

XML

Introduction

- XML stands for **EX**tensible **M**arkup **L**anguage.
- XML was designed to store and transport data.
- XML was designed to be both human- and machine-readable.

The Difference Between XML and HTML

- XML and HTML were designed with different goals:
 - XML was designed to carry data - with focus on what data is
 - HTML was designed to display data - with focus on how data looks
 - XML tags are not predefined like HTML tags are

How Can XML be Used?

- XML is used in many aspects of web development.
- XML is often used to separate data from presentation.

XML Separates Data from Presentation

- XML does not carry any information about how to be displayed.
- The same XML data can be used in many different presentation scenarios.
- Because of this, with XML, there is a full separation between data and presentation.

XML Document - example

```
<?xml version="1.0"?>
```

```
<contact-info>
```

```
  <name>Tanmay Patil</name>
```

```
  <company>TutorialsPoint</company>
```

```
  <phone>(011) 123-4567</phone>
```

```
</contact-info>
```

Document Prolog

Document Elements

Document Prolog Section & Elements

Section

- The **document prolog** comes at the top of the document, before the root element. This section contains:
 - XML declaration
 - Document type declaration
- **Document Elements :**
 - The building blocks of XML.
 - These divide the document into a hierarchy of sections, each serving a specific purpose.
 - Multiple sections so that they can be rendered differently, or used by a search engine.
 - The elements can be containers, with a combination of text and other elements.

XML Declaration

- Following syntax shows XML declaration:

```
<?xml
```

```
  version="version_number"
```

```
  encoding="encoding_declaration"
```

```
  standalone="standalone_status"
```

```
?>
```


parameter

Parameter	Parameter_value	Parameter_description
Version	1.0	Specifies the version of the XML standard used.
Encoding	UTF-8, UTF-16, ISO-10646-UCS-2, ISO-10646-UCS-4, ISO-8859-1 to ISO-8859-9, ISO-2022-JP, Shift_JIS, EUC-JP	It defines the character encoding used in the document. UTF-8 is the default encoding used.
Standalone	yes or no.	It informs the parser whether the document relies on the information from an external source, such as external document type definition <i>DTD</i> , for its content. The default value is set to <i>no</i> . Setting it to <i>yes</i> tells the processor there are no external declarations required for parsing the document.

Declaration Rules

- If the XML declaration is present in the XML, it must be placed as the first line in the XML document.
- If the XML declaration is included, it must contain version number attribute.
- The Parameter names and values are case-sensitive.
- The names are always in lower case.
- The order of placing the parameters is important. The correct order is: *version, encoding and standalone*.
- Either single or double quotes may be used.
- The XML declaration has no closing tag i.e. `</?xml>`

XML Declaration Examples

Following are few examples of XML declarations:

XML declaration with no parameters:

```
<?xml >
```

XML declaration with version definition:

```
<?xml version="1.0">
```

XML declaration with all parameters defined:

```
<?xml version="1.0" encoding="UTF-8" standalone="no" ?>
```

Element

Syntax

Following is the syntax to write an XML element:

```
<element-name attribute1 attribute2>  
....content  
</element-name>
```

XML Element Rules

- Following rules are required to be followed for XML elements:
 - An element *name can contain any alphanumeric characters. The only punctuation mark* allowed in names are the hyphen – , under-score _ and period . .
 - Names are case sensitive. For example, Address, address, and ADDRESS are different names.
 - Start and end tags of an element must be identical.
 - An element, which is a container, can contain text or elements as seen in the above example