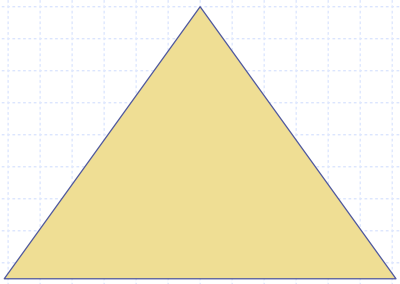
XML et BD

J. Guyot
ISI
Université de Genève

ISI MEDIA

Paradigme de la mobilité

DONNEE-BD



DOCUMENT-XML

CODE-JAVA

Généraliser la notion de documents

ISI MEDIA

I : Posons le problème

- Structuré / non Structuré
- Données / Documents
- BD / XML
- Avantage /Désavantage
- Solutions
- Problèmes

ISI MEDIA

J. Guyot BD/XML 2001 ISI - Université de Genève 3

Structuré versus Non structuré

structurée (BD)	**non-structurée(WEB)**
modélisable	indexable
domaine limité (nbr de catégories petit)	domaine illimité (nbr de catégories grand)
protocole connu (cycle de vie des objets)	pas de protocole explicité

ISI MEDIA

J. Guyot BD/XML 2001 ISI - Université de Genève 4

XML - semi structuré

- DTD ou schéma
  - Assure la modélisation du document
  - Nbr de catégories connues
  - Forme est modélisée
  - Contenu est non modélisé
- Cycle de vie des documents
  - Création
  - Conformité au schéma
  - Modification
  - suppression
- Assimilables à des données (mais différents!)

I S I

M E D I A

J. Guyot BD/XML 2001 ISI - Université de Genève 5

Problème des deux mondes

- Données structurées
  - + Existe des outils efficaces
  - - Rigidité et lourdeur de mise en œuvre
- Documents structurés
  - - Existe (pas encore) d'outils efficaces
  - + s'inscrit dans le paradigme de la MOBILITE
- Examinons les points forts

I S I

M E D I A

J. Guyot BD/XML 2001 ISI - Université de Genève 6

Base de données

- Problèmes bien gérés

- Interrogation
- Concurrence
- Maj de de masse
- Transaction
- Confidentialité
- Sécurité
- Robusteté
- Performance
- Gros volume

ISI MEDIA

J. Guyot BD/XML 2001 ISI - Université de Genève 7

Document XML

- Inscrit dans le paradigme de la Mobilité

- Standardisation du XML
- Publication des schémas (DTD sites de référence)
- Interopérabilité (schéma dépende pas d'un syst/lang)
- Communicabilité (données + schémas)
- Manipulation (XSLT)
- Édition (direct – Frame XML)
- Affichage (browser)

ISI MEDIA

J. Guyot BD/XML 2001 ISI - Université de Genève 8

Solution opportune:

- Profiter des deux mondes
- Pour les BD:Publier les données comme des doc. XML

BD -> Synthétiser -> XML

- Pour les doc XML :S'insérer dans un BD (modélisation)

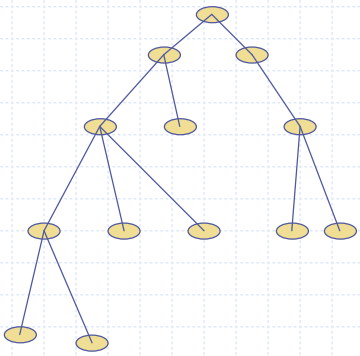
BD <- Stocker <- XML

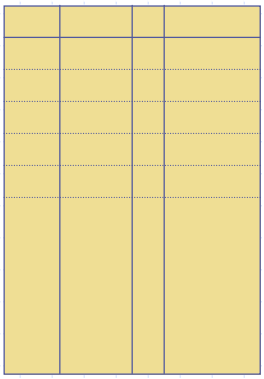
ISI MEDIA

J. Guyot BD/XML 2001 ISI - Université de Genève 9

Difficulté de la solution opportune

- XML - structure de l'arbre
- BD - structure de la table





ISI MEDIA

J. Guyot BD/XML 2001 ISI - Université de Genève 10

Autres solutions: la revanche des pointeurs ?

- Créer un système adhoc

- Stockage
- Indexation
- Résoudre tous les points BD

- Possible ? Oui
- Economiquement viable ? (peut être)
- Répétition des BD Objets ? Non (existe un marché)
- Compétion avec les acteurs SGBD ? certain

ISI

MEDIA

II: bref rappel sur XML

- Les objectifs annoncés pour XML dans la version 1.0 sont:
  - XML doit être directement utilisable pour Internet;
  - XML doit supporter un large éventail d'applications;
  - XML doit être compatible avec SGML;
  - Il doit être aisé d'écrire des programmes qui traitent des documents XML;
  - Le nombre de fonctions optionnelles dans XML doit être minimum;
  - Le nombre de fonctions optionnelles dans XML doit être minimum;
  - Le nombre de fonctions optionnelles dans XML doit être minimum;
  - Les documents XML doivent être humainement lisibles et raisonnablement clairs;
  - La conceptualisation de XML doit être formelle et concise;
  - La conceptualisation de XML doit être formelle et concise;
  - Les documents XML doivent simple à créer.

ISI

MEDIA

Un document XML

```
<?xml version="1.0" encoding='ISO-8859-1'?>
<atlas>
   <pays population="7.0"> Suisse
       <drapeau src="suisse.gif"/>
   </pays>
   <pays> France </pays>
   <pays> Autriche </pays>
</atlas>
```

- <élément> corps de l'élément </élément>
- <élément attribut="valeur"> … </élément>

ISI MEDIA

J. Guyot BD/XML 2001 ISI - Université de Genève 13

Document bien formé

- Un document est bien formé s'il respecte que:
  - Les balises sont fermées
  - L'imbrication est strictement arborescente
  - Les attributs sont obligatoirement entre guillemets
- Les documents HTML sont rarement bien formé!
- La conformité d'un document doit être testé par rapport à une DTD (ou schéma)

ISI MEDIA

J. Guyot BD/XML 2001 ISI - Université de Genève 14

DTD: standardiser les documents d'un domaine

```
<!-- DTD pour atlas-->
<!ELEMENT atlas (pays+)>
<!ELEMENT pays ( nom, capitale)>
<!ELEMENT nom (#PCDATA)>
<!ELEMENT capitale (#PCDATA)>
```

- La DTD type le document en :
- En circonscrivant les éléments possibles
- En fixant l'ordre, l'optionnalité, la répétition des éléments
- En typant les éléments

ISI

MEDIA

J. Guyot BD/XML 2001 ISI - Université de Genève 15

Valider un document par rapport à la DTD

```
<?xml version="1.0"  encoding='ISO-8859-1' ?>
<!DOCTYPE atlas SYSTEM "pays.dtd">
<atlas>
   <pays >
      <nom> Suisse </nom>
      <capitale> Berne </capitale>
   </pays>
   <pays >
      <nom> France </nom>
      <capitale> Paris </capitale>
   </pays>
   <pays >
      <nom> Autriche </nom>
      <capitale> Vienne </capitale>
   </pays>
</atlas>
```

ISI

MEDIA

J. Guyot BD/XML 2001 ISI - Université de Genève 16

Espace de noms : xmlns

```
<?xml version="1.0"  encoding='ISO-8859-1' ?>
<!-- ISI Information System Interface -->
<!DOCTYPE ISI:atlas SYSTEM "pays.dtd">
<ISI:atlas xmlns:ISI="http://cuiwww.unige.ch/isi">
   <ISI:pays >
      <ISI:nom> Suisse </ISI:nom>
      <ISI:capitale> Berne </ISI:capitale>
   </ISI:pays>
   <ISI:pays >
      <ISI:nom> France </ISI:nom>
      <ISI:capitale> Paris </ISI:capitale>
   </ISI:pays>
   <ISI:pays >
      <ISI:nom> Autriche </ISI:nom>
      <ISI:capitale> Vienne </ISI:capitale>
   </ISI:pays>
</ISI:atlas>
```

ISI

MEDIA

Feuille de style attachée au document: XSL

- Permettre de transformer un document XML en un autre document XML
- Description non-procédurale
- Règle basée sur le «matching» de patron dans le document XML
- La feuille XSL est aussi un document XML
- Applications
  - XML -> HTML
  - XML(DTD1) -> XML(DTD2)
  - XML -> XSL:FO -> PDF

ISI

MEDIA

Exemple: XSL

```
<xsl:stylesheet xmlns:xsl="http://www.w3.org/TR/WD-xsl">
    <xsl:template match="/">
            <HTML>
              <HEAD>
                <TITLE>Mon atlas ISI</TITLE>
              </HEAD>
              <BODY>
                <xsl:apply-templates/>
              </BODY>
            </HTML>
    </xsl:template>
    <xsl:template match="atlas">
          <h1> ATLAS ISI </h1>
           <xsl:apply-templates/>
    </xsl:template>
```

ISI MEDIA

J. Guyot BD/XML 2001 ISI - Université de Genève 19

Exemple: XSL (suite)

