



# 크롤링(Crawling)?

웹 상에 데이터를 긁어 모으는 작업



#### 크롤링 가능 사이트?

우리가 브라우저를 통해 볼 수 있는 모든 웹 사이트

#### |Python크롤링 |**크롤링**



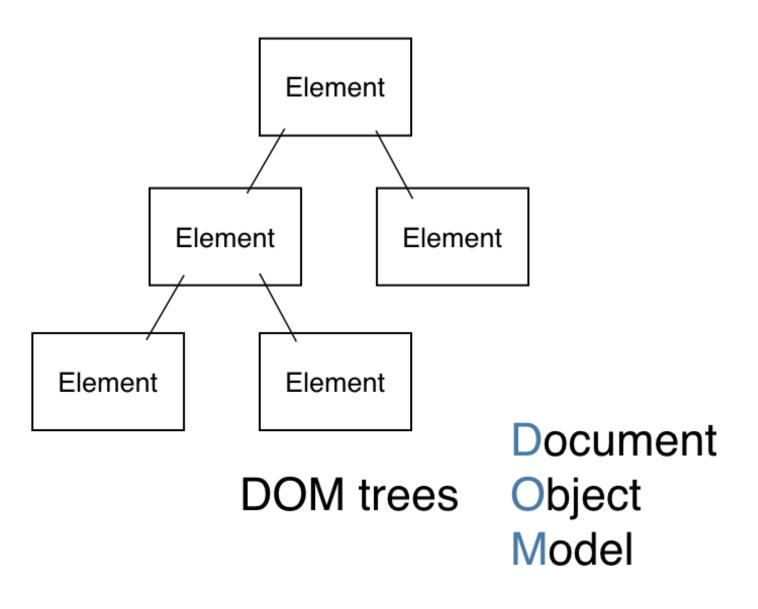
## **HTTP**













- 1. 웹 사이트 구조 분석
- 2. 크롤링 코드 작성
- 3. 파일 저장 및 데이터 분석



## Open API

기업(정부)에서 특정 부분에 대해 개발 권한을 주는 것



## Open API

기업(정부)에서 특정 부분에 대해 개발 권한을 주는 것



## Open API

- 많은 양의 데이터를 상대적으로 빠르게 가져올 수 있다.
- API로 제공하지 않는 데이터는 가져올 수 없다.
- 많은 양의 데이터가 필요할 경우에 요청 금액이 발생한다.



## 크롤링 시나리오

웹 상에 어떤 데이터를 어떻게 가져올지 정하는 작업



#### 크롤링 시나리오

웹 사이트마다 다른 방식으로 분석 필요 크롤링할 웹 사이트를 미리 접속하고 URL 등을 분석