

XML

outils de publication

Stéphane Bonhomme
Conférence Guilde
5 novembre 2003

Sommaire

- Introduction au langage XML
- Validité des documents – schemas XML
- Espace nominaux
- Les dialectes XML
- Applications
- DOM – Xpath - XSLT
- Tiwepen
- Démarche XML

Introduction

- XML (Extensible Markup Language)
- Recommandation W3C depuis 1998
- Successeur de SGML (html)

- Langage de représentation de données structurées. Support fichier textuels

Introduction

- XML est utilisable sans difficultés
- Permet de généraliser les outils d'exploitation
- Facilité de création
- Extensibilité
- Supporte tous les encodages de caractères

Exemple

```
<?xml version="1.0" encoding="utf-8" ?>
<!-- un exemple de compte -->
<compte monnaie="euro">
  <numéro>123456KL</numéro>
  <client>
    <nom>Dupont</nom>
    <prénom>Michel</prénom>
  </client>
  <solde>1683.25</solde>
  <opération id="789456125846">
    <date format="iso-xxx">03262000</date>
    <montant>+200</montant>
  </opération>
</compte>
```

Balises

Un document xml est composé de contenu textuel, encadré par des balises (tags) imbriqués

- Les balises sont nommées (GI)

Attributs

- liste de couples **nom="valeur"**
- Présents dans la balise de début d'éléments

On appelle élément la représentation objet d'une balise

Balises

Balise de début d'élément

<compte monnaie="euro">

Balise de fin d'élément

</compte>

Balise d'élément vide

<compte monnaie="euro"/>

L'imbrication des balises de début et de fin forme une chaîne bien parenthésée

<a> est correct

<a> est incorrect

Validité des documents

- Nom des balises, ordre d'apparition dans le document peuvent être contraintes par des règles
- Exprimées dans un schema XML ou une DTD
- Référencées dans l'en-tête du document
- Un document respectant les règles de structure qui lui sont associées est appelé un document valide

XML schema

- Définition des éléments pouvant apparaître dans l'ensemble des documents le référençant
- Un élément est décrit par :
 - les attributs qu'il peut porter
 - le modèle de contenu
 - éléments fils possibles
 - ordonnancement des fils
- Définition de types de base et types construits

XML schema - exemple

```
<?xml version="1.0" encoding="UTF-8"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema">
  <xs:element name="compte">
    <xs:complexType>
      <xs:sequence>
        <xs:element ref="numéro"/>
        <xs:element name="client" type="clientType"/>
        <xs:element ref="solde"/>
        <xs:element name="opération" type="opérationType"/>
      </xs:sequence>
      <xs:attribute name="monnaie" type="xs:string"
use="required"/>
    </xs:complexType>
  </xs:element>
  <xs:complexType name="opérationType">
    <xs:sequence>
      <xs:element name="date" type="dateType"/>
      <xs:element ref="montant"/>
    </xs:sequence>
    <xs:attribute name="id" type="xs:double" use="required"/>
  </xs:complexType>
  .../...
</xs:schema>
```

Espaces nominaux

- Préfixes sur les balises et noms d'attributs
- Permet d'identifier à quel dialecte appartiennent les balises dans les documents composites

```
<html:p>
```

```
  <math:math>
```

```
    <math:mfrac>....</math:frac>
```

```
  </math:math>
```

```
</html:p>
```

Documents composites

- Regroupent des balises de plusieurs dialectes
- Les namespaces sont déclarés dans une balise englobant leur utilisation

```
<html xmlns="http://www.w3.org/1999/xhtml"
xmlns:math="http://www.w3.org/1998/Math/MathML">
  .../...
  <math:math>
    <math:mfrac>....</math:frac>
  </math:math>
  .../...
</html:p>
```

« Dialectes » xml contenu

Xhtml

- 1.0 : représentation XML de html,
- 1.1 : modularisation

Docbook

- Format pour livres, articles
- Table des matières, préface, parties, chapitres, sections, annexes, références, glossaire, bibliographie, index.
- Modèle de tableaux CALS

« Dialectes » xml contenu

Xforms

- Nouvelle génération de formulaires pour les applications web.

Xlink

- Représentation de liens hypertexte et hyperdocuments

SVG

- Graphiques vectoriels en 2D

« Dialectes » xml contenu

SMIL

- Animation et scènes temporisées

MathML

- Représentations de formules mathématiques

« Dialectes » xml data

RDF

- Catalogues structurés de métadonnées

RSS

- Diffusion de canaux d'information, weblogs.

