



06

[AI활용텍스트분석]

뉴스 빅데이터 플랫폼의 활용과 텍스트 분석



학습 내용

1. 뉴스 빅데이터 플랫폼 실습
2. 응용 서비스 및 텍스트 분석 실습



학습 목표

- 온라인 상에서 활용 가능한 뉴스 빅데이터 플랫폼인 빅카인즈를 인지하고 활용법을 설명할 수 있다.
- 형태소 / 개체명 분석 서비스를 통해 텍스트 분석 결과를 실습할 수 있다.



1

뉴스빅데이터 플랫폼 실습

- 1) 빅카인즈의 이해
- 2) 빅카인즈 체험하기

1) 빅카인즈의 이해



세종사이버대학교



기본 검색어를 입력해주세요



상세 검색

검색도움말



로그인

뉴스 분석

기획 분석

뉴스 보기

빅카인즈 활용

빅카인즈 소개



금일 뉴스 수집

910건

전체 뉴스 수집 70,079,001건

오늘의 이슈 ?

뉴스듣기



분석대상뉴스 3,315건 · 뉴스클러스터 392건

이슈별

언론사별

나의 관심뉴스

서울시장 보궐선거 야권 단일화
여론조사 시작
151건

박범계 법무부 장관 "수사지휘
이행 과정 유감"
76건

예방접종전문위 "AZ 백신 접종
지속" 권고
74건

경찰, LH 부동산 투기 의혹 수사
58건

대통령·민주당 지지율 34% 역대
최저치 기록
47건

전체 뉴스보기 >



서울신문

단일화 조사 첫날 '윗 내국동' 저격한 안철수...이준
석 "이게 새정치냐"

단일화 앞두고 테마주 요동? 오세훈, 안철수 "관심 안 됨"

"오세훈이나 안철수냐"...지지층 응답률이 승패 가를 듯

안철수 "내국동 문제 확산...오세훈, 후보 사퇴할 수도" 직격

안철수-오세훈, 단일화 앞두고 날선 신경전

오세훈 "윤석열·김동연·홍정욱·금태섭 등과 개혁우파 플랫폼 만...

오세훈-안철수 회동 불발...당사 방문 제안에는 김종인 역정

안철수, 윗 직격 "내국동 '셀프 보상' 의혹... 문제가 확산하고 있다"

관계도 분석

중앙일보

매일신문

매일신문

대구일보

매일신문

YTN

세계일보

미디어 스타트업

공모주제

빅카인즈 뉴스콘텐츠를 활용한 서비스 및
뉴스 빅데이터 분석 기술 개발



공모 홈페이지 바로가기 >

1) 빅카인즈의 이해



세종사이버대학교

종합일간지, 경제지, 지역일간지, 방송사 등을 포함한
기사DB에 빅데이터 분석 기술을 접목해 만든 뉴스 분석 서비스

1990년부터 시작한 KINDS (Korea Integrated News Database)가 전신

빅카인즈는 2016년 4월19일 첫 선

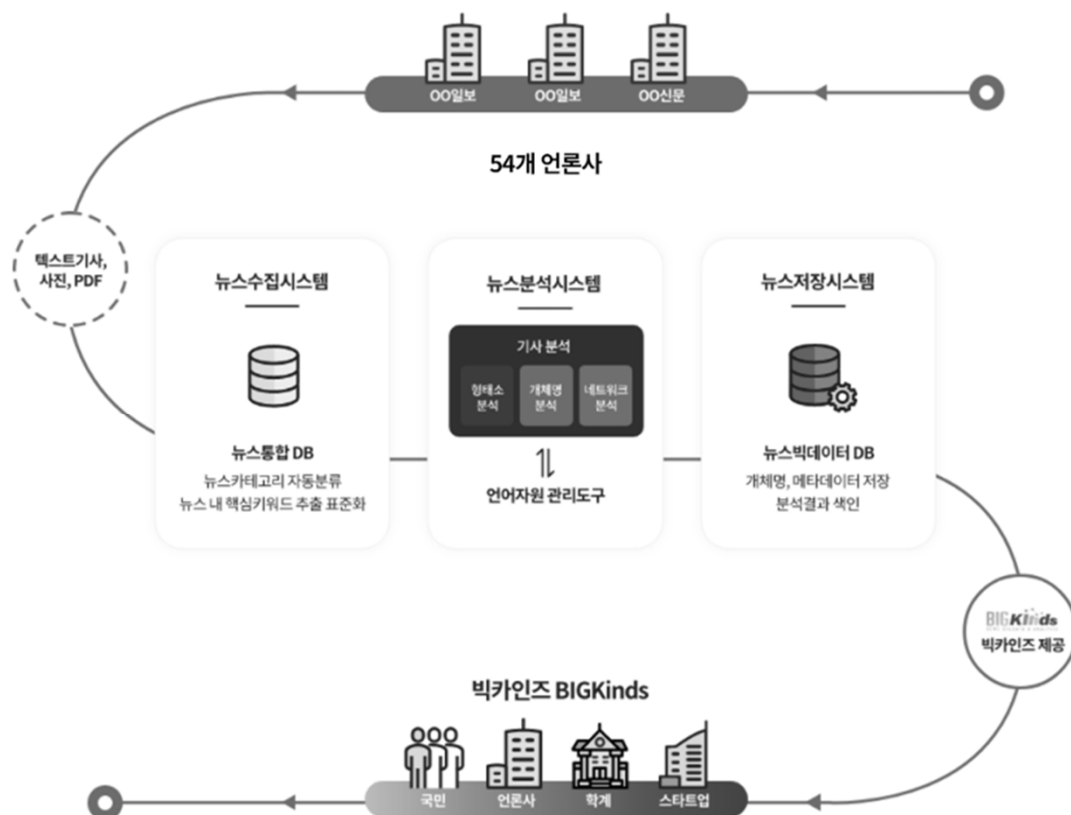
매일 54개 언론사의 콘텐츠가 자동으로 수집되고 분류됨

7천만건의 기사 집적(1990~2021)



서비스 개념도

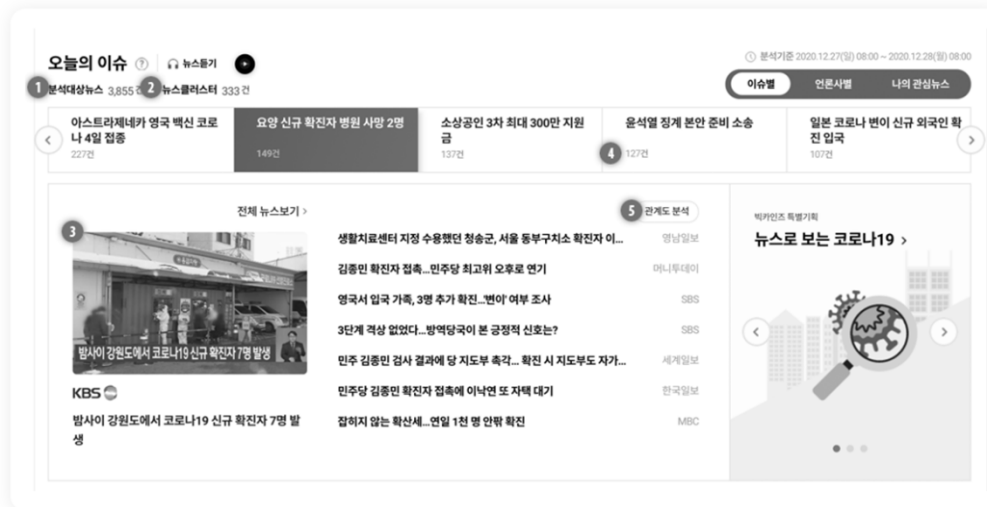
뉴스수집시스템, 분석시스템, 저장시스템 등으로 구성돼 있으며,
저장된 뉴스 분석정보는 국민, 언론사, 학계 스타트업 등이 활용할 수
있는 뉴스빅데이터 분석서비스 '빅카인즈(BIGKinds)'로 제공됩니다.





