

TEXTORY

투빅스 컨퍼런스 프로젝트

왜 역사일까?



- ✓ 최근 유명인의 역사 관련 발언의 문제점
- ✓ 어렵고 복잡한 역사
- ✓ 빅데이터를 활용해 요약하고 시각화해 보여주자!

사용한 프로그램

Step 1
데이터 수집 및
전처리
시각화



R

Step 2
Java Script를
이용한 시각화



D3.js

Step 3
결과물 구현



WEB

분석 및 구현 과정

데이터 수집 및 전처리



사건 키워드 추출



인물 관계 시각화



데이터 요약



웹에서 구현



데이터 수집 및 전처리 사건 키워드 추출

사전 구축 (국사편찬위원회 LOD)

한국사 LOD

Korean History Linked Open Data

한국사 콘텐츠

사태소개

콘텐츠페이지

Home

Data Search

SPARQL

Data Browser

Data Search

Datasets

Resources

4,072

검색

검색어

검색

전체

인물

사건

조직

유물

검색어: "" 조회 결과

입력어: 순

확인

30

확인

데이터셋	라벨명	설명	URI	업데이트 날짜
사건	강화 전도	"1222년(고종 19) 통문에 대항하기 위해 도읍지를 강화로 옮긴 사건."	http://lod.koreanhistory.or.kr/resources/사건_강화전도	2015-12-09
사건	강화도 문제의 속조	"문제는 조선시대 강화나 변방의 요지에 구축하여 충군을 설치하고 봉수시설을 갖추어 놓은 강화도를 조선 후기 호준대부터 진지를 설치하면서, 국군대에 설치 강화에 53개의 본대를 건설함."	http://lod.koreanhistory.or.kr/resources/사건_강화도_문제의_속조	2015-12-09
사건	강화도조약 체결	"1876년(고종 11) 2월 강화에서 조선과 일본 사이에 체결된 조약. 조약의 결과 강화도는 조일수호조약(韓日保護條約)이후, 강화조약(日韓條約) 또는 강화수호조약(韓日保護條約)이라고도 한다."	http://lod.koreanhistory.or.kr/resources/사건_강화도조약_체결	2015-12-09
조직	강화학회	"육촌 일년 강화도에 문화재 학문을 담은 강화학회(興學會)의 발명학적 학문을 이끈 조선 후기의 학회."	http://lod.koreanhistory.or.kr/resources/단체_강화학회	2015-12-09
인물	강화학	"고려 말 조선 초기 문신 분암은 진주(鎭州)의 문장으로는 최충안(崔忠安)이 있음."	http://lod.koreanhistory.or.kr/resources/인물_강화학	2015-12-09
인물	강화학	"조선전기기의 문신으로 일대의 뛰어난 문장가."	http://lod.koreanhistory.or.kr/resources/인물_강화학	2015-12-09
사건	강화방사선(興學放事案)	"조선시대 강화도에 있었던 일사(興學)에 대한 문헌적 증거의 존재."	http://lod.koreanhistory.or.kr/resources/사건_강화방사선	2015-12-09
인물	강화학	"조선전기기의 문신이며, 서 그림 공제에 뛰어난 인물. 화관 등과 더불어 3절(三絶)이라 불렀던 학자."	http://lod.koreanhistory.or.kr/resources/인물_강화학	2015-12-09