SOAP

- Contenus de messages d'une surcouche du protocole http
- webservices

EbXML

- Transaction commerciales, facturation.

Applications

OpenOffice.org, Gnumeric, Abiword

- Dialectes xml comme formats de documents bureautiques.

Glade, XUL

- Constructeurs d'interfaces graphiques

Sodipodi, Batik

- SVG : éditeurs, bibliothèque (apache)

Limsee2

- Atelier Multimédia

XML dans openoffice.org

- Format de stockage de fichiers sxi
- Réutilisation de formats standards
 - Svg, xsl (fo), xlink
- Projets de filtres XSLT
 - Docbook <-> sxi
 - Sxi -> xhtml

Glade - libglade

- Générateur d'interface GTK
- Produit un fichier xml de description de l'interface
- Libglade : liée à un programme, permet de lire le fichier xml et de construire dynamiquement l'interface utilisateur

XUL

- Générateur d'interface de mozilla
- Formulaires avancés :
 - Widgets fenêtres, menus, barres d'outils, onglets, barres de défilement, progressomètres...
- Interopérable avec xhtml et css

DOM

- API de manipulation de structures XML
- Le document xml est représenté sous forme d'une hiérarchie d'objets **node**.
- Interfaces **node, document, element, attribute, text-element**
- Méthodes : accès, insertion et modification des objets
- Disponible pour tous langages

XPath

- Langage d'expressions permettant de désigner des ensembles d'éléments
- Si on considère de document xml comme un arbre de noeuds, Xpath permet de désigner un sous-arbre

`//olist/item`

`/doc/chapter[5]/section[2]`

`employee[@secretary and @assistant]`

`chapter[title="Introduction"]`

XSLT

- Langage permettant l'écriture de processus de transformation de données xml
- Adapter les flux de données lors de la communication entre 2 applications
- Syntaxe xml, xpath
- S'intègre parfaitement dans la logique flux de données et processus (workflow – qualité)

Utilisation de XSLT

- Coté serveur
 - Adaptation de données
 - Liaison avec bd : exslt
 - Services web, rpc
 - Libxml2, sablotron, xalan.
- Sur le client
 - Présentation : conversion de sources xml en html
 - Mozilla >1.4 µ\$ie > 5.5

