XML
&
Allied Technologies

Presented by :

Eng. Mohamed A. Sayed

October 2022

Agenda

- ✓ What is XML?
- ✓ Why XML?
- ✓ Uses of XML.

What is XML?

- **eXtensible Markup Language.**
- A simple, standard way to interchange structured textual data between computer programs.
- XML describes a syntax for marking up documents so that **complex structures** may be easily described.

Before XML

- **Binary files.**
- **Text files.**
- **The rise of Markup.**

Markup Languages

- **Examples of Markup languages:**
  - ✓ SGML.
  - ✓ HTML.
  - ✓ XML.

SGML - History

- Standard Generalised Markup Language
- 1969 - GML from IBM
- 1980 SGML first published
- 1986 - ISO standard
- It is a standard for how to specify a document markup language or tag set.

SGML

- SGML is itself a document type definition (***DTD***). It is a standard for how to specify a tag set.
- SGML is **not** in itself a document language, but a description of how to specify Language. It is ***metalanguage***.
- SGML documents contain structural elements that can be described without consideration of how they are displayed.

HTML

- Hyper Text Markup Language.
- An application of **SGML** that could be used to create what we now know as web pages.
- It describes the structure of text-based information in a document .
- It describes the appearance of a document.

The need for extensibility

- **Problems with SGML:**
  - ✓ Complexity.
  - ✓ SGML specification is more than 150 pages that cover unlikely scenarios.
  - ✓ No S/W has fully implemented the specs.
- **Problems with HTML:**
  - ✓ Fixed set of tags.
  - ✓ HTML was not designed for current use (new devices).
  - ✓ Poor at representing specialised data: Maths, Music.

What XML is?!

- XML is derived as a subset of SGML, but is substantially simplified.

- XML is a metalanguage:
  - ✓ Doesn't have a fixed set of tags and elements.
  - ✓ Syntax may *optionally* be described by a DTD.

- Style and content are completely separate
  - ✓ XML documents contain content.
  - ✓ Style is specified by *stylesheets*.

What XML is?! (cont.)

- **XML is :**
  - ✓ Language independant.
  - ✓ Platform independent.
  - ✓ Application independent.
  - ✓ Foundation for several next-gen Web Technologies (XHTML , RSS, AJAX, Web services, XAML).

Advantages of XML

- **XML :**
  - ✓ Uses human ,not computer Language.
  - ✓ Is readable and understandable (by both humans and machines).
  - ✓ 100% portable.
  - ✓ Extensible.
  - ✓ Hierarchical Data Representation

XML Document Example

```
<?XML version="1.0"?>
<!--A very simple XML document -->
<ITITrainees>
        <Trainee Group=“1”>
                <Name> Nour Sherkawy </Name>
                <Friend Name=“Mohamed” Group=“1”/>
        </Trainee>
        <Trainee Group=“2”>
                 <Name> Hagar Ali </Name>
        </Trainee>

</ITITrainees>
```

What XML is not

- **No compiler** that generates executables.
- XML is **not a network protocol**, i.e another S/W has to do the sending while data can be stored as XML.
- XML is **not a database**, data can be stored in an XML format, but the engine has to exist.

Uses of XML

- **Configuration files.**
- **Web Services.**
- **Web content.**
- **Business Interoperability:**
  (Business-to-Business transactions &Business-to-Consumer transactions)
- **Database Systems.**
- **Image Representation.**

Content vs. Style

- XML tags contain **meaning *not* appearance**.

- Consider the example of the scientific names of animals.

  ✓ scientific names are in Latin and they are always printed in italics.

The *scientific* name of the domestic dog is *Canis familiaris*, and of the domestic cat is *Felis catus*.

Content vs. Style

In HTML :

```
<P>The <I>scientific</I>
name of the domestic dog
is <I>Canis familiaris</I>,
and of the domestic cat
is <I>Felis catus.</I></P>
```

NB:
there is no distinction between scientific names and emphasis.

The *scientific* name of the domestic dog is *Canis familiaris*, and of the domestic cat is *Felis catus*.

Content vs. Style

In XML :

```
The scientific
name of the domestic dog
is <Dog>Canis familiaris</Dog>,
and of the domestic cat
is <Cat>Felis catus.</Cat>
```

The *scientific* name of the domestic dog is *Canis familiaris*, and of the domestic cat is *Felis catus*.

Stylesheets

- Style in XML is defined by stylesheets.
- Stylesheets define the **physical appearance** of a document, and its **behaviour**.

- Stylesheet languages
  - ✓ **CSS** (Cascading StyleSheets)
    developed for HTML.
  - ✓ **XSL**
    developed specifically for XML.

XML Break.....

- **What are the main problems with HTML?**
- **How can we control the appearance of the XML document?**