



Caractérisation de l'information pédagogique pour le Web : principe d'application des recommandations du W3C dans un contexte de documents virtuels personnalisables

Christine Michel, Soufiane Rouissi

► To cite this version:

Christine Michel, Soufiane Rouissi. Caractérisation de l'information pédagogique pour le Web : principe d'application des recommandations du W3C dans un contexte de documents virtuels personnalisables. Actes du 5ème Colloque International sur le Document Électronique (CIDE'05)., Oct 2002, 20-23 octobre 2002, Hammamet, Tunisie. sic_00000327

HAL Id: sic_00000327

https://archivesic.ccsd.cnrs.fr/sic_00000327

Submitted on 22 Jan 2003

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Etude de l'organisation et caractérisation de l'information pédagogique pour construire des documents hypermédias adaptatifs diffusés sur le Web

Christine Michel, Soufiane Rouissi

CEM-GRESIC, Université Bordeaux 3

MSHA, 10, Esplanade des Antilles

33607 PESSAC Cedex

Christine.Michel@montaigne.u-bordeaux.fr

Soufiane.Rouissi@msha.u-bordeaux.fr

Résumé :

Dans cet article nous nous intéressons aux modèles descriptifs et aux méthodes de caractérisation des ressources pédagogiques de manière à les utiliser dans un contexte de *Document Virtuel Personnalisable* (DVP). Pour ce faire, nous analysons en premier lieu quels sont les objets pédagogiques qui ont été utilisés dans les systèmes tutoriels intelligents et dans les systèmes adaptatifs et observons les propriétés qui les caractérisent. Dans un second temps, nous observons les travaux actuellement en vigueur pour les caractériser au niveau international dans un contexte Internet ou dans les systèmes ouverts (plates-formes d'enseignement à distance par exemple) de manière à voir s'il y a une adéquation entre les deux.

MOTS-CLES : Caractérisation, Métadonnées, Systeme tutoriel intelligent - STI, E-Learning, Hypermédias adaptatifs, Pédagogie, Document Virtuel Personnalisable - DVP.

Introduction

L'E-learning est défini comme *"l'utilisation des nouvelles technologies multimédias et de l'Internet, pour améliorer la qualité de l'apprentissage en facilitant l'accès à*

"Caractérisation de l'information pédagogique pour le Web : principe d'application des recommandations du W3C dans un contexte de documents virtuels personnalisables" - Christine MICHEL et Soufiane ROUISSI- Article publié dans les actes du 5ème Colloque International sur le Document Electronique (CIDE'05). 20-23 octobre 2002, Hammamet, Tunisie. – Sous la direction de Mauro GAIIO, Christine VANOIRBEEK, Khaldoun ZREIK – Editeur Inria – 2002 - 308p - pp153-167.

des ressources et des services, ainsi que des échanges et la collaboration à distance" [ELE01]. Face à la croissance exponentielle des documents pédagogiques (cours, exercices, études de cas, résumés...) sur le Web, il convient de recourir à des procédés de description des données utilisés par le plus grand nombre. Ceci doit permettre de favoriser le caractère d'inter-opérabilité indispensable face à la prolifération de solutions logicielles : le développement des outils logiciels ne doit pas influencer sur le contenu créé aujourd'hui. Les contenus doivent être gérés indépendamment des modes d'exploitation qui seront mis en jeu pour les diffuser. De plus il convient d'utiliser des techniques qui caractérisent les ressources de la manière la plus précise possible, ceci de manière à pouvoir personnaliser l'offre de formation en fonction des utilisateurs.

Notre propos est de vérifier s'il est possible, en considérant les supports pédagogiques comme des documents virtuels personnalisables, de trouver des caractéristiques à la fois fines pour pouvoir personnaliser l'offre de formation aux différents apprenants mais aussi universelles pour être utilisées par le plus grand nombre, dans des systèmes ouverts comme des sites Web ou propriétaires comme les plates-formes. Il n'existe pas à l'heure actuelle de norme dans le contexte pédagogique, cependant des recommandations plus que fortes et largement suivies par les constructeurs apparaissent. Après avoir analysé les caractéristiques propres aux ressources pédagogiques, nous étudierons un modèle de description des données en passe d'être normalisé et un modèle d'organisation des ressources.

1. Pourquoi choisir l'hypothèse du DVP ?

"Un DVP (Document Virtuel Personnalisable) peut être considéré comme un ensemble d'éléments (que l'on peut appeler fragments) associé à des mécanismes de filtrage, d'organisation et d'assemblage sous contraintes, i.e. en respectant un modèle de l'utilisateur et des principes narratifs. La façon d'obtenir les éléments (segmentation de documents existants) et leurs caractéristiques (taille, média, etc.) sont à définir." (Texte introductif à DVP'02 <http://iasc.enst-bretagne.fr/DVP02/themes.htm>) Nous avons été sensibilisés à l'utilité du document virtuel personnalisable dans le cadre du projet Profil-Doc [LAI96] [MIC99]. En effet, l'idée de ce projet est d'utiliser les connaissances sur les usages de lecture de l'IST par les chercheurs [BEN97a et b] pour construire des liens entre certaines parties de documents caractérisés par des propriétés spécifiques et des profils d'utilisateurs. Dans ce contexte il a été montré que la qualité de description des ressources à des niveaux de granularité inférieurs (i.e. au niveau des unités documentaires), améliorerait la réponse du système [MIC00] si elle se faisait sous la forme d'un DVP.

Le découpage et la caractérisation des ressources pédagogiques présentent plusieurs intérêts. Tout d'abord, la modularité et la réutilisation des fragments ou briques d'information se trouvent facilitées à la manière d'un jeu de construction. Les

"Caractérisation de l'information pédagogique pour le Web : principe d'application des recommandations du W3C dans un contexte de documents virtuels personnalisables" - Christine MICHEL et Soufiane ROUISSI- Article publié dans les actes du 5ème Colloque International sur le Document Electronique (CIDE'05). 20-23 octobre 2002, Hammamet, Tunisie. - Sous la direction de Mauro GAIU, Christine VANOIRBEEK, Khaldoun ZREIK - Editeur Inria - 2002 - 308p - pp153-167.

