



CONTENTS

1.	로그인	 03 - 10
2.	검색	 11 - 25
3.	마이페이지	 26 - 33
4.	자료분석	 33 - 52
5.	OPEN API	 53 - 59
6.	자료열람	 60 - 64



주요기능

회원가입

로그인

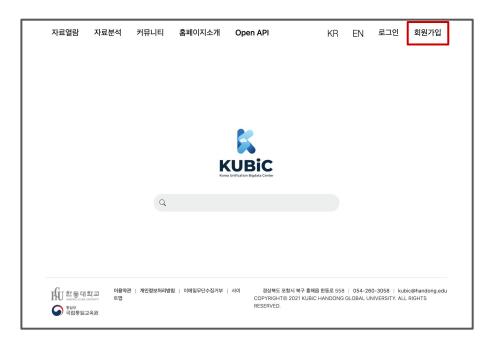




회원가입

Step1

메인화면 이동





회원가입

Step2

약관 확인







회원가입

Step3

신분 및 소속기관 입력



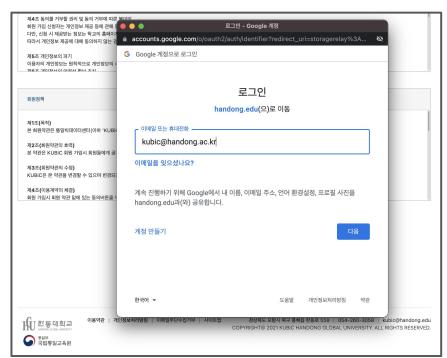




회원가입

Step4

구글 계정 로그인

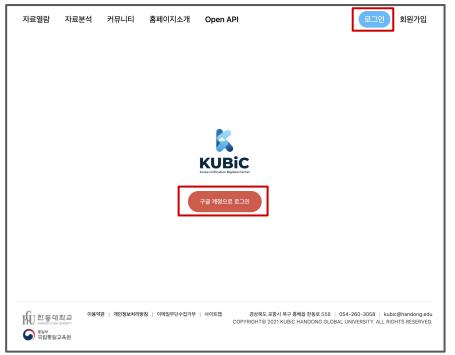






Step1

로그인 버튼

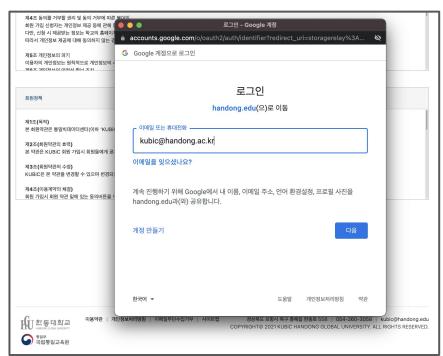






Step2

구글 계정 로그인







Step2

로그인 후 메인페이지





주요기능

키워드 검색

검색필터

검색 상세 페이지

문서 담기





키워드 검색

Step1

메인 화면 이동







키워드 검색

Step2

"올림픽" 키워드 입력





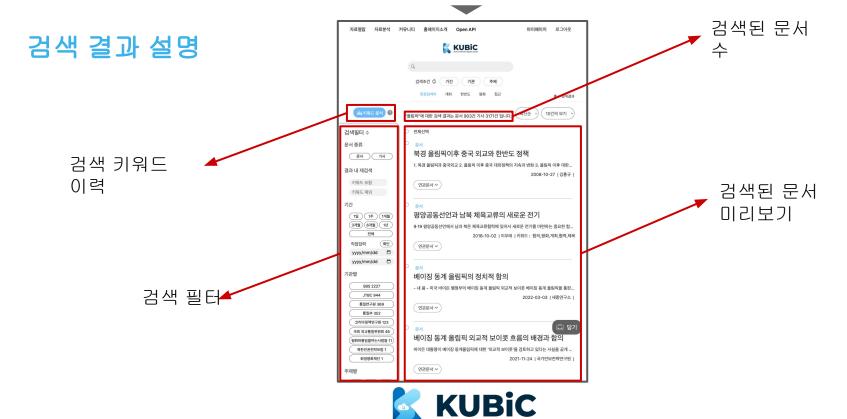
키워드 검색

Step2

"올림픽" 키워드 검색







검색 필터

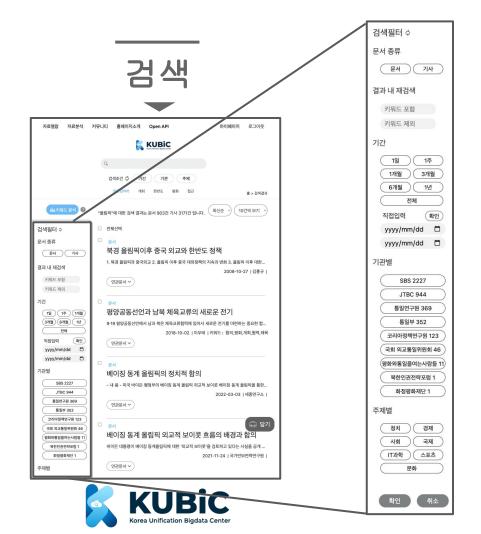
문서 종류

키워드 포함/제외

기간 검색

기관 검색

주제별 검색





검색 필터

Filter1 문서 종류

문서와 기사로 나누어서 검색할 수 있습니다. 오른쪽 사진은 [올림픽]

검색어에 [기사]필터를 적용한 결과입니다.









검색 필터

Filter2 키워드 포함/제외

문서 제목, 문서 정보, 첨부파일 본문을 대상으로 입력한 키워드가 포함되어있거나 제외하여 검색할 수 있습니다. 오른쪽 사진은 [올림픽] 검색 결과에서 [평창]을 포함하고 [베이징]을 제외한 결과입니다.