오늘의 이슈

매일 수집된 뉴스 중 뉴스클러스터링 결과 상위 8개의 이슈를 보여줍니다.



뉴스 클러스터링

항목설명

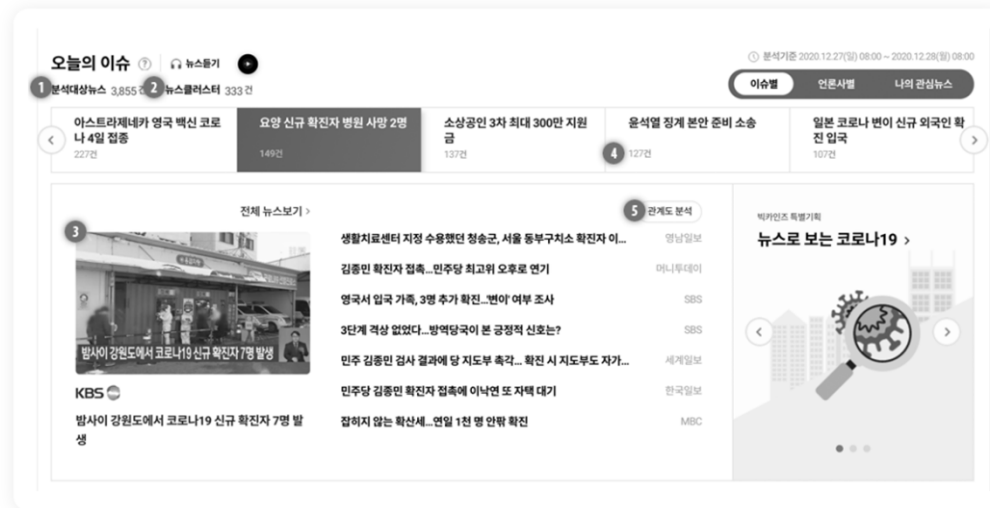
분석방법

- 형태소 분석과 문서유사도 분석을 사용하여 자동으로 유사한 뉴스를 클러스터링함
 - 형태소 분석과 문서 유사도 계산을 사용해 자동으로 유사한 뉴스를 묶어주는 기술임
 - 정확률 : 평균 82.43%
- 참고
 - 정확률 = (정확히 추출된 결과 정보의 수 / 전체 결과 정보의 수) × 100
 - 재현율 = (실제로 추출된 정확한 정보의 수 / 정확한 전체 결과의 수) × 100



오늘의 이슈

매일 수집된 뉴스 중 뉴스클러스터링 결과 상위 8개의 이슈를 보여줍니다.



뉴스 클러스터링

항목설명

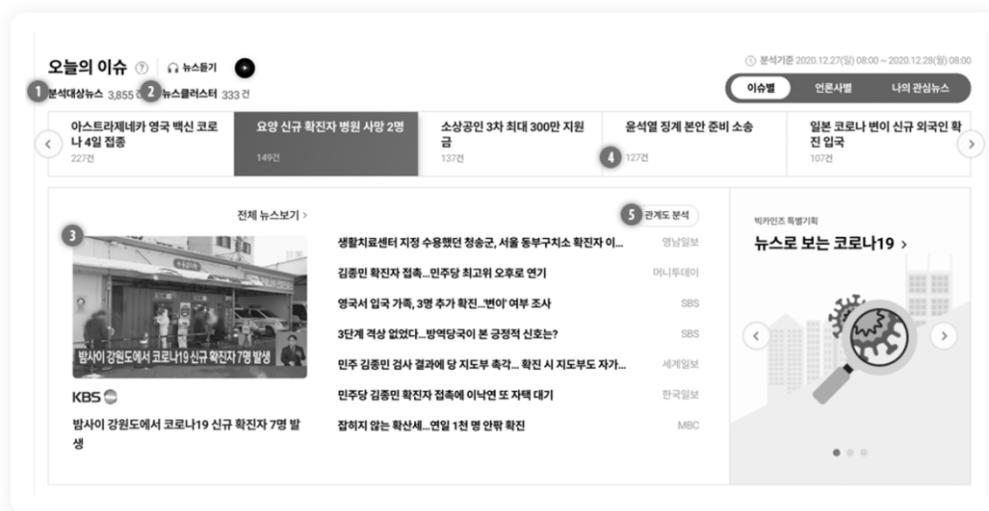
분석방법

- 오늘의 이슈를 도출하기 위해 분석된 뉴스 건수
- 뉴스클러스터의 개수(* 뉴스 클러스터 : 특정 이슈별로 묶인 뉴스의 그룹)
- 클러스터에 포함된 뉴스 건수 기준으로 최신순 8개의 이슈
- 특정 이슈에 클러스터로 묶인 뉴스 건수
- (관계도분석) 해당 이슈의 관계도 분석 등 시각화 분석



오늘의 이슈

매일 수집된 뉴스 중 뉴스클러스터링 결과 상위 8개의 이슈를 보여줍니다.



뉴스 클러스터링

항목설명

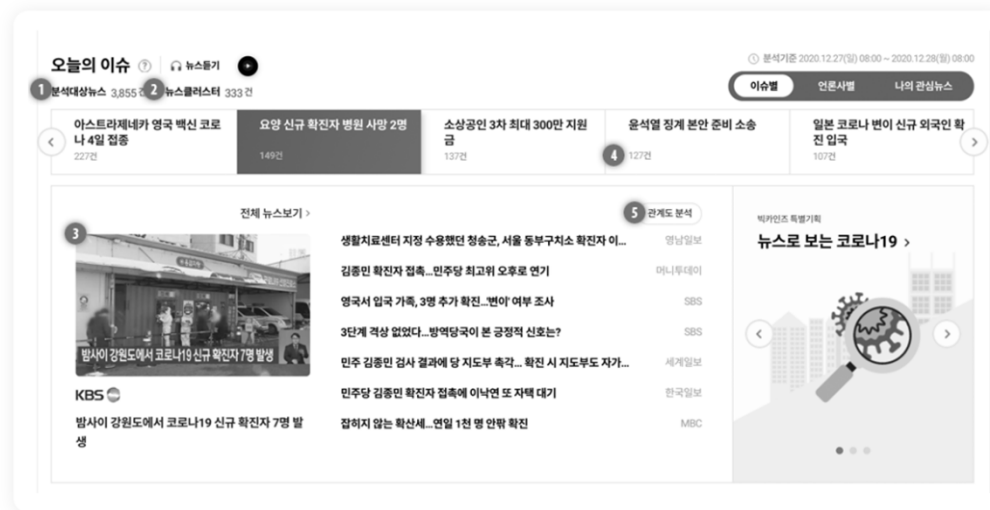
분석방법

- 분석주기 : 1일 2회 분석(08시, 17시)
 - 08시 : 전일 08시 ~ 금일 08시 수집뉴스 분석
 - 17시 : 금일 08시 ~ 금일 17시 수집뉴스 분석



오늘의 이슈

매일 수집된 뉴스 중 뉴스클러스터링 결과 상위 8개의 이슈를 보여줍니다.