"grains" ainsi créés se décomposent en deux catégories. Une ressource, dont la granularité est dite de type faible (*low granularity*), se décrit comme une ressource de taille importante et difficile à réutiliser dans d'autres contextes différents (*large resource wich is difficult to re-use in a different context*). Elle s'oppose à une ressource de granularité élevée qui est facile à réutiliser mais aussi à modifier pour être utilisée dans un autre contexte (*a resource wich is easy to re-use or edit for use in another context*). [TES01]

Le cours devenant assimilable à une collection d'objets permet une fabrication facilitée en puisant dans le gisement d'objets pédagogiques, une mise à jour plus simple, une diffusion personnalisée. Ainsi le document virtuel pédagogique est défini *"comme un ensemble de fichiers et de ressources réparties sur trois niveaux permettant à partir d'un contenu unique d'établir des activités variées présentées à des étudiants. Les niveaux différents autorisent l'introduction de méta-informations qui permettent des utilisations multiples du même contenu."* [MOU99]

"Les documents au format HTML peuvent être utilisés en tant qu'objet d'apprentissage lorsqu'ils sont placés sous le contrôle d'un système de tutorat intelligent reposant sur un navigateur Web en tant qu'interface. Ces documents peuvent être générés dynamiquement comme dans le cas par exemple d'une évaluation (test) en ligne permettant ainsi un gain de place ("gain of memory place") et une plus grande variété de tests". [NKA96b]

Le système KARINA [CHA00], que nous présenterons plus en détail dans la suite, propose sur la même idée de construire des supports de cours personnalisés en utilisant des ontologies de connaissances du domaine.

Dans le même esprit mais avec des techniques de filtrage différentes le *"Livre de Connaissances Électronique" (LCE)*, est un document hypermédia structuré comprenant un ensemble de diagrammes de connaissances et certains documents particuliers, depuis lequel il est possible de construire une requête d'interrogation sur des bases documentaires.../... Une architecture générique de Livre de Connaissances Électronique a été mise au point grâce à l'approche des Documents Virtuels Personnalisables. Elle utilise des composants modulaires pour définir dynamiquement l'interface en fonction de la nature et des besoins des différents utilisateurs. .../... Deux types d'éléments entrent dans la composition de l'architecture : d'une part, les éléments d'interface, qui sont indépendants du projet est forment la partie visible du système, et d'autre part, les éléments de modélisation qui sont spécifiques aux projets et constituent le cœur du système". [MED01]

La modélisation de l'utilisateur est relativement stable depuis quelques années.

En effet, l'utilisateur est généralement défini par [BRU01]:

- les objectifs/la tâche à réaliser
- l'état de son savoir
- ses centres d'intérêt (décomposés depuis l'avènement des systèmes de recherche d'information du Web en centres d'intérêt à court terme et à long terme)
- son passé
- ses préférences

"Caractérisation de l'information pédagogique pour le Web : principe d'application des recommandations du W3C dans un contexte de documents virtuels personnalisables" - Christine MICHEL et Soufiane ROUSSI- Article publié dans les actes du 5ème Colloque International sur le Document Électronique (CIDE'05). 20-23 octobre 2002, Hammamet, Tunisie. – Sous la direction de Mauro GAIQ, Christine VANOIRBEEK, Khaldoun ZREIK – Editeur Inria – 2002 - 308p - pp153-167.

- ses traits de caractère (sa personnalité introvertie ou extravertie, certains facteurs cognitifs, styles d'apprentissage). Ces caractéristiques sont généralement mises en évidence par des tests psychologiques.

Ces caractéristiques, à l'exception des traits de caractère, sont très largement utilisées et donnent des résultats probants dans la personnalisation des systèmes. Ceci est amplifié s'ils utilisent une technique de fragmentation du document.

Depuis 1996, deux caractéristiques supplémentaires apparaissent :

- **les données d'usage** qui sont significatives des interactions de l'utilisateur avec le système, mises en évidence par des études de profiling et de tracking. Elles sont généralement utilisées dans les systèmes de recommandation.
- **les données environnementales** où l'adaptation de l'interface se fait en fonction de la localisation et de la plate-forme de l'utilisateur c'est à dire de son type de matériel (ordinateur fixe, portable, palm Pilot) et du type de média qu'il peut voir.

Les autres éléments de modélisation concernent les ressources. Nous allons voir qu'en fonction des systèmes les objets pédagogiques sont variables.

1.1. Systèmes tutoriels simples

Les systèmes pédagogiques non intelligents n'utilisent que des ressources banales (au sens de Nkambou [NKA97]) et donc ne peuvent traiter l'interaction tutorielle avec l'apprenant. On peut considérer que ce sont des systèmes de sélection, comme les systèmes de recherche d'information peuvent l'être. L'apport qu'ils proposent est centré sur la personnalisation de la présentation en fonction de critères plus spécifiques à l'utilisateur : on parle d'hypermédias adaptatifs. Selon Brusilovsky ils sont de deux sortes [BRU01] :

- les systèmes qui adaptent le contenu présenté à l'utilisateur (*adaptive presentation*)
- les systèmes qui adaptent les liens présentés à l'utilisateur (*adaptive navigation support*).

Parmi les systèmes qui adaptent le contenu présenté se distinguent trois types de systèmes :

- ceux qui adaptent le texte :
 - Par une adaptation du langage
 - Par une adaptation des fragments : principalement par l'insertion, la suppression, le filtrage, l'ordonnancement, l'estompage de fragments
- ceux qui choisissent le média le plus approprié
- ceux qui adaptent le mode de présentation fragmentaire multimédia selon des critères technologiques.

Tous fonctionnent selon le même principe (énoncé par Buckland en 1994 [BUC94] et représenté dans la figure 1) : la succession de deux types d'opération (représenté

"Caractérisation de l'information pédagogique pour le Web : principe d'application des recommandations du W3C dans un contexte de documents virtuels personnalisables" - Christine MICHEL et Soufiane ROUISSI- Article publié dans les actes du 5ème Colloque International sur le Document Electronique (CIDE'05). 20-23 octobre 2002, Hammamet, Tunisie. - Sous la direction de Mauro GAIU, Christine VANOIRBEEK, Khaldoun ZREIK - Editeur Inria - 2002 - 308p - pp153-167.

par les boîtes en pointillé dans le graphique ci-dessous) ; des transformations de données et des traitements (*partitionning*) (tri, ordonnancement, sélection) permettant de créer des sous-ensembles de données (représentées par les boîtes en plein). Les opérations pourront être plusieurs fois répétées et alternées selon la complexité du traitement final.