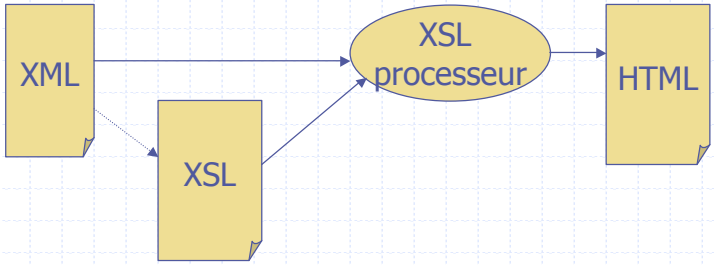
```
        <xsl:template match="pays">
      <xsl:apply-templates/>
                <br/>
    </xsl:template>
    <xsl:template match="nom">
      <b> <xsl:value-of select="."/> </b>
    </xsl:template>
    <xsl:template match="capitale">
      <xsl:value-of select="."/>
    </xsl:template>
</xsl:stylesheet>
```

ISI MEDIA

J. Guyot BD/XML 2001 ISI - Université de Genève 20

Document XML et son XSL

```
<?xml version="1.0"  encoding='ISO-8859-1' ?>
<!DOCTYPE atlas SYSTEM "pays.dtd">
<?xml-stylesheet type='text/xsl' href='pays.xsl'?>
<atlas>
   …
</atlas>
```



I S I

M E D I A

J. Guyot BD/XML 2001 ISI - Université de Genève 21

Exemple de résultat

```
<HTML>
      <HEAD>
         <TITLE>Mon atlas ISI</TITLE>
      </HEAD>
      <BODY>
          <h1> ATLAS ISI </h1>
          <b> Suisse </b>
           Berne
          </br>
          <b> France </b>
           Paris
          </br>
          <b> Autriche </b>
           Vienne
          </br>
      </BODY>
   </HTML>
```

I S I

M E D I A

J. Guyot BD/XML 2001 ISI - Université de Genève 22

III: bref rappel sur les BD

- BD relationnelles
- Schéma de la base
- Stockage – modèle de table
- Langage d'interrogation SQL (non procédural)
- Implémentation des structures de données (index, hash table, bit map)
- Modèle transactionnel
- Multi processing, parallèlisme
- Securité, confidentialité, performance …

ISI MEDIA

J. Guyot BD/XML 2001 ISI - Université de Genève 23

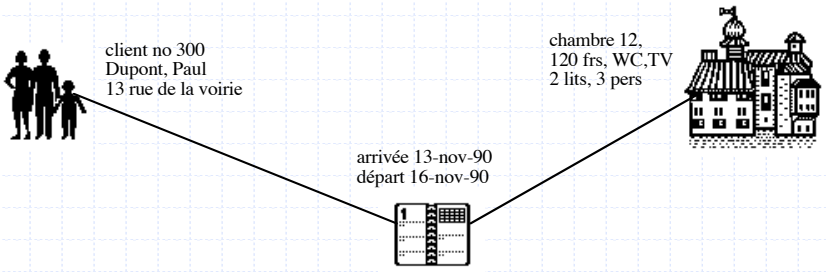
Schéma de la base

- modélisation du domaine (données / actions / contraintes)
- Catégorie d'objet (relation)
  - Propriété (attribut)
    - Type de valeurs (type)
- Processus de normalisation
  - Éviter les anomalies de mise à jour
  - Redondance des données
- Implémentation du modèle en SQL + struct des données

ISI MEDIA

J. Guyot BD/XML 2001 ISI - Université de Genève 24

Exemple d'un domaine



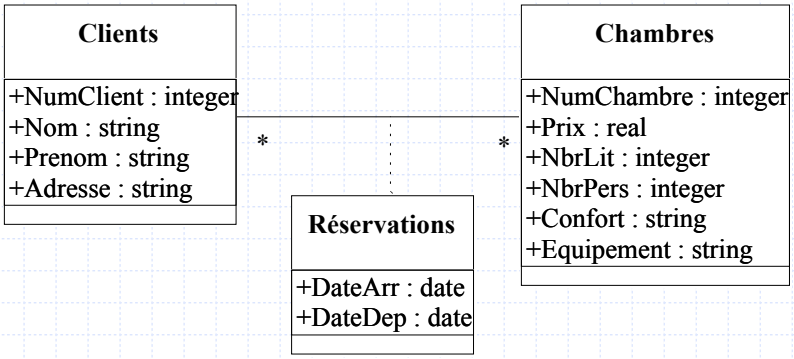
- Réservation dans un hotel
- Objets
  - Clients
  - Chambres
  - réservation

I S I

M E D I A

J. Guyot BD/XML 2001 ISI - Université de Genève 25

Exemple de schéma



- Schéma (notation UML)
  - 2 classes (clients - chambres)
  - 1 association n-n + info (réservation)

I S I

M E D I A

J. Guyot BD/XML 2001 ISI - Université de Genève 26

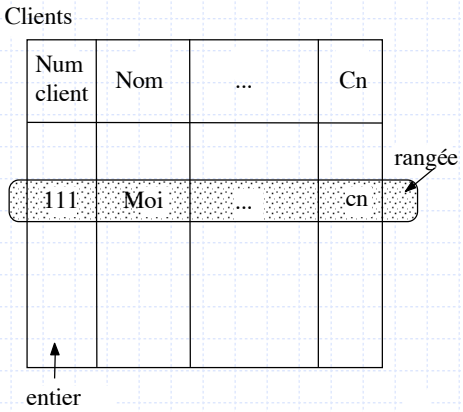
Schema SQL

- CREATE TABLE CLIENTS (
  NUM_CLIENT NUMBER (6) not null ,
  NOM CHAR (20),
  PRENOM CHAR (20),
  ADRESSE CHAR (40))
- CREATE TABLE CHAMBRES (
  NUM_CHAMBRE NUMBER (2),
  PRIX NUMBER (8,2),
  NBR_LITS NUMBER (1),
  NBR_PERS NUMBER (1),
  CONFORT CHAR(6),
  EQUIPEMENT CHAR(3))
- CREATE TABLE RESERVATIONS (
  NUM_CLIENT NUMBER (6),
  NUM_CHAMBRE NUMBER (2),
  DATE_ARR DATE,
  DATE_DEP DATE)

I S I

M E D I A

Modèle de table



- Table=relation
- Rangée=donnée
- Colonne=attribut

I S I

M E D I A

Interrogation avec SQL

- Basée sur une algèbre (selection,projection, composition, …)
- Standardisée avec SQL

```
SELECT C1, C2, ..., Cm
       FROM R1, R2, ... , Rn
       WHERE F
```

ISI MEDIA

Exemple d'interrogation: 1 table

- Les chambres coûtants au max. 80 francs ou ayant un bain et valant au max. 120

```
SELECT num_chambre, prix, confort
      FROM CHAMBRES
      WHERE  (prix<=80)
         OR ((confort='BAIN') AND (prix<=120))
```

NUM_CHAMBRE	PRIX	CONFOR
10	80	WC
20	80	WC
30	80	WC
40	80	WC
13	120	BAIN
23	120	BAIN
33	120	BAIN
43	120	BAIN

ISI MEDIA

Exemple d'interrogation: 2 tables + jointure

- Les noms des clients ayant réservés une chambre pour le 25-12-02 ?
- SELECT Nom
  FROM CLIENTS,RESERVATIONS
  WHERE Date_Arr<=to_date('25-dec-02')
  AND Date_Dep> to_date('25-dec-02')
  AND
  Clients.num_client=Reservations.num_client
- NOM
- ----------
- DUFOUR
- ZORO
- DUMAS
- ROMULUS

ISI J. Guyot BD/XML 2001 ISI - Université de Genève 31 MEDIA

Puissance et Limites SQL

- + opérateur de regroupement
- + non procédural
- + efficacité des implémentations

- - pas de récursion (parcours en arbre)
- - le résultat est toujours une table
- - l'ordre du résultat est fixé par le contenu (liste)

- Document XML (arbre ordonné)

ISI J. Guyot BD/XML 2001 ISI - Université de Genève 32 MEDIA

IV: Synthétiser du XML depuis une BD

- Pourquoi ?
- Interopérabilité: Echanger des données entre BD
  - Domaine = 1 DTD
  - Fédérer des BD ayant des schémas différents
- Afficher des données sur le WEB
  - XML (schéma + information)
  - XSL (transformation vers un média particulier)
- Exporter des données vers des applications
  - Rendre l'application indépendante du schéma de la BD
  - Eviter des drivers ODBC-XXXX

ISI

MEDIA

Idée la plus simple: utiliser les vues

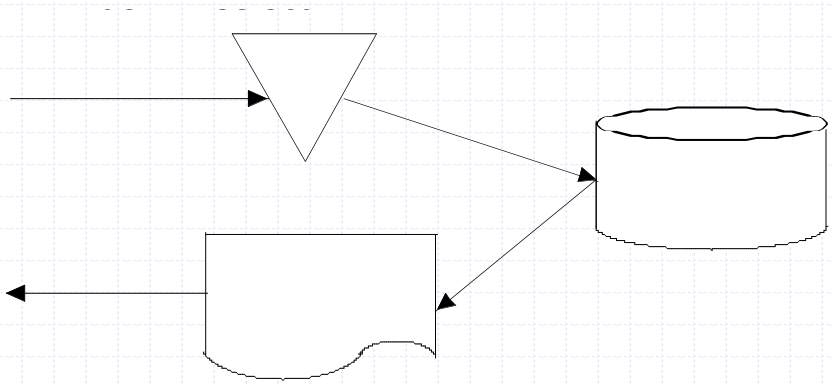
- Créer une vue au format du XML à générer
- Marche très bien pour une table
- Devient compliqué avec des imbrications

