

Cahiers **GUT** *enberg*

☞ MÉTADONNÉES, RDF ET DOCUMENTS PÉDAGOGIQUES

☞ Yolaine BOURDA, Marc HÉLIER

Cahiers GUTenberg, n° 35-36 (2000), p. 37-52.

<http://cahiers.gutenberg.eu.org/fitem?id=CG_2000__35-36_37_0>

© Association GUTenberg, 2000, tous droits réservés.

L'accès aux articles des *Cahiers GUTenberg*

(<http://cahiers.gutenberg.eu.org/>),

implique l'accord avec les conditions générales

d'utilisation (<http://cahiers.gutenberg.eu.org/legal.html>).

Toute utilisation commerciale ou impression systématique

est constitutive d'une infraction pénale. Toute copie ou impression

de ce fichier doit contenir la présente mention de copyright.

Métadonnées, RDF et documents pédagogiques

Yolaine BOURDA et Marc HÉLIER

*Supélec, Plateau de Moulon, F-91192 Gif-sur-Yvette CEDEX
{Yolaine.Bourda, Marc.Helier}@supelec.fr*

Résumé. Dans de nombreux domaines, dont l'enseignement, les documents électroniques ne sont pas « rentabilisés » (retrouvés, réutilisés. . .) comme ils le devraient. Pour remédier à cela, une solution possible passe par l'utilisation de métadonnées, de RDF et de XML.

Nous présentons la notion de métadonnées et insistons sur la nécessité d'une normalisation. Pour un même ensemble de métadonnées, on peut trouver différentes implémentations en XML. Cette multiplicité des DTDs possibles a de nombreux inconvénients. Une façon d'obtenir une implémentation unique est d'utiliser RDF. Le Dublin Core et les « métadonnées pédagogiques », dont la normalisation par les IEEE est en cours d'élaboration, nous servent d'exemples.

Mots-clés : Métadonnées, RDF, XML.

Abstract.*In many fields, like the educational field, electronic documents do not pay for themselves as they should (reused, found. . .). A possible solution is to rely on metadata, RDF and XML. The aim of this paper is to present the idea of metadata and to emphasize the importance of standardization. For a given set of metadata, many implementations using XML are possible. This multiplicity has its drawbacks. An unique implementation may be obtained by means of RDF. The Dublin Core initiative and the Learning Objects which are under construction by the IEEE, are given as examples of this process.*

Keywords: Metadata, RDF, XML.

1. Métadonnées

1.1. Contexte et définition

Trouver des informations sur le Web revient de plus en plus à chercher une aiguille dans une botte de foin [7]. En effet, le nombre actuel de documents accessibles via le Web est estimé entre huit cents millions et un milliard. Les moteurs de recherche les plus performants en indexent moins de 16%. De plus, de nombreux documents sont

générés automatiquement à partir d'informations stockées dans des bases de données et ne peuvent donc pas être retrouvés via un moteur de recherche.

Quant à trouver des informations pertinentes, c'est quasiment « mission impossible », les outils existants ne pouvant pas s'appuyer actuellement sur une description du contenu des documents.

La première idée qui vient à l'esprit est d'ajouter une information de nature sémantique aux documents de manière à en obtenir une description plus précise. Les métadonnées, comme l'a démontré Lassila [6], sont appropriées pour atteindre ce but. Qu'entend-on exactement par « métadonnées » ? Nous pouvons définir les métadonnées comme des « données relatives à des données ». Il s'agit d'informations décrivant les ressources dans le but de rendre plus performantes la recherche de contenus, leur gestion et leur utilisation. Comme l'a expliqué Tim Bernes-Lee [1] : « Les métadonnées sont des informations sur des ressources web (ou autres) compréhensibles par les êtres humains et traitables par des logiciels ». L'expression « traitable par des logiciels » est la clé : il s'agit d'informations exploitables par des outils, tels que, par exemple, des agents intelligents.

Un exemple de métadonnées qui vient tout de suite à l'esprit est le catalogue d'une bibliothèque car il contient des informations (mots-clés, auteur, ISBN...) sur les livres et leurs contenus. Un autre exemple, utilisé quotidiennement en informatique, est le répertoire contenant des informations (taille, protection...) sur les fichiers. Dans le domaine de l'éducation, des informations relatives aux auteurs de documents pédagogiques, à leurs champs d'intérêt, à leurs idées, à leurs objectifs pédagogiques, etc. sont des métadonnées.

Ainsi, comme Monsieur Jourdain faisait de la prose sans le savoir, nous utilisons tous des métadonnées sans en être conscients.

