

Architecture

1. 브라우저 플랫폼, 기본 셋팅



Data Collection
Data 저장/변환 } → 80% 차지.

Data 분석

Data 모델링
머신러닝.



하드웨어

jsoup is an open-source Java library designed to parse, extract, and manipulate data stored in HTML documents.

- ① Data Collection → HTML 파일
- 스크래핑 (웹 크롤링)
 - 마이닝 (텍스트마이닝)
- Not HTML 파일

파일
* .HTML → 인터프리터 언어 (보라넷)

! * .HTML → 자바어스럼



자바어스럼 → 마이닝.

인터프리터 언어 스럼 → 스크래핑 / 크롤링

JSoup : 자바 라이브러리.

= 한눈에 ↔ Casting

(parse (해킹/변환))

extract

manipulate → immutable

"Stored Data"

from HTML Document (가져옴) ⇒ DOM

Document 클래스



DOM의 데이터 클래스

Element 클래스



Dom<E>

JAVA의 Document 클래스
타입변환 해주어야 함.

지연된 로딩, 레이지 로딩 (lazy loading), 게으른 로딩은 필요 시점까지 객체의 초기화를 연기시키기 위해 컴퓨터 프로그래밍에 흔히 사용되는 디자인 패턴의 하나이다. 적절하게 사용될 경우 프로그램의 운영 차원에서 효율성에 기여할 수 있다. 그 반대의 경우 열정적인 로딩 (즉시 로딩, eager loading)이라고 부른다. 웹 페이지 등의 케이스에서 네트워크 콘텐츠가 접근되고 초기화 시간을 최소로 유지해야 하는 상황에서 이상적이다.

필요시점 객체 생성: `POJO = Member m = new Member()` 즉시 로딩, 디폴트.
미리 객체 생성: `Spring = 어노테이션`
→ 필요한 때만 만들 때 → @Lazy ↔ @Eager loading

문조건 X → 필요한 상황에서만 작성하거나 ex) 크롤링 (항상 해야 쓰면 되기 때문)
Because: 반응이 느려.

error vs exception.

① Error

서버를 Down 시키는 원인

1) 컴파일 에러

IDE에서 바로 감지.

2) 런타임 에러

가) 에러

→ 무조건.

서버를 완전히 Down

fly, catch throws

나) 예외

→ 막아준다

서버를 다음 시점에서 복구할 수 있도록