

# Vers une approche Web Sémantique dans la conception d'un système d'apprentissage.

## Revue du projet TRIAL SOLUTION

Michel Buffa<sup>1</sup>, Sylvain Dehors<sup>2</sup>, Catherine Faron-Zucker<sup>1</sup>, Peter Sander<sup>1</sup>

<sup>1</sup> MAINLINE, I3S, UNSA  
930 route des Colles, bât ESSI, BP 145, 06930 Sophia Antipolis cedex  
{buffa, faron, sander}@essi.fr

<sup>2</sup> ACACIA, INRIA  
2004 route des Lucioles, BP 93, 06902 Sophia Antipolis cedex  
sdehors@sophia.inria.fr

**Résumé :** Dans le cadre du projet européen TRIAL SOLUTION, nous avons travaillé sur la publication de documents électroniques personnalisés à partir de livres scientifiques existants. L'approche générale adoptée dans ce projet a consisté à découper des livres électroniques en ressources pédagogiques élémentaires et à annoter celles-ci avec des méta-données décrivant leurs contenus sémantiques, leurs caractéristiques pédagogiques et leurs interrelations. Ces ressources annotées sont disponibles sur un serveur dédié à la production, par des étudiants et des enseignants, de matériel d'enseignement ou d'apprentissage personnalisé, à l'aide d'outils dédiés. Les annotations sont le support d'une recherche *sémantique* des ressources pédagogiques lors la production d'un document personnalisé. Nous décrivons dans cet article l'ensemble du projet et nous nous attachons plus particulièrement à la construction de la mémoire des ressources pédagogiques annotées, pour laquelle nous avons développé un outil d'annotation. Nous dressons ensuite un bilan de ce projet dans la perspective du Web sémantique, qui n'était pas présente lors de la définition de ce projet.

**Mots-clés :** Web sémantique, Système d'apprentissage, Entrepôt de ressources pédagogiques (LOR), Annotation de ressources pédagogiques, Extraction de ressources à partir de texte.

## 1 Introduction

TRIAL SOLUTION est un mot-valise pour Tools for Reusable, Integrated, Adaptable Learning - Systems/Standards for Open Learning Using Tested, Interoperable Objects and Networking. Ce projet européen s'est intéressé à la publication de documents personnalisés à partir de livres scientifiques existants (essentiellement dans le domaine des mathématiques).

La conception de nouveaux documents électroniques « from scratch » est très coûteuse et, pour cette tâche spécialisée, les auteurs de livres imprimés possèdent rarement les compétences nécessaires. D'autre part, la plupart des livres imprimés ont maintenant des formats digitaux, e.g. en Microsoft Word ou en Latex, très répandus dans la communauté des mathématiques ; ces formats peuvent faciliter l'automatisation de la conception de documents électroniques à partir de livres imprimés. Les bénéfices de cette approche basée sur le re-engineering de matériel existant ont été décrits dans (DAHN, 2001) et le processus lui-même détaillé dans (DAHN *et al.*, 2001). Ces considérations ont présidé à la conception du projet TRIAL SOLUTION.

L'approche générale consiste à découper des livres électroniques en ressources pédagogiques élémentaires puis à raffiner ce découpage et à annoter les ressources avec des méta-données sur leurs contenus sémantiques, sur leurs caractéristiques pédagogiques et leurs interrelations, dans le but de permettre une recherche « intelligente ». L'ensemble des ressources annotées est disponible sur un serveur Web, dédié aux enseignants et aux étudiants. Des outils de recherche de ressources et de publication de documents ont été développés pour permettre à ces utilisateurs finaux de rechercher sur le serveur le matériel pédagogique qui les intéresse et de produire des documents personnalisés. Notre contribution, l'une des pierres angulaires du projet, est l'outil d'annotation des ressources pédagogiques, qui permet le raffinement de leur découpage et leur annotation.

Dans la section suivante, nous décrivons les processus d'exploitation des livres et d'extraction de ressources pédagogiques ; nous présentons les différents outils qu'intègre la plate-forme TRIAL SOLUTION, en détaillant plus particulièrement l'outil d'annotation que nous avons développé. Dans la section 2 nous dressons ensuite un bilan du projet, dans la perspective du Web sémantique ; nous mettons en évidence la similarité de l'approche adoptée dans le projet avec celles du Web sémantique qui visent à annoter automatiquement des ressources pédagogiques à partir d'une ontologie. Nous concluons en décrivant notre travail actuel de conception d'un système d'apprentissage qui tire partie de notre expérience du projet TRIAL SOLUTION et des technologies du Web sémantique.

## **2 Présentation du projet TRIAL SOLUTION**

La plateforme TRIAL SOLUTION intègre trois principaux services : l'extraction et l'annotation automatiques de ressources pédagogiques à partir de livres électroniques, le re-engineering des ressources pédagogiques extraites et la recherche dans l'entrepôt de ressources ainsi constitué, basée sur les annotations sémantiques des ressources. La figure 1 ci après présente une vue générale du processus.

