

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 (Educnets³) 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.