Need of XML (What is XML?)

- Solution for data Exchnage between Database Management System (DBMS)
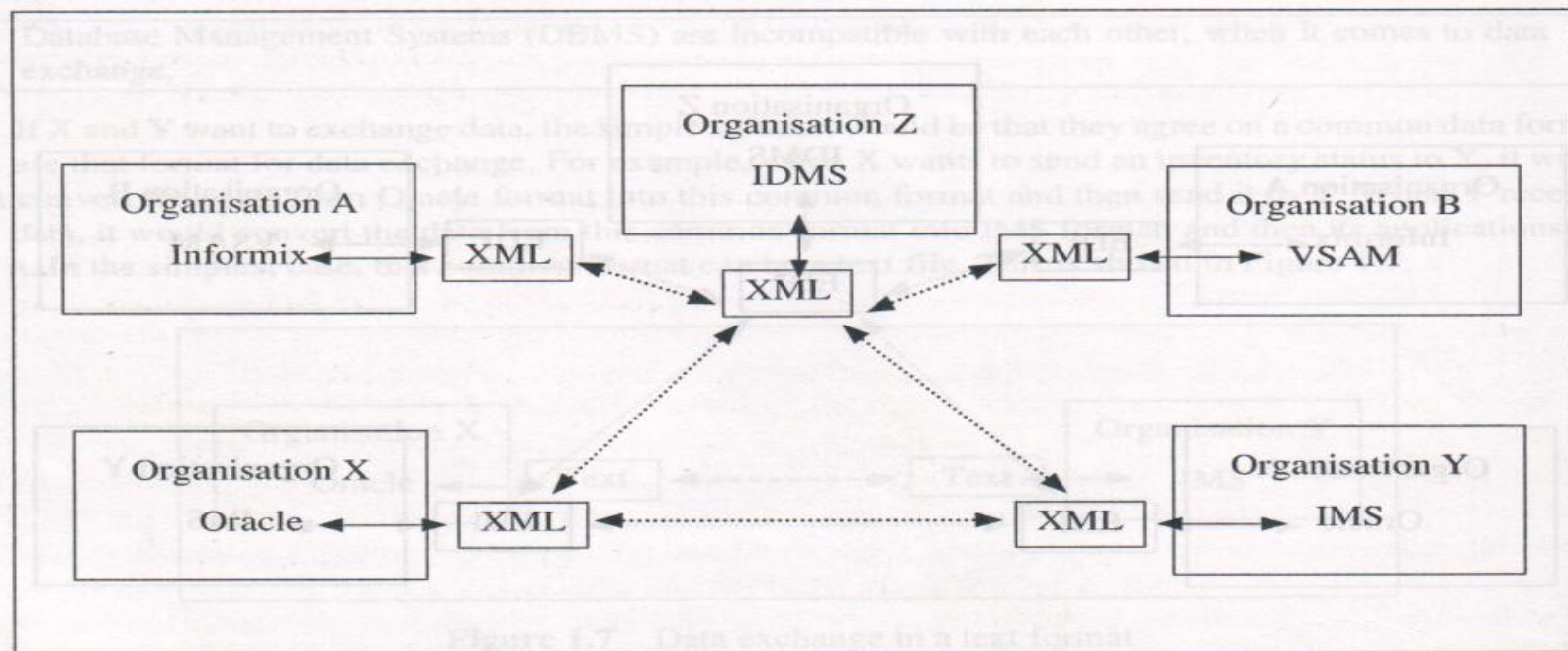

Figure 1.9 XML as the data exchange standard

Data Exchange using XML

XML Versus HTML

HTML

- HTML defines a set of tags describing how the web browser should display the content of a document.

- HTML is an information presentation language.

- HTML focuses on display of DATA

XML

- XML used tag to organise documents and the contents there in.