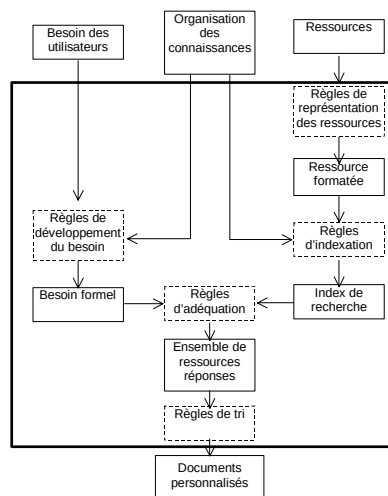


Figure 1

Le degré de personnalisation du système va dépendre donc de la qualité des traitements opérés en interne mais aussi de la qualité des descriptions des ressources, des utilisateurs et des connaissances.

1.2. Systèmes tutoriels intelligents

Les systèmes tutoriels intelligents (STI) doivent adapter leur offre de formation de manière dynamique, en fonction de règles pédagogiques mais aussi selon les réactions de l'apprenant qui l'utilise. Un très bon exemple est le STI CREAM de Nkambou [NKA96] qui permet de construire des cours qui sont adaptés à un public cible (défini par son niveau de connaissances). Le système est certes plus complexe que le système de sélection de Buckland mais cela est dû non seulement au fait que les combinaisons de transformation et de traitement sur les données sont plus nombreuses, mais aussi au fait que les ressources pédagogiques prises en compte sont plus variées et caractérisées de manière beaucoup plus précise.

2. Les objets pédagogiques et leurs caractéristiques descriptives

Les objets pédagogiques sont des documents comme les autres, la manière la plus usuelle de les caractériser n'est donc pas liée à leur aspect pédagogique. Ainsi, ils sont généralement décrits grâce à des standards comme XML (Extensible Markup Language) *"qui permettent de représenter la dimension logique des documents indépendamment de leur contenu et de leur dimension spatiale (forme), cette dernière faisant elle-même l'objet de standards comme DSSSL, CSS ou XSL. Cette décomposition de l'information portée par les documents a pour premier objectif de faciliter la portabilité des documents ainsi que leur traitement par des applications variées. Sur la base de cette approche, d'autres caractéristiques sont regroupées pour former de nouvelles dimensions : la dimension hypertexte, correspondant à l'ensemble des informations permettant de lier des documents (ou des fragments de documents) entre eux, et la dimension temporelle qui identifie le comportement dans le temps des documents (comme la durée des objets ou leur synchronisation). .../..."*

La diversité des besoins en matière de représentation de documents a conduit à une approche modulaire de ces spécifications et à la notion de profils (sous-ensemble de modules de spécifications pour un domaine applicatif donné). C'est une approche en cours de discussion dans le groupe de travail "XML Syntax" du W3C." [ROI99]

Nous allons voir que les objets pédagogiques sont de nature assez complexe et qu'ils nécessitent une règle de description propre.

2.1. Le curriculum

L'ensemble des ressources, au sens large, nécessaires à une activité pédagogique est regroupé dans un curriculum. Nkambou [NKA97] cite Legendre qui définit le curriculum comme *"l'ensemble structuré des expériences d'enseignement et d'apprentissage (objectifs de contenu, objectifs d'habileté, objectifs spécifiques, cheminements ramifiés et règles de progression, matériel didactique, activités d'enseignement et d'apprentissage, etc.) planifiées et offertes sous la direction d'une institution scolaire en vue d'atteindre des buts éducatifs prédéterminés"*.

Nkambou propose donc une définition formalisée du curriculum comme *"une représentation structurée de la matière à enseigner en terme de **capacité**¹ (capacités) au sens de Gagné, d'**objectifs** dont la réalisation contribue à l'acquisition de capacités et de **ressources didactiques** (exercices, problèmes, démonstrations, vidéos, simulations, etc.). Tous ces éléments sont organisés dans des structures de connaissances destinées à soutenir l'enseignement d'une matière."* [NKA97]. Il définit :

¹ Le terme *capacité* fait référence aux connaissances du domaine et habiletés que possède un individu.

C'est donc un terme qui couvre un sens plus large que celui de *connaissance du domaine*.

"Caractérisation de l'information pédagogique pour le Web : principe d'application des recommandations du W3C dans un contexte de documents virtuels personnalisables" - Christine MICHEL et Soufiane ROUISSI- Article publié dans *les actes du 5ème Colloque International sur le Document Electronique (CIDE'05)*. 20-23 octobre 2002, Hammamet, Tunisie. - Sous la direction de Mauro GAIIO, Christine VANOIRBEEK, Khaldoun ZREIK - Editeur Inria - 2002 - 308p - pp153-167.

- Une **capacité** comme une connaissance, acquise ou développée, permettant à une personne de réussir dans l'exercice d'une activité physique ou intellectuelle. Il s'agit d'une information qui est emmagasinée dans la mémoire à long terme de l'apprenant et qui rend possible un comportement.
- Un **objectif** d'enseignement comme une description d'un ensemble de comportements (ou performances) qu'un apprenant doit être capable de démontrer suite à un apprentissage. Par exemple : acquisition, compréhension, application, analyse, synthèse et évaluation.
- Une **ressource didactique** est un moyen tactique (exercice, problème, test, simulation, démonstration, document...), utilisé par le système d'enseignement pour supporter l'apprentissage, par un apprenant, de capacités le rendant capable des performances définies par les objectifs.

Nkambou [NKA96] indique qu'il y a trois grandes catégories de ressources:

- Les ressources intelligentes sont des ressources didactiques portant sur une activité précise à réaliser par un étudiant. Elles sont généralement assorties d'un modèle de connaissances leur permettant d'être implémentées à n'importe quel niveau du STI pour supporter l'interaction avec l'étudiant et de suivre la trace du raisonnement de ce dernier.
- Les ressources tutorielles sont capables de gérer l'interaction avec un étudiant durant une session tutorielle. Un certain nombre d'actions didactiques lui sont déléguées par le tuteur laissant ainsi le contrôle à la ressource tutorielle.
- Les ressources "banales" sont des ressources qui jouent plutôt un rôle passif dans le processus d'apprentissage dans la mesure où elles sont incapables de gérer l'interaction avec l'étudiant qui les utilise et ne peuvent pas comporter une évaluation sur les actions de l'étudiant contrairement aux ressources tutorielles et intelligentes.