Comme on vient de le voir, les métadonnées sont des données décrivant des données, donc ce sont aussi des données pouvant être elles-mêmes décrites par des métadonnées. Ainsi, le nom du créateur (documentaliste) des mots-clés associés à un livre est une métadonnée sur une métadonnée. Elle exprime une propriété de la description du livre et pas de son contenu.

Bien que la recherche d'informations soit une finalité essentielle des métadonnées, celles-ci peuvent aussi servir à de nombreuses autres applications comme la gestion des droits liés aux documents, l'administration des documents, leur archivage, etc.

Les métadonnées peuvent être incluses dans les ressources elles-mêmes ou enregistrées dans un fichier séparé, voire dans une base de données. Le type du contenu a bien évidemment une influence sur la façon de les stocker. Ainsi il n'est pas aisé d'inclure des métadonnées dans une image, par exemple. Les métadonnées sont structurées suivant des catégories ou champs sémantiques. Chaque champ représente une caractéristique particulière de la ressource, par exemple, son titre ou son résumé. Certains champs sont monovalués, d'autres multivalués. Certains sont obligatoires et d'autres optionnels.

La capacité des métadonnées à faciliter l'accès aux ressources en ligne dépend grandement de l'existence d'un standard, doté de propriétés. La première de ces propriétés est la pérennité des métadonnées définies afin que fournisseurs de ressources et utilisateurs puissent utiliser le même référentiel. La seconde est la possibilité d'évolutions ou d'extensions afin que les métadonnées puissent être adaptées à de nouveaux besoins. Certains organismes peuvent ainsi utiliser des champs qui leur seraient propres et non communiqués aux autres utilisateurs.

1.2. Exemple : le Dublin Core

Comme exemple d'ensemble complet de métadonnées, nous avons choisi le Dublin Core [2]. Le « Dublin Core Metadata Element Set », développé par une commission internationale et interdisciplinaire, est constitué de quinze éléments décrivant des ressources en ligne et devient un standard de facto. Chacun des éléments est optionnel et peut être répété autant de fois que nécessaire. Certains peuvent être associés à un nombre limité de qualificatifs afin d'affiner la description sémantique. Des terminologies précises sont utilisées pour certains d'entre eux, afin de favoriser l'interopérabilité.

Nous donnons ci-dessous, la version 1.1 « non qualifiée » du Dublin Core. Elle est constituée des éléments suivants :

Titre

Identificateur	title
Définition	Un nom donné à la ressource.
Commentaire	Généralement, un titre est un nom par lequel la ressource est connue.

Créateur

Identificateur	creator
Définition	Une entité principalement responsable de la création du contenu de la ressource.
Commentaire	Comme exemples de créateurs, on peut citer une organisation, un service, une personne.

Sujet et mots-clés

Identificateur	subject
Définition	Un thème de la ressource.
Commentaire	Généralement, le sujet sera décrit par des mots-clefs, des phrases ou des codes issus d'une classification qui précisent le thème de la ressource. Il est recommandé de choisir ces valeurs parmi des vocabulaires autorisés et des schémas de classification.

Description

Identificateur	description
Définition	Une note sur le contenu de la ressource.
Commentaire	Cet élément peut contenir, sans que cela soit restrictif, un résumé, une table des matières, une référence à une représentation graphique du contenu ou un texte libre.

Éditeur

Identificateur	publisher
Définition	Une entité responsable de la diffusion de la ressource.
Commentaire	Comme exemples d'éditeur, on peut citer une organisation, un service, une personne. Généralement, le nom de l'éditeur sera utilisé.

Contributeur

Identificateur	contributor
Définition	Une entité ayant contribué à la ressource.
Commentaire	Comme exemples de contributeur, on peut citer une organisation, un service, une personne. Généralement, le nom du contributeur sera utilisé.

Date

Identificateur	date
Définition	Une date associée à un événement du cycle de vie de la ressource.
Commentaire	Généralement, la date associée à la création ou à la mise à disposition de la ressource. Il est recommandé d'utiliser le format AAAA-MM-JJ (iso 8601).

Type

Identificateur	type
Définition	La nature ou le genre du contenu de la ressource.
Commentaire	Le type inclut des termes décrivant des catégories générales, des genres, etc. Il est recommandé d'utiliser une valeur choisie dans un vocabulaire autorisé (par exemple la version préliminaire des différents types).