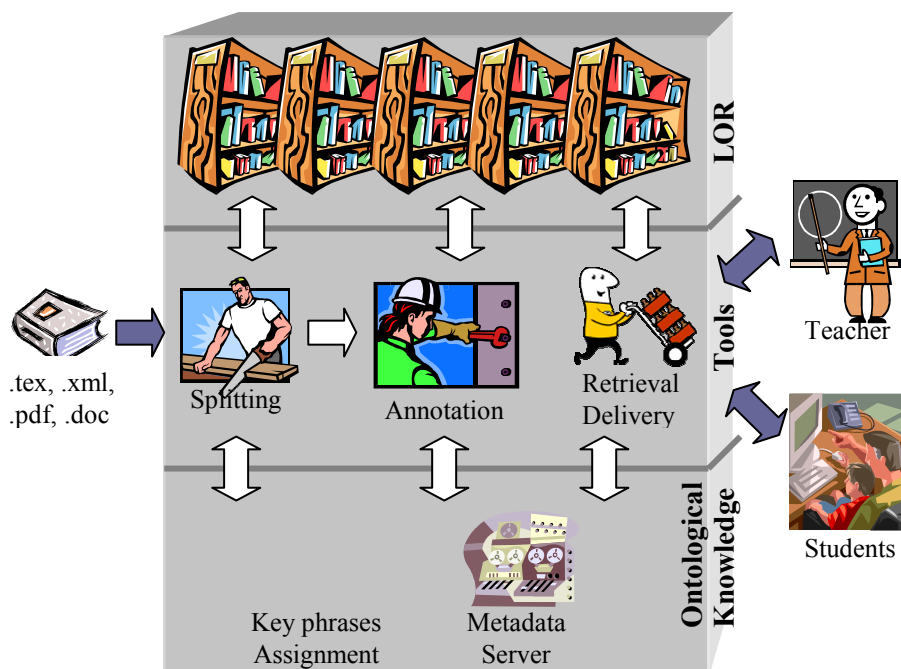


Fig. 1 – La plateforme TRIAL SOLUTION

## 2.1 Extraction de ressources pédagogiques à partir de textes

La première étape du processus consiste à désagréger automatiquement un document textuel en un ensemble de ressources, à l'aide d'un outil dédié, développé par la société Slicing Information Technology (SIT)<sup>1</sup>. Cet outil permet de traiter des documents dans des formats structurés tels que Word ou LaTeX. Il extrait de tels documents la structure hiérarchique sous la forme d'un arbre XML, dans lequel chaque noeud correspond à une ressource (une « tranche » découpée dans le document). La DTD qu'il utilise est détaillée sur le site du projet (TRIAL SOLUTION, 2001). Les sections, chapitres, tableaux, figures, exemples sont parmi les ressources les plus faciles à identifier automatiquement par cet outil et plus un document utilise de styles standards, meilleur sera son découpage. Le résultat de ce découpage est un arbre de ressources pédagogiques stocké dans un répertoire de fichiers ; ce répertoire est logé par un serveur Internet accessible par le protocole HTTP.

Cet outil annote automatiquement les ressources pédagogiques extraites avec des méta-données sur le contenu sémantique des ressources, leurs interrelations dans le livre d'origine et l'auteur de celui-ci. L'identification des relations entre les ressources est basée sur les hyperliens, les occurrences de « voir aussi », « cf. », etc. présents

<sup>1</sup> Infotech Slicing Technology GmbH, <http://www.slicing-infotech.de/de/index.php>

dans le document original. Les mots-clés, spécifiés parfois par l'auteur du document, aident à la classification du contenu des ressources, e.g. si le terme « sinus » apparaît dans le contenu d'une ressource, celle-ci est identifiée comme traitant de trigonométrie. A partir d'un thésaurus construit pour le domaine dont traite le document, il s'agit de rechercher dans celui-ci quels mots ou phrases clés du thésaurus apparaissent. Deux outils ont été conçus à cet effet : un outil de gestion du thésaurus, appelé serveur de meta-données, et un outil d'attribution de phrases clés qui gère une collection de phrases clés du domaine des mathématiques issues d'un manuel standard de mathématique (BRONSTEIN et al., 2000).

Cependant, ce processus automatique est presque toujours insuffisant pour annoter les ressources pédagogiques extraites, voire susceptible de produire des annotations incorrectes, comme cela est décrit dans (DAHN *et al.*, 2001). Les ressources pédagogiques extraites et leurs annotations sont affinées à l'aide d'un outil d'annotation.

## 2.2 Annotation de ressources pédagogiques

Nous avons développé l'outil d'annotation de la plateforme TRIAL SOLUTION qui permet à des experts du domaine traité de corriger et d'améliorer la base de ressources pédagogiques annotées et interconnectées produite automatiquement. Des expériences de mise en valeur d'un livre à l'aide de cet outil sont décrites dans (DAHN *et al.*, 2001). Les utilisateurs de cet outil doivent avoir une bonne connaissance du contenu du livre original ; il peut s'agir des auteurs eux-mêmes ou plus souvent d'enseignants utilisant ce livre comme matériel pédagogique.

Notre outil d'annotation est un client du serveur de ressources pédagogiques ; il communique avec celui-ci à l'aide d'un protocole où à la fois les requêtes (et leurs paramètres) et les valeurs retournées par le serveur sont basés sur un encodage XML défini dans la DTD de TRIAL SOLUTION.

