

# Understanding Markup Languages

## Chapter 01

### Topics

- Understanding markups
- Open, Closed markups
- Problems in markups
- Specific and generalized markup
- SGML

### Markups

Every day we use markup languages in design and printing world, and on media like web. Here we will try to gain an understanding of markups, where they are used and what purposes it is used for. We will also get introduced to different kinds of markups, their differences and so on.

### What is markup?

- It refers to marking up (tagging) document contents in the print and design world.

### What are the purposes of Markups?

- It is used for either of the two purposes:
  - To modify look and formatting of text
  - To describe the structure and meaning of a document for applications to process, for output to some medium such as printer or world wide web

### How Markups work?

- Markup consists of codes or tags - these are added to text content to change the look or meaning of the tagged text. Markup processor interprets the tag and acts on the content.
- For example - `<b>IDB-BISEW</b>` is a HTML tagged text which is used to change the text IDB-BISEW into bold faced. HTML processors (Web Browser) can interpret the tag and displays the tag content appropriately.
- Generally, a processor is used to interpret the marked-up document and display or process the content.

### Some common markups

- RTF - Rich Text Format
  - Word processors use RTF to format documents for print or display
- HTML - Hyper Text Markup Language
  - HTML is used to format and layout document in World Wide Web. HTML are generally processed by browsers.

### What is a markup processor?

- A markup processor interprets the meaning of the markup codes.
- The markup codes tell the processor just about everything about the application needs to know about it.
- A processor does not process all markups. For every different type of markup, you need a different type of processor. For example, HTML documents are processed by browsers, RTF documents are processed by word processors.

### What is the difference between an open markup language and a closed markup language?

- The Specifications for an open language is publicly available and any vendor can create a processor for it. For example - RTF, HTML.
- The specification for a closed language is not publicly available and vendors do not have access to the rules

of the markup language and cannot create a processor for it. For example - Microsoft Word uses a closed markup for its documents and Microsoft word processor only can process word documents.

### Problems in Markups like HTML

- Difficult to identify the structure of the document
- Looseness in structure makes it difficult to interpret
- Lack of rules makes difficult to create a processor for it

### Specific and generalized markup

- Specific markup languages
  - Specific markup languages are used to generate code that is specific to particular application and device.
  - They are often built to serve a particular need.
  - For example - RTF is used to format texts for printing and displaying, HTML is used to format documents for web.
- Generalized markup
  - Generalized markup languages are used to describe the structure and meaning of a document but they do not define how the document should be processed.
  - SGML, XML are generalized markup languages.

### SGML

- Stands for Standard Generalized Markup Language
- SGML is used to define the structure of a document but not its formatting
- It is extensible - it allows an author to create his/her own structure
- It is actually a meta-language - it allows you to create your own language.
- Both HTML and XML originated from SGML. XML is a subset of SGML. HTML is an application of SGML

### HTML vs. SGML

- SGML
  - Generalized language
  - Extensible
  - Used to describe the structure of a document
- HTML
  - Not generalized
  - Not extensible
  - Used to format and layout documents on web
  - There are also differences in coding