qui a été établi dès 2004 passe par l'exploitation des métadonnées disponibles dans les supports de cours publiés avec SCENARIchain. Lors de la publication des supports au format HTML et PDF dans la chaîne éditoriale Scenari, un descriptif LOM⁸⁷ (un fichier XML) est associé à chaque support produit. Ce descriptif LOM peut être récupéré et exploité par le portail documentaire lors du dépôt du support. En effet, les responsables de l'indexation des documents de la BUTC voient certains des champs d'indexation automatiquement préremplis, facilitant leur travail.

L'implémentation a suivi les étapes suivantes :

1. Rapprochement des métadonnées extraites du support publié par SCENARIchain et de celles exigées par le portail documentaire. La liste ci-dessous présente la valeur des métadonnées dans le portail documentaire lorsque le document issu de la chaîne éditoriale Scenari est déposé :

ÉTAT_DOC (à publier, à soutenir, publié...) = publié

TYPE_DOC (thèse, cours, rapport...) = cours

TITRE = titre du document publié

AUTRES_AUTEURS = auteurs des fichiers XML

ANNÉE_PUBLICATION = extraite de la date de publication

⁸⁶ <http://scenari.utc.fr/scenarisup/>

⁸⁷ LOM : Learning Object Metadata.

DESCRIPTION_1 = résumé du scénario⁸⁸ du document + plan du contenu

LANGUE = langue du document XML

2. Développement informatique sur le portail documentaire pour rechercher les informations dans le descriptif LOM des supports afin de les rendre disponibles pour le responsable de l'indexation.

Rôle des normes

La priorité pour la chaîne éditoriale SCENARIchain comme pour le portail documentaire est de proposer à ses utilisateurs des formats locaux et adaptés aux usages. Les normes facilitent une ouverture et une interopérabilité entre plusieurs systèmes. La chaîne éditoriale publie des supports LOM exploitables par n'importe quel système comprenant la norme et le portail documentaire comprend les documents respectant le LOM. Cela garantit une ouverture plus importante pour les deux systèmes.

Un standard local d'échange entre SCENARIchain et le portail aurait néanmoins pu être élaboré si le LOM n'avait pas répondu aux besoins d'interopérabilité. Cette option reste ouverte pour les évolutions futures et si les besoins évoluent.

Suite du partenariat

Le lien ainsi créé s'arrête au niveau des métadonnées. Cependant, il est possible d'aller plus loin en élargissant le partenariat à la récupération du document XML (et non plus uniquement des différents supports publiés) pour permettre un archivage optimisé sous format XML, une amélioration de la recherche dans le portail (lien entre le document logique et les supports physiques) et éventuellement un choix élargi de publication pour l'utilisateur (publication d'une partie version papier ou écran par exemple).

L'enjeu est d'améliorer la gestion documentaire des contenus de cours de l'UTC, de la création des documents à leur mise à disposition en passant par la phase de médiatisation.

⁸⁸ Scénario : Le scénario correspond à la structure et à l'ordonnancement dans lesquels se présentent les différentes parties d'un document de cours.

Synthèse

L'indexation des ressources pédagogiques numériques : un partenariat à créer entre les SCD et les services TICE au sein des universités

**Synthèse de la journée d'étude du 16 novembre 2004 à Lyon
rédigée par Monique JOLY à partir des notes prises en séance
Monique JOLY**

Note : tout au long de ce texte, nous utiliserons l'abréviation RPN pour désigner la notion de Ressource(s) pédagogique(s) numérique(s)

Cette présentation se concentre sur le concept d'objet d'apprentissage (Learning Object) qui focalise aujourd'hui un grand nombre d'attentes, que ce soit au niveau pédagogique, économique, politique ou culturel. Un des problèmes majeurs réside dans le manque de cohérence des différentes définitions, provenant principalement de la variété des points de vue adoptés pour aborder la question.

À travers différentes propositions (LOM, SCORM et IMS-LD), nous analysons trois des principales approches, apparues successivement : les langages d'indexation de données, les modèles de mise en œuvre informatique et enfin les langages de modélisation pédagogique. Ces derniers, qui font à l'heure actuelle l'objet d'études approfondies, présentent de réelles avancées en replaçant le point de vue pédagogique au centre du dispositif et en proposant de décrire l'ensemble des relations sémantiques liant les activités et les ressources de manipulation de connaissance. En particulier, cette nouvelle approche nécessite de s'intéresser aux usages effectifs des ressources d'apprentissage au sein de communautés de pratique (enseignants, formateurs et documentalistes), tant en termes de mutualisation, de réutilisation que de stratégies d'indexation.