- XML is an information description language

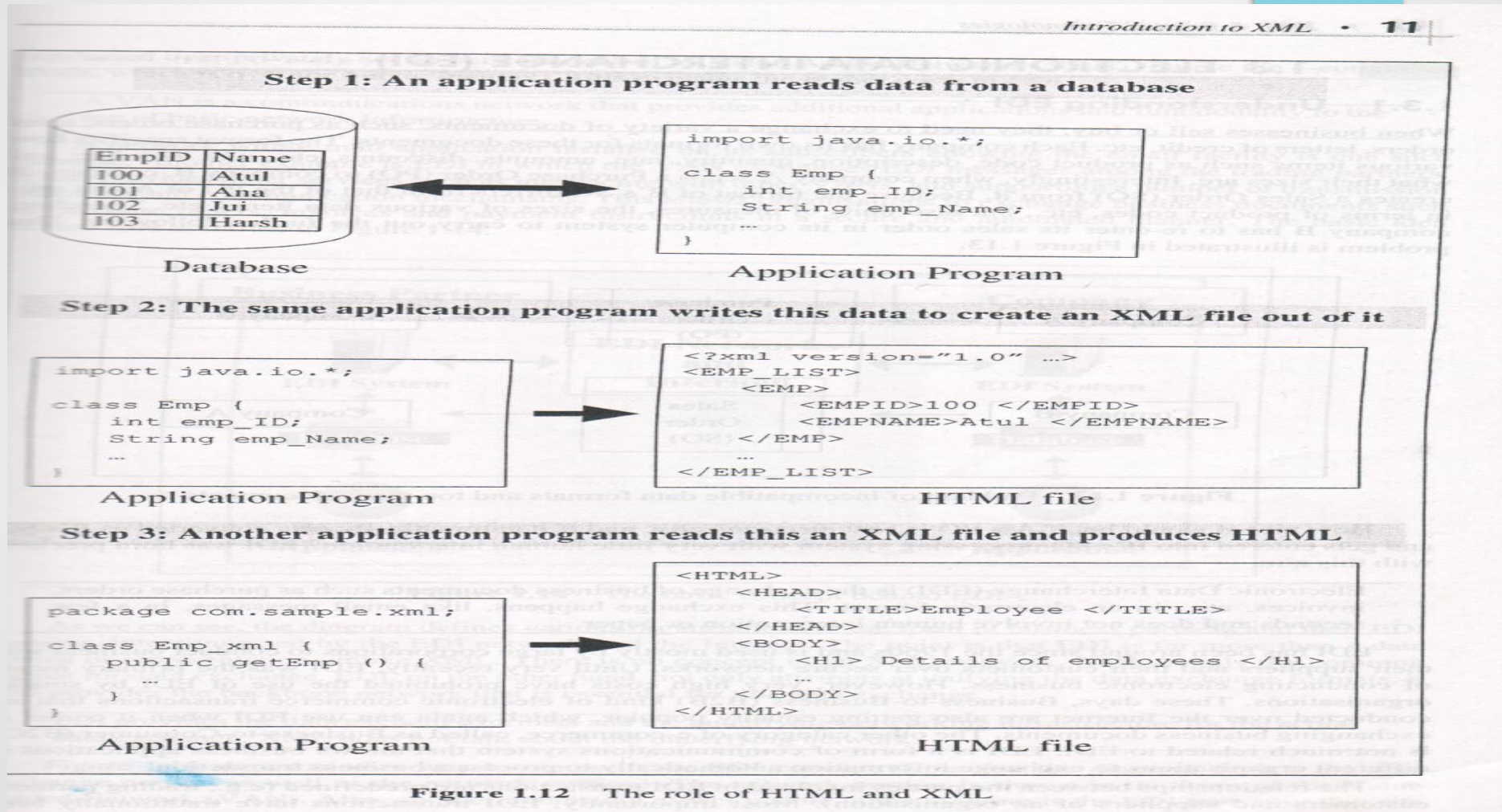- XML focuses on representation of DATA

XML Versus HTML

HTML

- HTML can not process, sort, encrypting data .

- HTML formate would convey nothing about meaning of the data.

XML

- XML can process, sort, encrypting data.

- XML describes the meaning of the data.

