Writing in XML

Was this written in XML?

Initially it was written in a word processor. But the word processor saved it in XML (the content) and other files which included an XSL (extensible style sheet) – so I suppose you could say it was written in XML after all! It was produced in OpenDocument format (rather than MS XML) which meant turning it into another form (this pamphlet) was very easy.

RSS

You might be wondering where we go from here. The gate is open and the horses are already loose – we could go anywhere from here! That's exciting. But it might help to go back a step. One of the simplest and most effective examples of this 'markup language' approach can be seen in RSS or Really Simple Syndication. RSS is based on a simple markup language that takes the news content from a website, declares that it is RSS, then gives the headline (which it calls <title>), a brief one or two line description (called, intuitively <description>) then provides a link (called <link>) which takes you to the whole article. The result is that you can catch up with a whole raft of news items by reading the headline, a line or two summarising the content – and then make the decision to read the whole item if you want to. This approach has revolutionised news reading. RSS does the hardwork to help you keep up with many websites in a very short space of time. RSS is the 'shiny apple' of Web 2.0

What do I use to write in XML?

If you use OpenOFfice Writer it does it for you. If you wish to 'get your hands dirty' with simple text writing, use any text editor you like - Notepad will do it! If you wish to be sure you are writing well-formed, so-called valid XML, then use an XML editor - text editor which shows your markup in colour and validates it for you. At that point you will want to do more - XML by itself does absolutely nothing! It preapres your text to be interpreted by something else. That's where XSLT comes in but that's another story.

What can you do to make what you do a witness to everyone?.

Direzione Generale
Opere Don Bosco
Via della Pisana 1111
00163 - Rome, Italy
Email:jbfox@sdb.org

Salesians of Don Bosco

Learning a new language - markup

With the courage of Don Bosco

Writing for Web 2.0

You may have heard people occasionally talking about Web 2. 0. What on earth is Web 2.0? Do we even know what Web 1.0 is? There possibly are not any absolute definitions of Web 2,0. but it is not hard to guess that it is supposed to be a step ahead of, an advance on Web 1.0. What we can say is that Web 2.0 emerged after the dot com bubble burst just a few years back – at least that's when the term was invented.
We can describe Web 2.0 though, by noting the differences between it and Web 1.0. Here are just a few of the differences:

Web 1.0	Web 2.0
Britannica online	Wikipedia
Personal websites	Blogging
Publishing	Participation
Content Management Systems	Wikis
Directories	Tagging

There's more, much more, but maybe the above gives you a clue. And common to all the above is markup language, but Web 2.0's markup is different to Web 1.0's markup. Web 2.0's is mostly semantic or descriptive markup. Web 1.0's is largely HTML which initially tried to be descriptive but became a mess.

Semantic markup?

If you were writing a dictionary, you would need a way to describe how each term and its etymology, then its definition, should be structured.
A way to do this, especially on a computer screen would be:
<entry>markup
<etym>from the verb 'mark', but the term mark up was already in use centuries ago by printers who made notes on the side of a text to prepare for printing
<def> making marks on paper or screen that indicate what a printer or publisher needs to do with the text.
Take no notice of the definition and whether it is accurate – but what you are looking at is semantic markup. Self-descriptive markup might be another way to describe it. A human being can already guess what <etym> means.
But it is also structural markup – note how the dictionary entry has been structured, and every entry will follow the same structure.

Are there other kinds of markup?

Many. Microsoft Word and other word processing programs mix structural/semantic markup with presentational markup – often code that says what the font should be, whether bold or italics and so forth. These codes are NOT readable by humans. They produce good pages. They can only be used to produce good pages – you cannot send them off in that format to other media.

So semantic, descriptive markup is a good thing?

It is a very good thing! It can be read by human beings to begin with – so if something goes terribly wrong with your computer, you have not lost your content. A computer can be taught to read the tags (those things with <....>) which can be collected, ordered in particular ways. A computer can be taught to interpret both their meaning and their order.
And best of all – if your text is marked up semantically and structurally, descriptively, that is, it can be sent to a range of different media: to print, to screen (small, very small, big, very big), to voice, to braille, to.....you name it. Descriptive markup might be the best term to include both meaning and structure. You or a computer can then add a stylesheet for presentation according to the media you are sending the content to.

Are there still other kinds of markup?

Yes. You've heard of postscript which has been the standard for the printing industry – this is procedural markup usually meaning it has lots of macros and routines built into it (pure computer programming stuff). There is HTML which built the web. The problem is that HTML was asked to do much more than it possibly could. HTML as such is now dead in the way Latin is dead – it exists, people still learn it and use it, but it is no longer under development. Instead.....

...Instead, the 'X' word!

Instead, there is XHTML, XML, XSL, XSLT... The 'X' stands for eXtensible. We can keep adding to it. HTML tended to mix content and presentation a bit too much, and used tags that were only partly clear. Take <h1> for example. That means Heading 1. XHTML is likely to avoid <h1> as not good descriptive language and prefers <title> which is very clear and will certainly be what an <h1> was. An<h2> would be or could be a <subtitle>.