Situer les enjeux de la journée dans les réseaux institutionnels et inter-personnels

Les discours d'ouverture de Messieurs **François DUPUIGRENET-DESROUSSILLES** (directeur de l'ENSSIB) et **Philippe GILLET** (président du PUL) situent la problématique de la journée :

- au sein des divers réseaux nationaux et locaux qu'ils animent à l'ENSSIB et au PUL,
- dans les expériences conduites par leurs établissements, en terme de démarches de mutualisation de la production des RPN,
- dans leurs préoccupations personnelles d'enseignants, en demande de RPN mutualisés, et, enfin, en signalant la convergence de la formation des professionnels de bibliothèque vers ce point.

Comprendre les enjeux de l'indexation des ressources pédagogiques numériques

Yolaine BOURDA, professeur d'informatique à SUPELEC et animatrice du groupe 4 de la CN36 (groupe de travail sur les métadonnées au sein de la commission de normalisation des technologies de l'information pour l'éducation, la formation et l'apprentissage à l'AFNOR), rappelle que nous sommes encore à une époque de pionniers, époque qui devrait rapidement déboucher sur des processus normalisés et simplifiés pour les technologies de l'éducation.

La production des RPN s'inscrit dans cette logique : coûteuses à produire et complexes à gérer dans leur cycle de vie, elles doivent être partageables et partagées (dans des environnements distribués), et pour cela elles doivent se conformer à certaines normes de description et d'indexation afin que :

- des « humains » apprenants et enseignants puissent les « re-connaître » intellectuellement,
- des ordinateurs puissent les traiter (en filtrage, en affichage, en échange) pour en augmenter l'accessibilité, la réutilisation et l'interopérabilité.

La description normalisée des RPN prend la forme de données à propos des données, autrement nommées métadonnées. Des formats déjà très élaborés existent comme LOM (Learning Object Metadata) et comme DC du DCMI (Dublin Core Metadata Initiative). Ces formats sont implémentés dans des applications concrètes sous la forme de profils d'application qui, dans certains cas, aboutissent à des interprétations divergentes. Il convient donc de continuer le patient

travail de normalisation pour fixer le plus possible un langage commun qui rendra le partage effectif et durable.

Laurent FLORY, ingénieur de recherche à l'université Lyon 1, dans le service PRACTICE en charge de la production et de la diffusion des ressources pédagogiques, décrit la RPN comme un objet pédagogique de dimension variable : le grain. Le grain peut être un très petit élément (métaphore du LEGO à une barre, comme par exemple une image ou un objet média simple) ou bien un élément beaucoup plus complexe (un assemblage de grains élémentaires, ou un gros grain indissociable en grains plus petits). L'étape d'assemblage transforme l'élément d'information (grain isolé) en une formation construite (grains assemblés). L'enjeu essentiel de la production des RPN se situe dans la personnalisation massive : il s'agit de faciliter le « réassemblage » de grains élémentaires pour créer de nouveaux cours (à l'image d'une construction LEGO).

Pour cela il faut être capable :

- de produire des RPN sous la forme de grains réutilisables,
- de mettre en place des systèmes permettant d'identifier et de retrouver ces grains,
- de construire un processus d'indexation qui prenne en compte une production nombreuse et foisonnante,
- d'indexer les RPN en temps réel pour rendre fluide le processus de production.

Vaste challenge car de nombreuses questions se posent :

- qui va indexer : des spécialistes de l'indexation ou des automates d'indexation ?
- selon quelle organisation : humaine, automatique, à la volée, avec quel contrôle qualité ? avec quel degré de réactivité ?
- avec quel langage partageable : un thesaurus local « maison », un vocabulaire plus universel mais nécessitant un important travail d'appropriation ?

Si les services TICE sont en attente des solutions que peuvent leur apporter les bibliothécaires dans le domaine de l'indexation des ressources pédagogiques numériques, le pragmatisme doit être de rigueur afin de répondre le mieux possible aux besoins des usagers, tout en n'obérant ni la réactivité ni la qualité.

Questions de l'assistance

- Qui cherche et qui trouve des RPN ? : les concepteurs de cours ? les apprenants de l'université ? d'autres apprenants ?
- Comme il y a plusieurs utilisateurs et plusieurs utilisations, n'y a-t-il pas plusieurs indexations ? À l'évidence, l'indexation doit être construite en fonction des destinations finales et probablement par des indexeurs distincts.