- Permet de garder des critères de sélection
- Permet de générer facilement des fichiers
- Doit être insérer avec un serveur HTTP avoir une génération dynamique

ISI

MEDIA

Architecture: vue



Exemple: Vues

```
drop view xmlpays;
create view xmlpays as
select '<nom>'||nom||'</nom>'||
 '<capitale>'||capitale||'</capitale>' ||
 '<population>'||population||'</population>' xml, nom,
  population, surface
from pays;
```

- Les pays avec une population de plus de 50 millons d'habitants

```
select '<pays>' from dual
union all
select xml from xmlpays where population>50
union all
select '</pays>' from dual;
```

Résultat: vue

```
<pays>
<nom>Allemagne</nom><capitale>Berlin</capitale>
<nom>France</nom><capitale>Paris</capitale>
<nom>Italie</nom><capitale>Rome</capitale>
<nom>Espagne</nom><capitale>Madrid</capitale>
<nom>Pologne</nom><capitale>Varsovie</capitale>
<nom>Suisse</nom><capitale>Berne</capitale>
<nom>Autriche</nom><capitale>Vienne</capitale>
</pays>
```

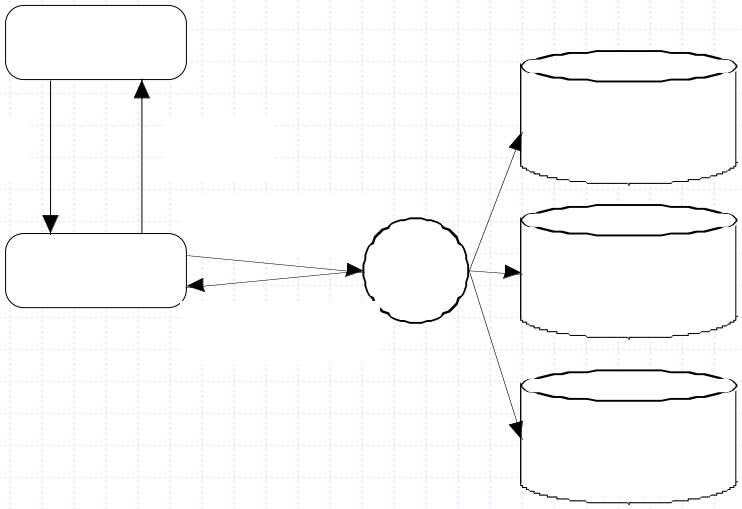
ISI MEDIA

Solution: application (java)

- Ecrire une classe qui:
  - Se connecte à la BD (ODBC)
  - Execute la requête SQL
  - Recupère résultat
  - (éventuellement fait autres requetes- imbrication)
  - Formatte le résultat <>
  - Retourne le XML
- Avantage:
  - Indépendant de la BD (relativement)
  - Réutilisable

ISI MEDIA

Architecture: application (java)



Cœur du code (pour les connaisseurs!)

```
while (pasfini)
    {if (typeligne.equals("ligne2"))typeligne="ligne1";
          else typeligne="ligne2";
     out.write("  <ligne style='"+typeligne+"'>"+"\r\n");
          for (i=1; i<=nbCol; i++) // pour chaque colonne
           { out.write("    <colonne>");
             out.write(results.getString(i));
             out.write("    </colonne>"+"\r\n");
           }
```

Résultat: application

```
<document>
 <tableau>
  <intro>Select nom, capitale from PAYS</intro>
   <ligne style='lignetitre'>
    <colonne>NOM    </colonne>
    <colonne>CAPITALE    </colonne>
</ligne>
   <ligne style='ligne1'>
    <colonne>Allemagne    </colonne>
    <colonne>Berlin    </colonne>
 </ligne>
 ▫
...
 ▫
<conclusion>7 lignes trouvees par la requete</conclusion>
 </tableau>
</document>
```

ISI MEDIA

J. Guyot BD/XML 2001 ISI - Université de Genève 41

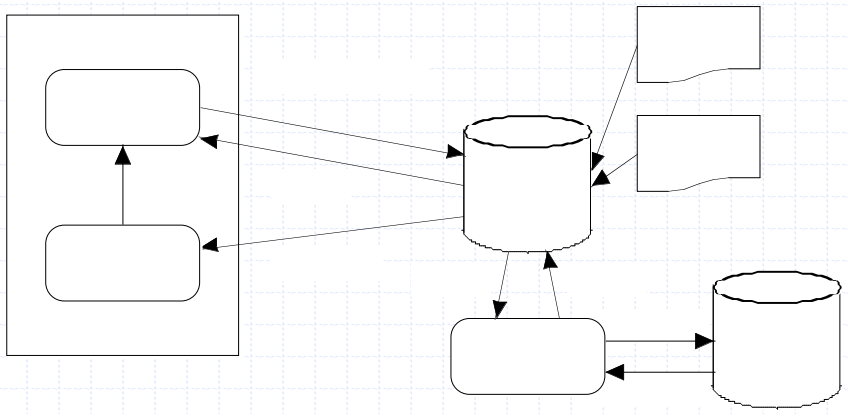
Serveur XML-BD

- Objectifs:
  - Servir du XML à partir d'un URL
  - Paramètres -> spécifie une requête à la BD
  - Afficher le résultat dans un Browser Web (IE>=5)
- Séparer:
  - Information (BD)
  - Sémantique (XML)
  - Forme (XSL)
  - Apparence (CSS)

ISI MEDIA

J. Guyot BD/XML 2001 ISI - Université de Genève 42

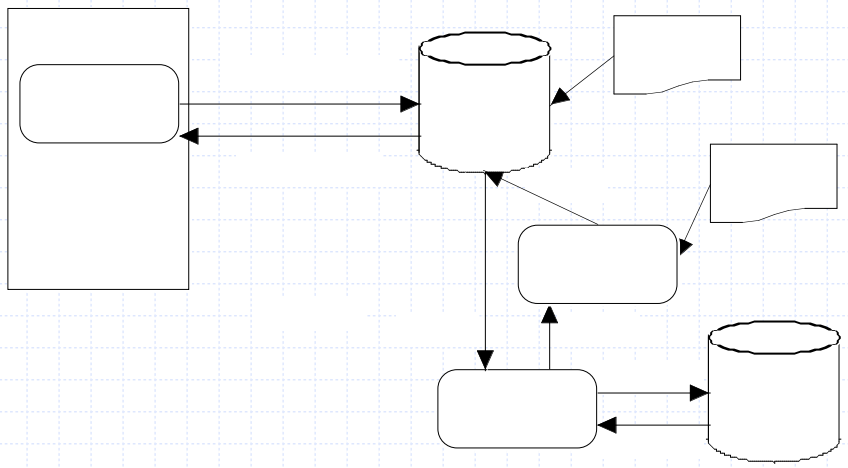
Architecture: Serveur XML-BD (CLIENT-SIDE)



ISI MEDIA

J. Guyot BD/XML 2001 ISI - Université de Genève 43

Architecture: Serveur XML-BD (SERVER-SIDE)



ISI MEDIA

J. Guyot BD/XML 2001 ISI - Université de Genève 44

Demo possible

- (serveur client side)..\Exemples\SQLXML-server-2 xml\RUN.bat
  - (ammorce) ..\Exemples
  - (voir le code source)
  - (modifier la requête)
  - (voir le code XSL - modifier XSL)
  - (voir le code CSS - modifier le CSS)
- (stoper le server)
- (serveur client side)..\Exemples\SQLXML-server-2 xml\RUN.bat
  - (ammorce) ..\Exemples
  - Refaire les points précédents
- (stoper le server)

I S I

M E D I A

J. Guyot BD/XML 2001 ISI - Université de Genève 45

indépendance:

- Contenu
  - Données
- Sémantique
  - <élément>
- Forme
  - Table, liste, ...
- Apparence
  - Couleur,fonte,...

un exemple de tableau modifie - Microsoft Internet Explorer

Fichier Edition Affichage Favoris Outils ?