검색 필터

Filter3 기간 검색

특정 기간에 게시된 문서를 검색할 수 있습니다. 오른쪽 사진은 [올림픽] 결과에서 [6개월] 동안 게시된 문서를 검색한 결과입니다. 사용자가 원하는 기간의 날짜를 직접 입력할 수 있습니다.













검색 필터

Filter4 기관검색

문서와 기사를 게시한 기관명으로 검색할 수 있습니다.

검색을 하면 기관명과 함께 해당 검색어로 검색된 기관의 문서수가 함께 서비스됩니다. 오른쪽 사진은 [올림픽] 검색어에 [통일부] 기관 필터를 적용한 결과입니다.











검색 필터

Filter5 주제별 검색

정치, 경제, 사회, 국제, IT과학, 스포츠 7가지 주제에 맞춰 문서를 검색할 수 있습니다.

오른쪽 사진은 [올림픽] 키워드 검색에 [국제] 주제 검색을 적용한 결과입니다.









검색 상세 페이지

Case1 : 문서를 선택할 때

문서를 선택하면 문서 상세페이지로 이동하여 <u>문서</u> 정보와 <u>연관문서</u> 그리고 워드클라우드를 열람할 수 있습니다.

Case2: 기사를 선택할 때

기사를 선택하면 기사를 게시한 언론사 페이지로 이동합니다.







복미국장 대형이 참석했습니다. 남북 또 복미 정상회담이 발표된 이후에, 남·복·미 관계자가 함께 만난 것은 처음이어서 여러 가지로 관심이 모아졌습니다. 이 회의에서 최각일 뜻한 점 무성 복미국장을 만나고 온 김준형 환동대 국제지역학과 교수를 참만한 등 연결해보겠습니



문서 담기

Step1

로그인 후 검색

ᄽ1장 로그인 참조





문서 담기

Step2

원하는 문서 체크







문서 담기

Step3

담기



저장된 문서는 [검색 키워드]와 [저장 시간] 으로 <u>마이페이지</u>에서 열람할 수 있습니다.

荐4장 마이페이지 참조

kubic.handong.edu says 문서가 내 보관함에 저장되었어요.