Concevoir des systèmes de description de ressources dans un processus de normalisation en cours d'avancement

État de l'art en matière de normes et approche critique

Elizabeth CHERHAL, de la Cellule MathDoc de Grenoble, apporte des éléments de contexte historique sur l'avènement des formats MARC et Dublin Core.

MARC, utilisé depuis 1965 pour indexer les documents des bibliothèques, existe aujourd'hui sous deux formats, après avoir connu une époque où chaque pays utilisait son propre format. Ce format ouvert permet des échanges industriels de données pour constituer des catalogues collectifs ; les migrations logicielles sont facilitées par la normalisation des données. Le contrepoint en est la lourdeur du format qui exige des compétences pointues pour la création des données (pas loin de 1000 éléments différents). Enfin ce format est inadapté pour décrire de grandes collections numériques.

DC (Dublin Core), qui a vu le jour en 1995 sous l'impulsion de bibliothécaires et d'informaticiens, a été conçu pour décrire simplement des ressources en ligne. Le revers de la simplicité est qu'il ne permet pas de tout décrire avec précision et qu'il nécessite donc des extensions qui le complexifient. Son succès est maintenant garanti : norme ISO, il est aussi reconnu par le protocole OAI-PMH de moissonnage (partage) de métadonnées.

MARC et DC cohabitent et savent se parler, c'est à dire échanger des données.

Jean-Philippe PERNIN, de l'INRP, se concentre sur l'évolution des représentations du rôle de l'enseignant :

- enseignant artisan qui est l'auteur du cours,
- enseignant éditeur qui conçoit et prescrit des ressources,
- enseignant documentaliste qui prospecte et référence
- et enfin, enseignant pédagogue centré sur l'activité pédagogique.

Ces divers rôles ont donné lieu à divers modèles :

- LOM axe tout sur la ressource et s'appuie sur un modèle de granularité des RPN ;
- SCORM lie les ressources aux activités, organise la traçabilité du parcours de l'apprenant et permet l'interopérabilité des RPN entre plates-formes ;
- IMS-Learning Design met l'accent sur les scénarios pédagogiques et non pas sur les RPN ; en conséquence il préconise la création de bibliothèques de ressources (numériques ou non, pédagogiques ou non) et la création de bibliothèques de scénarios pédagogiques. Cette nouvelle approche nécessite de s'intéresser aux

usages effectifs des ressources d'apprentissage, au sein des communautés, tant en termes de mutualisation, de réutilisation que de stratégies d'indexation.

Les normes et standards de description des ressources

Elizabeth CHERHAL, de la cellule MathDoc détaille le format Dublin Core, constitué de 15 éléments de base relatifs au contenu, à la propriété intellectuelle et à des instances techniques (date, type, identifiant...). Ces éléments peuvent être raffinés avec des « qualifieurs » qui sont soit des raffinements d'éléments, soit des schémas d'encodage (vocabulaires contrôlés). Il existe ainsi un raffinement « Educational Level ». Pour la compatibilité OAI-PMH seuls les 15 éléments de base sont concernés.

Rosa Maria GOMEZ DE REGIL, de Doc'INSA, présente le LOM (Learning Object Metadata), standard IEEE, détaille les 9 catégories d'éléments (General, LifeCycle, MetaMetadata, Technical, Educationnal, Rights, Relation, Annotation, Classification) et montre leurs parentés et leurs spécificités avec les formats MARC et DC.

De nombreux projets universitaires en Rhône-Alpes, en France et dans le monde utilisent la structuration LOM des métadonnées, mais souvent dans des profils d'application qui diffèrent. Dans le cadre du groupe 4 de la CN36, l'AFNOR élabore un profil d'application français en se basant sur le profil ManUel. La mise en œuvre de l'indexation est abordée avec le workflow d'indexation de l'INSA de Lyon, faisant intervenir à la fois des enseignants/auteurs/concepteurs, des développeurs et des documentalistes dans la chaîne de l'indexation. De nouveaux métiers se dessinent pour les bibliothèques, nécessitant beaucoup de relationnel pour coordonner l'ensemble du dispositif qui permet de constituer une bibliothèque de ressources numériques et d'assumer la gestion des droits associés aux ressources.