XML Versus HTML

The Role of HTML and XML

Figure 1.12   The role of HTML and XML

XML Terminology

- Every XML file has an extension of .xml

- Demo 1: Book.xml

XML Terminology

Terminology in XML

Figure 1.23 shows a short pictorial explanation of this XML document. A detailed explanation is provided in Table 1.1.

```
<?xml version="1.0"?>                                              [XML tag]
<?xml-stylesheet type="text/xsl" href="books_list.xsl"?>

<BOOKS>                                      [Root element]
    <BOOK pubyear="1929">                    [Start element indicator]
        <BOOK_TITLE>Look Homeward, Angel</BOOK_TITLE>
        <AUTHOR>Wolfe, Thomas</AUTHOR>
    </BOOK>                                   [End element indicator]
    <BOOK pubyear="1973">
        <BOOK_TITLE>Gravity's Rainbow</BOOK_TITLE>
        <AUTHOR>Pynchon, Thomas</AUTHOR>
    </BOOK>                                   [Element value]
    <BOOK pubyear="1977">        [Element name]
        <BOOK_TITLE>Cards as Weapons</BOOK_TITLE>
        <AUTHOR>Jay, Ricky</AUTHOR>
    </BOOK>                          [Attribute name]
    <BOOK pu byear="2001">                    [Attribute Value]
        <BOOK_TITLE>Computer Networks</BOOK_TITLE>
        <AUTHOR>Tanenbaum, Andrew</AUTHOR>
    </BOOK>
</BOOKS>
```
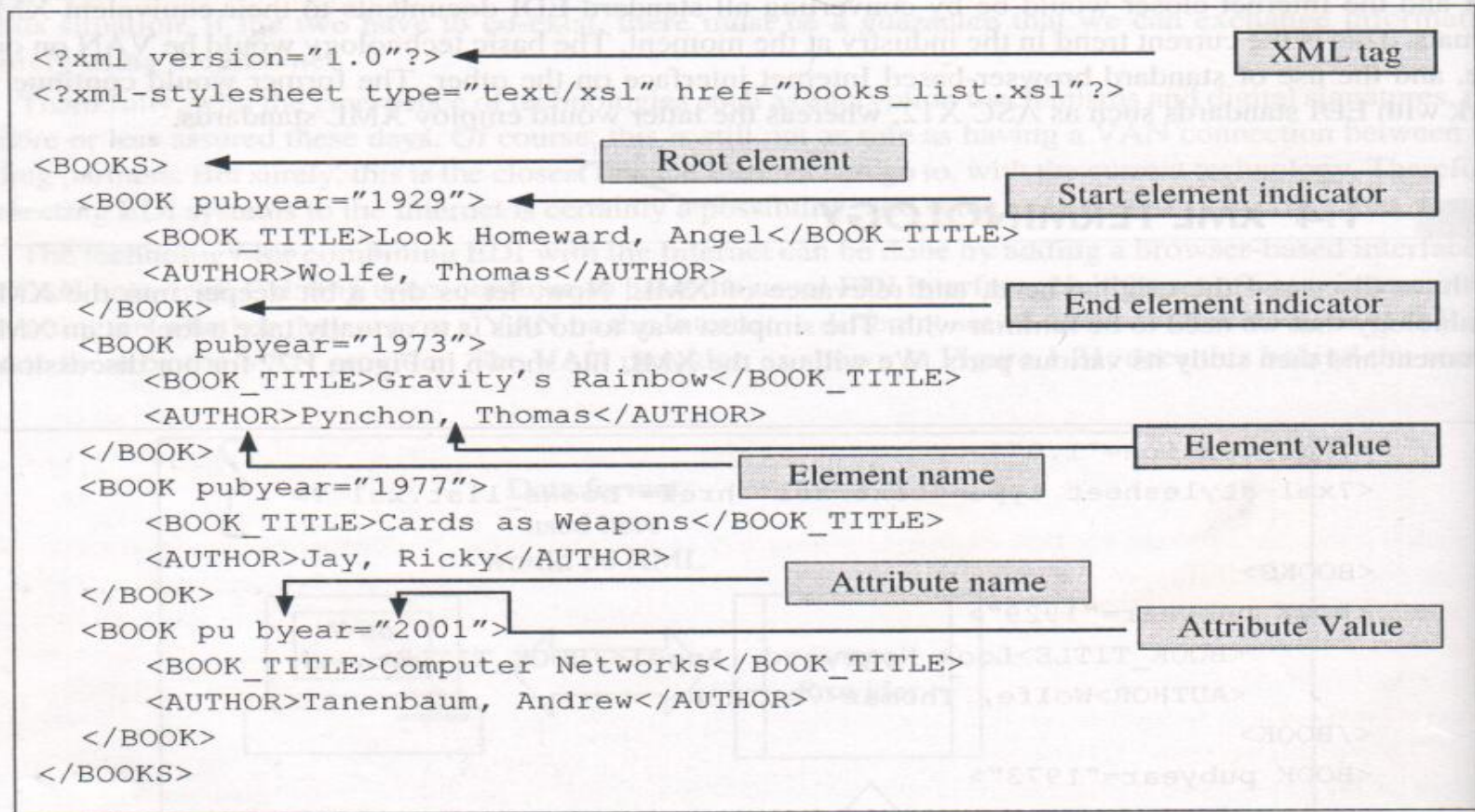
Figure 1.23   Terminology in XML – High level overview