Format

Identificateur	format
Définition	Le format physique ou numérique de la ressource.
Commentaire	Généralement, le format peut inclure le type de média ou la dimension de la ressource. Le format peut être utilisé pour identifier le logiciel ou le matériel qui peuvent être nécessaires pour afficher ou traiter la ressource. Comme exemples de dimensions, on peut citer la taille et la durée. Il est recommandé de choisir une valeur dans un dictionnaire autorisé.

Identificateur

Identificateur	identifier
Définition	Une référence non ambiguë, dans un contexte donné, à la ressource.
Commentaire	Il est recommandé d'identifier la ressource par une chaîne de caractères ou un nombre conformes à un système d'identification formel.

Source

Identificateur	source
Définition	Une référence à une ressource à partir de laquelle cette ressource est dérivée.
Commentaire	Cette ressource peut-être dérivée de la ressource source, en totalité ou en partie. Il est recommandé d'identifier la ressource par une chaîne de caractères ou un nombre conformes à un système d'identification formel.

Langage

Identificateur	langage
Définition	Un langage du contenu intellectuel de la ressource.
Commentaire	Il est recommandé d'utiliser les valeurs définies par la RFC 1766.

Relation

Identificateur	relation
Définition	Une référence à une ressource en liaison avec celle-ci.
Commentaire	Il est recommandé d'identifier la ressource par une chaîne de caractères ou un nombre conformes à un système d'identification formel.

Couverture

Identificateur	coverage
Définition	L'étendue du contenu de la ressource.
Commentaire	La couverture inclut, généralement, une localisation physique (coordonnées géographiques, nom de lieu, . . .), une période de temps (une date, un intervalle de dates, un nom de période, . . .). Il est recommandé de choisir une valeur dans un vocabulaire autorisé.

Gestion des droits

Identificateur	rights
Définition	Une information à propos des droits sur la ressource.
Commentaire	Généralement, cet élément contient une instruction sur la gestion des droits de la ressource ou une référence à un service fournissant ce type d'informations. Ces droits sont, entre autres, les droits de reproduction ou les droits d'utilisation. Si cet élément est absent, aucune hypothèse ne peut être faite sur les droits associés à la ressource.

La RFC 2731 explique comment coder les éléments du Dublin Core en HTML à l'aide des balises meta.

Un exemple d'utilisation des balises meta conformes au document précédent est donné ci-dessous.

```
<head>
<meta http-equiv="content-type"
      content="text/html;
      charset=iso-8859-1">
```

```

<meta name="dc.Title" content="Yolaine Bourda">
<meta name="dc.Description"
  content="Ceci est la page d'accueil de Yolaine Bourda">
<meta name="dc.Publisher" content="Supélec">
<meta name="dc.Creator" content="Yolaine Bourda">
<meta name="dc.Date" content="2000-03-10">
<meta name="dc.Language" content="fr">
<meta name="dc.Subject" content="Enseignement">
<meta name="dc.Subject" content="XML">
<meta name="dc.Subject" content="Bases de Données">
<meta name="dc.Rights" content="Copyright Yolaine Bourda">
<meta name="dc.Format" content="text/html">
<meta name="dc.Type" content="World Wide Web Home Page">
</head>

```

2. RDF

2.1. Introduction

Une fois défini un ensemble de métadonnées, la question importante qui se pose est celle de son implémentation. Un langage formel est alors nécessaire.

L'implémentation d'un ensemble de métadonnées peut se faire très simplement en XML. Ainsi les métadonnées « titre » et « auteur » d'un document quelconque peuvent être représentées de la façon suivante :

```

<document>
<titre>Le titre du document </titre>
<auteur>L'auteur du document </auteur>
</document>

```

Malheureusement, cette forme n'est pas unique, on peut aussi penser à la forme suivante :

```

<document auteur = "L'auteur du document"
  titre = "Le titre du document">

```

et à beaucoup d'autres.

Une façon d'éviter que chacun ait sa propre représentation des métadonnées est de leur associer, quand on les normalise, la façon de les représenter en XML. Mais cette solution, si elle évite les problèmes ci-dessus, n'est quand même pas suffisante car elle laisse subsister une perte de sémantique. Elle ne laisse percevoir que la structure. Or, ce à quoi il faut avoir accès, c'est aux propriétés (métadonnées) d'un document afin, entre autres, de pouvoir faire des recherches.

RDF (Resource Description Framework) propose un modèle pour la représentation des métadonnées sur le Web. RDF est défini par deux documents issus du W3C :

- « RDF Model and Syntax Specification » qui est une recommandation du 22 février 1999 [12];
- « RDF Schema Specification » qui est une recommandation proposée du 3 mars 1999 [11].

