What is XML?

* XML stands for EXtensible Markup Language
* XML is a markup language much like HTML
* XML was designed to carry data, not to display data
* XML tags are not predefined. You must define your own tags
* XML is designed to be self-descriptive
* XML is a W3C Recommendation

The Difference Between XML and HTML

XML is not a replacement for HTML.

XML and HTML were designed with different goals:

* XML was designed to transport and store data, with focus on what data is
* HTML was designed to display data, with focus on how data looks

HTML is about displaying information, while XML is about carrying information.

XML Does Not DO Anything

Maybe it is a little hard to understand, but XML does not DO anything. XML was created to structure, store, and transport information.

The following example is a note to Tove, from Jani, stored as XML:

<note>  
<to>Tove</to>  
<from>Jani</from>  
<heading>Reminder</heading>  
<body>Don't forget me this weekend!</body>  
</note>

The note above is quite self descriptive. It has sender and receiver information, it also has a heading and a message body.

But still, this XML document does not DO anything. It is just information wrapped in tags. Someone must write a piece of software to send, receive or display it.

With XML You Invent Your Own Tags

The tags in the example above (like <to> and <from>) are not defined in any XML standard. These tags are "invented" by the author of the XML document.

That is because the XML language has no predefined tags.

The tags used in HTML are predefined. HTML documents can only use tags defined in the HTML standard (like <p>, <h1>, etc.).XML allows the author to define his/her own tags and his/her own document structure.XML is Not a Replacement for HTML

**XML is a complement to HTML.**

It is important to understand that XML is not a replacement for HTML. In most web applications, XML is used to transport data, while HTML is used to format and display the data.

My best description of XML is this:

**XML is a software- and hardware-independent tool for carrying information.**

XML is a W3C Recommendation

XML became a W3C Recommendation on February 10, 1998.

XML is Everywhere

XML is now as important for the Web as HTML was to the foundation of the Web.

XML is the most common tool for data transmissions between all sorts of applications.

XML Separates Data from HTML

If you need to display dynamic data in your HTML document, it will take a lot of work to edit the HTML each time the data changes.

With XML, data can be stored in separate XML files. This way you can concentrate on using HTML/CSS for display and layout, and be sure that changes in the underlying data will not require any changes to the HTML.

With a few lines of JavaScript code, you can read an external XML file and update the data content of your web page.

XML Simplifies Data Sharing

In the real world, computer systems and databases contain data in incompatible formats.

XML data is stored in plain text format. This provides a software- and hardware-independent way of storing data.

This makes it much easier to create data that can be shared by different applications.

XML Simplifies Data Transport

One of the most time-consuming challenges for developers is to exchange data between incompatible systems over the Internet.

Exchanging data as XML greatly reduces this complexity, since the data can be read by different incompatible applications.

XML Simplifies Platform Changes

Upgrading to new systems (hardware or software platforms), is always time consuming. Large amounts of data must be converted and incompatible data is often lost.

XML data is stored in text format. This makes it easier to expand or upgrade to new operating systems, new applications, or new browsers, without losing data.

XML Makes Your Data More Available

Different applications can access your data, not only in HTML pages, but also from XML data sources.

With XML, your data can be available to all kinds of "reading machines" (Handheld computers, voice machines, news feeds, etc.), and make it more available for blind people, or people with other disabilities.

XML is Used to Create New Internet Languages

A lot of new Internet languages are created with XML.

Here are some examples:

* XHTML
* WSDL for describing available web services
* WAP and WML as markup languages for handheld devices
* RSS languages for news feeds
* RDF and OWL for describing resources and ontology
* SMIL for describing multimedia for the web
* All XML Elements Must Have a Closing Tag
* In HTML, some elements do not have to have a closing tag:
* <p>This is a paragraph.  
  <br>
* In XML, it is illegal to omit the closing tag. All elements **must** have a closing tag:
* <p>This is a paragraph.</p>  
  <br />
* **Note**: You might have noticed from the previous example that the XML declaration did not have a closing tag. This is not an error. The declaration is not a part of the XML document itself, and it has no closing tag.
* XML Tags are Case Sensitive
* XML tags are case sensitive. The tag <Letter> is different from the tag <letter>.
* Opening and closing tags must be written with the same case:
* <Message>This is incorrect</message>  
  <message>This is correct</message>
* XML Elements Must be Properly Nested
* In HTML, you might see improperly nested elements:
* <b><i>This text is bold and italic</b></i>
* In XML, all elements **must** be properly nested within each other:
* <b><i>This text is bold and italic</i></b>
* In the example above, "Properly nested" simply means that since the <i> element is opened inside the <b> element, it must be closed inside the <b> element.
* XML Documents Must Have a Root Element
* XML documents must contain one element that is the **parent** of all other elements. This element is called the **root**element.
* <root>  
    <child>  
      <subchild>.....</subchild>  
    </child>  
  </root>
* XML Attribute Values Must be Quoted
* XML elements can have attributes in name/value pairs just like in HTML.
* In XML, the attribute values must always be quoted.
* Study the two XML documents below. The first one is incorrect, the second is correct:
* <note date=12/11/2007>  
    <to>Tove</to>  
    <from>Jani</from>  
  </note>
* <note date="12/11/2007">  
    <to>Tove</to>  
    <from>Jani</from>  
  </note>