주요기능

내 보관함

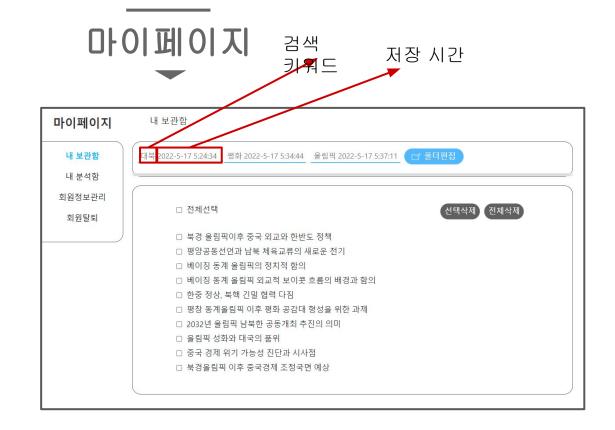
내 분석함

회원정보 관리 및 회원 탈퇴



내 보관함

담은 문서를 '마이페이지 > 내 보관함'에서 확인할 수 있습니다. 문서 제목을 선택하면 문서 상세보기로 이동합니다.

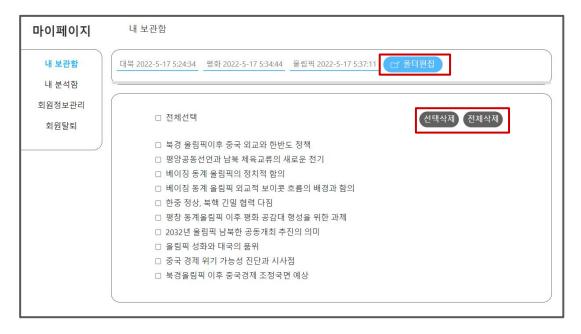




내 보관함

내 보관함에서 지원하는 문서 관리 기능은 아래와 같습니다.

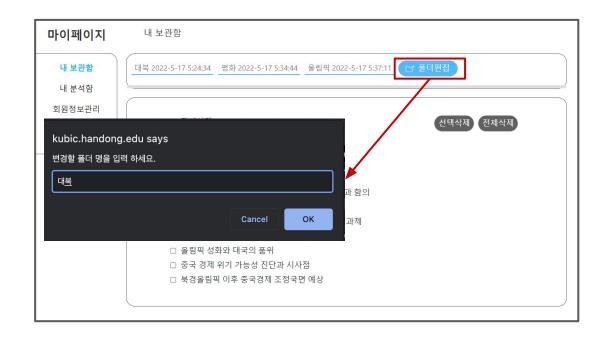
- → 폴더 편집
- → 선택삭제
- → 전체삭제





내 보관함

폴더 편집 버튼을 사용하여 저장된 키워드 명을 변경할 수 있습니다.







내 보관함

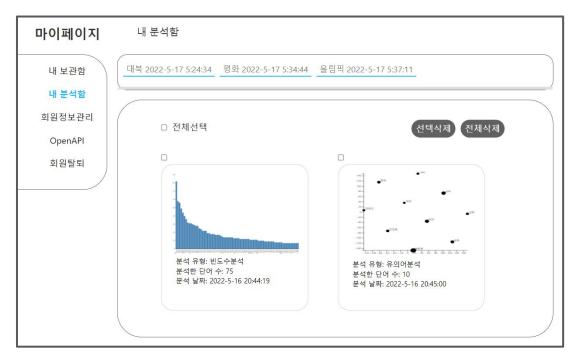
- → 선택삭제 버튼을 사용하여 체크된 문서를 삭제할 수 있습니다.
- → 전체 삭제 버튼을 누르면 해당 화면의 키워드와 해당된 문서가 모두 삭제됩니다.



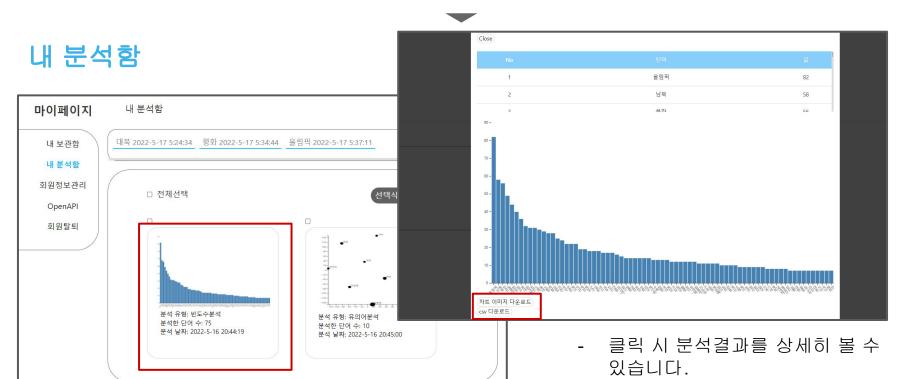


내 분석함

저장된 분석 결과를 열람할 수 있습니다.