Lorsqu'il est connecté au serveur de ressources pédagogiques extraites automatiquement à partir de livres électroniques, cet outil permet (1) d'attribuer aux ressources un titre ; (2) d'éditer leur contenu ; (3) d'éditer la structure arborescente de l'ensemble de l'entrepôt des ressources : les ressources peuvent être réorganisées, découpées, fusionnées, détruites et de nouvelles ressources peuvent être créées ; et (4) d'éditer les méta-données associées aux ressources. Trois types de méta-données sont distingués. Les « types » représentent le rôle pédagogique du contenu des ressources : définition, théorème, explication, etc. Les « mots-clés » représentent les sujets abordés dans le contenu des ressources. Les « relations » avec d'autres ressources constituent un réseau sémantique de ressources : référence, prérequis, etc.

Notre outil d'annotation permet l'édition non seulement des méta-données associées aux ressources pédagogiques mais aussi du thésaurus sur lequel reposent ces annotations : l'édition des termes du thésaurus, de la description associée à un terme, des relations hiérarchiques entre les termes, des relations de synonymie, de parenté.

La figure 2 ci-après présente l'interface de l'outil d'annotation pour l'édition des mots-clés. Une ressource intitulée « Zinsen und Zinsperiode » (intérêts et période d'intérêt) est sélectionnée ; elle a précédemment été automatiquement annotée avec



trois mots-clés : « Finanzmathematik » (mathématique financière), « Zinsen » et « Zinsperiode ». Un quatrième mot-clé lui est ici ajouté : « Nachschüssige Zinsen » (intérêts cumulés) ; ce mot-clé est sélectionné parmi une liste fournie par l’outil de gestion du thesaurus avec lequel communique l’outil d’annotation.

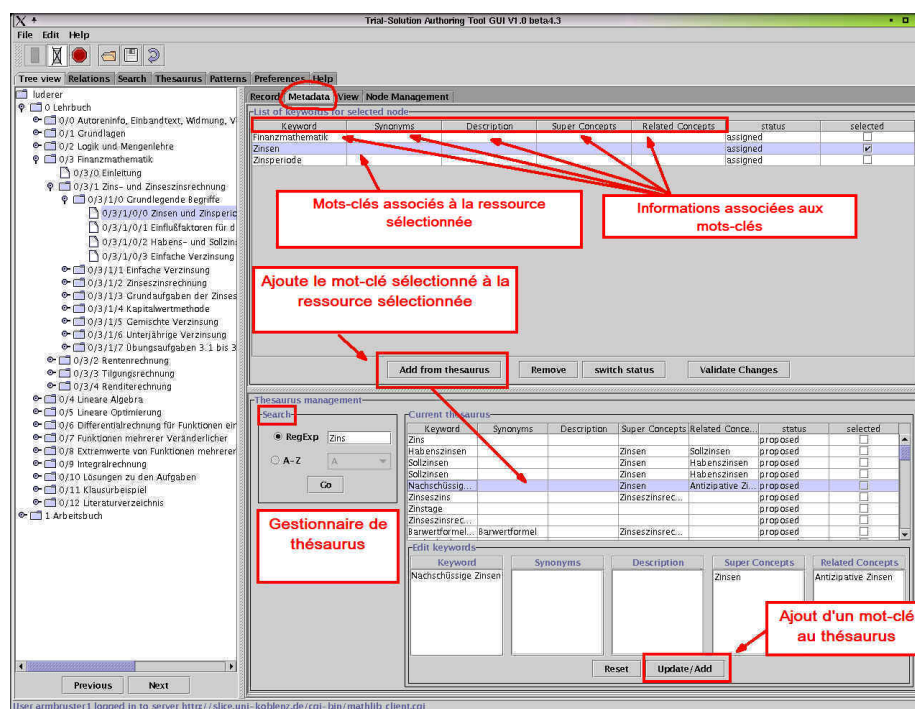


Fig. 2 – Edition de méta-données : les mots-clés du domaine

La figure 3 ci-après présente l’interface de l’outil d’annotation pour l’édition des types des ressources pédagogiques – types relatifs à leurs rôles pédagogiques. Sur cette figure, la ressource sélectionnée est annotée par les types « Definition », « Explanation » et « Mnemonic » sélectionnés dans une liste prédéfinie.

Les relations structurelles, les types et les mots-clés sont pour partie automatiquement extraits du document d’origine. En revanche les relations sémantiques entre les ressources sont plus délicates à extraire ; la plupart du temps elles sont acquises manuellement en utilisant l’outil d’annotation qui permet l’édition de relations, à choisir parmi une liste de relations prédéfinies, comme le montre la figure 4 ci-après.

La figure 5 ci-après montre un extrait du fichier XML dans lequel figurent les méta-données associées aux ressources. Dans cet extrait, la ressource intitulée « Chapter 2 » (1) est annotée par le type « Chapter » (2), par une relation de « Reference » deux autres ressources (3) et par deux mots-clés (4).