XML Terminology

Contentsof the XML File	Description
<?xml version="1.0" encoding="UTF-8"?>	This line idetifies that files is an XML document. Every XML document must begin with this line.
<BOOKS>	This line implicitly indicates the start of the actual contents in the XML file.
<BOOK Pubyear='1973'>	Pubyear='1973' is called as an attribute in XML. An Attribute serves the purpose of providing more information about an element.
<BOOK_TITLE>	This is another element declaration. Information in a hierachical manner comes into picture.
</BOOK>	This declaration indicates the end of the first <BOOK> element.

XML Terminology

Terminology in XML

Introduction to XML · 33

Exercise 3: Suppose we want to store information regarding employees in the following format in XML. Show such a file with one example:

Employee ID	Numeric	5 positions
Employee Name	Alphanumeric	30 positions
Employee Department	Alphanumeric	2 positions
Role	Alphanumeric	20 positions
Manager	Alphanumeric	30 positions

Solution 3:

```
<?xml version="1.0"?>
```

XML Terminology

- **<u>EXAMPLE</u>**

- Suppose we want to sotre information regarding employees in the following format in XML. Show such a file with example

  Employee Id   Numeric

  Employee Name   Alphanumeric

  Employee Department   Alphanumeric

  Role   Alphanumeric

  Manager   Alphanumeric

- Demo – 3 Employee.XML

XML STANDARDS

- The World Wide Consortium (W3C) has developed a number of standards for XML.

- Standards are :

  - XML

  - XML Namespaces

  - Document Type Definitions (DTD)

  - Cascading Style Sheets (CSS)

  - Extensible Stylesheet Language (XSL)

  - XML Schemas

XML STANDARDS

- Standards are :
    - XML Query Language (Xquery)
    - XLink
    - Xpointer
    - XPath
    - XML Digital Signatures
    - XHTML

XML STANDARDS - XML

- The first version of XML Standard was 1.0 finalised in feb, 1998.

- Which are as follow:

  - XML shall be **easily implemented** on the Internet

  - XML shall be **usable in a wide variety of applications**

  - XML shall **comply with SGML** (Standard Generalized Marked up Language)

  - XML documents **shall be readable to humans**

  - XML document **design shall be formal and compact**

  - XML document **design shall be done quickly**

  - XML documents shall be easy to create

  - XML markup is not required to be abrupt

XML STANDARDS –

- **XML Namespaces**

    – To avoid confusing situation in terms of tags duplications, Namespace help us resolve conflict.

- **Document Type Definitions (DTD)**

    – A DTD file allows us to specify the rules associated with an XML file.

XML STANDARDS - XMLNampespaces

- The XML Namespace allowing the users of XML to associate certain tags with their own identifiers.

- User A creating <book> tag in Book1.xml

- User B creating <book> tag in Book2.xml

- If Book1.xml and Book2.xml going to merge this would be an ambiguity.

XML STANDARDS –

- **Document Type Definitions (DTD)**

A Document Type Definition (DTD) file allows us to specify the rules associated with an XML file. Figure 1.25 shows the idea.

```
<?xml version="1.0"?>
<BOOKS>
        <BOOK>
        ...
    ...
</BOOKS>
```

The XML document contains the actual data, as usual. The file has an extension of .xml.

```
<!ELEMENT BOOKS (BOOK)>
<!ELEMENT BOOK
(BOOK_TITLE, AUTHOR)>
...
```

The DTD document contains the rules about the above XML document. The file has an extension of .dtd.