Ce modèle peut rendre compte des propriétés de n'importe quelle ressource, identifiée par une URI (Uniform Ressource Identifier), et des valeurs qu'elles prennent. Pouvant être représenté sous la forme d'un graphe orienté et valué, ce modèle a l'avantage de structurer naturellement les informations. Un ensemble de propriétés décrivant une ressource est appelé une « description » dans RDF. Celle-ci est composée d'un ensemble de propriétés, chacune d'entre-elles étant composée d'un type et d'une valeur, celle-ci pouvant être elle-même une ressource. L'association d'une ressource, d'un nom de propriété et de la valeur de cette propriété pour la ressource est une « assertion » pouvant être représentée par un triplet (propriété, ressource, valeur). Un ensemble d'assertions peut être modélisé par un graphe orienté dans lequel ressources et valeurs sont modélisées par des sommets. Les arcs, modélisant la propriété, partent de la ressource pour aller vers la valeur et ont comme étiquette la propriété.

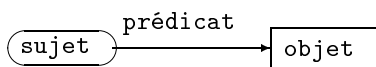
Une description RDF, c'est-à-dire un ensemble d'assertions modélisé par un graphe, peut être « sérialisée » en XML, bien que cela ne soit pas la seule possibilité. L'utilisation des espaces de nommage en XML permet de préciser, pour chaque propriété, à quel schéma elle appartient et ainsi, dans une même description, d'associer plusieurs ensembles de métadonnées.

2.2. Modèle de données RDF

Le modèle formel RDF est basé sur les concepts suivants :

1. Ressource (Resource) : entité physique (personne, livre, page web...) ou abstraite (couleur, qualité...). Une ressource peut être une composante d'une autre ressource. Toute ressource possède une URI.
2. Propriété (Property) : un aspect spécifique, une caractéristique utilisée pour décrire une autre ressource. Une propriété est elle-même une ressource.
3. Littéral (Literal) : une chaîne de caractères non interprétée par RDF. Elle ne peut pas avoir de caractéristiques.
4. Assertion (Statement) : un triplet de la forme {prédicat, sujet, objet} dans lequel **prédicat** est une propriété, **sujet** est une ressource et **objet** est, soit une ressource, soit un littéral. Un ensemble d'assertions peut être modélisé par un graphe orienté et valué dans lequel les sommets sont les ressources et les littéraux. Les arcs sont valués par les prédicats.

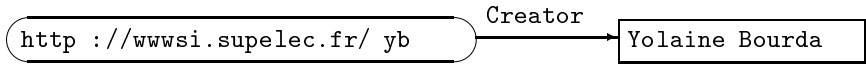
Ainsi, le triplet {prédicat, sujet, objet} peut être modélisé par le graphe suivant :



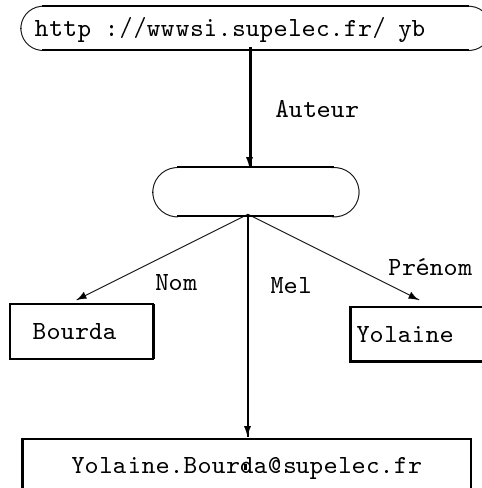
Par exemple, la phrase « Yolaine Bourda est l'auteur de la ressource `http://wwwsupelec.fr/yb` » peut être modélisée par le triplet :

```
{auteur, http://wwwsupelec.fr/yb, Yolaine Bourda}
```

et par le graphe :



Maintenant, il se peut que l'auteur de la ressource possède lui-même des propriétés. Par exemple, la phrase « L'auteur de la ressource `http://wwwsupelec.fr/yb` a comme nom Bourda, comme prénom Yolaine et comme mel `Yolaine.Bourda@supélec.fr` », peut être représentée par le graphe suivant :



La ressource représentant l'auteur est une ressource anonyme mais on peut aussi lui associer un identificateur unique comme par exemple un numéro d'employé. On obtient alors les triplets suivants :

```

{auteur, http://wwwsupelec.fr/yb,
  http://wwwsupelec.fr/employe/1515}
{prénom, http://wwwsupelec.fr/employe/1515, Yolaine}
{nom, http://wwwsupelec.fr/employe/1515, Bourda}
{mel, http://wwwsupelec.fr/employe/1515,
  Yolaine.Bourda@supélec.fr}
  
```

2.3. Syntaxe de RDF

Le modèle de données RDF fournit un cadre abstrait et conceptuel pour définir des métadonnée. Une syntaxe concrète est nécessaire pour créer et échanger ces métadonnées. Actuellement, RDF utilise XML comme syntaxe mais un autre choix est possible. RDF a aussi besoin des « espaces de noms XML » pour pouvoir associer chaque propriété au schéma qui la définit.