## Web sémantique pour le E-learning

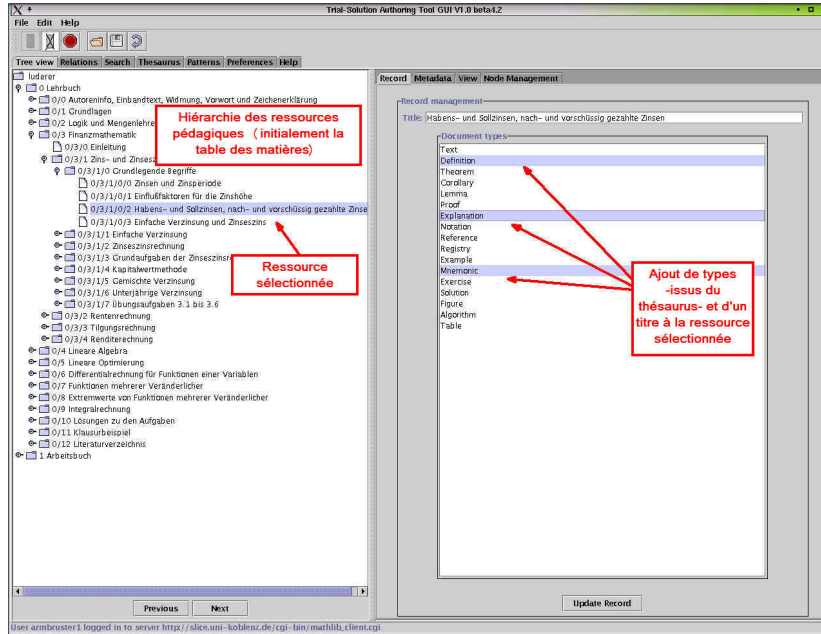


Fig. 3 – Edition de méta-données : les types pédagogiques

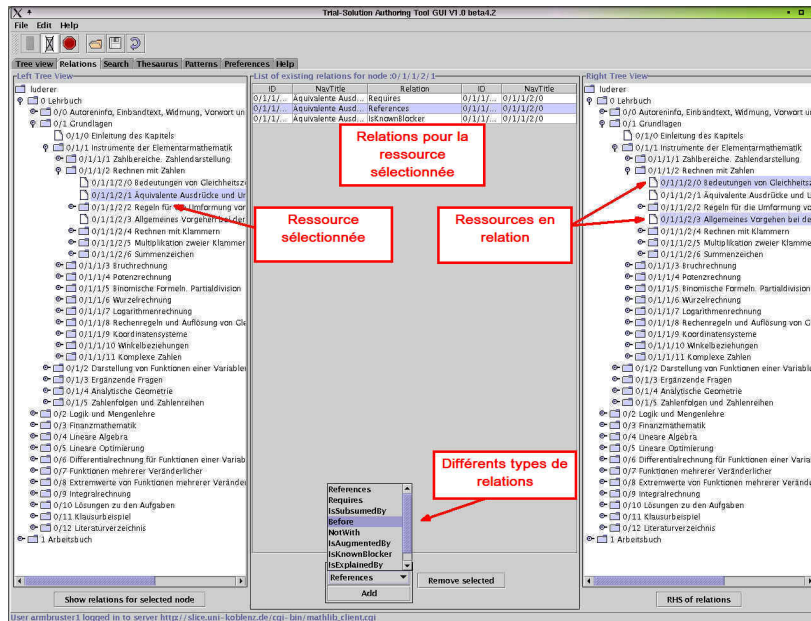


Fig. 4 – Edition de méta-données : les relations entre ressources pédagogiques

```
<?xml version="1.0" standalone="no"?>
<!DOCTYPE manifest SYSTEM "TS_IMS_CONTENT.dtd">
<manifest identifier="TSBook">
...
  <title> Complete Book</title>
  <item identifier="TSBook_all_i"
        identifierref="TSBook_r">
    <title>Title of Book</title>
...
  <item identifier="TSBook_all_i_2"
        identifierref="TSBook_r_2">
(1)   <title> Chapter 2</title>
      <metadata>
        <record>
          <sourcereference>2</sourcereference>
          <navtitle>Title of Chapter 2</navtitle>
(2)   <types>
        <type>Chapter</type>
      </types>
      <start/>
      <main href="slib_2_main.tex" />
      <end/>
(3)   <relation kind="References">
        <target href="TSBook_i_1_1" />
        <target href="TSBook_i_1_2" />
      </relation>
      <relation kind="IsAugmentedBy">
        <external href="http://www.trial-
solution.de">The Trial-Solution Project</external>
      </relation>
(4)   <keyphs>
        <keyph
          href= "slib_keywords_keys.db_i_260_s"
          inherits="down" author="ts_s"
          status="proposed" />
        <keyph
          href="slib_keywords_keys.db_s_41861_s" />
      </keyphs>
      </record>
    </metadata>
  </item> ... </item>
</manifest>
```

Fig. 5 – Les méta-données associées à la ressource intitulée « Chapter 2 »

## 2.3 Recherche de ressources pédagogiques

Une fois la version électronique d'un livre découpée et annotée, elle est disponible sous la forme d'un entrepôt de ressources pédagogiques annotées ; les utilisateurs finaux peuvent alors assembler sous la forme de documents personnalisés des ressources recherchées sur la base des méta-données qui leur sont associées. L'outil de recherche est une application Web destinée à des étudiants et des enseignants dans un environnement d'éducation.

