

1.XML

-Ứng dụng của XML

Truyền dữ liệu

Mọi ứng dụng đều cần một số cách để lưu trữ và truy xuất dữ liệu. Điều này thường xảy ra qua Internet bằng cách sử dụng API (Giao diện lập trình ứng dụng). [Back-End Engineers](#) tạo ra các API chạy trên các máy chủ web.

Nhiều ứng dụng có thể sử dụng cùng một API, bao gồm ứng dụng web, ứng dụng dành cho máy tính để bàn và ứng dụng dành cho thiết bị di động, để lưu và truy cập dữ liệu trong cơ sở dữ liệu. Một định dạng tiêu chuẩn cho dữ liệu này làm cho điều này trở nên khả thi.

XML là một định dạng mà các lập trình viên sử dụng để truyền dữ liệu trong một cấu trúc có thể được phân tích cú pháp bởi tất cả các ứng dụng đa dạng này và nó thường được sử dụng để tạo các API. SOAP và XML-RPC là hai loại API XML được sử dụng trong các dịch vụ web. Bất kỳ ứng dụng nào kết nối với một trong hai API này chỉ phải biết định dạng để sử dụng dữ liệu mà nó chứa.

Định dạng tài liệu

Các trang web là các tài liệu HTML và [HTML](#) rất giống với XML. HTML được xử lý bởi một trình duyệt web, trình duyệt này sau đó trình bày nó ở một định dạng trực quan dễ chịu. Các thẻ trong tài liệu HTML xác định các loại phần tử cụ thể, như tiêu đề, đoạn văn, hình ảnh và hơn thế nữa. Trình duyệt biết cách hiển thị các phần tử này dựa trên các thẻ này.

HTML cũng lưu trữ thông tin ngữ cảnh về dữ liệu mà nó chứa dưới dạng các thuộc tính bao gồm ID và lớp. [CSS](#) làm việc với HTML để áp dụng các kiểu cụ thể dựa trên các thuộc tính này, chẳng hạn như làm cho tiêu đề có màu đỏ hoặc đặt phông chữ cho một đoạn văn.

Đây chỉ là một ví dụ về việc sử dụng XML để định dạng. Tập PDF, tập PostScript, tài liệu Microsoft Word, tài liệu PowerPoint và tập văn bản RTF cũng được lưu trữ dưới dạng XML. Khi bạn mở các tệp này trong ứng dụng

mặc định của chúng, nó sẽ phân tích cú pháp XML này, định dạng và cung cấp cho nó phong cách mà bạn thấy trên màn hình máy tính của mình.

Tìm kiếm trên web

Công cụ tìm kiếm đã phát triển trong nhiều năm. Ban đầu, họ không làm gì nhiều hơn ngoài việc xác định xem một trang web có chứa cụm từ bạn đang tìm kiếm hay không. Giờ đây, họ sử dụng các thẻ HTML (XML) để làm cho các tìm kiếm chính xác hơn.

Một ví dụ sẽ là tìm kiếm một cuốn sách của tác giả yêu thích của bạn. Giả sử bạn đang tìm kiếm Mark Twain. Bằng cách phân tích cú pháp <author> thẻ trong các trang HTML, công cụ tìm kiếm có thể giới hạn kết quả tìm kiếm chỉ với những kết quả có Mark Twain trong thẻ này thay vì mọi trang có chứa tên của anh ấy.

Tạo bố cục

Mọi bố cục trong ứng dụng di động Android đều được tạo bằng XML. Các bố cục này xác định vị trí dữ liệu sẽ được hiển thị trên màn hình điện thoại. Các bố cục Android phổ biến bao gồm Bố cục Tuyến tính, cho ứng dụng căn chỉnh nội dung trên màn hình theo chiều ngang hoặc chiều dọc, Bố cục Khung, được thiết kế để chứa các bố cục khác một cách động và bố cục Danh sách, hiển thị các mục bạn có thể cuộn qua.