Les ressources intelligentes et les ressources tutorielles sont formellement écrites sous forme de règles de déduction ou de programmes. Elles ne peuvent donc pas être prises en compte telles quelles dans le contexte DVP. Les capacités et objectifs sont écrits sous forme de listes de procédures qui sont en fait un ensemble de règles. Toutes ces ressources sont de type opérationnel et peuvent s'inscrire, dans le schéma de Buckland, sous la forme de "boîte en pointillé". De plus, il serait absurde de vouloir les fragmenter pour leur donner plus de granularité car elles sont indivisibles. A ces deux titres, nous considérons qu'elles ne sont pas directement utilisables, en tant que fragments, dans le contexte des DVP.

Qu'en est-il des ressources banales ? Ces dernières sont fragmentables dans le contexte du DVP : plusieurs découpages sont possibles selon le rôle ou l'usage potentiel du fragment. Dans le système KARINA[CHA00], le critère est le rôle pédagogique qui a pour but de susciter l'action pédagogique chez l'apprenant. Ainsi les fragments (appelés briques d'information) sont caractérisés selon qu'ils sont : *l'appariement, la conclusion, la description, la définition, l'exemple, l'exercice cas*

"Caractérisation de l'information pédagogique pour le Web : principe d'application des recommandations du W3C dans un contexte de documents virtuels personnalisables" - Christine MICHEL et Soufiane ROUSSI- Article publié dans *les actes du 5ème Colloque International sur le Document Electronique (CIDE'05)*. 20-23 octobre 2002, Hammamet, Tunisie. - Sous la direction de Mauro GAIQ, Christine VANOIRBEEK, Khaldoun ZREIK - Editeur Inria - 2002 - 308p - pp153-167.

général, l'exercice d'ordonnancement, l'exercice de prononciation, l'explication, la formule, l'illustration, l'introduction, les questions à choix multiple, la référence, la résolution de problème, le résumé, le test vrai / faux, le théorème. Cette taxinomie correspond à une structure facilement repérable dans les documents pédagogiques et à des rôles présupposés :

- Appariement => *Mise en correspondance de certaines entités*
- Conclusion => *Lecture*
- Description => *Lecture et assimilation*
- Définition => *Lecture et assimilation*
- Exemple => *Lecture et assimilation*
- Exercice cas général => *Résolution exacte*
- Exercice d'ordonnancement => *Ordre exact*
- Exercice de prononciation => *Prononciation acceptable*
- Explication => *Lecture et assimilation*
- Formule => *Lecture et assimilation*
- Illustration => *Lire, regarder, écouter en fonction du média*
- Introduction => *Lecture, compréhension du contexte*
- Questions à Choix Multiple => *Choix exacts*
- Référence => *Lecture de l'ouvrage référencé*
- Résolution de problème => *Résolution exacte*
- Résumé => *Lecture*
- Test vrai / faux => *Sélectionner les propositions*
- Théorème => *Lecture et assimilation*

Il est intéressant de noter que, contrairement à Profil-Doc, les fragments pédagogiques peuvent être caractérisés selon plusieurs types tout en conservant le même rôle. A titre d'exemple, la même description de la fission nucléaire peut être une *définition* dans un cours sur le fonctionnement d'une centrale, une *introduction* dans un cours sur le progrès technologique et un *exemple* dans un cours de physique atomique et dans tous les cas son rôle sera d'être *lu et assimilé*. Sylvie Chabert-Ranwez propose, pour résoudre ce problème, de faire une description sous forme de vecteur d'état conceptuel qui est une représentation de plusieurs points de vue sur une brique d'information.

L'objectif du fragment n'est pas directement indiqué comme caractéristique de description dans le cas de KARINA car il est induit et correspond au rôle mais ce n'est pas toujours le cas comme nous allons le voir dans la suite.

2.2. Le cours

Le cours est l'ensemble des ressources choisies pour présenter une matière ou un savoir. Il est défini soit par des *objectifs d'enseignement* (ou d'apprentissage) ayant une finalité précise, soit par un *ensemble de connaissances* que l'étudiant doit acquérir. Une interface adéquate permet de personnaliser la présentation du cours à l'étudiant selon son profil.

"Caractérisation de l'information pédagogique pour le Web : principe d'application des recommandations du W3C dans un contexte de documents virtuels personnalisables" - Christine MICHEL et Soufiane ROUISSI- Article publié dans *les actes du 5ème Colloque International sur le Document Electronique (CIDE'05)*. 20-23 octobre 2002, Hammamet, Tunisie. – Sous la direction de Mauro GAIQ, Christine VANOIRBEEK, Khaldoun ZREIK – Editeur Inria – 2002 - 308p - pp153-167.

C'est un document structuré [NKA96] selon trois parties : une partie descriptive, un graphe de cours et une partie structurelle. **Le graphe de cours** comporte un ensemble d'objectifs retenus pour le cours considéré et un ensemble de liens vers des ressources pédagogiques définies dans le curriculum. **La partie structurelle** décrit, de manière structurée et hiérarchique l'ordonnancement des objectifs à réaliser pour atteindre les objectifs terminaux définis pour le cours. Ces ressources sont de type opérationnel et ne sont donc pas directement utilisables en tant que fragment dans le contexte DVP. La **partie descriptive** par contre l'est tout à fait. En effet, elle comprend :

- *le titre du cours*
- *sa description*
- *une finalité* (correspondant au but ultime d'un cours)
- un ensemble d'*objectifs* dans lesquels l'enseignant exprime son intention pédagogique c'est à dire l'ensemble des changements durables (cognitifs, affectifs ou psychomoteurs) qu'il souhaite voir se produire chez les apprenants.
- un *ensemble de thèmes* (un thème correspond à un élément du contenu de la matière dont la maîtrise passe par la réalisation d'un ou de plusieurs objectifs pédagogiques). Un thème peut être décomposé en sous-thèmes. Pour chaque thème ou sous-thème un ou plusieurs objectifs spécifiques et opérationnels peuvent être définis. Les objectifs spécifiques sont définis en fonction des capacités que l'on veut faire acquérir à l'étudiant.