Un utilisateur peut demander la recherche des ressources pédagogiques indexées par des mots-clés particuliers, e.g. « trigonométrie ». Il peut raffiner sa recherche en spécifiant le type des ressources à rechercher : définition, théorème, démonstration, etc. Le matériel « sur mesure » qui résulte d'une recherche peut être délivré en HTML ou PDF ; toutes les sections, figures et pages sont numérotées de sorte que le document final est autonome et cohérent.

La figure 6 ci-après présente l'interface de recherche de la plateforme TRIAL SOLUTION. Un utilisateur recherche les ressources de tous les types qui traitent d'« alphabet ». Le serveur propose deux mots-clés : « alphabet » et « alphabet d'un langage ». L'utilisateur choisit le second et demande également toutes les ressources qui sont des pré-requis aux ressources annotées par « alphabet d'un langage ». Il sauvegarde le document résultat de sa recherche dans un répertoire personnel.

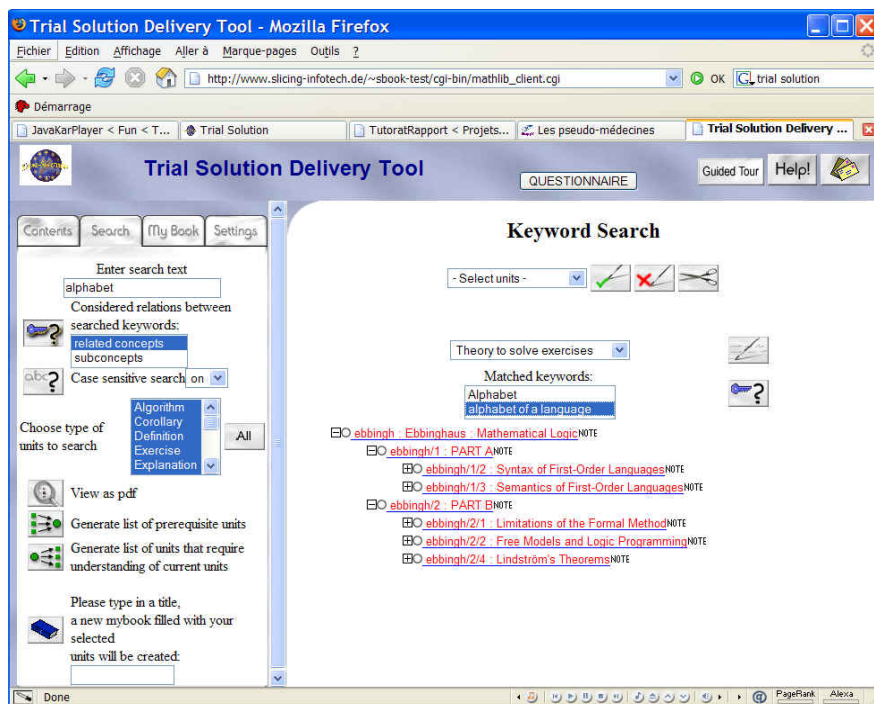


Fig. 6 – Recherche de ressources pédagogiques basée sur les méta-données

A l'issue de TRIAL SOLUTION, différents serveurs développés dans le cadre du projet sont maintenant en ligne, disponibles gratuitement sur le site Web du projet. D'autre part, des livres sont exploités commercialement par l'entreprise SIT.

### **3 Bilan du projet dans la perspective du Web Sémantique**

Nous dressons dans cette partie un bilan du projet Trial Solution dans la perspective du Web sémantique. Celle-ci ne présidait pas à la définition du projet, mais la description du projet met en évidence une démarche et des objectifs comparables.

#### **3.1 Comparaison de TRIAL SOLUTION avec d'autres travaux**

Une approche similaire de découpage automatique de document électronique a été mise en œuvre dans le projet SeLeNe (RIGAUX *et al.*, 2003). Les documents traités respectent la DTD DocBook qui est utilisée pour leur découpage. Dans le projet IMAT (DESMOULINS & GRANDBASTIEN, 2002) les documents sont issus de manuels techniques exportés en XML, puis découpés selon leur structure. Ils sont indexés selon leur contenu en identifiant la présence de termes issus d'une ontologie du domaine. L'indexation de l'aspect pédagogique est effectuée manuellement de même que la composition des documents finaux.

Dans les domaines des tutoriels intelligents (ITS) et des systèmes hypermédias adaptatifs (AHS) où l'utilisation de ressources pédagogiques élémentaires est très répandue, l'approche la plus commune pour construire des entrepôts de ressources est plutôt de concevoir le contenu de ces ressources à l'aide de systèmes auteurs dédiés (BRUSILOVSKY, 2003).

D'autres projets tels que Ariadne (DUVAL *et al.*, 2001) ou Memoraé (ABEL *et al.*, 2004) ont également proposé des entrepôts d'objets pédagogiques – ou Learning Object Repositories (LOR) - où les ressources sont stockées et annotées, avec des méta-données standard dans le premier cas, avec des ontologies dans le deuxième. Mais aucune hypothèse n'est faite sur l'origine des ressources, par exemple si elles ont été conçues spécialement pour constituer ces entrepôts ou bien si elles ont été extraites de documents existants. Dans l'entrepôt OLR (BRASE & NEJDL, 2003), les ressources pédagogiques sont annotées le langage RDF et sont organisées hiérarchiquement en tirant partie de la structure en chapitres du document dont elles sont tirées. Les ressources héritent ainsi les annotations de leurs « ancêtres » dans la structure. D'une manière générale tous ces LORs sont munis d'outils d'annotation qui permettent d'associer des méta-données aux ressources.