뉴스 클러스터링

항목설명

분석방법

■ 분석절차

- 뉴스 내 중요 키워드 추출(토픽랭크 알고리즘)
- 유사한 뉴스를 클러스터링(코사인 유사도)
- 클러스터 내 대표 키워드를 추출해 오늘의 이슈 제목 구성



오늘의 키워드

매일 수집된 뉴스 속에서 주요 인물·장소·기관을 분석해 해당 키워드가 포함된 뉴스 건수가 높은 순으로 보여줍니다.
수집된 뉴스는 내용에 따라 정치, 경제, 사회, 문화 등 분야별로 확인 가능합니다.



개체명 분석

이용방법

분석방법

- 홍길동(인물), 서울특별시(장소), 국회(기관)와 같이 특정 의미 범주를 가지는 단어를 개체명이라고 하며, 이를 식별해내는 기술을 개체명 분석이라 함
- 개체명 분석에는 BERT(Bi-directional Encoder Representations from Transformer) 알고리즘을 사용하고 있음
- 정확률 : 89.76% / 재현율 : 93.33%

오늘의 키워드

매일 수집된 뉴스 속에서 주요 인물·장소·기관을 분석해 해당 키워드가 포함된 뉴스 건수가 높은 순으로 보여줍니다. 수집된 뉴스는 내용에 따라 정치, 경제, 사회, 문화 등 분야별로 확인 가능합니다.



개체명 분석

이용방법

분석방법

- 뉴스 분류(정치, 경제, 사회, 문화, 국제, 지역, 스포츠, IT 과학)
- 뉴스 내 포함된 개체명(인물, 장소, 기관) 목록, 해당 개체명 클릭 시 관련 뉴스 조회
- 개체명 색상 분류이며, 분류명 클릭 시 해당 분류 키워드를 제외할 수 있음



매일 수집된 뉴스 속에서 주요 인물·장소·기관을 분석해 해당 키워드가 포함된 뉴스 건수가 높은 순으로 보여줍니다. 수집된 뉴스는 내용에 따라 정치, 경제, 사회, 문화 등 분야별로 확인 가능합니다.



분석방법

- 1일 9회 분석(0시, 3시, 6시, 9시, 12시, 15시, 18시, 21시, 23시 30분)
- 오늘의 키워드는 0시를 시작으로 3시간 간격으로 당일 뉴스를 누적 집계하여 분석을 진행하며 마지막 분석은 23시 30분에 수행함
- 분석절차
 - 뉴스 내 개체명 추출
 - 발생 빈도순으로 개체명 정렬

1) 빅카인즈의 이해



세종사이버대학교

뉴스분석

빅카인즈 내 수집된 뉴스를 다양한
방법으로 검색한 후 분석 결과를 활용 및
시각화할 수 있는 서비스임



1) 빅카인즈의 이해



세종사이버대학교

뉴스분석 - 3단계



1) 빅카인즈의 이해



세종사이버대학교

뉴스분석 - 3단계



- 검색 조건, 상세검색, 나의 검색식을 통해 뉴스를 검색하는 단계
 - 검색조건 : 기간, 언론사, 통합분류, 사건사고 분류를 설정할 때 사용함
 - 상세검색 : 검색어 처리방식, 범위, 기고자 명 검색 등을 통해 결과 범위를 줄이고자 할 때 사용함
 - 나의 검색식 : 직접 구성하여 저장된 검색식을 활용해 검색함
-

1) 빅카인즈의 이해



뉴스분석 - 3단계



-
- 다양한 검색 필터를 통해 검색 결과의 범위를 선정하는 단계
 - 검색 필터 : 검색 필터 조건을 이용해 결과 범위를 필터링할 수 있음
 - 결과 내 재검색 : 특정 키워드를 활용해 기본검색의 결과 범위를 줄이고자 할 때 사용함
 - 분석 제외 : 목록에서 뉴스를 선택하여 분석에서 제외할 수 있음

1) 빅카인즈의 이해



뉴스분석 - 3단계



-
- 검색식 저장 : 뉴스검색 조건과 필터링 조건을 저장하여 나의 검색식에서 활용할 수 있음
 - 리포트 생성 : 검색 결과를 관계도 분석, 키워드 트렌드, 연관어 분석 등 시각화 리포트로 만듦

1) 빅카인즈의 이해



세종사이버대학교

뉴스분석 - 3단계



-
- 뉴스 메타정보, 뉴스 분석 결과 등은 엑셀파일로 다운로드 할 수 있음
 - 다운로드한 데이터는 R, SPSS 등 프로그램을 활용해 추가분석이 가능함
-

1) 빅카인즈의 이해



세종사이버대학교



데이터 다운로드	관계도 분석	키워드 트렌드	연관어 분석	정보 추출
----------	--------	---------	--------	-------

검색한 뉴스의 메타데이터(언론사, 기고자, 제목 등)와 개체명(인물, 기관, 장소 등) 분석 데이터를 엑셀파일로 제공하는 서비스입니다.

데이터 다운로드는 최대 20,000건의 데이터가 다운로드 됩니다. 미리보기는 최대 20개까지 보여집니다.

‘키워드’ 항목은 본문 내에서 추출된 키워드 중 단순 숫자(1, 2, 2018, 2019 등), 이메일 주소, 시간을 뜻하는 단어(밤, 낮, 새벽 등)를 제외한 결과가 표시됩니다.