Si l'objectif est d'apprendre par cœur une liste de mots ou de concepts, le thème correspond au contenu effectif de la liste ou des définitions à apprendre ; si l'objectif est d'apprendre à réaliser un geste (faire une piqûre, faire un pansement), le thème correspond à la description exacte du geste (texte, film, ...). Ainsi, si l'on suppose que l'on fragmente en cours en briques élémentaires, le thème peut correspondre au contenu effectif de la partie de cours présentée dans le fragment et le titre, la description et l'objectif peuvent être les critères de méta description du fragment. La finalité, qui est une caractéristique d'un cours n'a plus vraiment de sens au niveau du fragment.

Comment renseigner la description et l'objectif ? Les critères *appariement*, *conclusion*, *description* sont généralement utilisés pour renseigner la description. Pour définir l'objectif Nkambou [NKA96] propose de prendre la nomenclature de Bloom qui les définit selon six catégories : *acquisition*, *compréhension*, *application*, *analyse*, *synthèse et évaluation*. A chaque catégorie est associée une liste de verbes d'action décrivant le comportement rendu possible par l'atteinte de l'objectif.

Nous allons voir dans la prochaine partie de cet article les modèles de description en cours d'élaboration pour modéliser ces deux types d'objets.

3. Vers la construction d'une norme descriptive des ressources

A l'heure actuelle, il n'existe pas encore de norme (au sens strict du terme) dans le domaine de l'e-learning. Les différents développements en cours consistent en un ensemble de spécifications proposées par différentes initiatives.

Nous retenons ici que seuls les aspects liés à l'organisation des ressources pédagogiques, leur caractérisation, aux moyens permettant leur transport c'est à dire l'échange entre systèmes logiciels nous concerne.

3.1. Les principaux groupes de travail

"La norme de métadonnées du Dublin Core est un ensemble d'éléments simples mais efficaces pour décrire une grande variété de ressources en réseau." [HIL01] Cet ensemble présenté en 1995 comprend quinze éléments (optionnels et pouvant être répétés) : titre, auteur, description, éditeur, autres collaborateurs, date, type de ressource, format, code d'identification de la ressource, source, langage, relation, portée, gestion des droits [DUB02].

Créé en 1988, l'AICC² [AIC02] est le comité chargé de l'enseignement par ordinateur de l'industrie aéronautique. Cette dernière, étant internationale, a depuis longtemps comme souci l'interopérabilité.

L'initiative ADL³ [ADL02] est connue pour le modèle SCORM⁴ lancé en 1999 (version 0.7.3) et propose une description de cours à l'aide de "grains" de telle manière que les déplacements de cours (contenu et données concernant les apprenants) et la répartition des cours en blocs réutilisables soient facilités. ADL est un programme du Ministère de la défense américaine et de l'Office of Science and Technology de la Maison-Blanche.

ARIADNE⁵ [ARI02] est un projet issu de l'Union Européenne mené entre 1996 et 2000 (projet de dépôt de documents pédagogiques mis en commun à travers le Knowledge Pool System). Le projet CANCORE⁶ (projet canadien) [CAN02] ou EdNA⁷ (projet australien) [EDN02] sont d'autres initiatives à échelle nationale menées actuellement.

² AICC : Aviation Industry CBT (Computer Based Training) Committee

³ ADL : Advanced Distributed Learning Initiative

⁴ SCORM : Shareable Course Object Reference Model

⁵ ARIADNE : Alliance of Remote Instructional Authoring and Distribution Networks for Europe

⁶ CANCORE : Canadian Core Learning Resource Metadata Application Profile

⁷ EDNA : Education Network Australia

"Caractérisation de l'information pédagogique pour le Web : principe d'application des recommandations du W3C dans un contexte de documents virtuels personnalisables" - Christine MICHEL et Soufiane ROUISSI- Article publié dans *les actes du 5ème Colloque International sur le Document Electronique (CIDE'05)*. 20-23 octobre 2002, Hammamet, Tunisie. - Sous la direction de Mauro GAI0, Christine VANOIRBEEK, Khaldoun ZREIK - Editeur Inria - 2002 - 308p - pp153-167.

Une convergence de ces différents projets est une première étape vers la future norme [MUR00]. Dans ce contexte l'IMS⁸ [IMS02] joue un rôle de centralisation des différentes propositions en les soumettant à IEEE⁹ qui a pour charge de soumettre les standards qu'il adopte à l'ANSI¹⁰. Les schémas ainsi définis peuvent avoir différentes utilisations selon que l'on soit fournisseur de contenus ou développeur de systèmes d'enseignement ou encore utilisateurs à la recherche de ressources (moteurs de recherche spécialisés). Il s'agit de spécifications ouvertes pour le repérage et l'exploitation de contenus à caractère pédagogique, le suivi des apprenants (administratifs et au niveau des évaluations). Ce consortium est également engagé dans d'autres processus de spécifications comme la description des évaluations ou la caractérisation des données concernant l'apprenant.

3.2. Un modèle de description des données : le LOM

Le LOM (*Learning Object Metadata*) est issu des travaux précédents mais se limite à l'ensemble minimal de caractéristiques indispensables pour gérer les objets pédagogiques. Il considère qu'un objet pédagogique est " *toute entité, sur un support numérique ou non (informatique), pouvant être utilisée pour l'apprentissage, l'enseignement ou la formation*" [LOM01]. Neuf catégories de descripteurs [IMS01a] [LOM01] sont prises en compte selon une structure hiérarchique (contrairement au modèle Dublin Core). Nous allons les présenter.

Description générale : "general" : La première catégorie **general** décrit l'objet pédagogique dans son ensemble en présentant les caractéristiques telles que l'identifiant de l'objet, son titre, sa description, la liste des langues utilisées, une liste de mots clés, l'étendue de la ressource (temps, géographie, culture...), le type de structure (collection, linéaire, hiérarchique...), son niveau de granularité (1 pour un niveau inférieur à 4 pour un cours entier).

Cycle de révision : "lifecycle" : Cette partie décrit la liste complète des modifications ou cycle de révision. Les éléments de l'ensemble de caractéristiques relatives à l'historique et à l'état courant de l'objet pédagogique (*draft, final...*), les personnes qui l'ont modifié, à quelle date ainsi que leur rôle (*author, publisher, instructional designer...*).

Métadonnées sur les métadonnées : "metametadata" : C'est un ensemble de métadonnées sur les métadonnées décrivant l'objet pédagogique. Cet ensemble décrit le schéma ou la spécification utilisée (metadatascheme). Il est possible de satisfaire à plusieurs schémas et de définir des liens dans un système de catalogage connu.