Dans TRIAL SOLUTION un document « sur mesure » est construit à partir des ressources « découpées », cette technique est également celle proposée par (CRAMPES & RANWEZ, 2000) où le parcours de lecture individuel est calculé à partir des annotations sur les documents. Ces annotations sont issues d'ontologies et exprimées dans un langage s'inspirant des graphes conceptuels.

L'outil d'annotation de TRIAL SOLUTION est comparable aux outils d'annotation des LORs; la spécificité de l'approche adoptée dans le projet est la combinaison d'un processus d'annotation manuelle des ressources avec une phase préliminaire d'extraction et d'annotation automatique des ressources à partir de documents électroniques.

### **3.2 TRIAL SOLUTION et les standards du Web**

La plateforme TRIAL SOLUTION est basée sur différents standards publics du web. Elle utilise le standard IMS Content Packaging (IMS, 2001) pour la modélisation des objets qui doivent être échangés entre les différents outils qui la composent. La spécification de ce standard est dédiée à la conception de formats d'échange, permettant en particulier l'agrégation et la désagrégation de ressources.

Les méta-données associées aux ressources pédagogiques sont conformes aux standards Dublin Core Metadata (DCMI, 1999), IMS Learning Resource Metadata (IMS, 2002) et LOM Learning Object Metadata (IEEE LTSC, 2002). Des adaptations ont cependant été nécessaires pour représenter de nouveaux matériels (LENSKI & WETTE-ROCH, 2001). Il a été nécessaire de spécifier un format de méta-données propre au projet, pour pallier certains défauts des spécifications des standards disponibles à l'époque de la conception du projet.

Les technologies ont depuis rapidement évolué. L'émergence des « Web services » (W3C, 2004c) répondrait aujourd'hui aux besoins d'échanges d'objets complexes entre client et serveur, aussi bien pour l'outil d'annotation que pour celui de recherche de ressources. De même, les langages du Web sémantique - RDF (W3C, 1999), RDFS (W3C, 2004a) et OWL (W3C, 2004b) – conviendraient davantage à la représentation des méta-données ; ils faciliteraient en outre leur échange et leur réutilisation.

### **3.3 Vers une approche « Web sémantique organisationnel »**

Le projet TRIAL SOLUTION au regard de l'approche adoptée dans la construction d'un entrepôt de ressources pédagogiques, s'accorde avec les enjeux du Web sémantique. Nous avons construit un entrepôt de ressources pédagogiques destinées à être partagées et réutilisées par une communauté d'utilisateurs. Ces ressources peuvent être efficacement recherchées grâce à leurs annotations sémantiques qui portent sur leurs contenus, leurs rôles pédagogiques, leurs interrelations structurelles et sémantiques. En somme, l'approche générale n'est pas différente de celle qui préside à la construction d'une mémoire organisationnelle pour une communauté donnée.

#### **3.3.1 Utilisation des standards du Web sémantique**

A l'époque de la conception du projet, les participants étaient étrangers à la communauté du Web sémantique et n'étaient pas conscients de son essor et des perspectives que cela ouvrait. A l'issue du projet, nous préconisons une représentation homogène de l'ensemble des méta-données, sous la forme d'annotations RDF basées

sur des ontologies RDFS ou OWL. Outre l'homogénéité de représentation qui facilite le partage et la réutilisation des annotations, ces langages offrent une puissance de représentation bien supérieure à celle de simples mots-clés pour décrire les ressources.

### **3.3.2 Extraction automatique d'annotations à partir de texte**

La seconde leçon que nous tirons de TRIAL SOLUTION concerne la démarche adoptée pour la population et l'annotation d'un entrepôt de ressources pédagogiques. L'expérience nous a montré que l'activité d'annotation manuelle des ressources est accablante pour des enseignants ou d'autres scientifiques, face à une importante quantité de ressources. Nous pensons qu'il est important d'automatiser autant que faire se peut ce processus ; notre expérience d'extraction automatique de ressources à partir de documents structurés est tout à fait concluante et les techniques mises en œuvre pour l'annotation automatique des ressources extraites pourraient être améliorées.

Cette approche n'exclue cependant pas l'utilisation d'un éditeur d'annotations. L'expérience montre que la mise au point des annotations ne se passe pas d'un expert du domaine et/ou d'un ingénieur de la connaissance et un outil d'édition des annotations reste nécessaire.

