

요구사항 정의서

검색엔진 유사문서 우회 API

프로그램 설명 :

구글,네이버 검색엔진의 검색결과를 크롤링하여 자연어 처리 API를 통해 가공한 후 새로운 문서를 만들어내는 프로그램.

프로그램 로직 :

1. 기존 DB는 구현되어 있습니다.
2. Database 접속 후 크롤링할 키워드를 참조합니다.
3. 검색엔진(구글,네이버)에 키워드 검색 후 검색결과(네이버-VIEW, 구글-웹문서) 5개를 랜덤으로 파싱합니다.
4. 파싱한 데이터를 자연어처리 API를 통해 혼합하여 새로운 문서를 생성합니다.
5. Return값을 반환합니다.

예시 API 문서

API 기본 정보

메서드	요청 URL	출력포맷	비고
GET	localhost/article	JSON	

요청 변수(매개변수)

변수 이름	타입	필수여부	기본값	설명
keyword	string	Y	-	검색할 키워드를 설정합니다.
engine	string	N	google(기본값), naver	검색결과를 크롤링->파싱할 검색엔진(구글,네이버)을 결정합니다.
mode	string	N	spin(기본값), clean	출력할 모드를 설정합니다.
display	Integer	N	1(기본값), 최대 100	출력할 문서 개수를 정합니다.
power	Integer	N	1(기본값), 최대 10	자연어 처리 강도를 정합니다. 낮을 수록 읽기 편하고 높을 수록 읽기 어렵습니다.

출력 결과

필드	타입	설명
articleNo	Integer	생성된 우회문서의 번호입니다.
articleText	text	생성된 우회문서 텍스트입니다.
articleLength	integer	생성된 우회문서의 텍스트 길이 입니다.

예시)

요청 example 1

<https://example.com/article?keyword=강남피부과&engine=google&mode=spin&display=2&power=5>

응답 결과(JSON)

```
{
  {
    "articleNo" : 0,
    "articleText": "우리의 피부는 {자외선|햇볕}에 노출되어 {있어|있기 때문에} 색소
{질환|질병}에 약할 수 밖에 없습니다.",
    "articleLength": 71
  },
  {
    "articleNo" : 1,
    "articleText": "{우리|저희} 피부는 {자외선|햇빛}에 노출되어 {있기 때문에|있어서}
색소질환에 약할 수 {있습니다|있기 때문입니다}.",
    "articleLength": 71
  }
}
```

요청 example 2

<https://example.com/article?keyword=강남피부과&engine=google&mode=clean&display=2&power=3>

응답 결과(JSON)

```
{
  {
    "articleNo" : 0,
    "articleText": "우리의 피부는 자외선에 노출되어 있어 색소질환에 약할 수 밖에
없습니다.",
    "articleLength": 71
  },
  {
    "articleNo" : 1,
    "articleText": "우리 피부는 햇볕에 노출되어 있기 때문에 색소질환에 약할 수 있
습니다.",
    "articleLength": 71
  }
}
```