

ELASTICSEARCH

1. ELASTICSEARCH 개요

2. ELASTICSEARCH 사용 사례

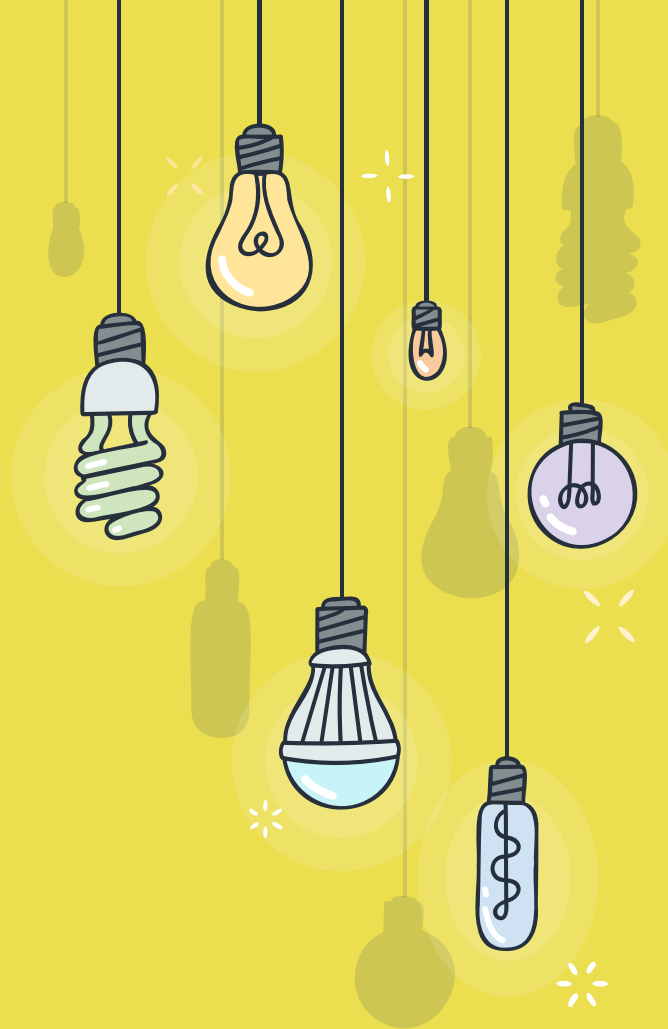
3. 주요 용어

4. 관계형 DB와 ELASTICSEARCH 구조 비교



1

ELASTICSEARCH



# \* ELASTICSEARCH 개념[1/2]

- + 실시간 데이터를 저장, 분석 및 검색하는 엔진
  - × 아파치 루씬(Appache Lucene) 사용하여 개발
    - ◆ 루씬 : 전문(full-text) 색인과 검색 기능을 지원하는 확장가능한 고성능 정보검색 java API
  - × JSON 문서 기반 저장 및 검색, HTTP 프로토콜의 RESTful API 지원

# \* ELASTICSEARCH 개념[2/2]

## + ELK 스택으로 제공

- × E (Elasticsearch): 데이터 저장 및 검색
- × L (Logstash): 데이터 수집
  - ◆ 로그스태시는 데이터의 입력, 변환, 출력을 실시간 파이프라인으로 처리하는 오픈 소스 데이터 수집 엔진
  - ◆ 다양한 입력 소스에서 동시에 데이터를 수집(Ingest)하여 변환한 후 자주 사용하는 "스태시(Stash) -보관소" (엘라스틱서치)로 전송
- × K (Kibana): 데이터 시각화 및 리포팅 분석
  - ◆ 키바나(Kibana)는 엘라스틱서치와 연동하여 동작하는 데이터 시각화 (리포팅) 및 분석

# \* ELASTICSEARCH 사용 사례

## + 위키피디아 (WIKIPEDIA)

- × 전문검색(FULL TEXT SEARCH)를 수행
- × 실시간 타이핑 검색
- × 추천 검색어 기능

## + 깃허브 (GITHUB)

- × 1,300억 줄이 넘는 소스 코드를 검색하는 데 사용

## + 더 가디언 (THE GUARDIAN)

- × 방문객의 로그 분석
- × 소셜 데이터 생성 및 분석으로 실시간 응대
- × 기사에 대한 반응 분석

# \* 관계형 DB VS ELASTICSEARCH

[관계형 데이터베이스]

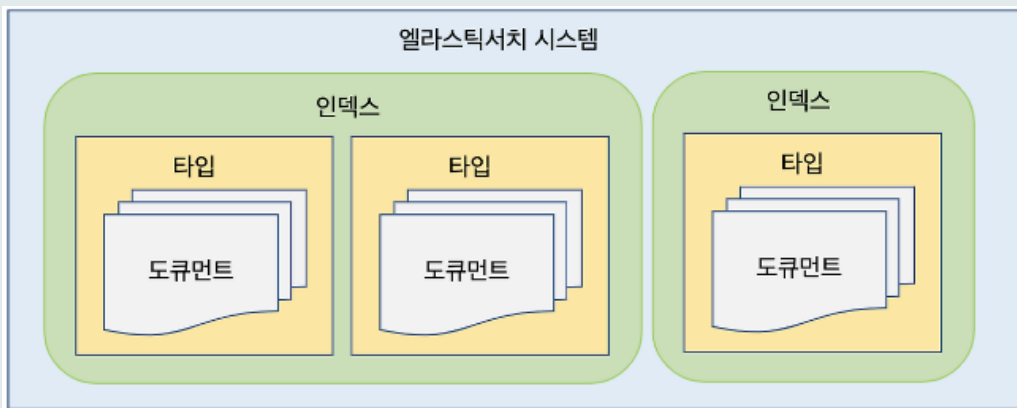
ID	addr
doc1	서울특별시 가산동
doc2	경기도 수원시
doc3	인천광역시 송도동
doc4	서울특별시 구로구
doc5	경기도 안양시

[엘라스틱서치]

Term	ID
서울특별시	doc1, doc4
가산동	doc1
경기도	doc2, doc5
수원시	doc2
인천광역시	doc3
송도동	doc3
구로구	doc4
안양시	doc5

# \* 주요 용어

- + 인덱스(index) : 명사
  - × 색인된 데이터를 다큐먼트(document)로 저장하는 장소
  - × 관계형 데이터베이스에서 데이터베이스에 해당
- + 색인(indexing) : 동사
  - × 원본 문서를 검색어 토큰들로 변환하여 저장
- + 데이터 구조는 인덱스, 타입, 다큐먼트 단위로 구성





# \* 관계형 DB와 엘라스틱서치 구조 비교

## + 구조 비교

관계형 DB	엘라스틱서치
데이터베이스(Database)	인덱스(index)
테이블(Table)	타입(Type)
행(Row)	도큐먼트(Document)
열(Column)	필드(Field)
스키마(Schema)	매핑(Mapping)

# THANKS!

+ Any questions?