### **3.3.3 Explicitation collaborative du rôle pédagogique des ressources**

Corollairement, il apparaît que le succès de l'étape d'extraction automatique de méta-données à partir du contenu textuel des ressources pédagogiques dépend largement de la collaboration des auteurs lors de la conception des documents originaux. Nous avons déjà souligné l'importance de l'utilisation de styles standards dans la conception des documents : ils permettent d'extraire facilement le type des ressources et le réseau sémantique selon lequel elles s'organisent. Une phase d'acquisition de connaissances auprès de l'auteur des documents permettrait de mieux expliciter la correspondance entre l'organisation pédagogique du document et sa mise en forme et d'améliorer ainsi l'interprétation de sa mise en forme, e.g. la sémantique des différents styles, du code des couleurs, etc. Les annotations sur les caractéristiques pédagogiques des ressources seraient ainsi enrichies.

Des méthodes de traitement automatique du langage naturel (TALN) pourraient quant à elles améliorer l'extraction du contenu sémantique des ressources pédagogiques. Nous n'avons pas davantage exploré cette direction de recherche ; cela dépasse le cadre de cet article. Nous préconisons plutôt de minimiser l'utilisation des techniques de TALN et l'extraction du contenu sémantique des ressources pour se focaliser sur l'extraction de l'organisation pédagogique des ressources. Nous pensons que cette approche est possible pour la construction d'entrepôts de ressources pédagogiques concernant un domaine spécifique (par opposition à des entrepôts généraux) : dans ce cas, la connaissance du domaine devient implicite ; une annotation minimale sur le domaine suffit et c'est l'annotation des ressources sur leurs caractéristiques pédagogiques qui va permettre de les rechercher efficacement et de les réorganiser dans des documents personnalisés.

### 3.3.4 Définition de scénarios d'apprentissage précis

Enfin, dans le projet TRIAL SOLUTION, les buts des utilisateurs finaux étaient très généraux ce qui a rendu la tâche d'annotation difficile. Pour composer un document personnalisé, l'utilisateur est susceptible de rechercher n'importe quelle ressource, dans n'importe quel contexte. Nous préconisons la définition de scénarios d'apprentissage plus spécifiques et de profils d'apprentissage pour améliorer l'adéquation du contenu des annotations aux requêtes des utilisateurs.

En somme, l'expérience du projet nous conduit vers une approche Web sémantique dans la conception d'un entrepôt de ressources pédagogiques, cet entrepôt étant dépendant d'un domaine, d'une communauté, tout comme une mémoire organisationnelle. L'utilisation des standards du Web sémantique est la première conséquence de ce bilan. Nous préconisons l'implication des auteurs des documents originaux en amont du processus d'extraction automatique d'annotations portant sur le rôle pédagogique des ressources à partir de la mise en forme du document. La prise en compte des buts précis des utilisateurs devrait permettre de mieux définir la nature des annotations intéressantes pour la recherche ultérieure des ressources annotées.

## 4 Conclusion et perspectives

Nous avons présenté le projet TRIAL SOLUTION dédié à la publication de documents personnalisés, basée sur des livres scientifiques existants. L'approche générale consiste à construire un entrepôt de ressources pédagogiques et pour cela à extraire et annoter des ressources à partir de documents textuels électroniques. Les annotations sont le support d'une recherche sémantique de ressources dans l'entrepôt, par les utilisateurs finaux, étudiants ou enseignants, dont l'activité est de composer des documents personnalisés. L'automatisation du processus d'annotation exploite la mise en forme des documents et l'outil dit de re-engineering que nous avons développé permet d'éditer le contenu des ressources et leurs annotations afin d'enrichir manuellement les annotations et le thesaurus à partir duquel ces dernières sont construites.

Nous avons dressé un bilan du projet dans la perspective du Web sémantique. Les standards du Web sémantiques émergeaient à peine au moment de la conception du projet, si bien que les annotations des ressources suivent LOM mais ni OWL ni RDF(S). Cependant, l'approche générale adoptée dans le projet s'accorde tout à fait avec les enjeux du Web sémantique. Nous avons construit un entrepôt de ressources pédagogiques annotées comme une mémoire organisationnelle. Ces ressources sont destinées à être partagées et réutilisées, leurs annotations sont basées sur une pseudo-ontologie et portent sur leur contenu sémantique, leurs rôles pédagogiques et leurs interrelations.

Les leçons que nous avons pu tirer de ce projet nous amènent à préconiser une approche « mémoire organisationnelle » dans la construction d'un entrepôt de ressources pédagogiques dédié à un domaine particulier ; à utiliser les standards du



Web sémantique pour annoter ces ressources ; à accentuer l'automatisation du processus d'extraction de ressources et d'annotation à partir de documents textuels ; à impliquer l'auteur du document dans ce processus d'annotation pour expliciter la correspondance entre sa stratégie pédagogique et la mise en forme de son document ; à définir des scénarios pédagogiques qui seront pris en compte dans le contenu des annotations.

Dans la continuité de ce projet, nos travaux actuels s'inscrivent dans le cadre de l'action scientifique Web-Learn dont le but général est d'étudier les bénéfices méthodologiques et techniques d'une approche Web sémantique pour le E-Learning. Nous tirons partie de notre expérience de TRIAL SOLUTION et travaillons à une méthode collaborative de construction d'une mémoire organisée de ressources pédagogiques à partir de documents existants (DEHORS *et al.*, 2005). Elle consiste dans un premier temps à acquérir auprès de l'auteur d'un document pédagogique la correspondance entre la mise en forme de son document et sa stratégie pédagogique. Nous construisons ensuite une ontologie permettant de représenter les caractéristiques pédagogiques explicitées. Les ressources pédagogiques et leurs annotations sont alors extraites du document en tirant partie de cette correspondance. Ces annotations sémantiques sont ensuite utilisées pour offrir différentes stratégies de navigation dans la mémoire des ressources pédagogiques d'un cours donné. Les ressources sont retrouvées en utilisant le moteur de recherche sémantique (CORBY *et al.*, 2004) ; la recherche est basée sur les annotations RDF des ressources et l'ontologie RDFS sur laquelle elles sont construites.