	일자	언론사	기고자	제목	통합 분류1	통합 분류2
1	20210126	전자신문	조정형	강훈식, 중소기업뉴딜 지원? 소상공인지원뉴딜 지원법 동시 발의	경제>취업_창업	경제>자동차
2	20210126	충청일보	이한영 기자	계룡시, 지역상품권 10% 할인 판매	경제>서비스_쇼핑	지역>경남
3	20210126	충북일보	윤호노	충주시청 배드민턴 김주원 선수, 태극마크 달아	스포츠>올림픽_아시안게임	스포츠>농구_배드민턴
4	20210126	헤럴드경제	장연주 기자	[속보]교육부, 3월부터 유아·초등 1~3학년 '매일등교' 추진	사회>교육_시험	지역>경기
5	20210126	매일경제	박윤균	한국로슈진단 "애자일 경영으로 환자 중심 혁신 주력"	사회>의료_건강	IT_과학>과학
6	20210126	매일경제	박윤균	KERI, AI 안전인식 기술로 출입 정보 실시간 관리 시스템 개발	IT_과학>과학	경제>경제일반
7	20210126	서울신문	신성운	[서울포토] 경기도 기본주택 토론회 참석한 박영선-우상호	정치>정치일반	경제>경제일반
8	20210126	서울신문	김주리	"육섯·폭행에 피해자 사진 찍어" 사시생요임, 박법계 특수폭행 혐의 고발(종합)	사회>미디어	사회>사건_사고
9	20210126	아시아경제	이승진	빙그레, 무라벨 패키지 키피 '아카데미 심플리' 판매 100만개 돌파	경제>유통	경제>자동차
10	20210126	아시아경제	나주석	이중배 "정의당 대표 성추행에 무관용 언급한 민주당, 적반하장도 유분수"	정치>국회_정당	정치>행정_자치
11	20210126	아시아경제	이영규	경기교육청, 지하수 이용 15개 학교 '상수도' 지원	사회>환경	지역>경기
12	20210126	한국일보	정한별	나상도, '우리말 겨루기' 우승 "민망해"	미분류	지역>경기
13	20210126	아시아경제	영남취재본부 이상현	김해시, 방역 수칙 위반 아간신고 대응반 운영	지역>충남	지역>강원
14	20210126	아주경제	장은영	이재용 부회장 "삼성, 제 상황 관계없이 가야 할 길 계속 가야"	미분류	지역>강원
15	20210126	중앙일보	이해준(lee.hayjune@joongang.co.kr)	[속보] 이마트, 1352억원에 프로야구 SK와이번스 인수	스포츠>야구	스포츠>야구>한국프로야구
16	20210126	머니투데이	주명호[기자]	폭스바겐코리아, 세스코와 안전 방역환경 구축 위한 MOU	경제>서비스_쇼핑	경제>자동차
17	20210126	머니투데이	김지훈[기자]	서울 104명 확진...19일째 100명대	사회>의료_건강	경제>자동차
18	20210126	서울경제	강동호 기자 kdhho@sedaily.com	국토부, 도시재생 스마트기술 지원사업 공모	지역>대전	지역>충남
19	20210126	머니투데이	김건우[기자]	인앤스드바이오, 유방암 치료제 韓·美 임상 1.2상 승인	사회>의료_건강	IT_과학>과학
20	20210126	매일신문	이영광	겨울에 주로 발생 '도로위 지뢰' 포트홀	지역>충북	지역>강원

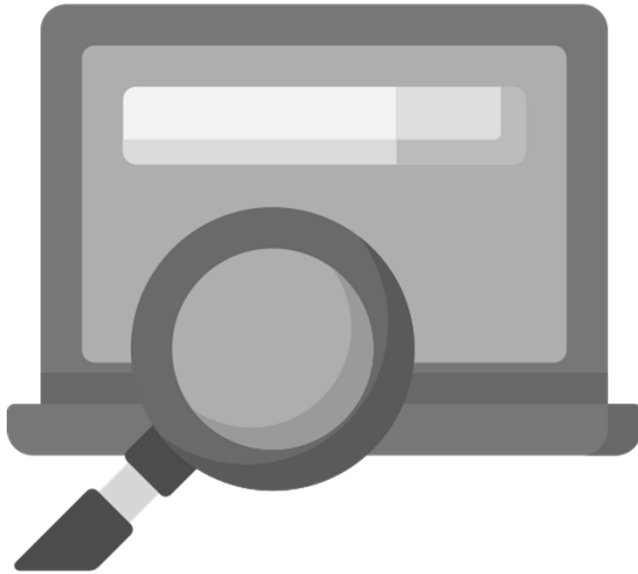
엑셀 다운로드

1) 빅카인즈의 이해



세종사이버대학교

시각화



◀ 검색된 뉴스를 분석해
다양한 시각화 차트를
제작할 수 있음 ▶

1) 빅카인즈의 이해

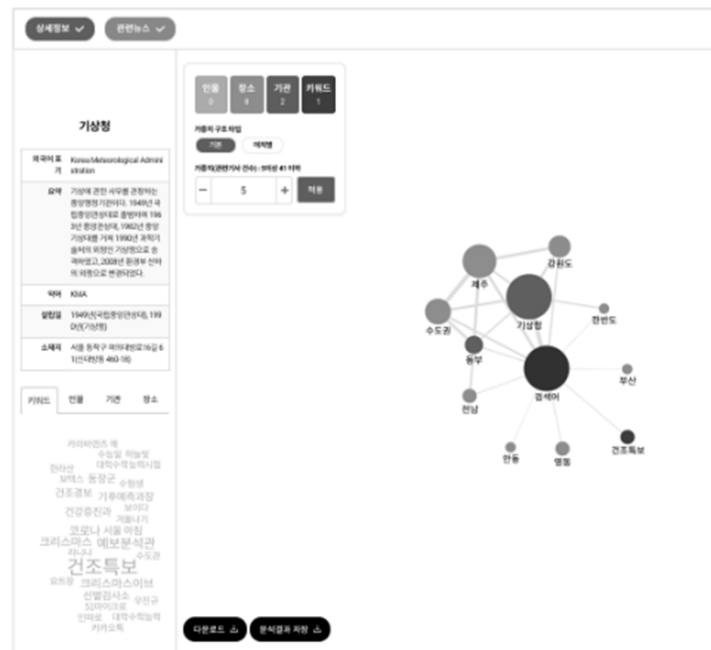


세종사이버대학교

시각화

▶ 관계도 분석

검색한 뉴스에서 추출된 개체(인물, 기관, 장소, 키워드) 사이의 관계를 확인하는 서비스





1) 빅카인즈의 이해

시각화

▶ 관계도 분석

분석방법

분석절차

- 오늘의 이슈, 키워드 검색결과를 분석대상으로 선택
 - 뉴스 내 개체명(인물, 기관, 장소, 키워드) 정보를 추출
 - 개체명 간의 연결관계 구성
 - 시각화 화면에 분석결과를 표
-

1) 빅카인즈의 이해

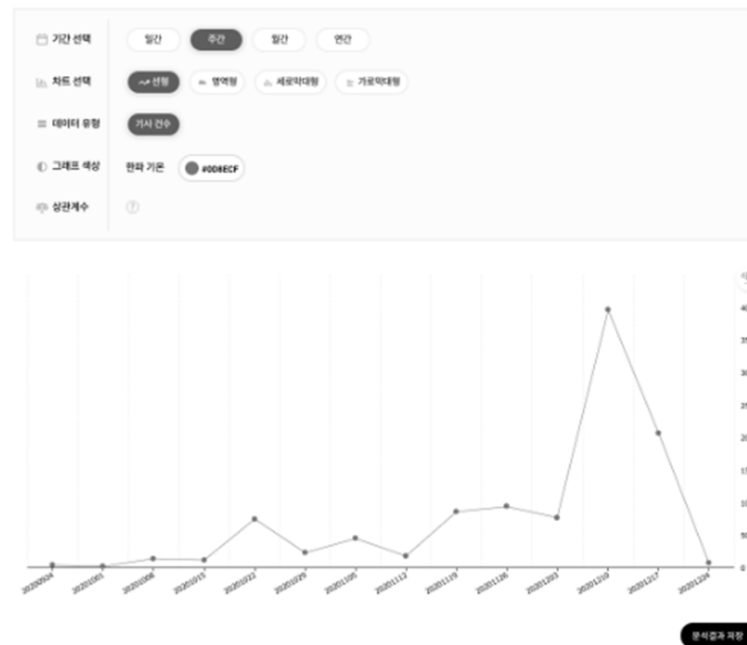


세종사이버대학교

시각화

▶ 키워드 트렌드

검색어와 관련된
기사의 수를
시계열 그래프로
확인하는 서비스



1) 빅카인즈의 이해



세종사이버대학교

시각화

▶ 키워드 트렌드

분석방법

분석절차

- 키워드로 뉴스를 검색함
 - 일, 월, 연 단위로 기사 발생건수 집계
 - 통계결과를 선형, 영역형, 세로막대형, 가로막대형 그래프로 표시
-