	A	B	C	D	E	F
1	hlod.titleKor	hlod.detailE	hlod.explain	hlod.hasEnd	hlod.hasEndDate	hlod.hasP
2	105인사건	[정의]	11911년 일제강점기	1913년 07월 15일	hlod:인악	
3	13도 창건	이인영은 원	1907년 고대한국	1909년 6월	NA	
4	28 독립 선언	제1차 세계	1919년 2월 일제강점기	1919년 02월 08일	NA	
5	3.1운동	원인과 배경	일본의 식민 일제강점기	1919년	hlod:2.8독	
6	6.10만세운동	순종의 인신	1926년 6월 일제강점기	1926년 6월	hlod:유월/	
7	가쓰라 태프트 밀약	1905년 7월	1905년 7월 대한제국	1905년 07월 31일	NA	
8	간도 합병(경신 합병)	3.1운동을	1920년 10월 일제강점기	1921년 5월	hlod:훈춘/	
9	간도 철막	간도 지역은	1909년 9월 대한제국	1909년 09월 04일	NA	
10	간도공산당사건	상해에 있던	공산주의의 일제강점기	1930년	hlod:임오/	
11	갑신정변	김옥균(호주)	1884년(고대조선)	1884년 10월 19일	hlod:임오/	
12	갑오개혁	1894년 봄	갑오개혁(무조선)	1896년 2월 11일	NA	
13	강제징용	강제징용은	일제강점기	1945년 08월 15일	hlod:중일	
14	강화도조약 체결	1876년 황	1876년(고대조선)	1876년 2월 27일	hlod:운요/	
15	개항 장안	천도교단(천도교에서)	일제강점기	1926년 8월	hlod:문화	
16	개항	[정의]	특히 세 여러 나라	대한제국	1899년	hlod:운요/
17	거문도 사건	거문도는 전	1885년 3월 조선	1887년 2월 5일	NA	
18	경복궁중건	세도정치로	경복궁 중건조선	고종 1872년	NA	
19	경부선 개통	1901년 8월	서울과 부산 대한제국	1905년 05월 25일	NA	
20	경성방송국 개국	1920년 미	1927년 2월 NA	미상	hlod:방송	
21	경원선 개통	경원선과 함서	특별특시 일제강점기	1914년 09월 06일	NA	
22	경의선 개통	1905년 일	서울과 평안 일제강점기	1911년 11월	NA	
23	경인선 개통	1896년(고대우리나라)	최대한제국	1900년 11월 12일	NA	
24	고부민란	당시 고부	1894년 1월 조선	1894년 3월	hlod:교조	
25	고종 친정 선포	철종이 죽	1873년 고대조선	1873년	NA	
26	공주 농민봉기	1862년 상	1862년(일제조선)	1862년	NA	
27	관동대진과 한국인학살사건	1920년대	1923년 일제NA	미상	NA	
28	광무개혁	정의	1891896년 아	대한제국	1904년 02월 08일	hlod:대한
29	광주학생 항일운동	광주 학생	1929년 민	일제강점기	1930년 3월	NA
30	교조신원운동	운동의 성격	1864년 동대조선	1893년 4월 3일	hlod:동학	
31	국가총동원법 실시	국가총동원	1938년 4월 일제강점기	1945년 08월 15일	hlod:총독	
32	국립중앙도서관	[정의]	1921년 10월	서울 일제강점기	1921년 06월 07일	hlod:국립

사전 구축(Cont.)



사전을 기준으로 문서 크롤링



키워드 추출

카이스트 형태소 분석기 SimplePos를 사용

```
install.packages("KoNLP")
```

```
library(KoNLP)
```

```
SimplePos09()
```