Figure 1.25   Concept of Document Type Definitions (DTD)

XML STANDARDS –

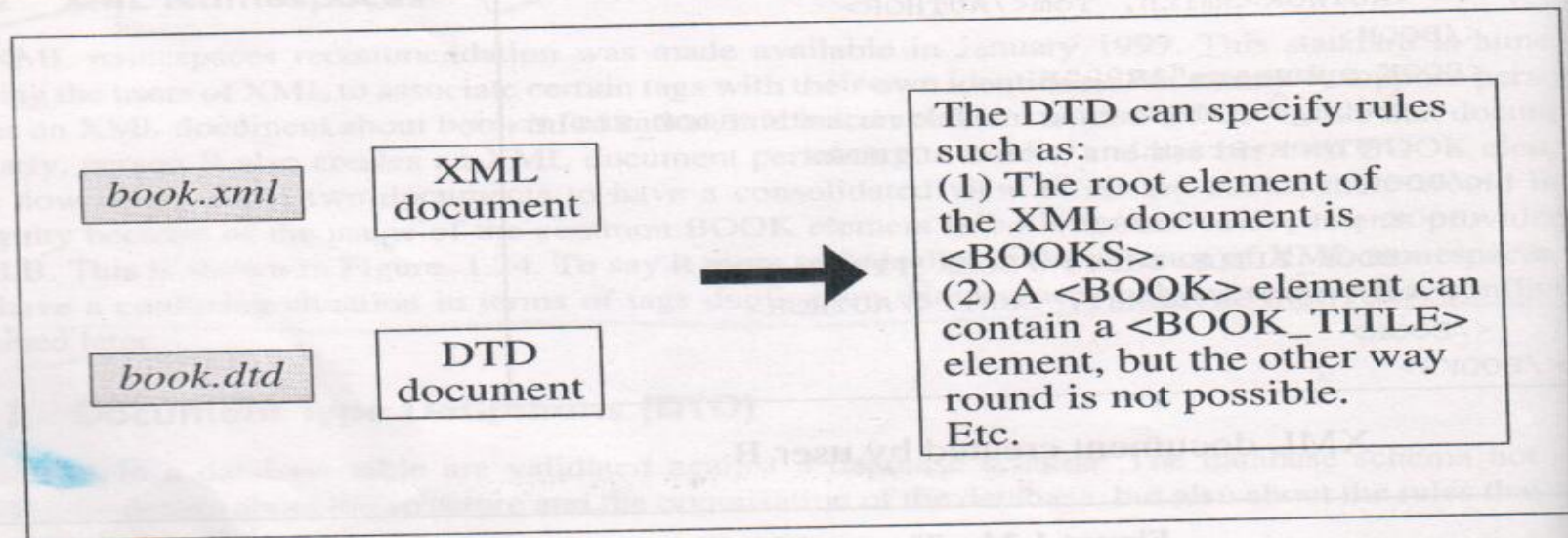