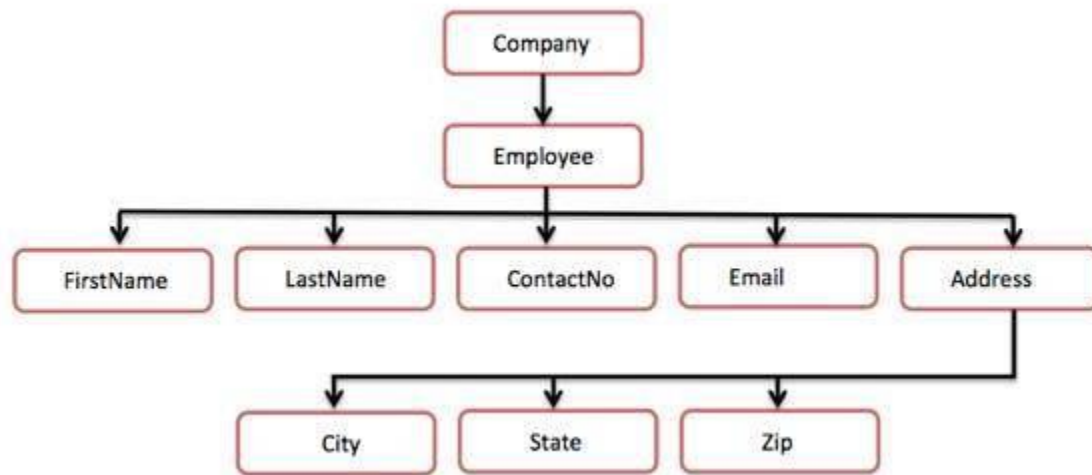
Lưu trữ dữ liệu cấu hình

XML cũng lưu trữ dữ liệu được sử dụng để cấu hình một ứng dụng. Trong Microsoft Excel, XML chứa tất cả thông tin có trong bảng tính. Không chỉ dữ liệu, mà còn cả định nghĩa của các cột, định dạng của các trường, bất kỳ phép tính nào chúng sử dụng và hơn thế nữa. Ứng dụng Android không chỉ sử dụng XML cho bố cục mà còn để lưu trữ màu sắc, kiểu và kích thước mà ứng dụng sẽ sử dụng.

-Cấu trúc của XML

+Là cấu trúc dạng Tree: Một tài liệu XML luôn luôn có tính mô tả. Cấu trúc cây (Tree Structure) thường được biết đến với tên gọi **Cây XML** và đóng một vai trò rất quan trọng để mô tả bất kỳ tài liệu XML nào một cách dễ dàng.

Cấu trúc cây chứa các phần tử gốc (hay phần tử cha), các phần tử con và ... Bằng việc sử dụng cấu trúc cây, bạn có thể biết tất cả các nhánh kế tiếp và các nhánh con bắt đầu từ phần tử gốc. Việc phân tích cú pháp bắt đầu tại gốc, sau đó di chuyển qua nhánh đầu tiên tới một phần tử, và tiếp tục tới các node khác.



Trong sơ đồ trên, có một phần tử gốc là `<company>`. Bên trong đó, có thể có nhiều hơn một phần tử `<Employee>`. Bên trong phần tử employee, có 5 nhánh là `<FirstName>`, `<LastName>`, `<ContactNo>`, `<Email>`, và `<Address>`. Bên trong phần tử `<Address>`, có ba nhánh phụ là `<City>`, `<State>` và `<Zip>`.

-Một số thư viện hỗ trợ XML trong java :SAX,DOM,StAX,JAXB

2.HTML

-HTML, hoặc Ngôn ngữ đánh dấu siêu văn bản, được sử dụng để tạo các trang web. Tác giả trang web sử dụng HTML để định dạng văn bản dưới dạng tiêu đề và tiêu đề, để sắp xếp đồ họa trên trang web, để liên kết đến các trang khác nhau trong một trang web và để liên kết đến các trang web khác nhau.

-Cấu trúc của một tài liệu HTML

Tài liệu HTML chủ yếu được chia thành hai phần:

- HEAD : Phần này chứa thông tin về tài liệu HTML. Ví dụ, Tiêu đề của trang, phiên bản HTML, Dữ liệu meta, v.v.
- BODY : Phần này chứa mọi thứ bạn muốn hiển thị trên Trang Web.

-Một số thư viện hỗ trợ phân tích HTML trong java như: DOMparser, Cheerio, jsdom, Parse5