1) 빅카인즈의 이해



세종사이버대학교

시각화

▶ 연관어 분석

검색어와 연관된 키워드를 워드 클라우드로 제공하는 서비스

분석 뉴스 건수: 30 50 100 300 500

차트 선택: 워드클라우드 막대그래프

데이터 유형: 가중치 기사 건수



다운로드



1) 빅카인즈의 이해

시각화

▶ 연관어 분석

분석방법

- 키워드로 뉴스를 검색함
- 정확도 순 상위 N개의 뉴스 내 키워드 추출
- 토픽랭크(TopicRank) 알고리즘을 활용해 키워드 간
↓ 연관관계 분석
토픽랭크는 특정 키워드와 동시에 발생한 키워드의 발생빈도와 중요도를 계산해 연관키워드를 추출하는 분석 방식임
- 연관도가 높은 키워드를 워드클라우드, 세로막대 형태로 표출

2) 빅카인즈 체험하기



세종사이버대학교

미세먼지 기사 2일치 다운로드

- 검색유형 : 뉴스
- 상세검색 : 미세먼지
- 검색기간 : 최근 2일
- 언론사 : 중앙지
- 주제분류 : 선택 없음
- 사건사고분류 : 선택 없음
- 검색
- 다운로드

BIGKinds
NEWS BIGDATA & ANALYSIS

뉴스 분석

기획 분석

뉴스 보기

빅카인즈 활용

빅카인즈 소개

로그인

뉴스검색 · 분석

홈 > 뉴스 분석 > 뉴스검색 · 분석

STEP 01 · 뉴스 검색 -

Q 미세먼지

검색도움말

기간

언론사

통합 분류

사건사고 분류

상세검색

☐ 서울 ☐ 경기 ☐ 강원 ☐ 충청 ☐ 경상 ☐ 전라 ☐ 제주

<input type="checkbox"/> 중앙지	경향신문	국민일보	내일신문	동아일보	문화일보	서울신문
<input type="checkbox"/> 경제지	세계일보	조선일보	중앙일보	한겨레	한국일보	매일경제
<input type="checkbox"/> 지역종합지	머니투데이	서울경제	아시아경제	아주경제	파이낸셜뉴스	한국경제
<input type="checkbox"/> 방송사	헤럴드경제	강원도민일보	강원일보	경기일보	경남도민일보	경남신문
<input type="checkbox"/> 전문지	경상일보	경인일보	광주매일신문	광주일보	국제신문	대구일보

2020-12-23 ~ 2021-03-23 ×

초기화

적용하기



2

응용 서비스 및 텍스트 분석 실습

- 1) 빅카인즈 응용 서비스 체험
- 2) 텍스트 분석 및 시각화 실습

1) 빅카인즈 응용 서비스 체험



세종사이버대학교



빅카인즈 체험하기

기타 서비스 뉴스로 보는 코로나 / 코로나19, 남북관계 뉴스, 국회의원 뉴스, 주간 이슈 등 의 뉴스를 심층적이고 입체적으로 볼 수 있음

검색한 뉴스의 메타 정보를 직접 다운로드 받아 R, SPSS와 같은 프로그램 을 통해 보다 정밀하게 분석할 수 있음

뉴스로 보는 코로나

**코로나19 지역별
확진자 정보 및
관련 기사를
확인할 수 있음**

SNU Fact Check 바로가기 >

SNU Fact Check ➤

질병관리청

[SNU 팩트체크] '중국 봉쇄'로 코로나19 막을 수 있
었다

[SNU팩트체크] 코로나19 사망자 수가 노로바이러스나 알라리아로 인한 사망자 수보다 훨씬 적다

| 시도별 확진자 현황

2021년 04월 25일 00:41

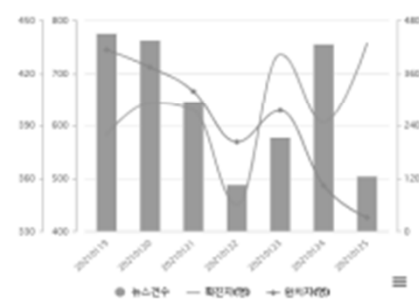
시도를 클릭하시면 상세 현황을 확인하실 수 있습니다.



서울시 코로나 현황 자세히 보기

| 전국 확진자 현황

알림확진자 · 국내 발생 405명 | 해외유입 32명



| 시도별 관련뉴스

서울 기사 전송 730전

2021-01-25 - 2021-01-26

“자랑” 뛰어난 문태충 “자랑 선물 연세행복으로 고르세요”

[illegible]

주-편자명 | 경제·경제일반 | 2025-08-26 | 위승목/가제



지역별 미세먼지

지역별 미세먼지 정보
및 관련 기사를
확인할 수 있음





남북관계 뉴스

- 1990년 이후 남북관계에 관련된 추천
- 키워드, 각국 정상 발언, 북한 전문가 발언, 히스토리 등의 뉴스 분석을 제공함

북한^① 북핵^② 김정은^③ 개성공단^④ 미국^⑤ 트럼프^⑥ 문재인^⑦ 북미정상회담^⑧
김정일^⑨ 비핵화^⑩ 방북^⑪ 대북정책^⑫ 김일성^⑬ DMZ^⑭ 이산가족^⑮ 판문점^⑯
남북정상회담^⑰ 대북제재^⑱ 주한미군^⑲ 금강산^⑳ 탈북^㉑ NLL^㉒ 남북경협^㉓ 경의선^㉔ 햇볕정책^㉕
평화협정^㉖ 북한인권^㉗ 경제협력^㉘ 경제협력^㉙ 국군포로^㉚ 국군포로^㉛

(((북한) OR (北) OR (北韓)), ((김정은 AND 북한 AND (장관 OR 위원장 OR 노총장 OR 국무위원장 OR 국방위원장 OR 총참모장 OR 부위원장) NOT ((수령인) OR (우익단체)) OR

* 검색을 원하는 키워드와 정부를 선택하세요(키워드는 복수 선택 가능하며 복수 선택 시 OR조건이 적용됩니다. 검색결과에서 AND조건으로 수정할 수 있습니다.)
원하는 검색어가 없는 경우 검색창에 직접 입력하세요.