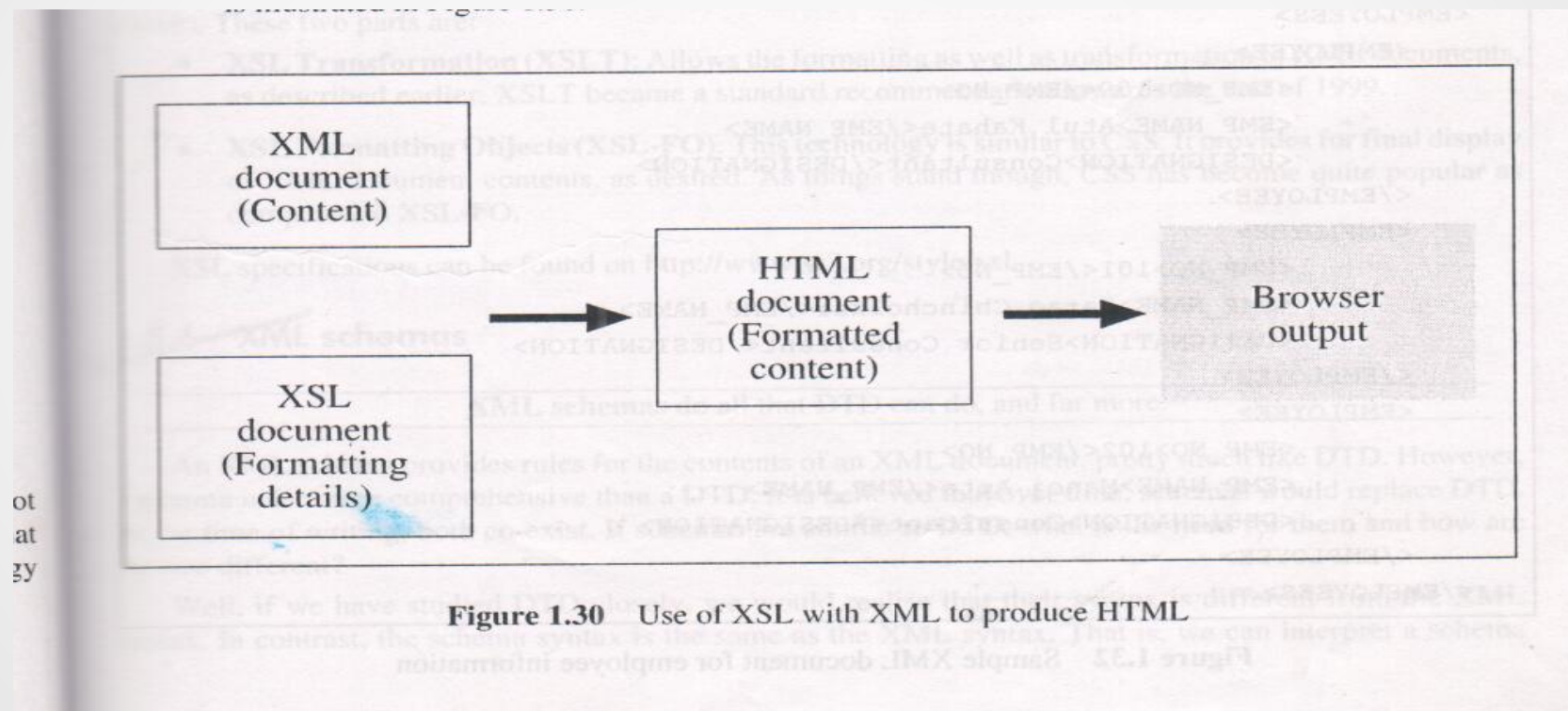- **Document Type Definitions (DTD)**

