Elastic Stack

- description : 엘라스틱 스택 관련 내용 정리
- author : 주레피
- email : dhan@repia.com
- lastupdate : 2021-10-12

Intro

검색 엔진인 동시에 데이터베이스(NoSQL)

데이터의 양과는 무관하게 1초 이내의 응답 속도 기대

News

벡터 검색

```
무차별 대입 벡터 유사성 지원(Brute force vector similarity) - 7.3 GA
 - 용량에 따른 복잡도 증가

HNSW(Hierarchical Navigable Small Words) & KNN (K-Neighbor Nearest) - 8.0
 - 대용량 처리를 위해 약간의 정확도(accuracy) 희생
 - 대용량 인덱스에서 쿼리 지연 감소

Lucene를 많이 사용함
```

허깅 페이스 (파이토치 모델)

```
Eland (엘란트) - 머신러닝 모델 관리

웹 크롤러도 있음
7.x와 8.x는 많은 변화가 있음
```

```
머신 러닝 지원 내용
 허깅 페이스에서 모델 다운로드
 -> 도큐먼트 학습
 -> 질문 쿼리를 벡터로 변경
  -> _knn_search 벡터로 검색
```

Case Study

Term

로그스태시 (요르단 시셀)

```
로그 전송 및 가공
200여개 이상의 확장 가능한 플러그인 사용
JRuby 언어로 작성
유사제품군
  트레저데이터의 플루언트디
```

비츠 (베를린 패킷비트)

```
데이터 전송 및 가공
Go 언어로 작성되었음, 경량 수집기
```

엘라스틱서치 (샤이 배넌)

```
특징
데이터 저장 & 분석 & 검색 엔진
모든 레코드를 JSON형태로 입력하고 관리
데이터의 양과 무관하게 1초 이내의 응답 속도 기대
연관도에 따른 정렬
아파치 루씬을 토대로 개발 되었음
단점
저장 공간이 크게 압축되지 않고 시스템 리소스를 많이 차지
DSL(Domain Specific Language)를 채택하여 JOIN 쿼리가 쉽지 않음, 반정규화를 기본으로 채택
Java 언어로 작성
유사제품군
  솔라
```

키바나 (라시드 칸)

```
데이터 시각화 및 모니터링
엘라스틱 스택의 UI, 관리 기능, API 콘솔, 모니터링...
키바나 때문에 엘라스틱 서치를 채택할 정도로 강력한 솔루션을 포함하고 있음
Javascript 언어로 작성
유사제품군
  그라파나, 태블로
```

엘라시틱스택

```
유사제품군
  Splunk(스플렁크)
```

Company

루시드웍스(https://lucidworks.com/) - 미국 실리콘밸리 AI검색 플롯폼 기업

> 2007년 아파치 솔라(Apache Solr) 오픈소스 검색 프로젝트의 컨트리뷰터가 모여 만든 회사

AI 검색 플랫폼 퓨전 (Fusion = Apache Sola + Spark)

```
Fusion Server (수집, 색인, SQL, Rule, 스케줄링)
Fusion AI (추천 검색, 의미 검색, 자동분류, 예외처리, 클러스트링, 토픽추출, 자동학습)
Fusion App Studio (Front-end)

Reddit, Lenovo, Morgan Stanley, goop, Rei

2020.11.29 - 클라우드 가상화 기업 틸론과 국내 총판 계약
2019.01.22 - 투비소프트, 루시드웍스 AI 검색솔루션 독점 총판계약 체결
```

아이브릭스 (사원수 70명, http://www.i-bricks.co.kr)

```
한국어 언어처리 전문업체, 엘라스틱 + 비츠, 로그스태시, 키바나 기반 openQuery SE (search edition) 제공
```

엘라스틱 (https://www.elastic.co/kr/)

```
루씬(Lucene)으로 시작하여 검색 엔진 -> NoSQL, 빅데이터 => 데이터 플랫폼
엘라스틱 클라우드 (요금 스탠더드(월 $16), 골드(월 $19), 플래티넘(월 $22), 엔터프라이즈(월 $30))
NETFLIX, Uber, Slack(기업용 메시지 앱, 협업 툴), twilio(트윌리오, 고객 커뮤니케이션 API 기업), Microsoft,
구글 클라우드, AWS, Azure에서 사용 가능
```

ElasticSearch AI(문서유사도) 관련

벡터 필드를 사용한 텍스트 유사도 검색(공식홈페이지)
엘라스틱서치로 문서 유사도 검색 구현하기
Elasticsearch를 이용한 문서 유사도 검색과 Springboot를 통한 구현

Install

Tip

Troubleshooting

Ref

- 지원 환경
- 라이선스별 기능 비교

주레피, 엘라스틱, 엘라스틱서치, 엘라스틱스택

From:
http://rwiki.repia.com/ - **문제를 잘 정의하는 것은 문제를 절반 해결한 것이다. - 2023.12**

Permanent link:
http://rwiki.repia.com/doku.php?id=wiki:3party:elasticsearch&rev=1661238070

Last update: **2022/08/23 16:01**