

Chapitre II

Notions de base

- 1 URI
- 2 XML
- 3 Espace de noms XML
- 4 XSLT

Chapitre II

Notions de base

- 1 URI
 - Définition
 - URIref
- 2 XML
- 3 Espace de noms XML
- 4 XSLT

URI

Uniform Resource Identifier

Définition (URI)

Une URI est une chaîne qui identifie une ressource ou un concept.

RFC 3986 :

<http://www.gbiv.com/protocols/uri/rfc/rfc3986.html>

Identifie (de façon unique) « tout ce dont il peut être question » sur le web.

URI

Uniform Resource Identifier

Exemple

- <http://www.info.univ-angers.fr>
Les URL (Uniform Resource Locator) sont des URI qui permettent en plus de localiser (d'accéder) à la ressource.
ftp, mailto, gopher
- `uuid:04b749bf-3bb2-4dba-934c-c92c56b709df`
Création d'un identificateur unique de façon totalement décentralisée : adresse MAC + aléatoire
- etc. Tout le monde peut créer des URI et des nouveaux schémas d'URI.

URI

Donc URI = URL ?

Non, car il peut être question de ressources qui ne sont pas disponibles (« localisables ») sur le web.

Exemple

URI pour une personne ?

- URL à partir de l'adresse e-mail
`mailto:gates@microsoft.com`
Pas satisfaisant : la même personne peut avoir plusieurs adresses (ou pas). Une personne *n'est pas* une adresse e-mail. Une personne *possède* une adresse e-mail (+ adresse postale, téléphone, etc).
- `http://www.microsoft.com/bgates` peut être une URI :
Une URI n'est pas obligatoirement un moyen de localiser une ressource.

URI et URIref

En fait, RDF n'utilise pas des URI mais des **URIref** (*URI references*).

Définition (URIref)

Une URIref est une URI accompagnée d'un identificateur de fragment, préfixé par #.

Exemple

`http://www.microsoft.com/directory#bgates`

Comme une URL, une URI peut être absolue ou relative. Par exemple `products` ou `#sballmer`.

Il peut n'y avoir aucun lien particulier entre plusieurs URIref de la même URI.

Chapitre II

Notions de base

- 1 URI
- 2 XML
 - Définition
 - Document bien formé, document valide
- 3 Espace de noms XML
- 4 XSLT

XML

La plupart des langages du web sémantique utilisent une notation XML.

XML (*eXtensible Markup Language*) permet à chacun de définir facilement son propre format de document, d'écrire des documents à ce format, et d'utiliser des outils logiciels pour faciliter l'utilisation de ce format.

Recommandation du W3C : <http://www.w3.org/TR/REC-xml>

- Les documents XML contiennent du texte...
- ...structuré à l'aide de balises.
- Contrairement à HTML, XML permet de définir de nouvelles balises, qui ne sont pas des balises de mise en forme.
- Un format de formats pour faciliter le partage de documents.

XML

Facilite l'interprétation par une machine :

Exemple

Je fais le cours de web sémantique aujourd'hui.



```
<phrase>
  <enseignant nom="D. Genest">Je</enseignant>
  fais le cours de
  <cours>web semantique</cours>
  <date j="18" m="9">aujourd'hui</date>.
</phrase>
```

On peut automatiser la détection d'une phrase, le Je repère un enseignant, et le nom du cours.

Possibilité d'ajouter des informations (*attributs*) aux éléments, un attribut étant composé d'un nom et d'une valeur.

XML

Document bien formé et document valide

Un document XML doit être *bien formé* : respect de conditions syntaxiques (tout élément ouvert doit être fermé, emboîtement des éléments, etc.).

Possibilité de définir des contraintes supplémentaires sur le format :

Un document peut faire référence à une DTD (*Document Type Declaration*) qui définit la structure du document : intitulés des éléments, comment ils peuvent s'emboîter, attributs autorisés.

Un document qui, en plus d'être bien formé, respecte sa DTD est dit *valide*.

