

Encoding is the process of converting unicode characters into their equivalent binary representation. When the XML processor reads an XML document, it encodes the document depending on the type of encoding. Hence, we need to specify the type of encoding in the XML declaration.

Encoding Types

There are mainly two types of encoding:

- UTF-8
- UTF-16

UTF stands for *UCS Transformation Format*, and UCS itself means Universal Character Set. The number 8 or 16 refers to the number of bits used to represent a character. They are either *8onebyte* or *16twobytes*. For the documents without encoding information, UTF-8 is set by default.

Syntax

Encoding type is included in the prolog section of the XML document. The syntax for UTF-8 encoding is as below:

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

Syntax for UTF-16 encoding

```
<?xml version="1.0" encoding="UTF-16" standalone="no" ?>
```

Example

Following example shows declaration of encoding:

```
<?xml version="1.0" encoding="UTF-8" standalone="no" ?>
<contact-info>
  <name>Tanmay Patil</name>
  <company>TutorialsPoint</company>
  <phone>(011) 123-4567</phone>
</contact-info>
```

In the above example **encoding="UTF-8"**, specifies that 8-bits are used to represent the characters. To represent 16-bit characters, **UTF-16** encoding can be used.

The XML files encoded with UTF-8 tend to be smaller in size than those encoded with UTF-16 format