Adresse http://localhost:1818/sql?req=Select+nom%2C+capitale%2C+surface OK

Précédente Suivante Arrêter Actualiser Démarrage

Liens Démarrage de Internet Guide des chaînes Hotmail

Exemple de lien SQL - XML

```
Select nom, capitale, surface,
population from PAYS
```

execute

resultat pour: Select nom, capitale, surface, population from PAYS

NOM	CAPITALE	SURFACE	POPULATION
allema	Berlin	356	82
France	Paris	551	58.80
Italie	Rome	301	56.70
Espagne	Madrid	504	37.50
Pologne	Varsovie	312	38.60
Suisse	Berne	41.30	7.20
Autriche	Vienne	83.80	8.10

7 lignes trouvees par la requete

Terminé Intranet local

I S I

M E D I A

J. Guyot BD/XML 2001 ISI - Université de Genève 46

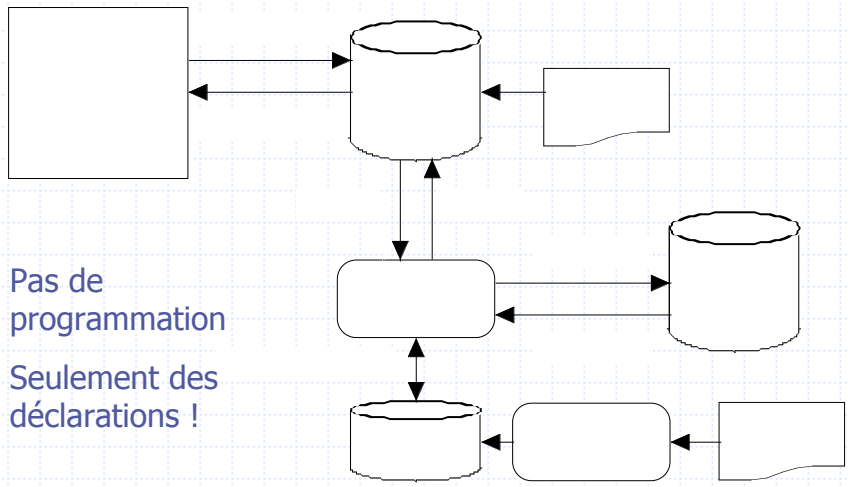
Un pas de plus pour les HyperTextes: Lazy

- Créer un schéma de l'hypertexte spécifiant:
- Contenu des nœuds
  - Sémantique (XML)
  - Données (BD)
- Navigation
  - Type de liens entre les nœuds:
    - Sauts
    - Inclusion
    - Expansion
- Forme (XSL)
- Apparence (CSS)

ISI MEDIA

J. Guyot BD/XML 2001 ISI - Université de Genève 47

Architecture: Lazy



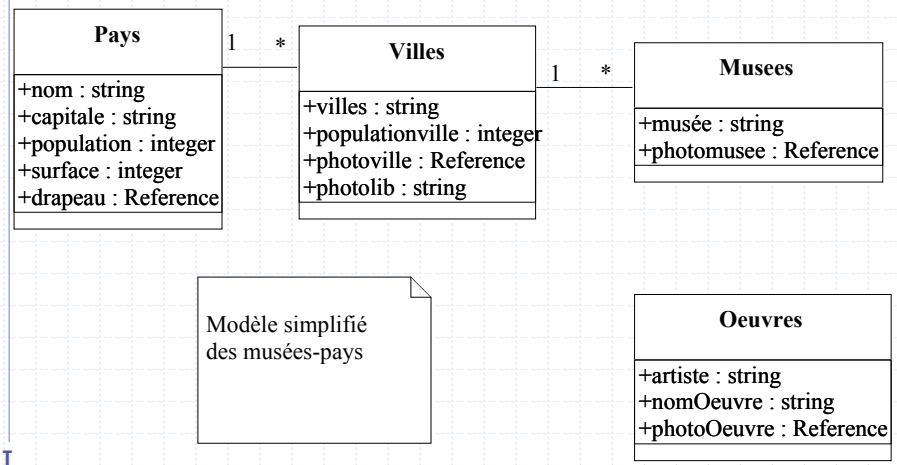
Pas de programmation

Seulement des déclarations !

ISI MEDIA

J. Guyot BD/XML 2001 ISI - Université de Genève 48

Exemple: Schéma des données Musée-pays



ISI MEDIA

Exemple de noeuds

```
node pays
items <pays>(
        <nom> ( expand href getdrapeau[nom] upper(nom) ),
        <capitale> ( capitale ),
        <ville> ( expand href villesparpays[nom] "?" ),
        <surface> ( surface ),
        <population> ( population )
        )
from pays selected by 1=1 order by nom

node villesparpays[selecteurpays]
items <villes>(
        <ville> ( expand href getphotoville[ville] ville,
                expand href museesparville[ville] " Musée ?"),
        <population> ( populationville )
        )
from villes
  selected by danspays=selecteurpays order by ville
```

ISI MEDIA

Démo Serveur de noeud

- (serveur node client side (démarrer)
- ..\Exemples
  - (voir le code source)
  - (explorer France - drapeau - ville, graphique, ...)
  - (voir le code XSL - modifier XSL)
  - (voir le code CSS - modifier le CSS)
- (stoper le server)

ISI MEDIA

J. Guyot BD/XML 2001 ISI - Université de Genève 51

Pays1

Mon atlas ISI - Microsoft Internet Explorer

Fichier Edition Affichage Favoris Outils ?

Nom	Capitale	Villes	Surface(1000Km2)2149.1	Habitants (Mio.)288.9
AUTRICHE	Vienne	?	83.8	8.1
ESPAGNE	Madrid	?	504	37.5
FRANCE	Paris	?	551	58.8
ITALIE	Rome	?	301	56.7
POLOGNE	Varsovie	?	312	38.6
SUISSE	Berne	?	41.3	7.2
ALLEMA	Berlin	?	356	82

**moyenne des surfaces: 307.0142857142857142857

Pays2

Mon atlas ISI - Microsoft Internet Explorer

Fichier Edition Affichage Favoris Outils ?

Nom	Capitale	Villes	Surface (1000Km2) 2149.1	Habitants (Mio.) 288.9
AUTRICHE	Vienne	?	83.8	8.1
ESPAGNE	Madrid	?	504	37.5
	Paris	ville / Habitants (Mio.): 6 Musée ?	551	58.8
ITALIE	Rome	?	301	56.7
POLOGNE	Varsovie	?	312	38.6
SUISSE	Berne	?	41.3	7.2
ALLEMA	Berlin	?	356	82

moyenne des surfaces: 307.0142857142857142857142857142857142286

GRAPHIQUE

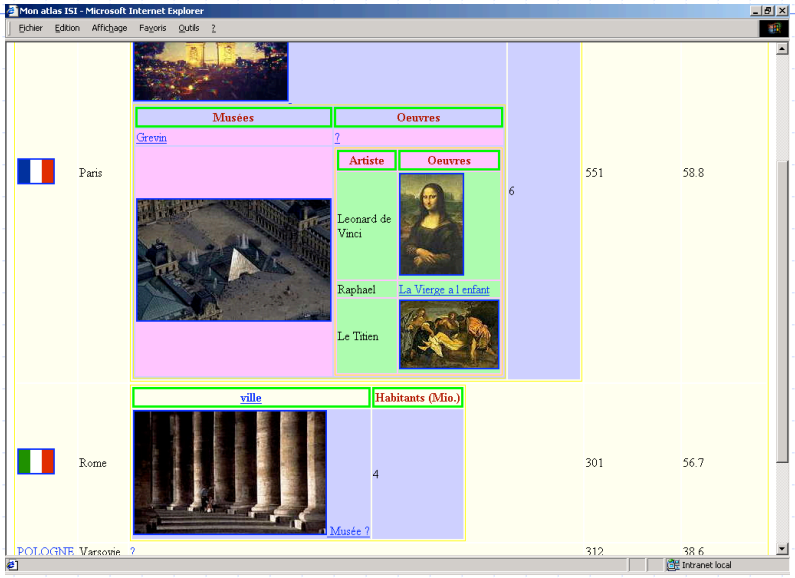
ISI J. Guyot BD/XML 2001 ISI - Université de Genève 53 MEDIA

Pays3

Mon atlas ISI - Microsoft Internet Explorer

Fichier Edition Affichage Favoris Outils ?

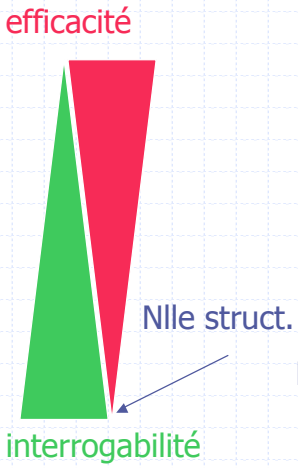
Musées	Oeuvres			
Grevin	?			
Paris	Artiste: Leonard de Vinci; Raphael; Le Titien	Oeuvres: La Vierge a l enfant	551	58.8
Rome	ville / Habitants (Mio.): 4 Musée ?		301	56.7
POLOGNE Varsovie	?		312	38.6



ISI J. Guyot BD/XML 2001 ISI - Université de Genève 54 MEDIA

V: stocker du XML - > BD

- Une question de granularité
- Stocker comme un tout
  - BLOB
  - I-file server (oracle)
- Stocker par blocs
  - Informations (auteur, titre)
  - Blocs (chapitre, …)
- Stocker par atomes
  - Dans une structure générique
  - Dans une structure du domaine



I S I

M E D I A

Exemple: de stockage globale

- Id = 1 document
- Line=numéro de séquence
- SVG= le texte XML de cette ligne

SVGSource
+Id : string +line : integer +svg : string

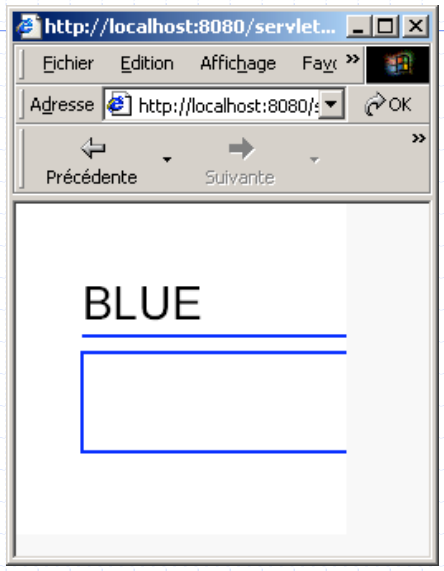
- Rapide pour le chargement
- Difficile de retrouver des infos
- (Like ???) souvent une mauvaise solution
- SVG pour le dessin vectoriel

I S I

M E D I A

Ex1: blue

- Un paramètre de l'URL définit le document XML (SVG) à afficher
- Document zoomable!