## Références

- M.-H. ABEL, C. BARRY, A. BENAYACHE, B. CHAPUT, D. LENNE, C. MOULIN (2004), Using an Organizational Memory for e-learning, in Proceedings of Workshop of Knowledge Management and Organizational Memories (ECAI'2004), Valencia, Spain, 2004
- J. BRASE & W. NEJDL (2003), Ontologies and Metadata for eLearning, Handbook on Ontologies, pp. 579-598, Springer Verlag, 2003.
- BRONSTEIN, SEMENDJAJEW, MUSIOL & MÜHLIG (2000), *Taschenbuch der Mathematik*, Verlag Harri Deutsch, 2000.
- P. BRUSILOVSKY (2003), Developing Adaptive Educational Hypermedia Systems: From Design Models to Authoring Tools. In T. Murray, S. Blessing and S. Ainsworth Eds.: *Authoring Tools for Advanced Technology Learning Environment*. Dordrecht: Kluwer Academic Publishers, pp. 377-409, 2003.
- O. CORBY, R. DIENG-KUNTZ & C. FARON-ZUCKER (2004), Querying the Semantic Web with the Corese Search Engine, in Proceedings of the European Conference on Artificial Intelligence (ECAI'2004), subconference PAIS, pp. 705-709, Valencia, Spain, 2004.
- M. CRAMPES & S. RANWEZ (2004) Ontology-Supported and Ontology-Driven Conceptual Navigation on the World Wide Web, in Proceedings of 11<sup>th</sup> ACM Conference on Hypertext, San Antonio, Texas, 2000.
- I.DAHN (2001), Slicing Book Technology – Providing Online Support for Textbooks, in Proceedings of the International Conference on Distant Education, ICDE'2001, Dusseldorf, Germany, 2001.

## Web sémantique pour le E-learning

I. DAHN, M. ARMBRUSTER, U. FURBACH & G. SCHWABE (2001). Slicing Books – The Authors' Perspective, in R. Bromme, E. Stahl Eds. : *Writing Hypertext and Learning: Conceptual and Empirical Approaches*, Pergamon, 2001.

DCMI (1999), Dublin Core Metadata Element Set, Version 1.1: <http://dublincore.org/documents/dces/>

S. DEHORS, C. FARON-ZUCKER, A. GIBOIN & J.P. STROMBONI (2005), Un web sémantique de formation par questionnement, in Actes de la journée EIAH, EGC2005, Paris, 2005.

C. DESMOULINS, M. GRANDBASTIEN (2002) Des ontologies pour la conception de manuels de formation à partir de documents techniques. In *Sciences et Techniques Educatives*, vol.9, n°3-4, Hermès.

E. Duval, E. Forte, K. Cardinaels, B. Verhoeven, R. Van Durm, K. Hendriks, M. Wentland-Forte, N. Ebel, M. Macowicz, K. Warkentyne & F. Haenni (2001), The ARIADNE Knowledge Pool System, in *Communications of the ACM*, 44 (5), pp. 73-78.

IEEE LTSC (2002), LOM, Learning Object Metadata: <http://ltsc.ieee.org/wg12/>

IMS (2001), IMS Content Packaging Specification: <http://www.imsproject.org/content/packaging/index.html>

IMS (2002), Learning Resource Metadata, <http://www.imslobal.org/metadata/index.html>

W. LENSKI, E. WETTE-ROCH (2001). The Trial-Solution Approach to Document Re-use, in *Proceedings of Workshop on Electronic Media in Mathematics*, Coimbra, Portugal.

P.RIGAUX & N.SPYRATOS (2003), Metada Management and Learning Object Composition in a Self e-Learning Network, in *Proceedings of Workshop on Information Search, Integration and Personalization*, Japan, 2003.

B. SIMON & J. QUEMADA (2002), A Reflection of Metadata Standards based on Reference Scenarios, 2002, <http://www.wu-wien.ac.at/usr/wi/bsimon/publikationen/SimonQuemada-ReflectionOnMetadataStandards.pdf>

TRIAL SOLUTION (2001), EU project, Information Society Technologies Program (IST): <http://www.trial-solution.de/>

W3C (1999), RDF, Resource Description Framework: <http://www.w3.org/RDF/>

W3C (2004a), RDFS, RDF Vocabulary Description Framework: <http://www.w3.org/TR/rdf-schema/>

W3C (2004b), OWL, Web Ontology Language: <http://www.w3.org/TR/owl-features/>

W3C (2004c), Web Services: <http://www.w3.org/2002/ws/>

WEB-LEARN (2004), Action Scientifique CNRS et Action Color INRIA: <http://www.lirmm.fr/~toutou/as-weblearn/>