Les informations techniques : "technical" : Cette partie technique définit les exigences techniques en terme de navigateur (type, version), de système d'exploitation ou les caractéristiques comme le type des données ou format

⁸ IMS : Instructional Management System

⁹ IEEE : Institute of Electrical and Electronics Engineers

¹⁰ ANSI : American National Standards Institute

"Caractérisation de l'information pédagogique pour le Web : principe d'application des recommandations du W3C dans un contexte de documents virtuels personnalisables" - Christine MICHEL et Soufiane ROUISSI- Article publié dans *les actes du 5ème Colloque International sur le Document Electronique (CIDE'05)*. 20-23 octobre 2002, Hammamet, Tunisie. – Sous la direction de Mauro GAIIO, Christine VANOIRBEEK, Khaldoun ZREIK – Editeur Inria – 2002 - 308p - pp153-167.

(permettant d'identifier les logiciels nécessaires pour les lire), la taille de l'objet numérique (en octets), sa localisation physique (URL Uniform Resource Locator ou URI UR Identifier), des informations pour installer l'objet pédagogique et sa durée (en particulier pour les fichiers de type son, animation ou vidéo).

La partie pédagogique : "educational" : Pour mieux exploiter le contenu pédagogique, il convient de définir les conditions d'utilisation de la ressource : comment celle-ci doit être utilisée, savoir quel est son type (exercice, figure, index...), son niveau d'interactivité (de très faible à très élevée), à qui s'adresse la ressource (apprenant, enseignant, auteur...), le contexte (université, formation professionnelle, école primaire...) ou la tranche d'âge à laquelle s'adresse la ressource.

La gestion des droits : "rights" : Cette partie concerne les droits (copyright) liés à la ressource pédagogique, éventuellement son coût.

L'aspect relationnel : "relation" : Cette catégorie couvre les relations ou liens avec d'autres objets pédagogiques en précisant le genre de relation ("... est requis par ...", "... est une partie de ...").

Annotation : Cette partie regroupe les annotations ou commentaires.

Classification : Ces classifications sont décrites selon un type (niveau de compétences, discipline, objectif pédagogique, pré-requis, type d'accès, niveau d'études...) en associant des mots clés et une description.

Ces différentes catégories peuvent être directement applicables pour faire de la méta description de ressources banales fragmentées de manière à pouvoir les exploiter dans un contexte DVP. Ainsi, nous avons défini précédemment que les fragments de cours devaient être au moins définis par un titre, une description et un objectif, le thème ne faisant pas partie des méta descriptions mais permettant de représenter le contenu du fragment. Les balises correspondantes sont :

- l'élément *title* de la balise *general* pour le **titre**
- l'élément *type* de la balise *educational* permet de renseigner suivant un vocabulaire normalisé (*exercise, simulation, questionnaire, diagram, figure, graph, index, slide, table, narrative text, exam, experiment, problem statement, self assesment*) la **description**
- l'élément *purpose* de la balise *classification* renseigné selon un vocabulaire normalisé (*LOMv1.0* par défaut) défini dans *educationalobjective* (*acquisition, compréhension, application, analyse, synthèse et évaluation*) pour l'**objectif**.

A ces dernières peuvent facilement s'ajouter des caractéristiques d'usage soit :

- des caractéristiques techniques avec les éléments *format* et *requirement* (type de matériel, système d'exploitation, format de fichier...) de la balise *technical*
- des caractéristiques de langue (*language* dans *general*)
- des caractéristiques de public visé comme *context*, et *typicalagerange* de la balise *educational*

"Caractérisation de l'information pédagogique pour le Web : principe d'application des recommandations du W3C dans un contexte de documents virtuels personnalisables" - Christine MICHEL et Soufiane ROUISSI- Article publié dans *les actes du 5ème Colloque International sur le Document Electronique (CIDE'05)*. 20-23 octobre 2002, Hammamet, Tunisie. – Sous la direction de Mauro GAIIO, Christine VANOIRBEEK, Khaldoun ZREIK – Editeur Inria – 2002 - 308p - pp153-167.

- la discipline concernée avec l'élément *purpose* renseigné avec le vocabulaire *discipline* dans balise *classification*.

A titre d'exemple, considérons la ressource banale suivante : un exercice en français destiné à stimuler la compréhension et la traduction de textes techniques en physique des surfaces dans lequel l'apprenant doit écouter et lire les mots et choisir en cliquant la bonne traduction parmi 3 choix. Sa méta description pourrait être :

```
<lom>
<general>
  <title><langstring xml:lang="fr">Traduction de textes techniques - Physique des surfaces
  </langstring></title>
  <description><langstring xml:lang="fr">Exercice en français destiné à stimuler la compréhension et la
  traduction de textes techniques écrits en anglais. Dans cet exercice vous devez écouter et lire les mots
  et choisir en cliquant la bonne traduction parmi 3 choix.</langstring></description>
</general>
<technical>
  <format>audio/realaudio</format>
  <format>text/html</format>
</technical>
<educational>
  <learningresourcetype>
    <source><langstring xml:lang="x-none">LOMv1.0</langstring></source>
    <value><langstring xml:lang="x-none">exercice</langstring></value>
  </learningresourcetype>
  <context>
    <source><langstring xml:lang="x-none">LOMv1.0</langstring></source>
    <value><langstring xml:lang="x-none">Professional Formation</langstring></value>
  </context>
  <typicalagerange>
    <langstring xml:lang="fr">Toute personne ayant au moins 3 ans d'expérience</langstring>
  </typicalagerange>
</educational>
<classification>
  <purpose>
    <source><langstring xml:lang="x-none">LOMv1.0</langstring></source>
    <value><langstring xml:lang="x-none">Educational Objective</langstring></value>
  </purpose>
  <taxonpath>
    <source><langstring xml:lang="fr">Taxinomie de Bloom</langstring></source>
    <taxon><entry><langstring xml:lang="fr">compréhension</langstring></entry></taxon>
  </taxonpath>
</classification>
<classification>
  <purpose>
    <source><langstring xml:lang="x-none">LOMv1.0</langstring></source>
    <value><langstring xml:lang="x-none">Discipline</langstring></value>
  </purpose>
  <taxonpath>
    <source><langstring xml:lang="fr">Classification Décimale de Dewey</langstring></source>
    <taxon>
      <id>420</id>
      <entry><langstring xml:lang="fr">Langues anglaise et anglo-saxonne</langstring></entry>
    </taxon>
  </taxonpath>
</classification>
</lom>
```

Les vocabulaires utilisés pour certains éléments ne sont pas tous complètement figés ni normalisés. A titre d'exemple, l'élément *discipline* lié à *classification* peut être défini suivant les vocabulaires de la DDC (Dewey Decimal Classification), de l'UDC (Universal Decimal Classification), de la LCC (Library of Congress Classification), etc., mais l'élément *context* lié à *educational* est défini par le vocabulaire du DoL (Departement of Labor) qui est à l'heure actuelle instable et immature [IMS01a]. L'étude de toutes les potentialités et possibilités est énorme et doit faire l'objet d'un travail ultérieur.