Figure 1.26 DTD concept illustrated further

XML STANDARDS –

- **Cascading Style Sheets (CSS)**

  – The **Cascading Styling Sheets technology allows the formatting of HTML documents** in a standard uniform manner.

- **Extensible Stylesheet Language (XSL)**

  – The Extensible Stylesheet Language standard provides means for formatting and **transforming XML documents in the desired manner.**

  – XSL technology consists of two parts.

    - XSL Transformation (XSLT)

    - XSL Formatting Objects (XSL - FO)

XML STANDARDS –

- **Extensible Stylesheet Language (XSL)**

Figure 1.30 Use of XSL with XML to produce HTML

XML STANDARDS –

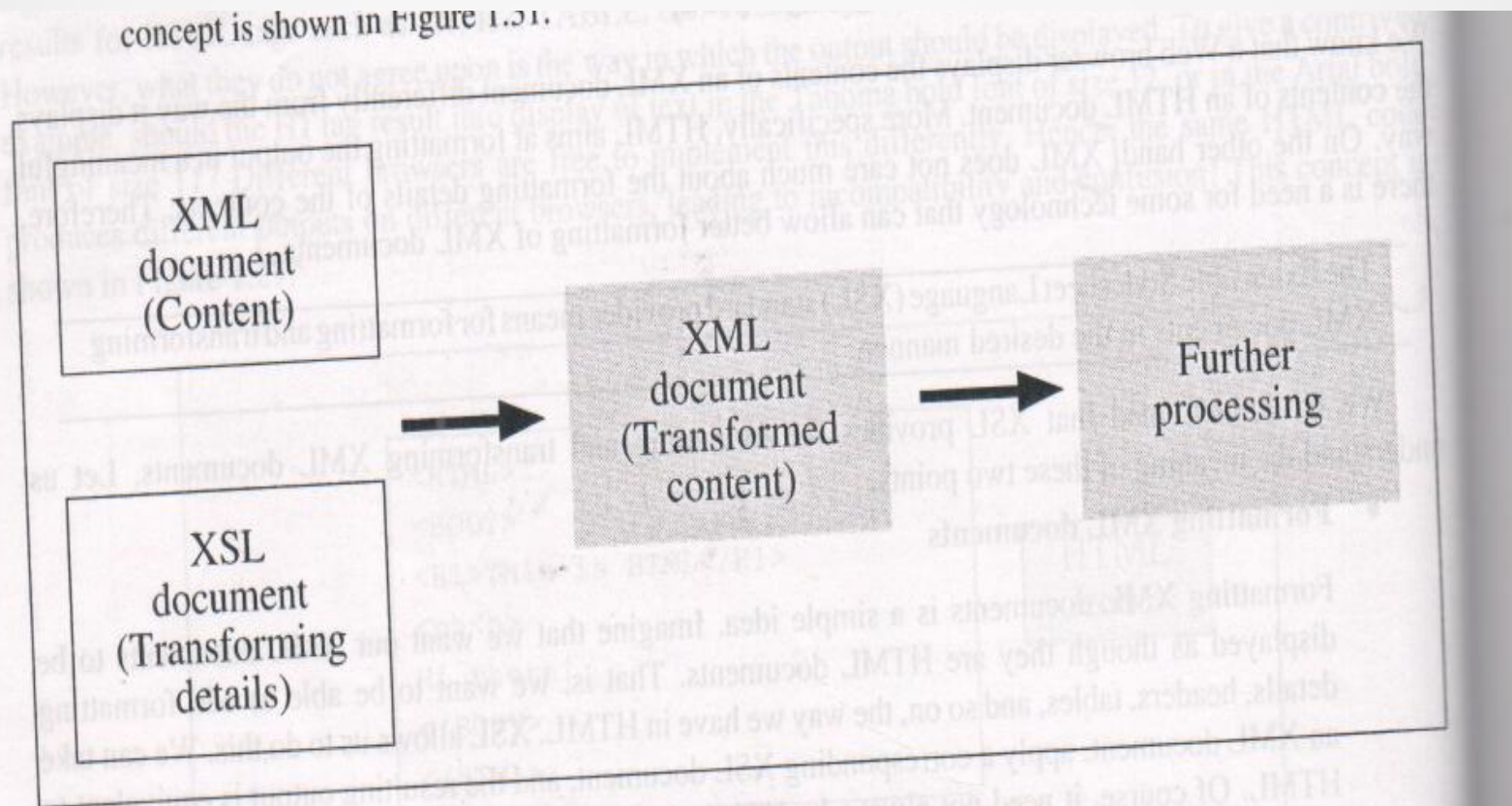- **Extensible Stylesheet Language (XSL)**

Figure 1.31   Use of XSL to transform XML

XML STANDARDS –

- **XML Schemas**

    - An XML schema provides rules for the contents of an XML document, pretty much like DTD.

    - It is far more than DTD.

- **XML Query Language (Xquery)**

    - The XML query language, also called as Xquery, is still being developed.

    - Xquery specifications can be found on

    - http:// www.w3.org/tr/xmlquery-req/.

XML STANDARDS –

- **Xlink**

  - Xlink defines a standard mechanism for **creating hyperlinks** in XML documents.

    - Xlink is a language for creating hyperlinks in XML doc.

    - Xlink is similar to HTML links.

    - Any elements in an XML document can behave as an Xlink

    - Xlink supports two types of links

      - Simple link

      - Extended links

XML STANDARDS –

- **Xpointer**

  – Xpointer allows the **hyperlinks to point to more specific parts in the XML** documents.

- **Xpath**

  – Xpath is used **to refer to specific portions of an XML document using XSLT and Xpointer.**

- **XML Digital Signatures**

  – XML digital signature **provide a means of message integirty** and non- repudiation for XML documents.

  –

XML STANDARDS –

- **XHTML**
  - It is a revised version of HTML with rules from XML.

The Idea of Markup

- In XML, tag containg values. This **tags specify certain ruels.**

- Therefore, we can say that **XML is nothing but a set of rules.**

- XML has been written in s**uch a manner that it can be extended easily, depending on the business domain, particular sets of requirements or technology.**

The Idea of Markup

- XML is based on yet another language, called as Standard Generalised Markup Language (SGML).

- SGML is the parent of almost all important modern markup languages.

```
          Standard Generalised Markup Language (SGML)
                              |
          +-------------------+-------------------+
          |                                       |
Hyper Text Markup Language (HTML)      Extensible Markup Language (XML)
```