La spécification du W3C définit deux syntaxes pour encoder XML : la « syntaxe de sérialisation », régulière et la « syntaxe abrégée » qui permet une forme plus compacte. Les deux syntaxes doivent être acceptées par les interpréteurs XML.

Dans la suite, les exemples que nous donnons utilisent la syntaxe régulière qui paraît plus claire.

Ainsi, la phrase « Yolaine Bourda est l'auteur de la ressource <http://wwwsu.supelec.fr/~yb> » peut être représentée en RDF/XML par :

```
<?xml version="1.0" encoding="iso-8859-1" ?>
<rdf:RDF
  xmlns="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:edu="http://eduml.org/schema/">
  <rdf:Description
    about="http://wwwsu.supelec.fr/~yb">
    <edu:Auteur>Yolaine Bourda</edu:Auteur>
  </rdf:Description>
</rdf:RDF>
```

La phrase « L'auteur de la ressource <http://wwwsu.supelec.fr/~yb> a pour nom Bourda, pour prénom Yolaine, pour mel Yolaine.Bourda@supelec.fr et pour identificateur <http://wwwsu.supelec.fr/employe/1515> » peut être représentée en RDF/XML par :

```
<?xml version="1.0" encoding="iso-8859-1" ?>
<rdf:RDF
  xmlns="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:edu="http://eduml.org/schema/">
  <rdf:Description about="http://wwwsu.supelec.fr/~yb">
    <edu:Auteur
      rdf:resource="http://wwwsu.supelec.fr/employe/1515"/>
  </rdf:Description>
  <rdf:Description
    about="http://wwwsu.supelec.fr/employe/1515">
    <edu:PreNom>Yolaine</edu:PreNom>
    <edu:Nom>Bourda</edu:Nom>
    <edu:Mel>Yolaine.Bourda@supelec.fr</edu:Mel>
  </rdf:Description>
</rdf:RDF>
```

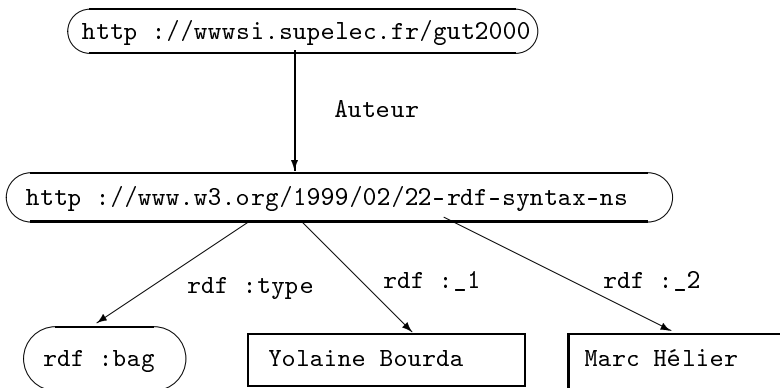
2.4. Les collections en RDF

Il arrive fréquemment que, dans une description, on doive se référer à des collections de ressources, comme dans la phrase : « les auteurs de <http://wwsi.supelec.fr/gut2000> sont Yolaine Bourda et Marc Héliér ». Ceci peut se faire en RDF en utilisant les trois types suivants de collections :

- La séquence (Sequence) : une liste ordonnée de ressources ou de littéraux. On l'utilise pour indiquer qu'une propriété a plusieurs valeurs et que l'ordre des valeurs est important. On peut dupliquer les valeurs.
- Le sac (Bag) : une liste non ordonnée de ressources ou de littéraux. On l'utilise pour indiquer qu'une propriété a plusieurs valeurs et que l'ordre des valeurs est sans importance. On peut dupliquer les valeurs.
- l'alternative (Alternatives) : une liste de ressources ou de littéraux qui représentent les différentes possibilités pour la valeur (unique) d'une propriété.