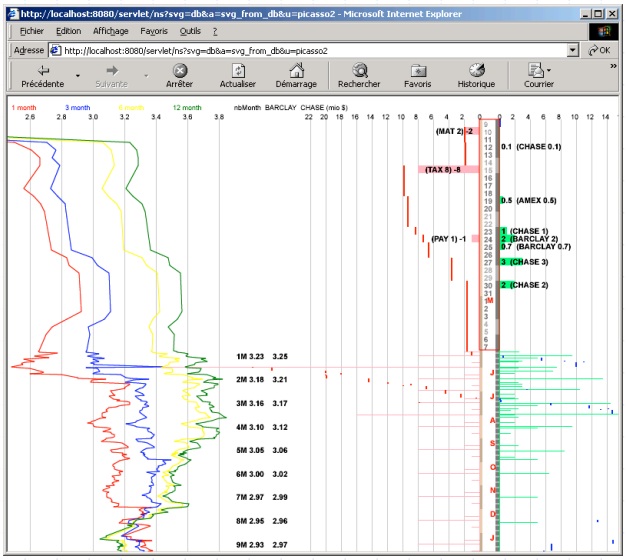


ISI

J. Guyot BD/XML 2001 ISI - Université de Genève 57

MEDIA

Ex1 (6000 lignes): cash-flow analysis (fish eye)



ISI

J. Guyot BD/XML 2001 ISI - Université de Genève 58

MEDIA

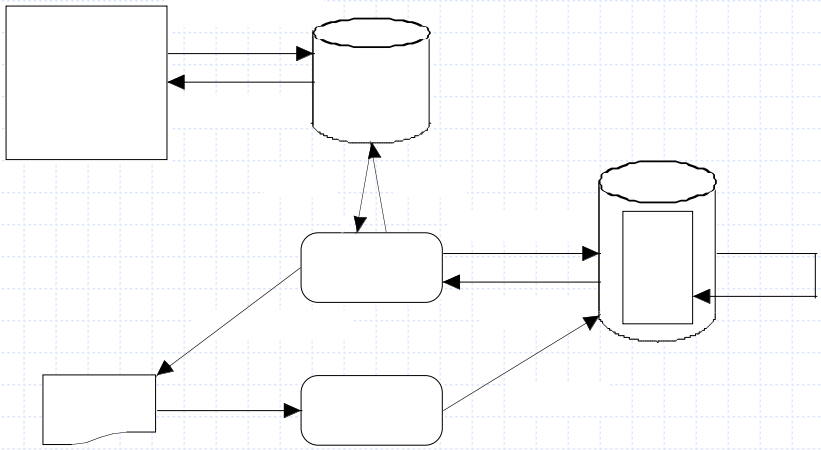
Stockage atomique de documents XML

- Utilisation d'une structure générique stocker:
  - Une arborescence
  - L'ordre des branches
  - Valeur sur chaque nœud
  - Valeur des attributs
- Indépendant de la dtd
- Retrouver chaque élément-valeur
- Retrouver chaque élément-attribut-valeur
- Spécfier des chemins e1/../e2/e3[val]
- Problème: coût de la reconstitution

ISI MEDIA

J. Guyot BD/XML 2001 ISI - Université de Genève 59

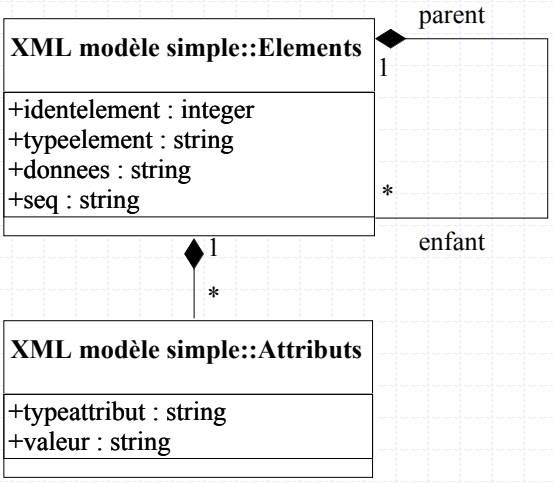
Architecture XML->BD et retour



ISI MEDIA

J. Guyot BD/XML 2001 ISI - Université de Genève 60

Modèle XML



Exemple: théatre

- A Midsummer Night's Dream
- XML statistique:

TYPEELEMENT	COUNT(*)
ACT	5
FM	1
GRPDESCR	2
LINE	2159
P	4
PERSONA	23
PERSONAE	1
PGROUP	2
PLAY	1
PLAYSUBT	1
SCENE	9
SCNDESCR	1
SPEAKER	500
SPEECH	500
STAGEDIR	136
TITLE	16

Reconstruire le document : attention long!

Mon atlas ISI - Microsoft Internet Explorer

Fichier Edition Affichage Favoris Outils ?

Adresse http://jg.gva.rolex:8080/servlet/ns?a=racine&u=4001

Précédente Suivante Arrêter Actualiser Démarrage Rechercher Favoris Historique

- COBWEB
- MOTH
- MUSTARDSEED

Description: fairies.

- Other fairies attending their King and Queen.
- Attendants on Theseus and Hippolyta.

Description scene: SCENE Athens, and a wood near it.

Sous titre de la piece: A MIDSUMMER NIGHT'S DREAM

ACT I

SCENE I. Athens. The palace of THESEUS.

Enter THESEUS, HIPPOLYTA, PHILOSTRATE, and Attendants

THESEUS

Now, fair Hippolyta, our nuptial hour
Draws on apace; four happy days bring in
Another moon: but, O, methinks, how slow
This old moon wanes! she lingers my desires,
Like to a step-dame or a dowager
Long withering out a young man revenue.

Terminé Intranet local

I S I J. Guyot BD/XML 2001 ISI - Université de Genève 63 M E D I A

Interro tous les acteurs test

Mon atlas ISI - Microsoft Internet Explorer

Fichier Edition Affichage Favoris Outils ?

Adresse http://localhost:8080/servlet/ns?a=actor&u=

Précédente Suivante Arrêter Actualiser Démarrage Rechercher

- THESEUS, Duke of Athens en grece.
- EGEUS, father to Hermia.
- LYSANDER
- DEMETRIUS
- PHILOSTRATE, master of the revels to Theseus.
- QUINCE, a carpenter.
- SNUG, a joiner.
- BOTTOM, a weaver.
- FLUTE, a bellows-mender.
- SNOUT, a tinker.
- STARVELING, a tailor.
- HIPPOLYTA, queen of the Amazons, betrothed to Theseus.
- HERMIA, daughter to Egeus, in love with Lysander.
- HELENA, in love with Demetrius.
- OBERON, king of the fairies.
- TITANIA, queen of the fairies.
- PUCK, or Robin Goodfellow.
- PEASEBLOSSOM
- COBWEB
- MOTH
- MUSTARDSEED
- Other fairies attending their King and Queen.
- Attendants on Theseus and Hippolyta.

Node: actor[] -- Lazy node XML server v2.5b2

I S I J. Guyot BD/XML 2001 ISI - Université de Genève 64 M E D I A

Les dialogues de LYSANDER

Mon atlas ISI - Microsoft Internet Explorer

Fichier Edition Affichage Favoris Outils ?

Adresse http://localhost:8080/servlet/ns?a=dialogue&u=LYSANDER

Précédente Suivante Arrêter Actualiser Démarrage Rechercher

LYSANDER

You have her father's love, Demetrius;
Let me have Hermia's: do you marry him.

LYSANDER

I am, my lord, as well derived as he,
As well possess'd; my love is more than his;
My fortunes every way as fairly rank'd,
If not with vantage, as Demetrius';
And, which is more than all these boasts can be,
I am beloved of beauteous Hermia:
Why should not I then prosecute my right?
Demetrius, I'll avouch it to his head,
Made love to Nedar's daughter, Helena,
And won her soul; and she, sweet lady, dotes,
Devoutly dotes, dotes in idolatry,
Upon this spotted and inconstant man.

LYSANDER

How now, my love! why is your cheek so pale?
How chance the roses there do fade so fast?

ISI J. Guyot BD/XML 2001 ISI - Université de Genève 65 MEDIA

La mise à jour de XML dans la BD test

Mon atlas ISI - Microsoft Internet Explorer

Fichier Edition Affichage Favoris Outils ?

Adresse http://localhost:8080/servlet/ns?eip=ZYX0XYZ678&a=enfantexp&u=4001

Précédente Suivante Arrêter Actualiser Démarrage Rechercher

THEATRE ISI !