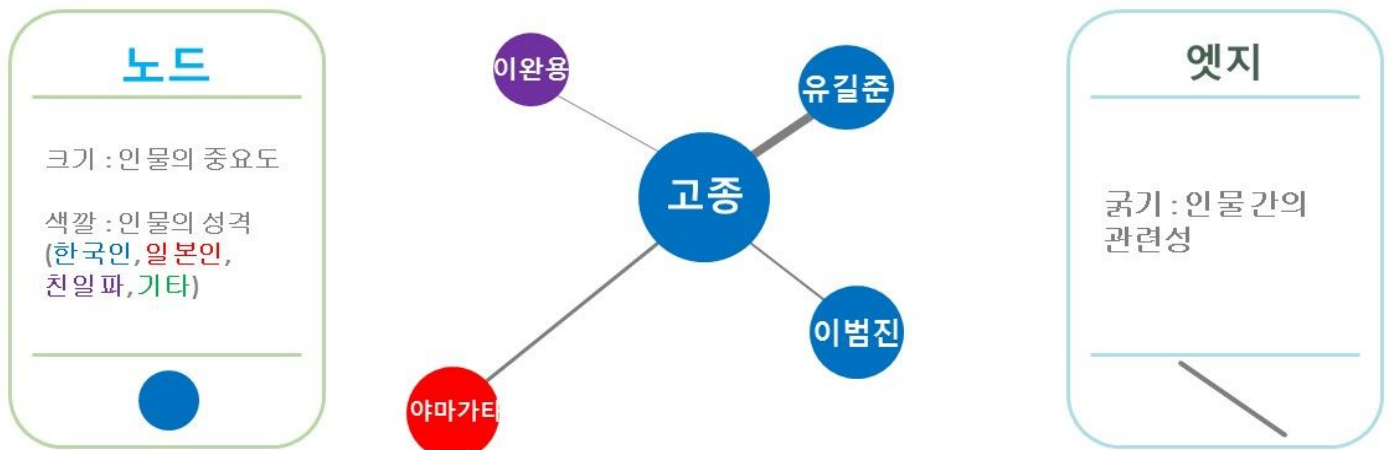
```
> SimplePos09('너는 정말 잘생겼다')
$`너는`
[1] "너/N+는/J"

$정말
[1] "정말/M"

$잘생겼다
[1] "잘생기/P+었다/E"
```

인물 네트워크

인물 네트워크



노드와 엣지

노드 : 사건에서 해당 인물의 중요도와 성격

엣지 : 두 인물 간의 관계의 정도

인물 네트워크 (Cont.)

을미사변(乙未事變)은 1895년 음력 8월 20일(양력 10월 8일) 경복궁(景福宮)에서 명성황후 민씨가 조선 주재 일본 공사 미우라 고로(三浦梧樓)의 지휘 아래 일본군 경성 수비대 미야모토 다케타로(宮本竹太郎)^[1] 등에게 암살된 사건이다.

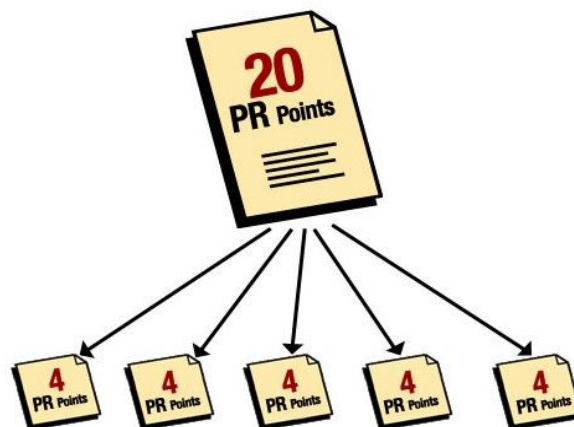
일본은 조선을 침략하는 데 가장 큰 걸림돌인 명성황후를 살해하려 했고, 조선 측에서도 흥선대원군, 우범선(禹範善), 이두황(李斗璜), 이진호(李軫鎬), 전 군부협판 이주회(李周會), 개화파 문신 유길준, 친일파 송병준의 사위 구연수(具然壽) 등 다양한 계층의 조선인 지도자들이 내통, 협조했다.

문서 요약

PAGE RANK
TEXT RANK

PAGE RANK란?