Pour représenter une collection, il faut utiliser une ressource additionnelle identifiant la collection et, en lui donnant une propriété « type ». Le lien entre la collection et ses constituants est modélisé par des propriétés définies pour cet usage et de noms « `_1` », « `_2` », « `_3` », etc.

Ainsi, la propriété précédente pourra être représentée par le graphe :



En RDF/XML, on utilise la notation « li » pour éviter d'avoir à numéroter explicitement les éléments de la collection. Le graphe précédent peut être sérialisé par :

```

<?xml version="1.0" encoding="iso-8859-1" ?>
<rdf:RDF
  xmlns="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:edu="http://eduml.org/schema/">
  <rdf:Description
    about="http://wwsi.supelec.fr/gut2000">
    <edu:Auteur>
      <rdf:bag>

```

```

    <rdf:li>Yolaine Bourda</rdf:li>
    <rdf:li>Marc Hélier</rdf:li>
  </rdf:bag>
</edu:Auteur>
</rdf:Description>
</rdf:RDF>

```

2.5. Assertions sur des assertions

Non seulement RDF permet de faire des assertions sur des ressources, mais il permet aussi de faire des assertions sur des assertions.

Considérons la phrase suivante : « Marc Hélier dit que Yolaine Bourda est l'auteur de la ressource `http://wwsi.supelec.fr/yb` ». Elle ne dit rien sur la ressource mais sur une assertion faite par Marc Hélier. Pour pouvoir exprimer des faits de ce genre en RDF, il faut modéliser l'assertion de départ comme une ressource ayant quatre propriétés. Ce processus est nommé « réification ». Un modèle d'assertion est nommé « assertion réifiée ».

Les propriétés à définir pour une assertion que l'on cherche à modéliser sont les suivantes :

- Sujet (subject) : la ressource décrite par l'assertion dans l'exemple : `http://wwsi.supelec.fr/yb` ;
- Prédicat (predicate) : la propriété décrite par l'assertion dans l'exemple : `auteur` ;
- Objet (object) : la valeur de la propriété de l'assertion dans l'exemple : Yolaine Bourda ;
- Type (type) : le type de la nouvelle ressource. Toutes les assertions réifiées sont des instances de `rdf:Statement` : elles ont donc une propriété `type` dont l'objet est `rdf:Statement`.

Une ressource, avec les quatre propriétés précédentes, représente l'assertion de départ et peut donc être utilisée comme objet d'autres assertions. Elle peut aussi avoir d'autres assertions faites sur elle. Une telle ressource est un modèle de l'assertion de départ, elle n'est pas l'assertion elle-même. Une assertion et l'assertion réifiée correspondante peuvent, toutes les deux, être dans un graphe RDF indépendamment l'une de l'autre.

Ainsi, l'assertion suivante (notée sous forme de triplets) :

```
{auteur, http://wwsi.supelec.fr/yb, Yolaine Bourda}
```

peut être réifiée de la façon suivante (YB représente l'identificateur de la nouvelle ressource) :

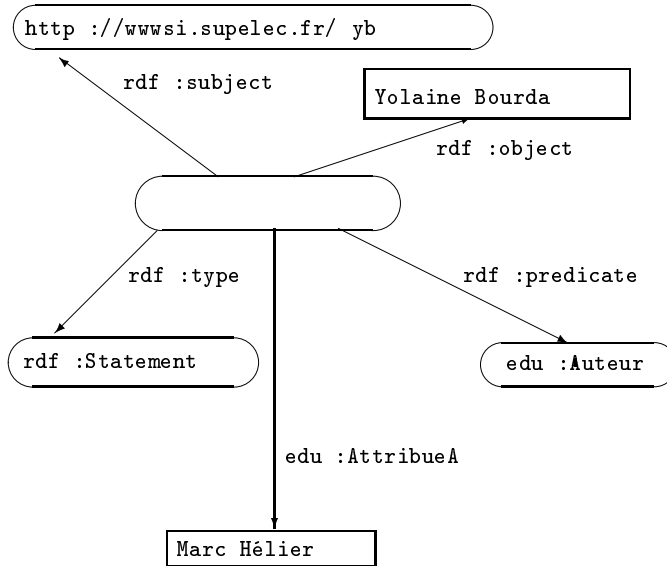
```

{type, YB, rdf:statement}
{predicate, YB, auteur}
{subject, YB, http://wwsi.supelec.fr/yb}
{object, YB, Yolaine Bourda}

```

Pour modéliser complètement notre exemple ci-dessus, on peut attacher à l'assertion réifiée une propriété (AttribueA) ayant la valeur appropriée (Marc Héliier).