A Midsummer Night's Dream up

Sources: up

Personnages, up

Description scene: SCENE Athens, and a wood near it. up

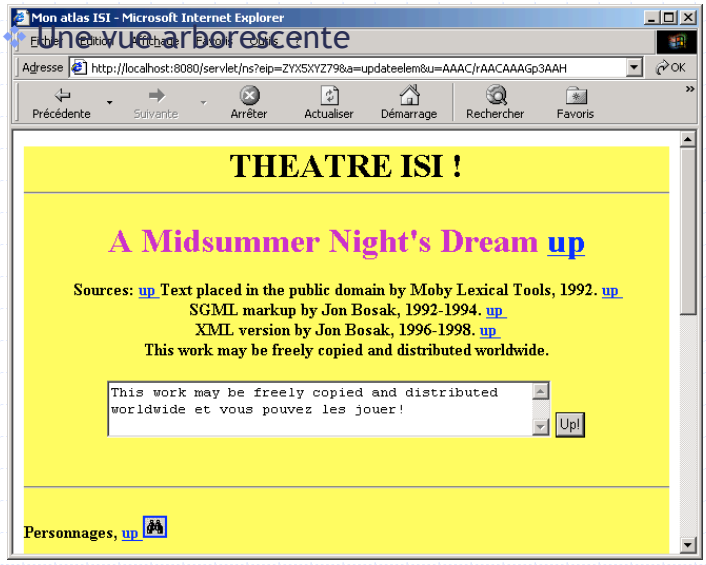
Sous titre de la piece: A MIDSUMMER NIGHT'S DREAM up

up

ISI J. Guyot BD/XML 2001 ISI - Université de Genève 66 MEDIA

Mise à jour et confirmation

Une vue arborescente

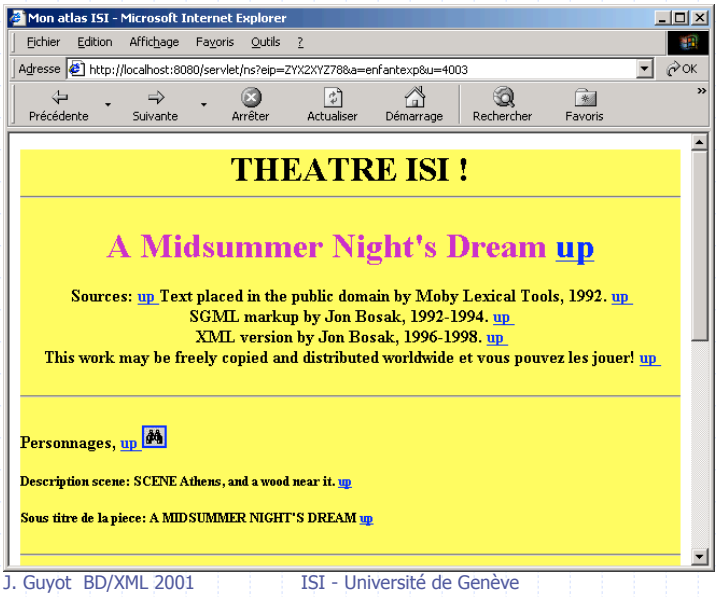


ISI

J. Guyot BD/XML 2001 ISI - Université de Genève 67

MEDIA

C'est OK!



ISI

J. Guyot BD/XML 2001 ISI - Université de Genève 68

MEDIA

Modélisation du domaine

- DTD -> UML (voir article de rational)
- DTD -> schéma relationnel
- Idée: rapprocher les données (éviter des jointures)
- Problèmes
  - Éléments -> relations (si multiples ref, si attribut)
  - Eléments -> attributs (si 0-1)
  - Attributs -> attributs
  - CDATA -> attributs
- Quelques propositions de traduction

ISI MEDIA

J. Guyot BD/XML 2001 ISI - Université de Genève 69

Traduction du ou

```
<!DOCTYPE racine [
    <!ELEMENT racine (cas1 | cas2 | cas3)>
    <!ELEMENT cas1 CDATA>
    <!ELEMENT cas2 EMPTY>
    <!ELEMENT cas3 (….)
]>
```

Solution : on fait l'expansion que du cas différent de NULL

Table racine (id,seq, cas1, cas2, cas3)

- Id est utilisé dans les autres tables comme référence
- Seq est utilisé pour ordonner les éléments
- PCDATA,EMPTY, (…) doivent être traduit en fonction de leur type

ISI MEDIA

J. Guyot BD/XML 2001 ISI - Université de Genève 70

Traduction de la sequence

- <!DOCTYPE racine [
- <!ELEMENT racine (seq1 , seq2 , seq3)>
- <!ELEMENT seq1 EMPTY>
- <!ELEMENT seq2 EMPTY>
- <!ELEMENT seq3 EMPTY>
-]>

Solution 1 : on fait l'expansion que dans l'ordre des attributs

Table racine (id, seq, Seq1, seq2 seq3)

- Id est utilisé dans les autres tables comme référence
- PCDATA,EMPTY, (...) doivent être traduit en fonction

ISI MEDIA

Traduction des modificateurs

- Elément (rien) min=1 max=1
- Element? min=0 max=1
  - -> bon sujet pour devenir des attributs (si eux mêmes sont simples (finissent par PCDATA ou EMPTY)
  - NOT NULL si (rien)
- Elément* min=0 max=n
- Elément+ min=1 max=n
  - -> bon sujet pour devenir des attributs (si eux mêmes sont simples (finissent par PCDATA ou EMPTY) d'une table

    Table Element(idracine, seq, PCDATA)

    NOT NULL si +
  - Sinon on applique récursivement la règle

    Table Element(idracine, seq, id)
- Pour les EMPTY on choisit abitrairement une valeur (O/N)

ISI MEDIA

Traduction des attributs

- Solution 1
  - On ajoute les attributs en fin de liste à la table de l'élément (ou éventuellement dans la table le contenant!)
- Solution 2
  - On gère une table unique pour tout les attributs! Comme dans le cas générique
  - (perd de la performance!)

ISI MEDIA

Exemple 1

```
<!DOCTYPE moi[
    <!-- DTD pour une
généalogie-->
    <!ELEMENT moi
(pers)>
    <!ELEMENT pers
(nom, prénom+,
diplôme*, (père,
mère)?) >
    <!ELEMENT père
(pers)>
    <!ELEMENT mère
(pers)>
    <!ELEMENT prénom
(#PCDATA)>
    <!ELEMENT nom
(#PCDATA)>
    <!ELEMENT diplôme
```

Table moi
(id : number,
nom : string not null,
table_prenom : dummy,
table_diplôme : dummy,
id_pere : number
id_mere : number)

Table prenom
(id_moi:number,
seq : number,
prenom:string)

Table diplome
(id_dip:number,
seq : number,
prenom:string)

ISI MEDIA

Exemple 2

```
<!DOCTYPE atlas[
    <!ELEMENT atlas
  (pays)*>
    <!ELEMENT pays
  EMPTY >
    <!ATTLIST pays
nom CDATA #REQUIRED
capitale CDATA
  #REQUIRED >
]>
```

Table atlas_pays
(id : number,
seq : number,
nom_pays:string
capitale_pays:string
)

ISI MEDIA

Remarques générales

- La dtd (schéma) doivent être stable
- Sinon -> évolution du schéma de la BD
- La modélisation doit être guidée par des objectifs du système à créer (interro, opération, …)

ISI MEDIA

Oracle initiative

- Propositions:
- Depuis du SQL -> génère du XML
- XML -> insert / update /delete dans les tables
- Utilisation des objets oracles (imbrications)

- XDK tool kit (Java, C++, PL/SQL)

- 9i stockage natif du XML ?

ISI MEDIA

J. Guyot BD/XML 2001 ISI - Université de Genève 77

XSU: XML SQL Utilities - select * from emp

- Générer du XML depuis un SQL (Version - java appl)

```
C:>java OracleXML getXML -user "scott/tiger"  "select * from emp"
<?xml version = '1.0'?>
<ROWSET>
   <ROW num="1">
      <EMPNO>7369</EMPNO>
      <ENAME>SMITH</ENAME>
      <JOB>CLERK</JOB>
      <MGR>7902</MGR>
      <HIREDATE>12/17/1980 0:0:0</HIREDATE>
      <SAL>800</SAL>
      <DEPTNO>20</DEPTNO>
   </ROW>
   <ROW num="2">
      <EMPNO>7499</EMPNO>
….
    </ROW>
</ROWSET>
```

ISI MEDIA

J. Guyot BD/XML 2001 ISI - Université de Genève 78

XSU: avec DTD

```
C:java OracleXML getXML -user "scott/tiger" -withDTD "select * from dept"
<?xml version = '1.0'?>
<!DOCTYPE ROWSET [
<!ELEMENT ROWSET (ROW)*>
<!ELEMENT ROW (DEPTNO, DNAME?, LOC?)>
<!ATTLIST ROW num CDATA #REQUIRED>
<!ELEMENT DEPTNO (#PCDATA)>
<!ELEMENT DNAME (#PCDATA)>
<!ELEMENT LOC (#PCDATA)>
]>
<ROWSET>
  <ROW num="1">
    <DEPTNO>10</DEPTNO>
    <DNAME>ACCOUNTING</DNAME>
    <LOC>NEW YORK</LOC>
  </ROW>
  …
</ROWSET>
```