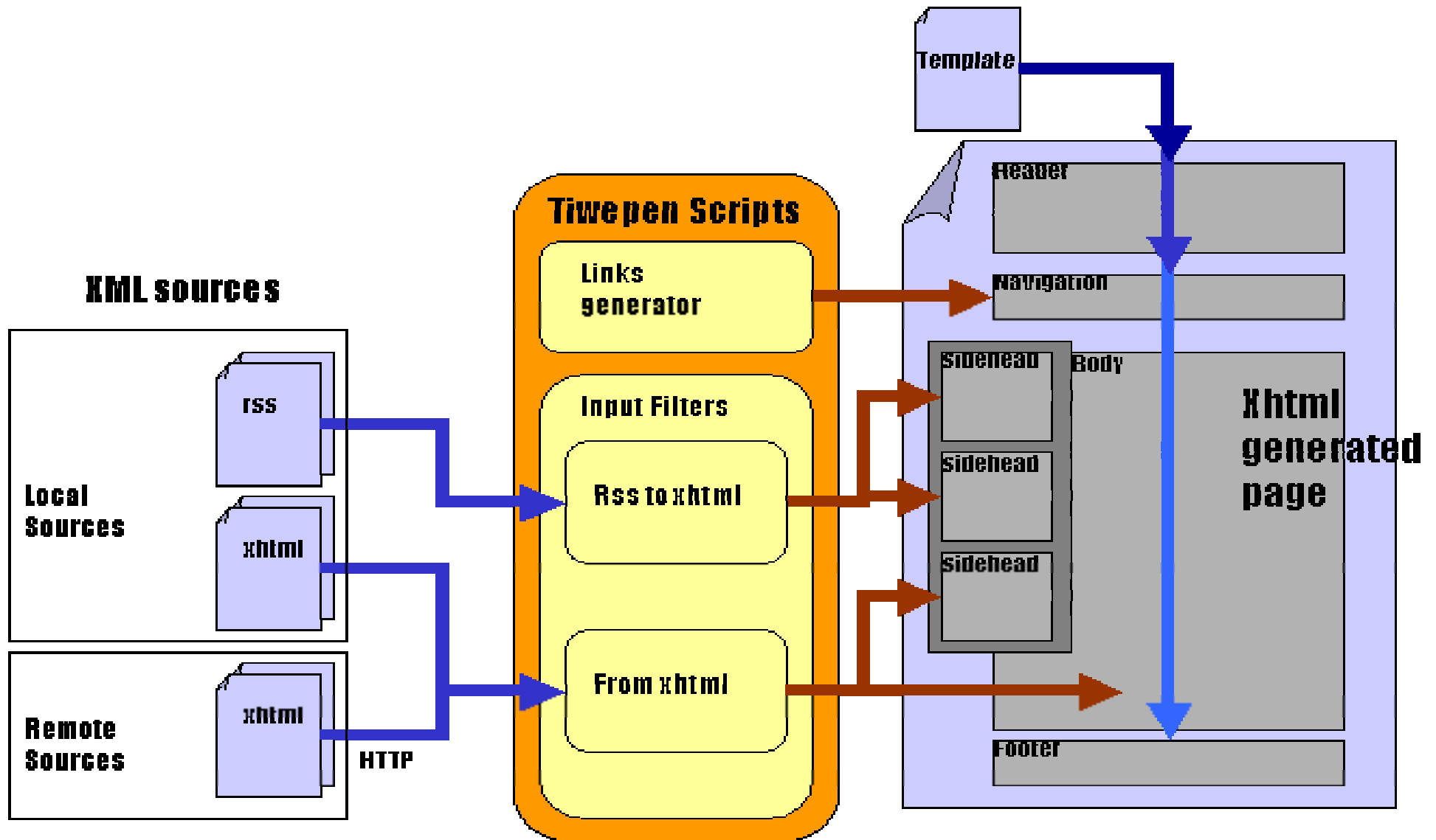
XSLT - exemple

```
<?xml version="1.0"?>
<xsl:stylesheet version="1.0"
  xmlns:xsl="http://www.w3.org/1999/XSL/Transform"
  xmlns="http://www.w3.org/1999/xhtml"
  <xsl:template match="compte">
  <div class="cmpt">
    <h2>compte ( <xsl:value_of select="@monnaie"/> )</h2>
    <xsl:apply-templates/>
  </div>
</xsl:template>
<xsl:template match="client">
  <p><xsl:value_of select="nom"/>
    <xsl:value_of select="prénom"/></p>
</xsl:template>
</xsl:stylesheet>
```

Tiwepen

- Un système de publication web écrit en xslt
- Composition de pages xhtml avec intégration de flux XML (rss)

Tiwepen : principe général



Démarche XML

- Pourquoi utiliser XML ?
 - Besoin d'un format de stockage de données
 - Communication avec d'autres SI
 - Intégration multi-formats
- Quel format utiliser ?
 - Après analyse des besoins :
 - Recherche w3c oasis.org si standard existant
 - Définition des schémas ou spécialisations

Démarche XML

- Définition des process xslt
 - Protocoles et adaptation des flux de données
- Développement des composants
 - Nouveaux composants : dom
 - Composants existants : définir des interfaces (soap/xmlrpc)
- Toujours penser inter-opérabilité
- Toujours penser réutilisation

Références WWW

- Site w3c: normes, applications, tutoriels
 - <http://www.w3.org/>
- Tutoriels
 - <http://zvon.org>, <http://w3schools.com>
- la bible : applications, articles...
 - <http://xml.coverpages.org/>

Références WWW

- **Applications**

- **Openoffice.org** :

- <http://xml.openoffice.org/>

- **Gnumeric** :

- <http://www.gnome.org/projects/gnumeric/index.shtml>

- **Abiword** : <http://www.abisource.com/>

- **Glade** :

- <http://developer.gnome.org/doc/API/libglade/gladexml.html>

Références WWW

- **Applications**

- **XUL :**

- <http://www.mozilla.org/projects/xul/joy-of-xul.html>

- **Sodipodi :** <http://sodipodi.sourceforge.net/>

- **Batik :** <http://xml.apache.org/batik/>

- **Limsee2 :**

- <http://wam.inrialpes.fr/software/limsee2/>

- **Tiwepen :**

- <http://waloo.homelinux.net/tiwepen>

Références WWW

- **Bibliothèques XSLT**

- **Libxml2** : <http://xmlsoft.org>

- **Xalan** : <http://xml.apache.org>

- **Sablotron** :

- http://www.gingerall.com/charlie/ga/xml/p_sab.xml

- **Portails, info**

- xmlfr.org, xml.com, waloo.homelinux.net ;)