Le graphe correspondant est :



et la forme RDF/XML est :

```

<?xml version="1.0" encoding="iso-8859-1" ?>
<rdf:RDF
  xmlns="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:edu="http://eduml.org/schema/">
  <rdf:Description>
    <rdf:subject
      resource="http://wwsi.supelec.fr/~yb"/>
    <rdf:predicate
      resource="http://eduml.org/schema/auteur"/>
    <rdf:object>Yolaine Bourda</rdf:object>
    <rdf:type resource
      ="http://www.w3.org/1999/02/22-rdf-syntax-ns#Statement"/>
    </rdf:AttribueA>Marc Héliier</rdf:AttribueA>
  </rdf:Description>
</rdf:RDF>
  
```

2.6. Dublin Core

Le Dublin Core fournit un guide [3] sur la façon d'exprimer son ensemble de métadonnées en RDF.

En l'utilisant, on peut décrire une page web de la façon suivante :

```
<rdf:RDF
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:dc="http://purl.org/metadata/dublin_core#"
<rdf:Description about="http://wwsi.supelec.fr/yb.html">
<dc>Title>Yolaine Bourda</dc>Title>
<dc>Description>
    Ceci est la page personnelle de Yolaine Bourda
</dc>Description>
<dc:Publisher>Supélec</dc:Publisher>
<dc>Date>1999-10-09</dc>Date>
<dc:Subject>
<rdf:Bag>
    <rdf:li>Enseignement</rdf:li>
    <rdf:li>XML</rdf:li>
    <rdf:li>Bases de données</rdf:li>
</rdf:Bag>
</dc:Subject>
<dc>Type>World Wide Web Home Page</dc>Type>
<dc:Format>text/html</dc:Format>
<dc:Language>fr</dc:Language>
</rdf:Description>
</rdf:RDF>
```

3. Applications pédagogiques

Les travaux autour des métadonnées pédagogiques sont actuellement en pleine expansion. Certains sont issus des milieux académiques, d'autres des milieux industriels, certains sont motivés par la recherche et l'altruisme, d'autres par des motifs plus marchands.

Des travaux sont en cours actuellement au sein du Dublin Core pour étendre l'ensemble de base par des éléments spécifiques au domaine éducatif. Ils ne sont pas encore suffisamment avancés pour que la structure se dégage de façon claire.

ARIADNE (Alliance of Remote Instructional Authoring and Distribution Networks for Europe) et IMS (Instructional Management Systems), deux consortiums impliqués depuis quelques années dans le développement d'outils et de métadonnées pédagogiques, ont décidé de collaborer pour définir un ensemble commun de métadonnées. Tous deux sont profondément impliqués dans le processus de standardisation du « Educational Metadata Learning Technology Standards Committee » des IEEE. Ils

convergent vers un standard unique et général, applicable à un très grand nombre de situations éducatives.

Ce standard est construit au-dessus du Dublin Core en le complétant par des extensions propres au domaine éducatif. Il spécifie la syntaxe et la sémantique des métadonnées pédagogiques et définit les attributs nécessaires pour une description adéquate et complète des « objets pédagogiques ». Ce standard de métadonnées pédagogiques, dit LOM (Learning Object Metadata), se limite à l'ensemble minimal de caractéristiques indispensables pour gérer les « objets pédagogiques », les rechercher et les évaluer.

Le LOM définit, dans sa version provisoire, neuf catégories de descripteurs :

General	caractéristiques indépendantes du contexte comme Identifier (un identificateur global unique) ou Title (le nom de la ressource) ou Language (la langue utilisée principalement par la ressource pour communiquer avec l'utilisateur) ;
LifeCycle	caractéristiques relatives au cycle de vie, comme Version ou Status (Draft, Final, Revised, Unavailable) ;
Meta-metadata	caractéristiques de la description elle-même comme Contribute (personnes ayant participé à l'élaboration des métadonnées) ;
Technical	caractéristiques techniques comme Format (du logiciel nécessaire pour accéder à la ressource) ;
Educational	caractéristiques pédagogiques Interactivity Type : le type d'interaction entre la ressource et l'utilisateur typique (Active, Expositive, Undefined) ;
Learning Resource Type	le type pédagogique (Exercise, Simulation...) ;
InteractivityLevel	degré d'interactivité ;
SemanticDensity	densité sémantique (Very Low, Low, Medium, High, Very High) ;
Intended end user role	utilisateur de la ressource ;
Context	environnement d'utilisation de la ressource ;
Typical Age Range	âge de l'utilisateur ;
Difficulty	difficulté de la ressource ;
Typical Learning Time	temps approximatif ou typique pour travailler avec la ressource ;
Description	commentaires sur l'utilisation de la ressource ;
Language	la langue de l'utilisateur ;
Rights	caractéristiques exprimant les conditions d'utilisation comme cost (ressource payante ou non) ;