문재인 정부
2017.05.10 ~ 현재



박근혜 정부
2013.02.25 ~ 2017.03.10



이명박 정부
2008.02.25 ~ 2013.02.24



노무현 정부
2003.02.25 ~ 2008.02.24



김대중 정부
1998.02.25 ~ 2003.02.24



김영삼 정부
1993.02.25 ~ 1998.02.24



기업으로 보는 뉴스

300대 기업, 30그룹사의 뉴스 이슈 키워드, 뉴스 관계도, 키워드 트렌드, 최신 뉴스, 관련 뉴스 등을 종합적으로 파악할 수 있음

전체 300건

 삼성전자 Samsung Electronics Co., Ltd. 업종: 통신 및 방송 장비 제조업 대표: 이진희, 이재용, 권오현, 윤부근, 이종균	 현대자동차 Hyundai Motor Co. 업종: 승용차 및 기타 여객용 자동차 ... 대표: 정주영, 정몽구, 정희선, 윤갑현, 이원희	 포스코 POSCO 업종: 제철, 제강 및 합금철 제조업 대표: 최정우(대표이사 회장), 오인환(대표이사 사장), 장인환(대표이사 사장)	 LG전자 LG ELECTRONICS INC. 업종: 통신 및 방송 장비 제조업 대표: 조성진, 정도현
 한국전력공사 KOREA ELECTRIC POWER CORPORATION 업종: 발전업 대표: 조환익	 기아자동차 KIA MOTORS CORPORATION 업종: 승용차 및 기타 여객용 자동차 ... 대표: 정몽구, 이철근	 한화 HANWHA CORP. 업종: 화학 및 불꽃제품 제조업 대표: 김연철, 이만석, 옥경석	 한국산업은행 KDB 업종: 개발금융기관 대표: 이동길

분석결과 시각화



총 > 뉴스 분석 > 분석결과 시각화

- 사용자가 입력한 데이터의 빈도수를 추출하여 빈도수 결과 및 시각화를 제공하는 서비스입니다.

④ 도출법

◎ 초기화

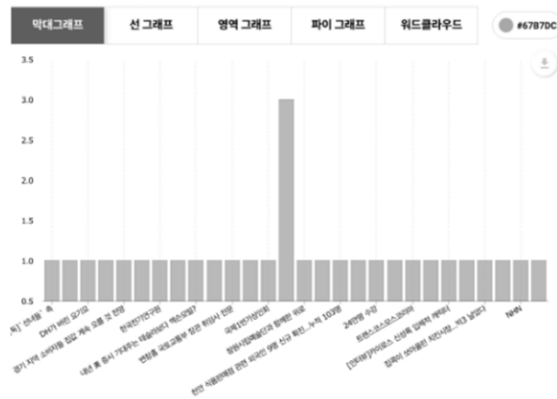
[단독] '선언문'... 성민석 사사는 높은 표절 의혹에 "사실 관계 파악 중"
 "DNA 검사 요요, 불발 결과" 밝혀내
 "경기 지역 소비자들" "집값 계속 오를 것" 전망
 한국인도 외국국 2020년 가뭄... 첫 학년 졸업반에 100명 넘
 "한글서체 연구회, 무의미하고 성과에" "정신적 육체적 소용없음" 선정
 내년 美 중시 기대하는 테슬라코리아 CEO 오일모?
 대한 국의 교차점 단말기... 5주에 58명
 변장형 국토교통부 장관 취임사 전문
 유석열... 본문... 이 때문에 2020년 다 하는 것
 '국제1번가' 신인화, 부산 충무영화 이문경 영화 '집값'
가사노동시간에 다 가사노동 시간 가사노동

사용자 PC 환경에 따라 텍스트 처리량의 차이가 있을 수 있습니다.

(0 / 65,536자)

빈도수 추출

1	용어
2	1. [단행] 선내륙 축
3	2. 삼천리에서 논은 뽕의 열매와 사색과 산경과 풍
4	3. DM이 바뀐 요리요
5	4. 분할 매각 나사나
6	5. 경기 지역 소비자를 집값 계속 오를 것 전망
7	6. 천안에서 또 외국인 구두가 감. 적 작전 수출입액 100억
8	7. 한국전직업연
9	8. 2020년 최고 성과에 찬가.특히 육상산악 스포츠 문호
10	9. 내년 총선 시 가져오는 대승과라도 예소반요
11	10. 대한 천문교 회.집단연구 5.로서 58명
12	11. 변형률 곡률도분류 부를 결정지.전문
13	12. 윤석열 봉봉.이.이.정.윤.윤.2020년 다.이.는.것
14	13. 국제1번가상인회
15	14. 부산 중구청에 이우정.김.정.열
16	15. 대한민국예술재단과 함께
17	16. 갈사.2020년
18	17. 천안 시흥문화원 관원.이.우.정.2020.신.취.진.부.작.103
19	18. AI 영어.선생님.이.재.원.참.사.로.올.려.온.김.효.순
20	19. 24만명 수감
21	20. 청정어 한미당으로 갈무리하



사용자가 입력한
데이터의 빈도수를
추출하여 빈도수
결과 및 시각화를
제공하는 서비스

2) 텍스트 분석 및 시각화 실습



세종사이버대학교

분석결과 시각화

뉴스룸 | 최신기사

[디지털스토리] 문 대통령 취임사 단어 분석...'대통령' 34회 사용

키워드	특성값	빈도수	빈도수비중
국민	0.00037	58	6.98
시대	0.00016	23	2.768
행복	0.00008	21	2.527
문화	0.00015	20	2.407
경제	0.00009	20	2.407
대한민국	0.01313	12	1.444
사회	0.00012	12	1.444
국가	0.0001	12	1.444
희망	0.00007	10	1.203
개인	0.00006	10	1.203
창조	0.00003	10	1.203
발전	0.00008	9	1.083
창조경제	0.00979	8	0.963
국민행복	0.00705	8	0.963
신뢰	0.0001	8	0.963
정부	0.00009	8	0.963
세계	0.00008	8	0.963
미래	0.00009	7	0.842
사람	0.00006	7	0.842
대통령	0.00009	6	0.722
산업	0.00005	6	0.722
국민들	0.00763	5	0.602
경제부흥	0.00643	5	0.602
북한	0.00007	5	0.602
가치	0.00007	5	0.602
능력	0.00006	5	0.602
개개인	0.00006	5	0.602
융합	0.00006	5	0.602

키워드	특성값	빈도수	빈도수비중
대통령	0.00053	34	8.134
국민	0.00029	25	5.981
대한민국	0.01733	9	2.153
나라	0.00016	8	1.914
선거	0.00014	7	1.675
국민들	0.01328	6	1.435
역사	0.00012	6	1.435
정치	0.00011	6	1.435
세상	0.00012	5	1.196
해결	0.00012	4	0.957
지지	0.00009	4	0.957
대화	0.00009	4	0.957
권력	0.00008	4	0.957
약속	0.00008	4	0.957
존경	0.00007	4	0.957
문재인	0.00821	3	0.718
광화문	0.00691	3	0.718
갈등	0.00009	3	0.718
한반도	0.00008	3	0.718
말씀	0.00008	3	0.718
머리	0.00008	3	0.718
감사	0.00007	3	0.718
평화	0.00006	3	0.718
소통	0.00005	3	0.718
시대	0.00005	3	0.718
대통령 선거	0.02014	2	0.478
국방력	0.00642	2	0.478
19대 대통령	0.00619	2	0.478

2) 텍스트 분석 및 시각화 실습



세종사이버대학교

빅카인즈 뉴스 메타정보 활용하기

뉴스 검색 · 분석 > Step03. 분석 결과 및 시각화 > 데이터 다운로드에서 다운받은 뉴스 메타정보 엑셀파일을 활용하여 데이터 입력이 가능함