Gilles BERTIN, projet Ancoly du PUL, présente SCORM, standard IEEE, qui se propose de rendre les RPN interopérables sur diverses plates-formes pédagogiques (autrement nommées LMS : Learning Management System), tout en offrant un environnement d'exécution centré sur les ressources et sur les activités pédagogiques. Pour cela SCORM définit :

- un modèle d'agrégation de contenus qui permet de créer, de décrire et d'échanger des ressources indépendamment du système qui les a créées ; l'agrégation se définissant en trois niveaux : ressources, grain de formation, cours,
- un environnement d'exécution qui permet une communication avec les plates formes pour l'exécution de l'enseignement.

Il existe des outils logiciels permettant de « scormer » rapidement des RPN. SCORM complète LOM – qui continue d’être utilisé pour décrire le niveau d’agrégation de la ressource – et est bien supporté par les plates-formes, avec une bonne interopérabilité.

Les évolutions du format LOM

Yolaine BOURDA, SUPELEC, animatrice du groupe 4 CN36, présente les travaux d’évolution des normes de métadonnées pour les RPN. LOM, standard IEEE a été proposé en norme ISO en procédure accélérée avec appel à commentaires en 2002, mais l’IEEE ayant refusé de discuter une évolution du LOM en fonction des commentaires (dont ceux de la France), la norme ISO a été refusée. LOM est rarement utilisé tel que le définit le standard IEEE de 2002 ; de nombreux profils d’application viennent raffiner les définitions, ajouter des éléments, spécifier des valeurs, rendre obligatoires des éléments facultatifs – ce qui rend extrêmement difficile la préservation de l’interopérabilité avec le schéma de base. Plusieurs exemples d’interprétations divergentes du sens des éléments et de listes de valeurs non compatibles, sont montrés.

Il y a donc un réel besoin de revenir à une réflexion fondamentale sur la ressource et sa description : ceci fait l’objet du projet MLR : Metadata for Learning Resources de l’ISO. La France est co-éditrice du MLR avec la Chine. Deux niveaux de réflexion prévalent : conceptuel/abstrait et réalisation/implémentation. La norme de conception ISO/IEC 11179 est utilisée comme méthode de conception, afin d’éviter des confusions de sens dans l’élaboration du MLR, confusions qui n’ont pas été évitées dans LOM. Ce travail est de longue haleine, sur plusieurs mois.

En attendant le groupe 4 CN36 travaille sur un profil d’application français du LOM, dans le but d’élaborer une norme AFNOR expérimentale. Le profil d’application ManUel de CampuSciences sert de base à la réflexion. Un premier appel à commentaires est dépouillé en novembre et décembre 2004, pour une enquête élargie en janvier 2005 et une publication expérimentale en mars 2005.

Tous les participants à cette journée sont conviés à se manifester lors de cette consultation : c’est en recueillant les avis de tous que l’on tendra vers un profil d’application adapté au plus grand nombre.

L’indexation des RPN : questions transversales

Ghislaine Chartron, INRP, URSIDOC/enssib, s’interroge :

- Sur les niveaux actuels d’utilisation des standards pour les RPN, et confirme qu’ils sont utilisés :
 1. Identification des documents avec DOI,
 2. Description des documents avec DC norme ISO,

3. Échange de données avec des protocoles OAI-PMH.

- En ce qui concerne les RPN, l'indexation sert à décrire un contexte pédagogique d'usage (niveau, méthode, évaluation, ...) et un contexte technique pour la réutilisation dans la conception (plate-forme, format, liens avec d'autres objets...).
- Sur la création des métadonnées des RPN : elles peuvent se présenter selon 4 niveaux : métadonnées de repérage d'un contenu (notice bibliographique renseignées par un bibliothécaire), métadonnées de conception pédagogique (renseignées par un enseignant), métadonnées sur la conception technique (renseignées par les concepteurs/développeurs) et enfin les métadonnées permettant de gérer les droits (renseignées par les bibliothèques).
- Sur la réalité de la mutualisation des ressources pédagogiques : certaines études montrent que les RPN sont encore peu partagées.
- Sur la réalité de l'utilisation du LOM : d'après l'enquête OCLC, sur 35 projets, les métadonnées les plus renseignées sont celles de DC et les moins renseignées sont celles du bloc pédagogique.

L'indexation des RPN doit donc se concevoir comme une activité à pratiquer de façon raisonnable en tenant compte du facteur temps et du facteur coût, et en fonction des contextes d'usages.