KUBIC

32

분석 결과를 이미지 또는 csv

파일로 저장할 수 있습니다.



주요기능

전처리

자료분석



전처리

Step1

문서 선택





전처리

- 유의어 사전: 같은 의미의 다른 표현들을 대표 단어로 변경합니다.
- 불용어 사전: 분석 시 필요 없거나, 의미 없는 단어들을 등록해, 전처리 과정에서 제외할 수 있습니다.
- 복합어 사전: 신조어나 커스터마이징한 단어들을 분석에 포함시키고 싶을 때,
 복 합어 사전에 그 단어들을 등록하여 분석을 진행할 수 있습니다



전처리

유의어

첫번째 컬럼(A)에 대표 단어들을 작성하고, 다음 컬럼들(B,C..)에는 각각의 첫번째 컬럼에 해당하는 단어의 유의어들을 작성합니다. 첫번째 컬럼의 단어들이 대표어가 되며, 나머지 컬럼의 단어들은 첫번째 컬럼의 단어로 대체됩니다

4	Α	В	С	D
1	김정은 위원장	김정은	김위원장	
2	한동대학교	한동대	한동대학	
3	북한	북조선	조선민주주의인민공화국	
4				
5				

불용어

불용어로 등록하고 싶은 단어들을 첫번째 컬럼(A)에 작성합니다.

1	A	В
1	첨부	
2	중략	
3	하지만	
4		

복합어

첫번째 컬럼(A)에 복합어로 등록하고 싶은 단어를 쓰고, 다음 컬럼(B)에는 그 단어의 품사를 메캅품사태그를 참고하여 작성합니다.

4	Α	В	С
1	한동대학교	NNP	
2	보호법	NNG	
3	김정은 체제	NNG	
4	네브라스카	NNP	
5	남북 회담	NNG	
6			



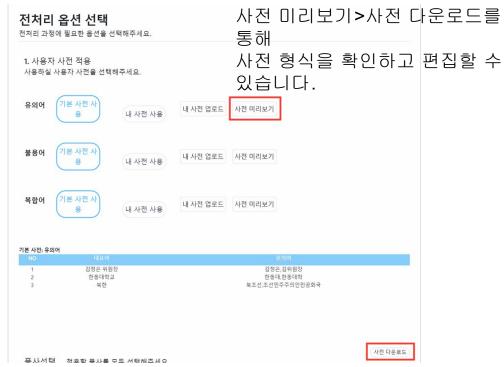
전처리

Step2

사용자 사전 편집을 위해 기본 사전 다운로드

자료분석







전처리

Step3

유의어 사전에

"오륜기, 올림픽기" 추가

올림픽기의 유의어를 오륜기로 설정합니다.

À	Α	В	С	D
1	김정은 위원장	김정은	김위원장	
2	한동대학교	한동대	한동대학	
3	북한	북조선	조선민주주의인민공화국	
4	오륜기	올림픽기		
5				
6				



전처리

Step4

CSV 형식으로 파일 저장

'다른 이름으로 저장'에서 파일 형식을 'CSV UTF-8'로 지정하고 파일을 저장합니다.

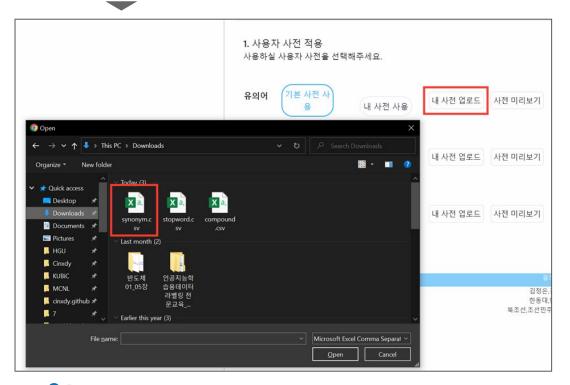


전처리

Step5

내 사전 업로드

저장한 csv파일을 내 사전으로 업로드합니다.