2) 텍스트 분석 및 시각화 실습



세종사이버대학교



빅카인즈 뉴스 메타정보 활용하기

1

뉴스 메타정보 엑셀파일 내 빈도수 추출을 원하는 컬럼을 선택 후 복사함

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	일자	언론사	기고자	제목	통합 분류1	통합 분류2	통합 분류:사건/사고	사건/사고	사건/사고	사건/사고	인물	위치	기관	키워드
2	20190331	서울경제	박준호	[머니+트렌드] 10명 중 경제>취업,창업	경제>취업,창업	경제>취업,창업	경제>취업,창업					캐나다,사람인,미국,베트남		키워드 머니,트렌드
3	20190330	헤럴드경제	김대우	구직자 10명 중 경제>취업,창업	경제>중남미	경제>중남미	경제>중남미					캐나다,사람인		구직자,해5
4	20190329	동아일보	신호정	구직자 10명 중 경제>취업,창업	경제>중남미	경제>중남미	경제>중남미					캐나다,사람인,베트남		의향,구직자
5	20190329	한국경제	권유환	청년내일채움공제>취업,창업	경제>금융, 경제>무역	경제>금융, 경제>무역	경제>금융, 경제>무역				김도영	고용노동부,사람인,종합정보대행사,정부,중소벤처기업부		청년내일채움공제
6	20190329	머니투데이	고석홍	구직자 10명 중 경제>취업,창업	경제>일본, 경제>무역	경제>일본, 경제>무역	경제>일본, 경제>무역					캐나다,미국,사람인,사람인에이치알,대우		구직자,해5
7	20190328	중앙일보	김상진	[국경] 조철한 AI '국경'>중국	경제>일본, IT, 과학>IT, 과학일반	경제>일본, IT, 과학>IT, 과학일반	경제>일본, IT, 과학>IT, 과학일반				스가 요시히데,미국,일본,고등전문학교,중국,AI,아마존,이과,카네기멜런대,정부,민관,영국,넷케이,백아관,청화대,국경,조철한		국경,조철한	
8	20190326	울산매일	김정아	[기고] 통치오로 경제>취업,창업	사회>노동, 경제>자동차	사회>노동, 경제>자동차	사회>노동, 경제>자동차					울산분 민주노동,정부,IMF,대한민국,이대,통치오로		통치오로
9	20190326	중앙일보	서유진	[장남] 취업특혜 2정치>국회,정당	정치>선거	정치>선거	정치>선거				김,이양,고연봉,한국해양대학교,해수부,한국선급,국회,유엔,교육지원청,자유한국당,세계해사대학,농림축산식품해양수산,취업,장남		장남,취업	
10	20190326	세계일보	구직자기	가장 2경제>취업,창업	IT, 과학>IT, 과학일반	IT, 과학>IT, 과학일반	IT, 과학>IT, 과학일반					대우,잡코리아		정보,연봉
11	20190325	매일경제	배준희	[여의도 Talk] 신 경제>금융,재테크	경제>증권, 경제>무역	경제>증권, 경제>무역	경제>증권, 경제>무역					미국,시프,프랑스,레미,핀트,사무소		신입,눈길
12	20190324	경향신문	이혜인	'장시간 노동' 및 경제>일본	사회>노동, 사회>여성,사회>사건/사고>산업사고>붕괴	사회>노동, 사회>여성,사회>사건/사고>산업사고>붕괴	사회>노동, 사회>여성,사회>사건/사고>산업사고>붕괴				히로시,일본,타,경찰청,과로사,법원,도쿄대,과로		선택,장시간	
13	20190323	헤럴드경제	김대우	구글코리아 11년 경제>취업,창업	경제>취업, 경제>자동차	경제>취업, 경제>자동차	경제>취업, 경제>자동차					한국코카,사람인,넷플릭스,코리아,해보,나이키,코리아,코리아,지연스,코리아		구글코리아
14	20190321	아시아경제	박소연	3년만에 전경련, 경제>취업,창업								허,허창,일,미국,국채부,전국경제인연합회,전경련,산업부		3년,전경련
15	20190320	한국경제	김일규	'열정' 잠재력 및 경제>취업,창업				사고>스포츠사고				김은정,서트,서아와소프트,인사기획팀,정부,중앙대,인턴,밸류시스템,육군,고려사,동덕여대,해병대,인사지원팀		열정,잠재력
16	20190320	아주경제	현상철	'기대보다 800만' 경제>취업,창업	경제>무역, 사회>노동, 복지	경제>무역, 사회>노동, 복지	경제>무역, 사회>노동, 복지					환산,해중소기업중앙회,통계청,아이클릭아트		800만,중소
17	20190320	아주경제	정세희	별론 '복합형' 기 경제>금융,재테크	경제>취업, 문화>방송,연예	경제>취업, 문화>방송,연예	경제>취업, 문화>방송,연예					세미나,기안84,MBC		별론,복합형
18	20190320	경인일보	오경택	'경향' 난 양평공시 경제>취업,창업	경제>산업, 지역>경7,사회>사회갈등>시위	경제>산업, 지역>경7,사회>사회갈등>시위	경제>산업, 지역>경7,사회>사회갈등>시위				이모	양평공 한국미래산업연구원		경향난,양평
19	20190319	한국경제	공태윤	(공기업) 인사담당자 경제>취업,창업	사회>교육, 사회>여성	사회>교육, 사회>여성	사회>교육, 사회>여성					정은정,인성,일,사람인,인턴,한국산업인력공단,산업부,한국남부발전,국민건강보험공단,한수원,동아대,한국수력원자력		Q&A,공기
20	20190319	한겨레	장나래	'귀' 않은 사람 공 경제>취업,창업	경제>금융, 경제>부동산	경제>금융, 경제>부동산	경제>금융, 경제>부동산					장나래,양도양,우리로,wf,O,오,영(M		사람,공력
21	20190319	한겨레	장나래	[단독] 자영업자 경제>유통	경제>취업,창업	경제>취업,창업	경제>취업,창업					김,2천만,일해,정내일		자영업자
22	20190316	아시아경제	박병희	[신간 안내] <일 문화> 미술,건축	문화>출판	문화>출판	문화>출판					김보경, <일본, 민병대,마이나비,마이나비그룹		일본,일본
23	20190315	매일경제	조성진	코어교육, 유행 경제>취업,창업	사회>교육, 시험	사회>교육, 시험	사회>교육, 시험					CS		코어교육
24	20190314	서울경제	이상훈	다국 종아지는 4 경제>취업,창업	경제>산업, 경제>유통	경제>산업, 경제>유통	경제>산업, 경제>유통					추광호, 인턴,한국경제연구소,한국경제연구원		취업문,채
25	20190314	파이낸셜뉴스	파이낸스	12월 취업자수 2 경제>취업,창업	경제>산업, 기업	경제>산업, 기업	경제>산업, 기업					한경연, 함양		채용,민
26	20190313	충청일보	배명식	대기업 5곳 중 1 경제>취업,창업	경제>산업, 경제>유통	경제>산업, 경제>유통	경제>산업, 경제>유통					전국경제인연합회,인턴,한국경제연구소,한국경제연구원,산학협력		1곳,상반기
27	20190313	국민일보	백승연	대기업 20% '상' 경제>취업,창업	경제>산업, 기업	경제>산업, 기업	경제>산업, 기업					추광호 함양		감축,20%
28	20190313	한국경제	김보형	대기업 절반, 아 경제>취업,창업	경제>산업, 경제>유통	경제>산업, 경제>유통	경제>산업, 경제>유통					이상인		절반,개척
29	20190313	서울경제	이상훈	다국 종아지는 4 경제>취업,창업	경제>산업, 경제>유통	경제>산업, 경제>유통	경제>산업, 경제>유통					추광호, 인턴,한국경제연구소,한국경제연구원		취업문,채
30	20190313	한국경제	공태윤	(한화시스템, 한화경제) 경제>취업,창업	IT, 과학>IT, 경제>산업, 기업	IT, 과학>IT, 경제>산업, 기업	IT, 과학>IT, 경제>산업, 기업					신동일,방산,정,한화정보통신,이화여대,산업부,기술직,미래혁신센터,석박사,ICT,일해,경영기획실		한화시스템
31	20190313	머니투데이	송정훈	고객 돈 끌리는 경제>산업, 기업	경제>금융, 경제>유통	경제>금융, 경제>유통	경제>금융, 경제>유통					삼성자산운용,미래에셋,미래에셋자산운용		고객,일직
32	20190313	헤럴드경제	유재준	한경연 '대기업' 경제>취업,창업	경제>산업, 기업	경제>산업, 기업	경제>산업, 기업					추광호 함양		한경연,절
33	20190313	국민일보	김준연	대기업 46% 상 경제>취업,창업	경제>유통	경제>유통	경제>유통					추광호		계획,46%
34	20190313	조선일보	한동희	대기업 46% 상 경제>취업,창업	경제>산업, 경제>유통	경제>산업, 경제>유통	경제>산업, 경제>유통					추광호 함양		대기업,20%
35	20190313	저지시도	로조요	'대기업' <고 조 경제>취업,창업	경제>산업, 경제>유통	경제>산업, 경제>유통	경제>산업, 경제>유통					한겨레, 함양		1,고 조요

