Introduction  

1.) Context  

2.) Database  
   2.1) Definition  
   2.2) The different models  

3.) XML (Extensible Markup Language) and relational database  
   3.1) What is XML?  
   3.2) Storage of XML documents in relational database  
   3.3) XML and query language  
   3.4) Publication of XML documents in relational database  

4.) XML and federated database  

Conclusion  

Bibliography  

Glossary
Introduction

The explosion of internet and the progress of technologies give some new approaches of databases used in information system. It has some consequences on the management of information which becomes more complicated. Firms are some good consumers of big and heterogeneous information. Also in order to answer to their needs, some new standards such as XML: extensible markup language are trying to make and to represent information.
1.) **Context**

Today web applications become more and more sophisticated and need to access to a lot of different type of information. These applications are the diffusion of information for example publicity, the exchange of information for example news, the search of information which uses some search web engine and finally the web business such as purse activities or transactional commercial activities which develop and increase every day.

Now a lot of firm open their system of information to the internet, and so some new kind of architecture appear.

The development of internet and firms' technologies show that the web applications need to couple internet technologies and databases.

Today clients access to the databases from the web and so it needs a special architecture in order to answer to their requests. In parallel, Web site are developing every day. The web sites become very numerous and very heterogeneous. Their contents are very complex for example we can find several HTML files with different links and it has became very difficult to modify it when it is necessary that's why database is a good solution to solve this problem. It gives a method to represent information and organize a logical hierarchy between information.

2.) **Database**

2.1) **Definition:**

2.2) **the different models.**
Information will be obtained by asking some requests on this table by using a query language SQL: Structured Query Language.

For example, the relational table CLASS is:

<table>
<thead>
<tr>
<th>NumClass</th>
<th>Nbstudent</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>100</td>
</tr>
</tbody>
</table>

The advantages of this model are that the query language is based on a solid theory algebra rules and it is simple to build.

But the inconvenience of this model is that the schema is fixed and it is difficult to modify it.

So the second model is a model of files. It has a simple structure and so it is easier to modify it than the first model. This model consists to collect files and it is organizing as the system of files of an operating system such as we have on a computer. In this model, the big advantage is the simplicity of management of files. But there are no information on the structure and no schema so the consequences are that information can be redundant. But the real problem is that there is no query language for this model.

The third model is the semi-structured data which must answer to the facts of that:

• The data can be not compatible with a schema.
• The structure of the data evolves and needs to be updated frequently.
• The data can be structured slightly.

It is a compromise between the relational model and the model of files. So there is a supple schema and a query language. This model appeared in order to manage heterogeneous and structured data.
3.) **XML (Extensible Markup Language) and relational database.**

3.1) **What is XML?:**
3.2) Storage of XML documents in relational database:

<table>
<thead>
<tr>
<th>root</th>
<th>origin</th>
<th>target</th>
<th>label</th>
<th>ltype</th>
<th>ntype</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
<td>name</td>
<td>AGGR</td>
<td>STRING</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>3</td>
<td>manager</td>
<td>AGGR</td>
<td>INTEGER</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>4</td>
<td>address</td>
<td>AGGR</td>
<td>NODE</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>5</td>
<td>number</td>
<td>AGGR</td>
<td>INTEGER</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>6</td>
<td>street</td>
<td>AGGR</td>
<td>STRING</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>7</td>
<td>drive</td>
<td>ASSC</td>
<td>NODE</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>8</td>
<td>mark</td>
<td>AGGR</td>
<td>STRING</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>9</td>
<td>color</td>
<td>AGGR</td>
<td>STRING</td>
</tr>
</tbody>
</table>
3.3) **XML and query language.**

- XPATH 2.0 works on an XML document. It gives the possibility to extract nodes and put a reference on each node of the tree.
- XSLT 2.0 which can translate an XML document to XML, HTML, or text.
- XQUERY 1.0 comes from XPATH 2.0. It is used for access to the database and it is a real query language.

3.4) **Publication of XML documents in relational database.**

```
XPATH/XQUERY---------->XML VIEW------
Æ
SQL------------
Æ
database
Å
data---------
```

Today XML is using for publication and exchange data. In parallel, database is using for transaction and for modifying data in order to be updated. In a database all the data are structured and they have got a type, but in XML there are simple structure, no type, and a lot of text. So the advantage is that XML can publish a part of the data from a database and can be used for generating different document formats.
4) XML and federated database.

**Conclusion:**
Bibliography:

Online resources

http://www.e-xmlmedia.fr/site_francais/documentation_livres_blancs.htm

Georges, Gardarin Les datawebhouses arrivent


BSTAT Generalites sur les bases de donnees.

www.blois.univ-tours.fr/~mirian/mi-BD3.html

cedric.cnam.fr/vertigo/Cours/Valeur-C/bamain-valc-4x.pdf

Bern Amann et Philippe Rigaux Données semi structurés et XML

lithwww.epfl.ch/teaching/docmul/15-bd/bd-slides.pdf

Dr Yanne Rikik XML et les bases de données

www.httr.ups-tlse.fr/pedagogie/cours/bd/bdrel.pdf

M.Boughanem Bases de données relationelle

Text resources

B.Defude :XML et Bases de données
Glossary:

- DTD: document type definition
- HTML: HyperText Markup Language
- SGML: Standard generalized markup language
- XML: Extensible Markup Language
- XML-QL: Extensible Markup Language-Query Language
- XQuery: Extensible Query Language
- XSLT: Extensible Stylesheet Language transformation