전처리

Step6

업로드 된 내 사전 확인

내 사전 사용으로 지정한 뒤 오륜기, 올림픽기가 추가된 것을 확인합니다.

*엑셀 파일 사용 시 글자 깨짐 현상이 일어나는 경우도 있으니 꼭 '사전 미리보기'로 사전이 성공적으로 업로드 되었는지 확인합니다.



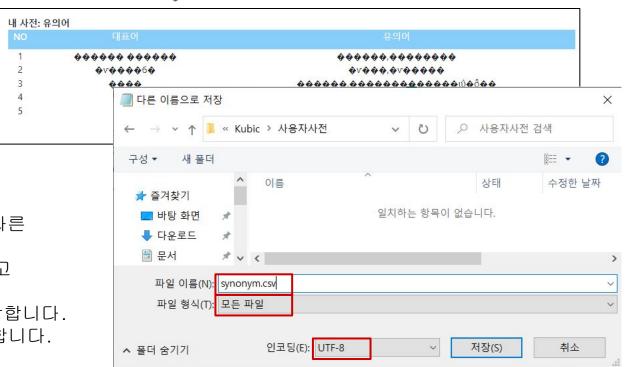


전처리

Step6

업로드 된 내 사전 확인

다음과 같이 글자가 깨질 시, 해당 파일을 <mark>메모장</mark>으로 열어 '다른 이름으로 저장'하여 파일이름 뒤에 '.csv'를 붙여 넣고 파일 형식은 '모든 파일' 인코딩 'UTF-8'로 지정하고 저장합니다. 이렇게 저장한 파일을 업로드 합니다.





전처리

Step7

품사 선택 한 후 전처리 실행

동사, 명사, 형용사 중 어떤 품사들을 수집하고 싶은지 선택할 수 있으며, 복수 선택이 가능합니다.





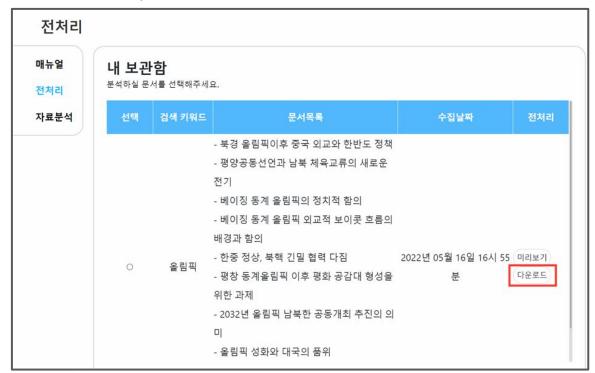


전처리

Step8

전처리 결과 다운로드

전처리 미리보기에서는 일부 결과만 출력되며 전체 내용은 다운로드하여 csv파일로 확인하실 수 있습니다.





자료 분석

8가지 종류의 자료 분석 지원

분석 대상에서 많이 등장하는 단어의 등장 빈도를 분석함으로써 핵심어를 추출하는 분석방법

빈도수 분석

의미연결망

분석 대상들이 속한 모집단 또는 범주에 대한 사전 정보가 없는 경우 관측값들 사이의 유사성을 이용하여 하위군집이 상위군집에 포함되도록 묶는 분석 방법 특정 말뭉치에 등장하는 어휘의 빈도를 계산해 단어가 특정 문서에서 얼마나 중요한 핵심어인지 추출하는 분석방법

TFIDF

N-gram

연속한 n개의 토큰을 하나의 분석단위로 두고 진행하는 분석방법 석 대상들이 속한 모집단 또는 범주에 대한 사전 정보가 없는 경우 관측값들 사이의 유사성을 이용하여 원소가 겹치지 않는 몇 개의 부분집합으로 묶는 분석방법

