

XML

XML stands for Extensible Markup Language.

Extensible means capable of being stretched-out or extended.

Markup is nothing but the instructions which are often called as tags.

XML is subset of SGML (Standard Generalized Markup Language)

It is used to describe the structure of a document not the way that is presented.

HTML

- Predefined tags define how to present the data

- Missing end tags are permitted. Ex: `
`

- Attributes don't require quotation marks (" ")
eg ``

- HTML tolerate non-nested tags
`<p><h1>abc</p></h1>`

XML

User defined tags to identify the data.

It requires matching closing tags Ex: `
...</br>`

Attributes require quotation marks - either single or double
eg ``

It requires properly nested tags.
`<p><h1>abc</h1></p>`

5 - Browsers may do a well-formedness violation best guess on not well-formed html causes parser to generate error.

6 - HTML is not case-sensitive - XML is case sensitive.
<h1>xyz </h1>

Basic Rules of XML

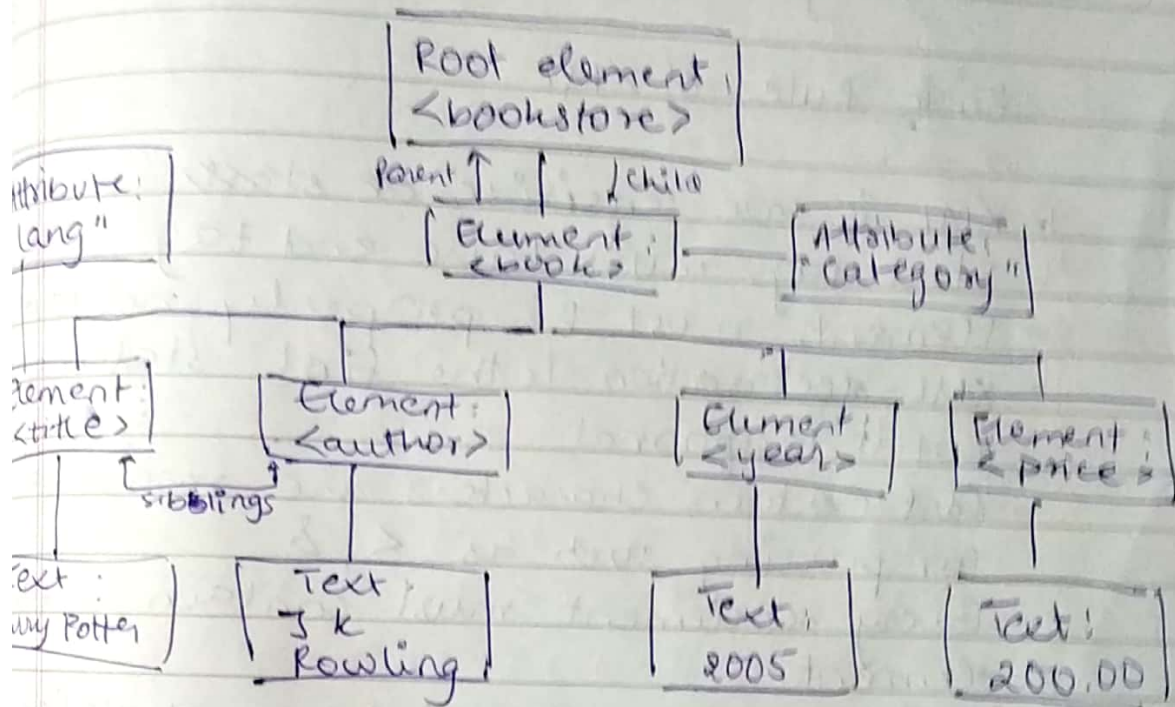
- 1) - XML is case-sensitive ~~or start~~
- 2) - Start tags must have end tag
- 3) - Elements must be properly nested.
- 4) - XML declaration is the first statement but it is optional.
- 5) - ~~But~~ Certain characters are reserved for parsing such as <, &
- 6) - Every document must contain a root element.
- 7) - Attribute value must have quotation marks.

Advantages of XML

- 1) - It provides a basic syntax that can be used to share information between different types of computers in different applications & organizations.
- 2) - It provides a gateway for all communication b/w applications of different systems.
- 3) - It supports ~~the~~ unicode, allowing ~~code~~ any information written in any human language.
- 4) - It can ~~be~~ represent CS, DS, & records list & trees.
- 5) - The ~~strict~~ strict syntax & parsing req. make the necessary parsing also extremely simple, efficient & consistent.

- It is mainly used as a format for document storage & processing it online.
- The hierarchical structure is suitable for most type of documents.
- It is platform independent.

XML Tree structure



XML well formed valid & invalid document

If all tags in a document are correctly formed & follow xml guidelines then the doc. is considered as well formed doc.

Valid XML doc

It is defined in the XML specifications as a well formed xml doc which also conforms to the rules of doc. type definition.