XSU: parameter

```
OracleXML getXML
 [-user "username/password"]    -- the user name and password
 [-conn "JDBC_connect_string"]  -- JDBC connect string
 [-withDTD]                     -- include the DTD in the output
 [-useDOMDoc]                   -- use DOMParser to create the doc
 [-rowsetTag <rowsetTag>]       -- document tag name
 [-rowTag    <rowTag>]          -- row element tag name
 [-rowIdAttr <attrName>]        -- row-id attribute name
 [-rowIdColumn <column_name>]   -- db-column to use for the row id
 [-collectionIdAttr <attrName>] -- collection element-id attribute
 [-useNullAttrId]               -- use a null attribute
 [-styleSheet <URI>]            -- stylesheet processing instruction header
 [-styleSheetType <type>]       -- stylesheet header type (e.g.text/xsl)
 [-setXSLT <URI>]               -- XSLT to apply to XML doc
 [-setXSLTRef <URI>]    -- external entity reference for the XSLT doc
 [-useLowerCase| -useUpperCase] -- the case of the tag names
 [-errorTag <errorTagName>]     -- error tag name
 [-raiseException]              -- raise exceptions for errors
 [-raiseNoRowsException]        -- raise exception if no returned
 [-maxRows   <maxRows>]         -- maximum rows in output
 [-skipRows  <skipRows>]        -- rows to skip in output
 [-encoding <encoding_name>]    -- encoding to be used
 [-dateFormat <date format>]    -- date format to be used
 [<sql_string>| -fileName <sqlfile>] -- the SQL query string
```

Créer une vue objet pour l'imbrication

```
CREATE OR REPLACE TYPE employee_t AS OBJECT (
    empno    NUMBER (5),
    ename    VARCHAR2 (20),
    sal      NUMBER (9, 2),
    job      VARCHAR2 (20),
    deptno   NUMBER (2) );
CREATE TYPE employee_list_t AS TABLE OF employee_t;
CREATE VIEW dept_view1 OF departement_t
  WITH OBJECT OID (deptno) AS
      SELECT d.deptno, d.dname,
            CAST( MULTISET (
                        SELECT e.empno, e.ename, e.sal, e.job, e.deptno
                        FROM emp e
                        WHERE e.deptno = d.deptno)
                      AS employee_list_t)
   FROM   dept d;
```

ISI

MEDIA

XSU: sélection d'une vue objet

```
C:java OracleXML getXML -user "scott/tiger"  "select * from dept_view1"
<?xml version = '1.0'?>
<ROWSET>
  <ROW num="1">
    <DEPTNO>10</DEPTNO>
    <DNAME>ACCOUNTING</DNAME>
    <LIST_OF_EMP>
      <LIST_OF_EMP_ITEM>
        <EMPNO>7782</EMPNO>
        <ENAME>CLARK</ENAME>
        <SAL>2450</SAL>
        <JOB>MANAGER</JOB>
        <DEPTNO>10</DEPTNO>
      </LIST_OF_EMP_ITEM>
…
      <LIST_OF_EMP_ITEM>
…
      </LIST_OF_EMP_ITEM>
    </LIST_OF_EMP>
  </ROW>
<ROW num="3">
…
  </ROW>
</ROWSET>
```

ISI

MEDIA

Création d'un objet departement (…,{emp})

- CREATE TABLE xmldept (
- deptno NUMBER (2),
- dname VARCHAR2 (20),
- list_of_emp employee_list_t
-) NESTED TABLE list_of_emp STORE AS store_for_list_of_emp;

- Possible de modifier directement à partir du XML

ISI MEDIA

J. Guyot BD/XML 2001 ISI - Université de Genève 83

XSU: Insertion direct de XML

- Dept.xml contient le résultat de la vue dept_view1

```
C:>java OracleXML putXML -user "scott/tiger" -fileName "dept.xml" "xmldept"

 successfully inserted 4 rows into xmldept

C:\XSU>java OracleXML getXML -user "scott/tiger"  "select * from xmldept«¤

<?xml version = '1.0'?>
<ROWSET>
  <ROW num="1">
    <DEPTNO>10</DEPTNO>

…
  </ROW>
</ROWSET>
```

ISI MEDIA

J. Guyot BD/XML 2001 ISI - Université de Genève 84

VI: XML et BD au cœur des architectures complexes

- Objectifs: Créer des architectures ouvertes

- Les éléments communiquent avec XML
- Les éléments sont interconnectables
- Les éléments sont orthogonaux (imbricables)

- 3 exemples:
  - Une fabrique de documents imprimables
  - Une fabrique de graphiques
  - Une fabrique de sens (ontologie)

ISI MEDIA

J. Guyot BD/XML 2001 ISI - Université de Genève 85

Une fabrique de documents imprimables

- XSL:FO – la partie de XSL qui permet de formatter des documents

- Existe des visualiseurs XSL:FO
- FOP 0.15 ! Un logiciel qui permet de convertir en pdf
- Permet de transformer aussi du SVG en PDF
-
- XML +XSL → FO → PDF

ISI MEDIA

J. Guyot BD/XML 2001 ISI - Université de Genève 86

Exemple document XML

```
<doc>
<header>XML BD Une sortie pour tous les documents - p. <page-
   number/></header>
<body>
<title level="1">XSL:FO une possibilite de generer du texte
   imprimable</title>
<title level="2">Resume</title>
<para>
Ici du bla bla etc ...
</para>
<para>
C'est important. C'est important.C'est important.C'est
   important.C'est important.C'est important.C'est
   important.C'est important.C'est important.C'est important.
</para>
<title level="2">Importance reelle</title>
<para>
Il n'y plus de difference entre l'ecran et le papier: Tout est
   question de style!
</para>
</body>
</doc>
```

ISI MEDIA

J. Guyot BD/XML 2001 ISI - Université de Genève 87

extrait de XSL qui écrit du XSL:FO

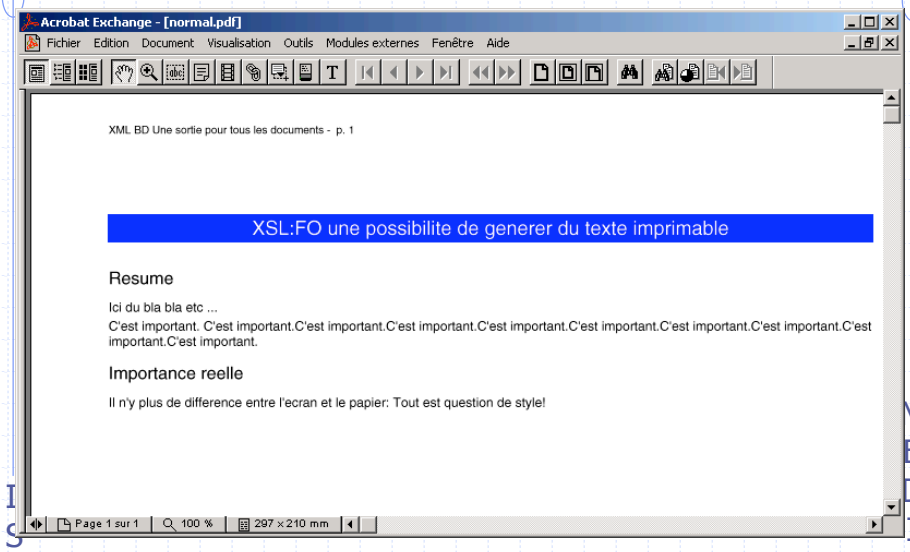
```
<xsl:template match="para">
   <fo:block font-size="12pt"
         font-family="sans-serif"
         line-height="15pt"
         space-after.optimum="3pt"
         text-align="start">
         <xsl:apply-templates/>
   </fo:block>
</xsl:template>

<xsl:template match="title[@level='1']">
   <fo:block font-size="18pt"
         font-family="sans-serif"
         line-height="24pt"
         space-after.optimum="15pt"
         background-color="blue"
         color="white"
         text-align="center"
         padding-top="3pt">
         <xsl:apply-templates/>
   </fo:block>
</xsl:template>
```