"Caractérisation de l'information pédagogique pour le Web : principe d'application des recommandations du W3C dans un contexte de documents virtuels personnalisables" - Christine MICHEL et Soufiane ROUISSI- Article publié dans *les actes du 5ème Colloque International sur le Document Electronique (CIDE'05)*. 20-23 octobre 2002, Hammamet, Tunisie. – Sous la direction de Mauro GAIU, Christine VANOIRBEEK, Khaldoun ZREIK – Editeur Inria – 2002 - 308p - pp153-167.

3.3. Un modèle d'organisation des ressources : Le content framework

L'IMS préconise d'organiser les ressources nécessaires aux systèmes tutoriels selon un modèle, le *content framework* qui se divise selon deux blocs distincts [IMS01b] :

- le *package* regroupe l'ensemble des ressources d'apprentissages proposées par les auteurs
- le *data store* organisant les données sur les utilisateurs (apprenants, écoles, entreprises) et le *run time environnement* organisant les règles d'apprentissage et d'interactions sont regroupés dans le *content management scope*.

Le *package* a un modèle spécifique. Il s'organise en deux parties :

- un document XML qui décrit l'organisation du contenu ainsi que les ressources (ce document est appelé *manifest*)
- les fichiers physiques (fichiers textes, html, graphiques, vidéos) caractérisés par des métadonnées conformes aux spécifications de l'IMS c'est à dire au LOM.

Ce package est autonome et constitue une ressource ré-utilisable. Il est par exemple incorporable dans une plate-forme d'enseignement ou peut servir dans d'autres "conteneurs".

Le *content framework* est donc un modèle pouvant supporter le curriculum. La partie *content management* regroupe les ressources didactiques intelligentes et tutorielles, les capacités et les objectifs seront programmées dans le *run time environnement* et le *content packaging* permet de coder les ressources didactiques "banales", et aussi d'écrire des cours comme défini ci-dessus. En effet, le graphe de cours et la partie structurelle peuvent être écrits en XML dans le *manifest* et la partie descriptive est décrite grâce aux éléments du LOM.

Conclusion

Nous avons vu que les objets pédagogiques manipulés dans le cadre des systèmes tutoriels intelligents pouvaient être très diversifiés et plus ou moins interactifs. Un modèle général permet de les représenter : c'est le curriculum. Il regroupe l'ensemble des ressources utiles comme les capacités, les objectifs et les ressources didactiques intelligentes, tutorielles et banales. Nous avons vu que les ressources banales pouvaient facilement être exploitées dans un contexte DVP et que les organismes internationaux proposaient avec le LOM un modèle de description assez directement applicable. Pour les autres ressources qui font partie des ressources opérationnelles du système, le modèle DVP a des limites puisqu'il n'est pas applicable tel quel. Les questions qu'ouvre cette recherche sont de différentes natures : elles sont *structurelles* tout d'abord puisque des réflexions restent à mener sur l'adéquation du modèle DVP avec celui du *content framework*, et sur la complétude du vocabulaire utilisé pour définir au mieux les ressources. Elles sont *fonctionnelles* ensuite et portent sur la possibilité de faire non seulement de la

"Caractérisation de l'information pédagogique pour le Web : principe d'application des recommandations du W3C dans un contexte de documents virtuels personnalisables" - Christine MICHEL et Soufiane ROUISSI- Article publié dans les actes du 5ème Colloque International sur le Document Electronique (CIDE'05). 20-23 octobre 2002, Hammamet, Tunisie. - Sous la direction de Mauro GAIIO, Christine VANOIRBEEK, Khaldoun ZREIK - Editeur Inria - 2002 - 308p - pp153-167.

fragmentation automatique de supports de cours complets, mais aussi de la caractérisation automatique de fragments. En effet, l'obstacle majeur lié à l'utilisation de cette technique est que ces tâches restent à l'heure actuelle manuelles et il a été montré qu'elles étaient difficilement réalisables de manière automatique sur de l'information textuelle structurée [LAI00] ou des ressources de type multimédia [TES01].

Bibliographie

- [ADL02] <<http://www.adlnet.org>> (site consulté en mars 2002)
- [AIC02] <<http://www.aicc.org>> (site consulté en mars 2002)
- [ARI02] <<http://www.ariadne-eu.org>> (site consulté en mars 2002)
- [BEN97a] Ben Abdallah Nabil, Michel Christine, Lainé-Cruzel Sylvie "Caractérisation et découpage de textes scientifiques pour la construction de systèmes de requête personnalisés." In *Journées Scientifiques et Techniques du réseau Francil de l'AUFELF-UREF "L'ingénierie de la langue : de la recherche au produit"*. Avignon, 15-16 avril 1997.
- [BEN97b] Ben Abdallah Nabil – "*Analyse et structuration de documents scientifiques pour un accès personnalisé à l'information utile : vers un système d'information évolué*" – Thèse soutenue le 7 juillet 1997 à l'Université Lyon I.
- [BRU01] Brusilovsky Peter – "Adaptive Hypermedia" in *User Modeling and User-Adapted Interaction 11*- Kluwer Academic Publishers - pp87-110, 2001.
- [BUC94] Buckland Michael K. , Plaunt Christian "On the Construction of Selection Systems In *Library Hi Tech*, 12:4:15-28, 1994
- [CAN02] <<http://www.cancore.ca>> (site consulté en mars 2002)
- [CHA00] Chabert-Ranwez Sylvie – "*Composition Automatique de Documents Hypermédia Adaptatifs à partir d'Ontologies et de Requêtes Intentionnelles de l'Utilisateur*" - Thèse soutenue le 21 décembre 2000 à l'Université de Montpellier II
- [DUB02] <http://purl.oclc.org/metadata/dublin_core/> (document consulté en mars 2002)
- [EDN02] <<http://www.edna.edu.au>> (site consulté en mars 2002)
- [ELE01] eLearning - Penser l'éducation de demain, communication de la Commission des Communautés européennes, COM(2001) 172 final, 28 mars 2001, <http://europa.eu.int/eur-lex/fr/com/cnc/2001/com2001_0172fr01.pdf> (consulté en mars 2002)
- [HIL01] Guide d'utilisation du Dublin Core, Diane Hillmann, 2001, <<http://www.bibl.ulaval.ca/DublinCore/usageguide-20000716fr.htm>> (document consulté en mars 2002)
- [IMS01a] IMS Learning Resource Meta-Data Best Practice and Implementation Guide, Version 1.2.1, September 2001, Final Specification, <http://www.imsglobal.org/metadata/imsmdv1p2p1/imsmd_bestv1p2p1.html>
- [IMS01b] IMS Content Packaging Information Model, Version 1.2.1, Thor Anderson, Mark McKell, August 2001, Final Specification, <<http://www.imsglobal.org/content/packaging>> (document consulté en mars 2002)

