

사용자 매뉴얼		프로젝트 명 : 뉴스 빅데이터 분석 시스템 구축	
		OpenAPI 서비스	
단계명 : 운영	활동명 :		작업명 :
작성자 : 왕경문	작성일 : 2023-06-21		버전 : 1.49

topic_rank	내부적으로 순위를 부여하기 위한 일련 번호입니다. 숫자가 작을수록 중요한 이슈입니다.
topic_keyword	토픽랭크 알고리즘으로 뉴스에서 추출한 키워드

5. 연관어 분석 API

※ (기존) 워드 클라우드 → (변경) 연관어 분석

입력 키워드와 연관이 높은 키워드들을 지정한 검색 조건에 해당하는 문서들에서 실시간으로 추출하여 제공합니다.

※ 연관어 분석은 문서 내에서 사용자가 입력한 질의어를 기반으로 연관 단어를 추출하여 문서의 특징을 직관적으로 확인할 수 있도록 하는 분석기법입니다. (TOPIC RANK 알고리즘을 사용하며 하단 TOPIC RANK 알고리즘 참조)

5.1 요청 URL

https://tools.kinds.or.kr/word_cloud

5.2 요청 형식

```
{
  "access_key": "d3a80ae9-482c-41d0-9c39-146fe526e09d",
  "argument": {
    "query": "청와대",
    "published_at": {
      "from": "2019-01-01",
      "until": "2019-01-31"
    },
    "provider": [
      "경향신문",
    ],
    "category": [
      "정치>정치일반",
      "스포츠일반",
    ],
    "category_incident": [
      "범죄",
      "교통사고",
    ],
    "byline": "",
    "provider_subject": [ "사회" ]
  }
}
```

사용자 매뉴얼		프로젝트 명 : 뉴스 빅데이터 분석 시스템 구축
		OpenAPI 서비스
단계명 : 운영	활동명 :	작업명 :
작성자 : 왕경문	작성일 : 2023-06-21	버전 : 1.49

query(필수값)

관련된 키워드를 추출하기 위한 키워드를 지정합니다.

published_at

연관 키워드를 추출하기 위한 검색 기간을 지정합니다.

- 날짜 형식은 한국 시간을 기준으로 YYYY-MM-DD 형식으로 지정합니다.

예약어	설명
from	조회 시작일을 지정합니다. 지정한 일자를 포함합니다.
until	조회 종료일을 지정합니다. 지정한 일자를 제외합니다.

provider

지정한 언론사에서 게재한 뉴스만을 분석합니다.

- 명칭 또는 코드 값으로 지정할 수 있습니다.
 - 여러 언론사들을 지정한 경우 OR 조건으로 같이 검색하여 결과를 제공합니다.
- ※ 코드 값은 하단 코드 테이블 참조

category

한국언론진흥재단에서 지정한 뉴스 통합 분류체계 값을 조건으로 지정합니다.

- 명칭 또는 코드 값으로 지정할 수 있습니다.
 - 여러 분류 값을 지정한 경우 OR 조건으로 같이 검색하여 결과를 제공합니다.
- ※ 코드 값은 하단 코드 테이블 참조

category_incident

한국언론진흥재단에서 지정한 뉴스 사건/사고 분류체계 값을 조건으로 지정합니다.

- 명칭 또는 코드 값으로 지정할 수 있습니다.
 - 여러 분류 값을 지정한 경우 OR 조건으로 같이 검색하여 결과를 제공합니다.
- ※ 코드 값은 하단 코드 테이블 참조

byline

기자 이름을 조건으로 지정합니다.

- 기자 이름을 여러명 지정하기 위해서 OR 연산자를 사용 가능합니다.
- ※ 언론사에서 기자 이름을 제공한 News인 경우에만 검색이 가능합니다.

provider_subject

언론사에서 제공한 News의 주제 분류 값(코드)을 조건으로 지정 합니다.

- ※ 해당 주제 분류 값(코드)는 기사를 제공한 언론사와 시점에 따라 다르므로 통일된 값을 제공하지 않습니다.

사용자 매뉴얼		프로젝트 명 : 뉴스 빅데이터 분석 시스템 구축	
		OpenAPI 서비스	
단계명 : 운영	활동명 :		작업명 :
작성자 : 왕경문	작성일 : 2023-06-21		버전 : 1.49

※ 뉴스검색 API 2.2. 요청형식을 참고하세요.

5.3 출력 결과 형식

```
{
  "result": 0,
  "return_object": {
    "nodes": [
      {
        "id": 2,
        "name": "청원",
        "level": 3,
        "weight": 1.33
      },
      {
        "id": 3,
        "name": "유씨",
        "level": 3,
        "weight": 0.86
      },
      {
        "id": 4,
        "name": "누리꾼",
        "level": 3,
        "weight": 0.67
      },
      {
        "id": 5,
        "name": "명예훼손 혐의",
        "level": 3,
        "weight": 0.67
      }
    ]
  }
}
```

nodes

지정한 검색어와 연관된 키워드들을 반환합니다.

필드명	설명
name	토픽명칭을 반환합니다.
weight	토픽의 중요도를 수치로 반환합니다.(하단 TOPIC RANK 알고리즘 참고)
id	토픽들을 구분하기 위한 식별자입니다.
level	지정한 검색어와 토픽간의 연결 단계를 나타냅니다. 그래프 분석을 위한 예약어로 워드 클라우드에서는 무시하셔도 됩니다.