2) 텍스트 분석 및 시각화 실습



세종사이버대학교



빅카인즈 뉴스 메타정보 활용하기

2

데이터 입력창에 복사한 데이터를 붙여 넣은 후,
용어개체 빈도 추출 버튼을 클릭함

- 입력한 데이터는 쉼표(,)와 엔터(↵)를 기준으로

뉴스 분석 기획 분석 뉴스 보기 빅카인즈 활용 빅카인즈 소개 로그인

분석결과와 시각화

홈 > 뉴스 분석 > 분석결과와 시각화

• 사용자가 입력한 데이터의 빈도수를 추출하여 빈도수 결과 및 시각화를 제공하는 서비스입니다.

데이터 입력 도움말 초기화

대우, SK이노베이션, 통일, 개진
구글, 카카오, 네이버, 대우, FAANG, 넷플릭스, 라인플러스, 트위터, 페이스북, MAGA, 해커튼, 야후, 애플, 유튜브, MAGA(마이크로소프트, 쿠팡)
수출입은행, 중앙은행, 연방준비은행, 산업은행, 한국은행, 금융감독원, 코로나19
노사, 기재부, 기은, 산업은행, 산은, 국책은행, 국책, 기업은행, 수출입은행, 기획재정부, 한창, 국민의힘
노사, 기재부, NH농협은행, 기은, 산업은행, 산은, 부산은행, 국책은행, 기업은행, 수출입은행, SC제일은행, 기획재정부, 한창, 국민의힘
국민권익위원회, 수성구, 법무사, 한국감정원, 부동산 직거래, 대구
알리바바, 구글, 비바리퍼블리카, 한국교육개발원, 애플, 스타, 네이버, 기술직, 노동통계국, 코드스테이츠, 소프트웨어정책연구소
정부, 에이엠티, 더불어민주당, 충남지방중소벤처기업청장, 사무소
삼진, 비바리퍼블리카, 중앙일보, 유일, 네이버, 토스뱅크, 토스인슈어런스, 위라벨, 신청 · 조회, 토스증권
영도조선소, KDB, 케이스트파트너스 컨소시엄, 한진해운, 한국토지신탁, 한진중공업, SM그룹, 산업은행, 산은, 구조조정본부, 수비조선소, 정부, 채권단, 최대주주, 해운사, KDB인베스트
먼트(KDBI), 투자은행, 대우건설

사용자 PC 환경에 따라 텍스트 처리량의 차이가 있을 수 있습니다.

(844 / 65,536자)

빈도수 추출

2) 텍스트 분석 및 시각화 실습



세종사이버대학교

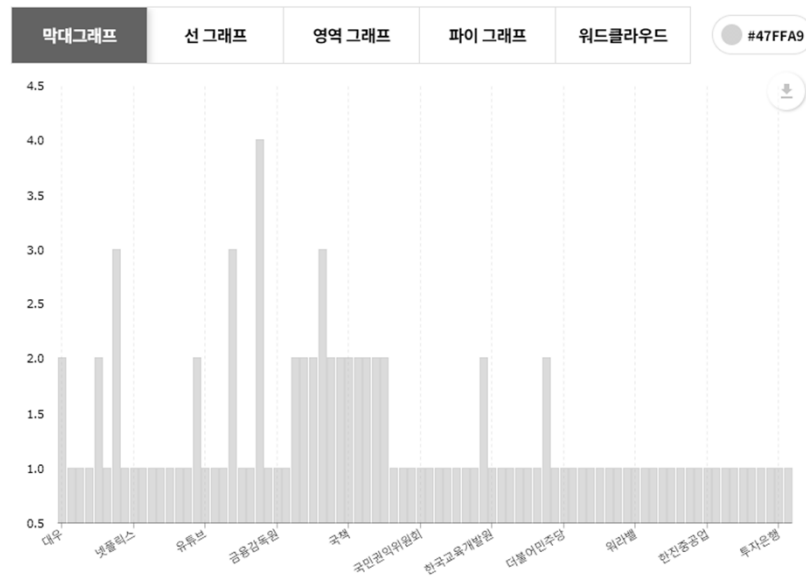
빅카인즈 뉴스 메타정보 활용하기

3 입력한 데이터 빈도 추출 결과 및 시각화 화면이 나타남

■ 비디스 커리어즈 더브 크리치며 내리카스 /

빈도 추출 결과

	용어	빈도수
1	대우	2
2	SK이노베이션	1
3	통일	1
4	개진	1
5	구글	2
6	카카오	1
7	네이버	3
8	FAANG	1
9	넷플릭스	1
10	라인플러스	1
11	트위터	1
12	페이스북	1
13	MAGA	1
14	해커톤	1
15	아후	1
16	애플	2
17	유튜브	1
18	MAGA(마이크로소프트)	1
19	쿠팡	1





학습 정리

- 뉴스 빅데이터 플랫폼의 활용과 텍스트 분석
 - 빅카인즈를 통해 다양한 뉴스 빅데이터 분석을 실행 할 수 있으며 형태소 분석결과를 통한 시각화 실습 등 비정형 텍스트 결과를 손쉽게 추출 및 활용할 수 있음