분할군집분석

유의어 분석

A라는 단어가 B라는 단어와 얼마나 많이 연결되어 있는지를 나타내는 분석 기법 분석 대상이 서로 어떻게 관련을 맺고 연결망을 구성하는가에 주목하여, 대상들간의 관계를 연결망 구조로 표현하고 이를 계량적으로 제시하는 분석기법

계층군집분석

토픽 모델

문서를 이루는 핵심어를 바탕으로 문서에서 토픽, 즉 주제를 추출해 주는 확률모델 알고리즘



자료 분석

각 분석 종류 별로 분석 결과로 제공하는 차트와 표

분석할 단어 수 : 빈도수 상위 몇개의 단어를 시각화할지 설정

빈도수 분석

분석할 단어 수: 75

분석 실행

분석할 단어 수

: TF-IDF 상위 몇개의 단어를 시각화할지 설정

TFIDF

분석할 단어 수: 100

분석 실행

군집 개수

: 문서를 몇개의 군집으로 분석할지 설정

분할군집분석

군집 개수: 5 분석 실행

분석 실행

→ Tip1

"연결 강도"는 100% 에서 점점 줄여가면서 분석해보세요!

. →Tip2

"분석할 단어 수" 와 "분석할 ngram 개수"는 점점 늘려가면서 분석해보세요!

* 분석할 대상의 수가 늘어날 수록 분석 시간이 오래걸리니 주의해주세요!

의미연결망

분석할 단어 수: 50

연결 강도(%) 75

분석 실행

분석할 단어 수

: 상위 몇개의 단어를 시각화 할지 설정 연결 강도

: 상위 몇%의 연결을 시각화 할지 설정

N-gram

분석할 ngram 개수:

3

연결 강도(%) 100

분석 실행

분석할 ngram 개수

: 상위 몇개의 ngram을 시각화 할지 설정 연결 강도

: 상위 몇%의 연결을 시각화 할지 설정

유의어 분석

분석할 단어 수: 10

분석 실행

토픽 모델

옵션 없음

계층군집분석

토픽 수: 10

분석 실행

토픽 수

: 선택한 문서 전체를 몇개의 토픽으로 토픽모델링할지 설정

분석할 단어 수

: 빈도수 상위 몇개의

분석을 할지 설정

단어를 선택하여 유의어

자료 분석

각 분석 종류 별로 분석 결과로 제공하는 차트와 표

분석 종류	빈도수 분석	TFIDF	분할군집 분석	의미연결망	N-gram	계층군집 분석	유의어분석	토픽 모델링
차트 종류	Bar		Network		Tree	Scatter for word2vec		LDA
丑	Index 단어 값		Category Title	Index 사이중심성 근접중심성 빈도수 연결중심성 Eigenvector	X			
비고				D3로 구현됨				오픈 소스 이용

자료 분석

Step1

문서 선택





자료 분석

Step2 분석 선택 & 변수 입력

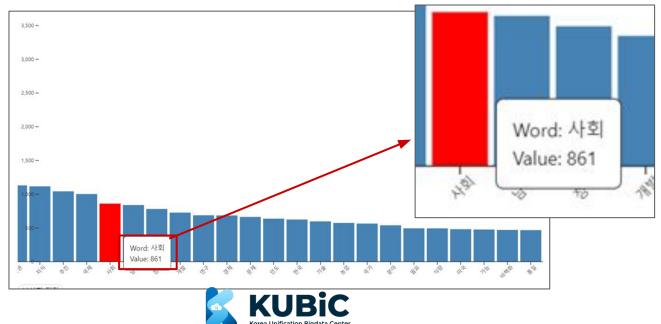




자료 분석

Step3 분석 결과 표와 차트 확인

마우스 휠을 통한 Zoom in& Zoom out과 마우스 오버를 통한 레이블 확인이 가능합니다.