Relation	caractéristiques exprimant les liens avec d'autres ressources comme <i>kind</i> (nature de la relation) ;
Annotation	commentaires sur l'utilisation pédagogique de la ressource ;
Classification	caractéristiques de la ressource décrites par des entrées dans des systèmes de classification.

Le LOM n'étant pas encore finalisé, il n'existe pas encore de recommandations officielles sur la façon dont il doit être exprimé en RDF, mais on trouve déjà des versions « officieuses » du LOM en RDF.

4. Conclusion

Le recours aux métadonnées apparaît aujourd'hui comme l'une des meilleures solutions pour rationaliser l'utilisation du Web et en tirer un profit plus satisfaisant, pour la plupart des communautés d'utilisateurs de ressources.

Dans l'enseignement et la formation, la valeur est de moins en moins dans le contenu lui-même, mais de plus en plus dans la capacité de rechercher ce contenu, de le trouver et d'en assembler des fragments pour apporter des informations pertinentes et une aide efficace aux utilisateurs, enseignants ou étudiants : pour cela, les métadonnées sont indispensables, car elles permettent, comme on l'a vu, de donner du sens (contenu sémantique) à une quantité considérable, et auparavant inexploitable, d'informations présentes sur le Web. Pour cette raison, les métadonnées portent en elles la capacité de transformer radicalement la manière de produire, de gérer et d'utiliser des documents pédagogiques.

Néanmoins, leur succès potentiel reste très dépendant de la définition, de la diffusion et de l'adoption de standards, requis pour parvenir à une réelle interopérabilité dans la création, l'édition, l'assemblage, l'échange et la diffusion des contenus. Les travaux entrepris dans le domaine des métadonnées pédagogiques par le groupe ARIADNE (Europe) et les IEEE (États-Unis) montrent qu'il est aujourd'hui possible de converger vers un standard unique, construit au-dessus du Dublin Core. Parallèlement, l'apport de RDF, associé à XML, semble décisif pour la mise en œuvre d'un tel standard car il devrait faciliter le développement et la diffusion des outils nécessaires en proposant une technique de description unifiée des métadonnées.

Cet important effort de standardisation ne peut que déboucher sur une métamorphose des documents électroniques avec, pour conséquence, une forme de (r)-évolution pédagogique. On peut en attendre qu'elle libère les enseignants d'une partie des contraintes liées à la production de leurs documents et qu'elle les aide à en améliorer le contenu et la qualité, qu'elle simplifie la mise en place de formations à distance, via le Web et donc tout au long de la vie, mais on peut pressentir aussi qu'elle sera l'alliée de l'industrialisation des activités de formation, déjà significative aux États-Unis.

5. Références

- [1] Tim BERNERS-LEE, *Metadata architecture*, 1997. <http://www.w3.org/DesignIssues/Metadata.html>
- [2] *Dublin core metadata initiative*. <http://purl.org/dc/>
- [3] *Guidance on expressing the dublin core within the resource description framework (rdf)*. <http://www.ukoln.ac.uk/metadata/resources/dc/datamodel/WD-dc-rdf/>
- [4] Yolaine BOURDA et MARC HÉLIER, *Eduml : un langage pour la description de ressources pédagogiques*. In NTICF98, Rouen, poster, 1998.
- [5] *IEEE p1484.12 learning objects metadata working group*. <http://ltsc.ieee.org/wg12>
- [6] O. LASSILA, Web metadata : A matter of semantics, *IEEE Internet Computing*, July-August :30-37, 1998.
- [7] Steve LAWRENCE and C. LEE GILES, Accessibility of information on the web, *Nature*, 400 :107-109, 1999.
- [8] *XML, extensible markup language*. <http://www.w3.org/XML>
- [9] *Metadata at W3C*. <http://www.w3.org/Metadata/>
- [10] *W3C resource description framework*. <http://www.w3.org/RDF/>
- [11] *RDF schema specification, March 1999*. <http://www.w3.org/TR/PR-rdf-schema>
- [12] *RDF model and syntax specification, February 1999*. <http://www.w3.org/TR/REC-rdf-syntax>