**Web:**

là một tập hợp các trang web con, bao gồm văn bản, hình ảnh, video, flash v.v.. WEBSITE chỉ nằm trong một tên miền hoặc tên miền phụ lưu trữ trên các máy chủ chạy online trên đường truyền World Wide của Internet.

Một trang web là tập tin HTML hoặc XHTML có thể truy nhập dùng giao thức HTTP hoặc HTTPS.

WEBSITE có thể được xây dựng từ các tệp tin HTML (trang mạng tĩnh) hoặc vận hành bằng các CMS chạy trên máy chủ (trang mạng động). WEBSITE có thể được xây dựng bằng nhiều ngôn ngữ lập trình khác nhau (PHP,.NET, Java, Ruby on Rails...).

Các website hoạt động trên giao thức HTTP

Tập hợp những quy tắc, quy ước truyền thông đó được gọi là giao thức

Client: Có thể là trình duyệt web, ứng dụng khác (crawler), chrome, firefox, ie

Server: Ứng dụng trên máy tính lưu trữ và xử lý website

GET và POST là các Giao thức HTTP để truyền dữ liệu từ client lên server.

**Flask**

Flask là một micro web framework khá nổi tiếng viết bởi Python. Nhanh, nhỏ nhẹ, đơn giản, nếu tập tành viết web application với Python thì Flask là 1 sự lựa chọn thích hợp.

Cài đặt flask:

$ pip3 install flask

**HTML**

HTML là chữ viết tắt của HyperText Markup Language, có nghĩa là ngôn ngữ đánh dấu siêu văn bản. Nó dùng để định dạng bố cục, các thuộc tính liên quan đến cách hiển thị của một đoạn text và được hiển thị trên một chương trình đặc biệc ta gọi là Browser. Hiện nay có khá nhiều Browser như Firefox, Chrome, Cốc Cốc, ...Tất cả Browser đều có điểm chung là giúp người dùng thao tác với website và nó đều có khả năng biên dịch những đoạn mã HTML, CSS và Javascript.

Web tĩnh:

- Sử dụng hoàn toàn HTML, chỉ hiển thị nội dung không tương tác

- Dùng khi nội dung ít cập nhật, tiết kiệm chi phí

- Ưu điểm: Tốc độ truy cập nhanh, rẻ, nhẹ

- Nhược điểm: Khó quản lý nộ dung, bảo trì, nâng cấp

Web động:

- Là một trang web có hình ảnh, âm thanh, văn bản, các phần truy xuất và xử lý thông tin. Luôn có thông tin mới được cập nhật.

- Dùng làm web tin tức, blog cá nhân, thương mại điện tử, bán hàng

- Ưu điểm: Dễ nâng cấp và bảo trì, xây dựng web lớn, tương tác cao

- Nhược điểm: Chi phí xây dựng bảo trì cao

Một số note khác:

Crawler

Là:

    1 module hoặc phần mềm, tự động bóc tách phân tích dữ liệu từ nguồn, lấy những gì ta cân lấy thôi

    từ 1 dữ liệu k có cấu trúc thành có cấu trúc

    so sánh dữ liệu, phân tích đối thủ

    là Ngôn ngữ đánh dấu k phải là ngôn ngữ lập trình

    tạo thông tin trên world wide web

Thư viện request (tập hợp các hàm màn mình dùng luôn), cần import để dùng

Lấy HTML

Bóc tách dữ liệu (tập trung những gì cần lấy)

Lưu trữ dữ liệu (để hiện thị)

pip lấy một đoạn code trên internet kéo về

dùng beautiful soup: tìm kiếm các thành phần

Json = mô tả bằng chuỗi

id - thuộc tính của thẻ

thêm thuộc tính sẽ là dạng từ điển

.text (lấy hết các thẻ bên trong, + dồn các thẻ khác **và** .string (chỉ lấy text khi k có gì khác bên trong)

html parser là công cụ: (XML - ngôn ngữ mô tả dữ liệu)

    báo cho thằng kia là dữ liệu add vào là html chứ k phải xml

    xml là dạng chuẩn

    xpath - cách tìm kiếm trên xml search trong f12

//input[@id='q']

Test tự động, chỉ ng thật tương tác ms hiện ra (lazada)-selenium (tự động hoàn toàn như ng thật-test tự đông, hợp cho các trang phức tạp) - dạng exe

**Python - phân tích dữ liệu**

**PHP - chuyên giao diện**

**FRONT END: HTML (CSS + JS) - Hiển thị**

**BACK END: PYTHON - Xử lý nghiệp vụ**

Host: địa chỉ nhà

Port: cửa đi vào nhà 1-65355

int not callable - k call dc, đầu vào là chuỗi đầu ra là chuỗi

"""  để ghi j thì ghi bên trong đó """

Jinza2

{{}} dùng để chiết xuất các biến

Web Động, có tương tác, thông tin đc cập nhật

Tạo 1 trang web động?

    Client - Server (Máy khách - máy chủ)

* Máy chủ: xử lý các yêu cầu
* Máy khách: chrome, mobile

**-> Trả về html để hiển thị**

http protocal: giao thức truyền tải html giống trong chuyển mail

GET - POST (Dữ liệu truyền)

    read data rồi insert data

làm sao để post: dùng form (các ô để nvien điền vào -> nhân viên để họ làm cho khách hàng)

route k có gì thì mặc định là GET, vừa có POST vừa có GET

MongoDB Atlas, Python3