Theme 1: XML

Introduction of XML

World Wide Web

- There are over 1 billion websites on the World Wide Web today
- There is an uncountable amount of data on the web.
- Most of that data is in the form of HTML docs.
- We use HTML to display data in a web browser.
- HTML is too **limited** to manage the vast growth of info.

SGML

- **S**tandard **G**eneralized **M**arkup **L**anguage (SGML) - 1986
- SGML is a text language that enables you to describe your content in a way that is independent of hardware, software, formats, or operating system.
- HTML was used to create markup and display it using browser technology
- It could not semantically describe what the content was…thus encouraged the development of XML

XML and HTML

- XML, with HTML, act as the **foundational** web language.
- **Today, XML is…**
  - ...a popular method for **storing** data and...
  - ...the most popular method for **transmitting** data between all sorts of systems/apps.

Why both?

- HTML was designed to **display** data, not manage it.

- XML was designed to take over data management.

- **In this module we will cover…**
  - The basics of XML
  - Additional and supporting languages and systems of XML
  - Application of XML
  - Other related technologies

What is XML?

- XML stands for e**X**tensible **M**arkup **L**anguage.

- XML is a specification for...
  - ...**storing** info and...
  - ...**describing the structure** of that info.

- XML is a **markup language** like HTML.

What is XML?

- An HTML author must use HTML tags, e.g. <body>.
- XML has **no tags of its own**.
- The XML author creates whatever tags they need.
- The new tags must follow the rules in the **XML specification**.
- We will cover these rules later.

XML Example

- Look at the following example, then ask yourself:
  - What **data** does the XML document store?
  - What is the **structure** of the data?
  - What **tags** were created to describe the data and its structure?

```
<?xml version="1.0"?>
<contacts>
  <contact>
    <name>
   <first>Daddy</first>
   <middle>Long</middle>
   <last>Legs</last>
    </name>
    <location>
   <latitude>50.7218</latitude>
   <longitude>-3.5462</longitude>
    </location>
    <phone>+2778 312 3341</phone>
    <related>Gary the Snail, Marley the Blob</related>
    <description>
   He likes long walks on beaches
   He likes romantic movies on Tuesdays
   He likes candle lit bath sessions
    </description>
  </contact>
</contacts>
```

Daddy Long Legs
50.7218:3.5462
(+27)78 312 3341
Gary the Snail; Marley the Blob

He likes long walks on beaches
He likes romantic movies on Tuesdays
He likes candle lit bath sessions

What is XML?

- The **XML specification** is...
  - ...a set of rules...
  - ...for defining custom markup languages.

- A **custom markup language** is...
  - ...an author-created **XML tag set** that...
  - ...describes and structures **specific info**.

The Power of XML

- Why use XML?
  - What does XML do that existing technologies don’t?

- **Reason 1:**
  - XML was explicitly designed for data storage & transmission.

    - HTML is unsuitable because it was designed only to display info.

The Power of XML

- **Reason 2:**
  - XML is easily **extended** and **adapted**.
    - You can define your own custom tag set (markup language).
    - The tags can be re-used in other XML docs, scaled back, added to, etc.

The Power of XML

- **Reason 3:**
  - You can use XML to share data between **disparate** systems.
    - XML is platform-independent because it is plain text.
    - It is well-structured, easy to understand, easy to parse (i.e. machine-readable), easy to manipulate, and human-readable.
    - Interoperability

The Power of XML

- **Reason 4:**
  - XML is **non-proprietary** and free for anyone to use.
    - The W3C created XML and controls it, but nobody owns it.
    - XML data sharing not only bypasses platform barriers but also governmental barriers.

Extending XML

- Use a **schema** to standardise a custom XML language.
- A schema is a doc that contains **structural definitions**.
- It specifies what tags a custom XML language may use and what content and attributes they may contain.
- A schema ensures that different people using the same XML language create docs with the exact same structure.

Extending XML

- Apps can be programmed to understand and retrieve data from docs of a standardised XML language.

- Write a schema using a **schema language**.
  - You can use either **DTD** or **XML Schema**.

- You can use **XML Namespaces** to extend XML Schemas.

- We will cover schemas and namespaces in detail later.

Extending XML

- You can use **XSL** to specify how to **display** an XML doc.

- **XSL** stands for e**X**tensible **S**tylesheet **L**anguage.

- **XSL comprises three languages:**
  - **XSLT** for transforming XML docs;
  - **XSL-FO** for formatting an XML doc.
  - **XPath** for identifying different parts of an XML doc;

Extending XML

- With XSL, you can convert XML info into another format.

- **XML info is most often converted into…**
  - …HTML to display the data, or…
  - …an XML doc with a different structure from the original.

- We will cover XSL in detail later.

Theme 1: XML Introduction

- TO BE CONTINUED…