XML

Exemple (note12.xml)

```
<?xml version="1.0" encoding="ISO-8859-1" ?>
<!DOCTYPE note SYSTEM "note.dtd">
<note priorite="urgent" id="n12">
  <dest>Caroline</dest> <exp>Roger</exp>
  <msg>Coucou</msg>
</note>
```

Exemple (note.dtd)

```
<?xml version="1.0" encoding="ISO-8859-1" ?>
<!ELEMENT note (dest*, exp, msg)>
<!ATTLIST note priorite (normal | urgent)
  id CDATA #REQUIRED>
<!ELEMENT dest (#PCDATA)>
<!ELEMENT exp (#PCDATA)>
<!ELEMENT msg (#PCDATA)>
```

Chapitre II

Notions de base

- 1 URI
- 2 XML
- 3 Espace de noms XML
- 4 XSLT

Espace de noms XML

Problème. Plusieurs formats basés sur XML peuvent utiliser les mêmes noms d'éléments pour repérer des concepts différents (exemple : note à un examen). Il est dans ce cas impossible d'effectuer un traitement faisant intervenir des documents utilisant différents formats.

Définition (Espace de noms XML)

Un espace de noms XML (*XML namespace*) identifie une partie du web qui est utilisée pour identifier un ensemble de noms (d'éléments, d'attributs).

Recommandation W3C :

<http://www.w3.org/TR/REC-xml-names>

Un espace de nom est une URI qui est utilisée pour préfixer un nom.

Espace de noms XML

Exemple (note12.xml)

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<mn:note xmlns:mn="http://www.mynoteapp.org/"
  mn:priorite="urgent" mn:id="n12">
  <mn:dest>Caroline</mn:dest> <mn:exp>Roger</mn:exp>
  <mn:msg>Coucou</mn:msg>
</mn:note>
```

Dans un élément XML, on peut déclarer un espace de nom `identificateur=URI`. Dans cet élément et les éléments emboîtés, l'identificateur pourra être utilisé pour préfixer des noms : `identificateur:nom`.

`xmlns=URI` peut être utilisé pour définir l'espace de nom par défaut, qui est appliqué aux noms qui ne sont pas préfixés.

Chapitre II

Notions de base

- 1 URI
- 2 XML
- 3 Espace de noms XML
- 4 XSLT

XSLT

XML permet de représenter des données sous une forme facilement manipulable par des machines. Mais un document XML n'est pas lisible par un navigateur.

Recommandation W3C : <http://www.w3.org/TR/xslt>

XSL (eXtensible Stylesheet Language) permet de définir des mécanismes de transformation et de présentation de documents XML.

XSLT (XSL Transformation) permet de transformer un document XML en document XML. Par exemple XHTML.

Donc XML peut être utilisé comme langage de base pour le web sémantique ?

- XML pour les machines
- XML + XSLT → XHTML pour les humains

Non

XML + XSLT = Web sémantique ?

Problème. Des données peuvent être représentées de façons très différentes en XML.

Exemple

```
<note prority="urgent" id="n12">
  <dest>Caroline</dest> <exp>Roger</exp>
  <msg>Coucou</msg>
</note>
```

et

```
<messages exp="Roger">
  <dest nom="Caroline">
    <msg priority="urgent">
      <id>n12</id> <texte>Coucou</texte>
    </msg>
  </dest>
</messages>
```

représentent les mêmes données : le message n12 a pour expéditeur Roger, pour destinataire Caroline et pour texte Coucou.

XML + XSLT = Web sémantique ?

Comment comparer des documents XML écrits sur différentes DTD ?...

... En utilisant XSLT.

Oui, mais. XSLT est Turing complet, et compliqué à utiliser...

XML est bien adapté pour décrire des **documents** : structure particulière et complexe.

Pour le Web sémantique, les données à manipuler sont « simples » : sujet - prédicat - objet.

Fusionner deux documents ?

XML : Complexe. Alors qu'une union de triplets est bien plus simple.