Questions de l'assistance

- Quelle organisation mettre en place pour une indexation en « strates » avec plusieurs acteurs ? quelle dimension humaine ?
 - oui, il est important d'associer plusieurs acteurs à cette création,
 - oui il est important de passer de pratiques individuelles à des pratiques coopérantes,
- La convergence des divers formats présentés est frappante ; vont-ils tous converger ?
 - ces formats ne sont pas des sur-ensembles : ils doivent co-exister,
 - il semble que les spécificités du LOM soient peu utilisées,
 - LOM et DC Educational Level ont signé pour une convergence commune.

Observer les partenariats entre SCD et TICE : rôle et positionnement des acteurs

Cas de l'UTC : Annie BERTRAND, SCD, et Xavier HENNEQUIN, SCENARI

L'UTC vient d'ouvrir son portail documentaire, conçu comme un point d'entrée unique sur les ressources documentaires et pédagogiques, réalisation portée par la stratégie de l'UTC en matière de diffusion des ressources TICE. Le portail est enrichi avec des technologies de workflow permettant de matérialiser cette relation entre l'enseignant et la bibliothèque au travers des fonctions suivantes : déposer des documents, proposer des sujets et des mots clés, gérer les autorisations de consultation pour des groupes et des durées ; le SCD ayant la charge de finaliser la description et de la publier.

SCENARI est une chaîne éditoriale qui produit et gère des contenus ainsi que des métadonnées propres (pédagogiques, documentaires, d'indexation sujet, de formation). Les contenus et leurs métadonnées sont exportés pour être publiés sur le portail documentaire.

Il est prévu un renforcement du partenariat.

Cas de l'université de Lille 1 , avec Jean-Bernard MARINO, SCD

[NDLR : C'est avec une immense douleur que nous avons appris le décès de Jean-Bernard Marino quelques jours seulement après son intervention au sein de cette journée d'études.]

La bibliothèque numérique GRISEMINE propose 1100 documents issus de la production pédagogique et scientifique de l'université. Le retour d'expérience, de plusieurs années, est très encourageant sur la visibilité des ressources numériques. La description est actuellement basée sur la DTD Marc. Une évolution technique de l'application est souhaitée pour mieux répondre aux utilisateurs : recherche type Google, métadonnées DC, compatibilité OAI-PMH, diversification des contenus (évolution vers une politique d'archives ouvertes pour l'université), signalement de l'unité documentaire la plus fine, validation des contenus pas les responsables de communautés. Le SCD a choisi d'implanter DSpace du MIT qui sera décliné en deux parties : Ariane pour la recherche et Iris pour le SCD. Ce sont 48 éléments de métadonnées DC qui sont retenus. La Dewey a été implémentée.

Questions de l'assistance

- Les bibliothécaires deviennent-ils des documentalistes ?
 - oui, en utilisant des outils spécialisés comme Rameau, et d'autres Thésaurus,
 - en quoi l'utilisation de Rameau transforme-t-il des bibliothécaires en documentalistes ?
- Les documents sont-ils validés ? et si oui, se trouve-t-on dans une réappropriation du rôle éditorial de l'université ?
 - oui, c'est bien le cas, l'université se positionne comme un éditeur.

Identifier les enjeux du partenariat SCD-TICE – Point de vue des tutelles et des réseaux nationaux

Table-ronde animée par Jean-Michel SALAUN, ENSSIB avec :

- Sabine BARRAL, ABES,
- Gilles BRAUN, Direction de la Technologie, sous-direction des TICE,
- Michel ROLAND, Couperin,
- Jean-Émile TOSELLO-BANCAL, Direction de l'Enseignement Supérieur, Sous Direction des Bibliothèques.

Le débat s'articule autour de 3 thèmes : Qui doit indexer ? Quelle politique centralisée/distribuée ? Qui prend la responsabilité de cette édition ?

Thème 1 : Qui doit indexer ? avec quels moyens ? avec quelle motivation ?

Michel ROLAND, souhaite qu'un tel point soit discuté au sein des comités de pilotage des systèmes d'information des établissements, car il faut éviter que ce travail d'indexation des RPN ne devienne une tâche supplémentaire sans moyens supplémentaires ; il convient donc qu'une prise de conscience globale habite une telle décision.

Jean-Émile TOSELLO-BANCAL rappelle que les bibliothécaires sont déjà naturellement impliqués dans la pédagogie (glossaires, bibliographies...) et qu'ils doivent être indexeurs, mais en association avec les enseignants. Tout ceci doit être réalisé dans un schéma de système d'information d'établissement comportant des représentants des CRI, TICE et SCD.