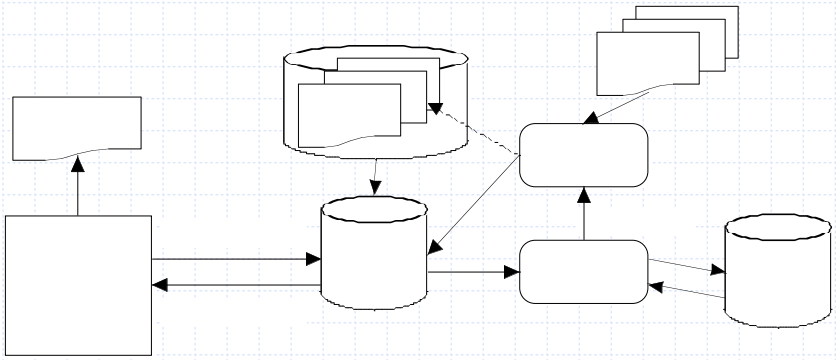
ISI MEDIA

J. Guyot BD/XML 2001 ISI - Université de Genève 88

Résultat en PDF



Architecture de la fabrique



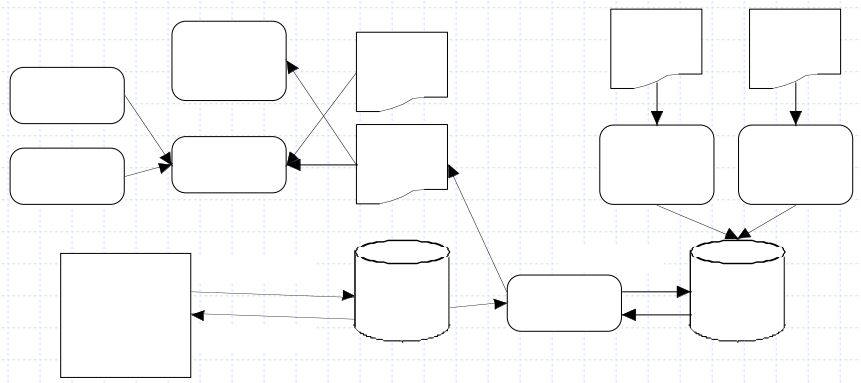
Une fabrique de sens

- Utilisation de la sémantique dans la gestion des documents
- Ontologie
- Visualiser
- Éditer
- Exporter
- Utiliser dans d'autres applications

I S I

M E D I A

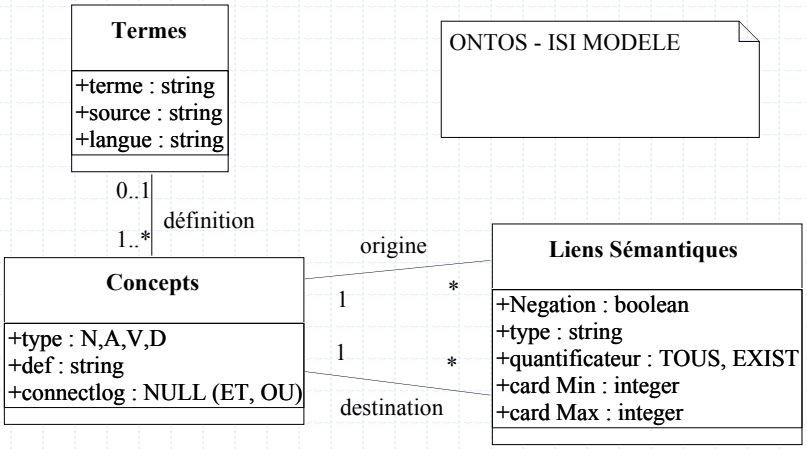
Architecture de la fabrique



I S I

M E D I A

Modèle ontos BD

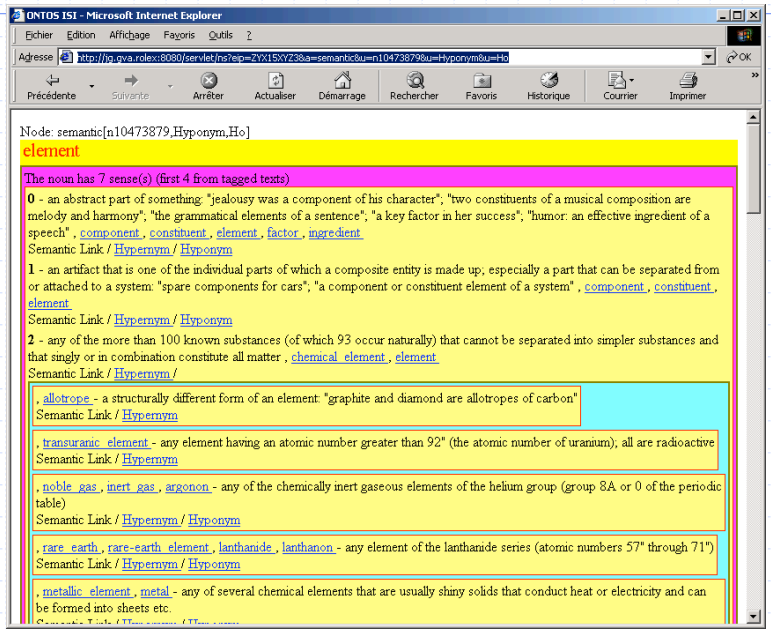


I S I

M E D I A

J. Guyot BD/XML 2001 ISI - Université de Genève 93

Visualisation de l'ontologie: element

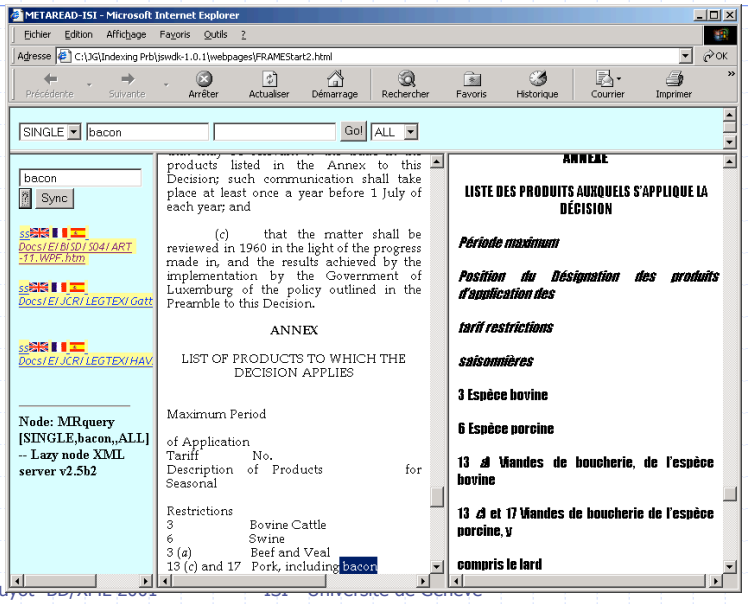


I S I

M E D I A

94

Metaread-ISI: indexation doc + XML +ontologie



I S I

J. Guyot BD/XML 2001 ISI - Université de Genève 95

M E D I A

Une fabrique de graphique

- Problème:
  - Extraction des données
  - Spécification des paramètres du graphe
  - Insertion dans un document (commentaire)
  - Impression de qualité
  - Répétabilité du processus (quotidien)
- Excel n'est pas adapté !

I S I

J. Guyot BD/XML 2001 ISI - Université de Genève 96

M E D I A

Solutions

- Extraction des données
  - BD- nœud lazy ->XML
- Spécification des paramètres du graphe
  - BD- nœud lazy ->XML (insérant le XML des données)
- Impression de qualité
  - Utiliser du SVG
- Insertion dans un document (commentaire)
  - Ok pour HTML (plugin SVG)
  - Ok pour PDF (FOP)
- Répétabilité du processus
  - Quand on veut tout est spécifié, donc la génération est automatique

ISI MEDIA

J. Guyot BD/XML 2001 ISI - Université de Genève 97

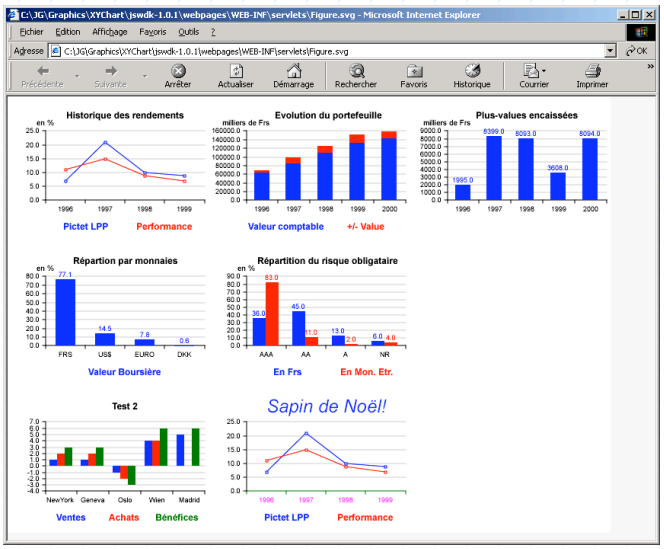
Un document XML de specif de graphique

```
<graphic>
  <title>ici le titre du graphique
    <colorText>red</colorText>
  </title>
  <axisX>ici le libelle de l'axe X
    <colorText>red</colorText>
    <colorAxis>yellow</colorAxis>
  </axisX>
  <axisY>ici le libelle de l'axe Y
    <colorText>red</colorText>
    <colorAxis>green</colorAxis>
  </axisY>
  <dataSet>
    <showValues>on</showValues>
    <set>Pictet LPP
      <colorText>blue</colorText>
    </set>
    <set>Performance
      <colorText>red</colorText>
    </set>
    <x>1996
      <y>7</y>
      <y>11</y>
    </x>
    …
  </dataSet>
</graphic>
```

ISI MEDIA

J. Guyot BD/XML 2001 ISI - Université de Genève 98

Un grapheur XML -> SVG !



I S I

J. Guyot BD/XML 2001 ISI - Université de Genève 99

M E D I A