Ngôn Ngữ XML

XML là gì

XML là viết tắt của từ **eXtensible Markup Language**, hay còn gọi là ngôn ngữ đánh dấu mở rộng do [W3C](http://www.w3.org/) đề nghị với mục đích tạo ra các ngôn ngữ đánh dấu khác. Đây là một tập hợp con đơn giản có thể mô tả nhiều loại dữ liệu khác nhau nên rất hữu ích trong việc chia sẻ dữ liệu giữa các hệ thống. Ví dụ khi bạn xây dựng một ứng dụng bằng C# và một ứng dụng bằng PHP thì hai ngôn ngữ này không thể hiểu nhau, vì vậy ta sẽ sử dụng XML để trao đổi dữ liệu.

Các tính chất của XML:

* XML là ngôn ngữ có tính chất đánh dấu
* Được thiết kế để lưu trữ và truyền tải dữ liệu
* Xuất hiện cuối thập niên 90
* Theo chuẩn W3C từ năm 1998
* Có thể tự mô tả dữ liệu
* Có thể chứa dữ liệu, nhưng lại không thể hiển thị
* Không được định nghĩa trước. Bạn phải tự định nghĩa các thẻ XML
* Là một ngôn ngữ và nền tảng độc lập

Để việc học tập có hiệu quả, và để giúp bạn hứng thú hơn với bài học, trước khi bắt đầu loạt bài này, bạn nên có sự hiểu biết nhất định về HTML và JavaScript.

**Khai báo XML**

## Cú pháp

Cú pháp cơ bản của khai báo XML là:

<?xml

version="version\_number"

encoding="encoding\_declaration"

standalone="standalone\_status"

?>

## Qui tắc khai báo XML

Một khai báo XML phải tuân theo các qui tắc sau:

* Nếu có khai báo XML, nó phải được đặt trong dòng đầu tiên của tài liệu XML.
* Nếu có khai báo XML, nó phải chứa thuộc tính version number.
* Tên và giá trị các tham số phải là phân biệt kiểu chữ.
* Tên luôn luôn ở dạng chữ thường.
* Thứ tự đặt xác tham số là quan trọng. Thứ tự đúng là: version, encoding và standalone.
* Trích dẫn đơn hoặc kép có thể được sử dụng.
* Khai báo XML không có thẻ đóng, ví dụ: </?xml>

### Ví dụ khai báo XML

Dưới đây là một số ví dụ về khai báo XML:

Khai báo XML không có tham số:

<?xml >

Khai báo XML với định nghĩa phiên bản:

<?xml version="1.0">

Khai báo XML có tất cả các tham số:

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

Khai báo XML có tất cả các tham số trong dấu trích dẫn đơn:

<?xml version='1.0' encoding='iso-8859-1' standalone='no' ?>

## Qui tắc thẻ trong XML

### Qui tắc 1

Thẻ trong XML là phân biệt kiểu chữ. Dòng code sau là một cú pháp sai </Address>, bởi vì hai thẻ này có kiểu chữ khác nhau.

<address>This is wrong syntax</Address>

Dòng code sau là cú pháp đúng, với thẻ mở và thẻ đóng có cùng kiểu:

<address>This is correct syntax</address>

### Qui tắc 2

Các thẻ trong XML phải được đặt trong một thứ tự chính xác, ví dụ: một thẻ XML được mở trong phần tử khác phải được đóng trước khi phần tử bên ngoài đó được đóng. Ví dụ:

<outer\_element>

<internal\_element>

This tag is closed before the outer\_element

</internal\_element>

</outer\_element>