Sabine BARRAL, nous rappelle, qu'il y a 8 ans, il n'y avait pas de bibliothécaires dans les colloques sur les TICE et que cela a changé. Pour le reste chacun fait son métier et si on veut mutualiser, la normalisation est très importante.

Gilles BRAUN note que l'indexation de la ressource est partie intégrante de la ressource et qu'elle en est même à l'origine ; il faut donc intégrer des bibliothécaires dès l'élaboration de la ressource. Des exemples très concrets comme ceux de la numérisation des ressources vidéos de l'INA sont là pour en témoigner.

Jean-Michel SALAUN prend acte de ce consensus commun et convient que les pouvoirs publics doivent prendre la mesure de l'extension des métiers des SCD à l'indexation des RPN, avec l'affectation de moyens qu'il convient de prévoir.

Thème 2 : Quelle politique centralisée/distribuée privilégier pour donner un moteur à toutes ces actions ?

Sabine BARRAL indique que la production des RPN et leur description se fait le plus souvent au niveau local, ce qui n'exclut pas la logique d'un partage dans des réseaux nationaux. Les réseaux nationaux pouvant eux produire et mettre à disposition des fondamentaux utilisables par tous.

Jean-Emile TOSELLO-BANCAL ne pense pas que le paradigme des grands réseaux bibliographiques de bibliothèques puisse être transposé aux RPN, car la production des RPN est plutôt locale et qu'elle doit garder un fort potentiel d'évolutivité, au gré des mises à jour des RPN et des changements de programmes.

Gilles BRAUN plaide fortement pour que les producteurs de RPN se conforment à la normalisation, ceci prenant la forme de mesures financières incitatives pour soutenir la production de RPN ; dans ce cas, l'intervention incitative est nationale.

Michel ROLAND considère que le modèle des réseaux bibliographiques ne peut être totalement transposé, car la RPN est un objet mouvant, avec des statuts de validation très divers. Le niveau d'intervention est donc plutôt local dans une logique de partenariat fort entre les services TICE et les SCD, qui seront à même de spécifier les meilleures méthodes de travail.

Thème 3 : Qui prend la responsabilité de la publication ? Qui gère les droits ?

Sabine BARRAL considère deux cas : validation locale du contenu par l'auteur enseignant et/ou l'auteur personne morale (un établissement), validation plus centralisée (nationale, régionale...) en cas de construction d'archives pérennes.

Gilles BRAUN considère qu'il faut évaluer les ressources publiées sur le web, et l'établissement doit donc prendre ses responsabilités d'éditeur pour valider les RPN et la description LOM des RPN.

Jean-Émile TOSELLO-BANCAL considère que la validation scientifique est plurielle et à faire par les acteurs impliqués dans la production des RPN. La validation pourrait aussi être

complétée en dehors de l'établissement pas un comité scientifique, à l'instar de ce qui se passe dans les comités scientifiques de numérisation.

Questions de l'assistance

- Les comités de pilotage des systèmes d'information sont-ils très présents dans les établissements ?
 - oui, l'on commence à en rencontrer et la CPU pousse fortement les universités dans ce sens.
 - oui, dans le cadre des contrats quadriennaux, certains projets d'établissements les signalent.
- Un témoignage sur l'évaluation des ressources pédagogiques : attention à la question de l'évaluation car évaluer la ressource conduit à évaluer indirectement la qualité des enseignements et c'est un sujet très sensible.
- Quelle est la synergie entre les portails du CERIMES et de l'ABES ?
 - Le CERIMES met en place des systèmes de distribution de contenus spécialisés comme des vidéos à destination des UNT.
 - L'ABES met en place un système de fédération de portail, donc un point d'entrée sur des ressources.

Synthèse

Monique JOLY souligne la qualité des interventions et des débats montrant :

- qu'il reste encore un important travail de spécifications pour concevoir des méthodes de travail homogènes, afin de rendre parfaitement interopérables les systèmes de partage de RPN,
- mais que cela ne doit pas retarder l'engagement des actions coopératives TICE-SCD car l'on peut considérer cette étape de finalisation comme un ensemble de détails de mise au point qui seront rapidement résolus,
- parce que les acteurs concernés (TICE, SCD, DT, DES/SDB) se réunissent dans une dynamique récente mais très active, sous la pression des utilisateurs de ces futurs systèmes.

Ce sont eux qui auront le dernier mot en jugeant la qualité, la fiabilité et la pérennité des systèmes que nous aurons su mettre en place à leur intention.