"Caractérisation de l'information pédagogique pour le Web : principe d'application des recommandations du W3C dans un contexte de documents virtuels personnalisables" - Christine MICHEL et Soufiane ROUSSI- Article publié dans *les actes du 5ème Colloque International sur le Document Electronique (CIDE'05)*. 20-23 octobre 2002, Hammamet, Tunisie. – Sous la direction de Mauro GAI0, Christine VANOIRBEEK, Khaldoun ZREIK – Editeur Inria – 2002 - 308p - pp153-167.

- [IMS02] <<http://www.imsproject.org/>> (site consulté en mars 2002)
- [LAI96] Lainé-Cruzel S., Lafouge T., Lardy J.P., Ben Abdallah N – "Improving information retrieval by combining user profile and document segmentation" in *Information Processing and management* -1996- vol 32 n 3 - pp 305-315.
- [LAI00] Lainé-Cruzel S., Guinet E – "Fragmentation et enrichissement de textes scientifiques sous forme électronique" In *Document numérique, numéro spécial sur l'indexation*, 2000, vol.4, n°1-2, pp. 59-84
- [LOM01] Draft for Learning Object Metadata (LOM 6.1) , LTSC-IEEE P1484.12 LOM Working Group, 2001, <<http://ltsc.ieee.org/wg12/>> (document consulté en mars 2002)
- [MED01] Medini Lionel : "Accès à l'information orienté-utilisateur par recontextualisation et articulation du besoin d'information. Application à la maîtrise du risque de criticité." Thèse soutenue le 7 Septembre 2001 à l'Université Paris XI (Orsay))
- [MIC00] Michel Christine - "Diagnostic evaluation of a personalized filtering information retrieval system. Methodology and experimental results" in *Actes du congrès RIAO 2000 "Content based multimedia information access"*- Collège de France, Paris, 12-14 avril 2000 - pp1578-1589.
- [MIC99] Michel Christine, Lainé-Cruzel Sylvie - "Profil-Doc : Un prototype de système de recherche d'information personnalisé selon le profil des utilisateurs" - *Ateliers A2 11ème conférence francophone Interaction Homme Machine IHM'99* L'interaction pour tous" Montpellier, 22-26 novembre 1999.
- [MOU99] Moulin Claude, Pazzaglia Jean-Christophe – "Documents pédagogiques adaptatifs dans un environnement d'apprentissage distribué" in *Atelier : Documents virtuels personnalisables, de la définition à l'utilisation, 11ème conférence francophone Interaction Homme Machine IHM'99* "L'interaction pour tous" Montpellier, 22-26 Novembre 1999, <<http://www.eerie.fr/~multimedia/ihm99/articles/moulin.doc>> (document consulté en mars 2002)
- [MUR00] Mc Murray Elaine - "Des normes pour les technologies de la formation" in *Journal informatique de l'EPF de Lausanne*, Mai 2000, <<http://sawwww.epfl.ch/SIC/SA/publications/FI00/fi-4-00/4-00-page3.html>>
- [NKA96] Nkambou Roger : " *Modélisation des connaissances de la matière dans un système tutoriel intelligent, modèles, outils et applications*" Thèse présentée à l'Université de Montréal en Avril 1996
- [NKA96b] Roger Nkambou - Gilles Gauthier - "Integrating WWW resources in an Intelligent Tutoring System" in *Journal of Network and Computer Applications* 1996 N° 19 pages 353-365, Academic Press Limited
- [NKA97] Nkambou Roger, Gauthier Gilles et Frasson Claude - "Un modèle de représentation des connaissances relatives au contenu dans un système tutoriel intelligent" in *Sciences et Techniques Éducatives, International Journal*, Vol. 4, No. 3, pp. 299-330, HERMES, Paris.1997
- [ROI99] Roisin Cécile – "Documents structurés multimédia" - Habilitation à diriger les recherches présentée le 22 septembre 1999 à L'Institut national polytechnique de Grenoble.

"Caractérisation de l'information pédagogique pour le Web : principe d'application des recommandations du W3C dans un contexte de documents virtuels personnalisables" - Christine MICHEL et Soufiane ROUISSI- Article publié dans *les actes du 5ème Colloque International sur le Document Électronique (CIDE'05)*. 20-23 octobre 2002, Hammamet, Tunisie. – Sous la direction de Mauro GAIIO, Christine VANOIRBEEK, Khaldoun ZREIK – Editeur Inria – 2002 - 308p - pp153-167.

[TES01] *Testing IMS in real contexts, Implementing IMS specifications*, University of Glasgow, November 2001, <<http://www.gla.ac.uk/rcc/projects/tircs/tircs-full.pdf>> (document consulté en mars 2002)

"Caractérisation de l'information pédagogique pour le Web : principe d'application des recommandations du W3C dans un contexte de documents virtuels personnalisables" - Christine MICHEL et Soufiane ROUSSI- Article publié dans *les actes du 5ème Colloque International sur le Document Electronique (CIDE'05)*. 20-23 octobre 2002, Hammamet, Tunisie. – Sous la direction de Mauro GAIÒ, Christine VANOIRBEEK, Khaldoun ZREIK – Editeur Inria – 2002 - 308p - pp153-167.