자료 분석

Step4 차트 다운로드

차트 이미지 다운로드: 차트를 png파일로

다운로드하는 기능

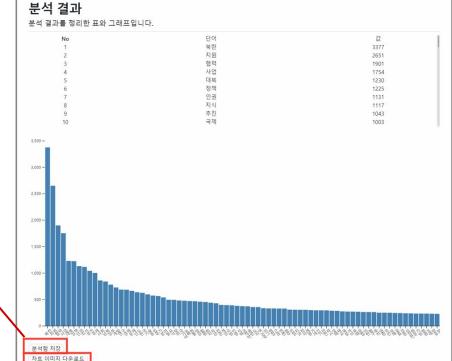
분석함 저장: 표와 차트를 마이페이지> 내 분석함에

분석함 저장

차트 이미지 다운로드

저장하는 기능

♣4장 마이페이지 참조





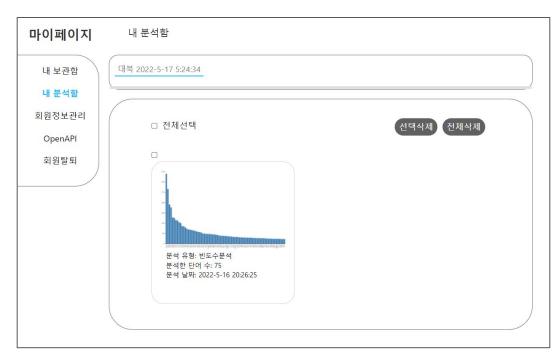
자료 분석

Step5 내 분석함 확인

저장한 분석결과는 '마이페이지 > 내 분석함'에서 열람할 수 있습니다.

자료분석









주요기능

회원가입

활용 신청

활용 관리

API 이용



회원가입

Step1 약관 동의 및 회원가입





활용 신청

Step1 활용 관리 > 활용신청

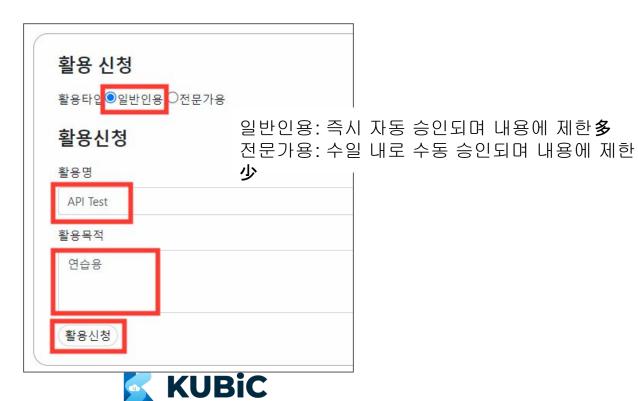




활용 신청

Step2 정보 입력

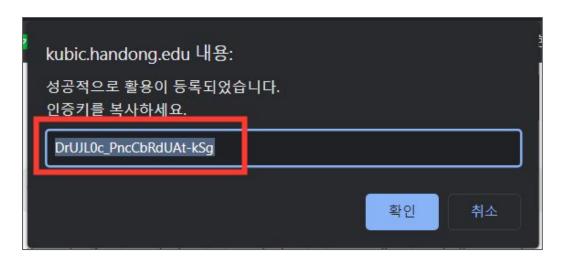
활용타입 선택, 활용명, 활용 목적을 입력합니다.



활용 신청

Step3 인증키 복사

API를 이용하기 위해 발급되는 인증키는 최초 1번만 보이기 때문에 복사하여야 합니다.



※ 전문가용의 경우 메일로 인증키를 알려드립니다.



활용 관리

신청한 활용을 확인

할 수 있습니다. 활용관리

활용 관리

API 이동



Open API 활용을 신청하거나 인증키 재발급을 받을 수 있습니다.

 No.
 활용명
 활용무업
 등록일
 만료일
 데이터량
 관리

 1
 일반인용
 API Test
 연습용
 2022년 05월 17일 10시 33분 2023년 05월 17일 10시 33분 00년 05월 10년 05월 10년



API 이용

신청한 활용을 확인

할 수 있습니다.





주제별 열람

사전편찬별 열람

기관별 열람



자료열람 이동







주제별 열람

7개의 주제로 열람을 지원합니다.

정치경제사회국제IT과학스포츠문화





사전편찬별 열람

문서를 가나다 순으로 열람할 수 있습니다.





기관별 열람

문서를 게시한 기관명으로 문서를 열람할 수 있습니다.