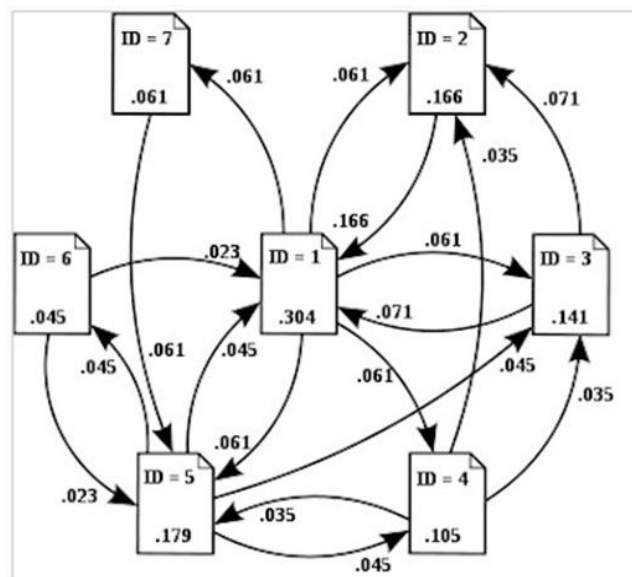
Google의 사이트 추천 알고리즘



20 PR Points / 5 Links

$$PR(u) = \sum_{v \in B_u} \frac{PR(v)}{L(v)},$$

$$PR(A) = \frac{1-d}{N} + d \left(\frac{PR(B)}{L(B)} + \frac{PR(C)}{L(C)} + \frac{PR(D)}{L(D)} + \dots \right).$$



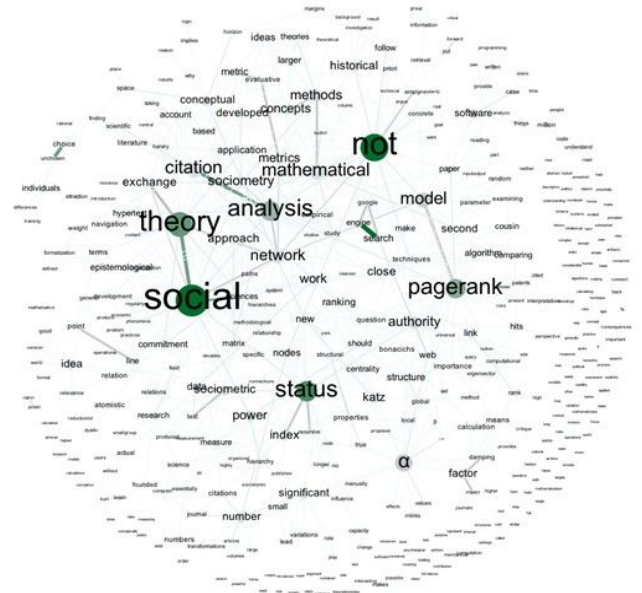
TEXT RANK란?

PAGE RANK를 문서에 적용

사이트 -> 단어

링크 -> 동시출현

문장 점수 계산



시각화

TEXT RANK란?(Cont.)

PAGE RANK를 문서에 접목

두산백과

105인 사건

[百五人事件]

‘테라우치 총독암살미수사건’, ‘선천사건(宣川事件)’ 등으로도 불리며, ‘105인 사건’이라는 명칭은 제1심 재판에서 105명이 유죄 판결을 받은 데서 비롯되었다. 일제가 조선을 강점한 직후에 민족의식이 높았던 황해도와 평안도 지역의 민족운동을 탄압하기 위해 테라우치 총독에 대한 암살미수사건을 날조하여 일으킨 사건이다.

사건의 배경

1910년 8월, 일본이 조선의 국권을 빼앗아 강점한 뒤에도 평안도와 황해도 등 서북지역에서는 독립운동이 꾸준히 확산되고 있었다. 특히 1907년 안창호(安昌浩), 양기탁(梁起鐸), 신채호(申采浩) 등이 국권회복을 목적으로 조직한 항일비밀결사인 신민회(新民會)는 평양과 정주에 대성학교와 오산학교를 세워 민족의식을 고취시키고자 하였으며, 국외에 무관학교와 독립군 기지를 만들어 독립전쟁을 준비하려 했다.

1910년 10월 1일 조선총독부의 초대 총독으로 부임한 테라우치 마사타케(寺內正毅)는 헌병경찰제도에 기초한 ‘무단통치(武斷統治)’를 행하는 한편, 민족의식이 높았던 황해도와 평안도 지역에 대한 대대적인 탄압을 계획하였다. 그는 1910년 12월 압록강 철교 개통식에 참석하기 위해 평양(平壤), 선천(宣川), 신의주(新義州) 등을 시찰했는데, 이때 조선인들이 그를 암살하려는 계획을 추진하였으나 실패로 끝났다는 소문이 돌았다. 조선총독부는 안중근(安重根)의 사촌인 안명근(安明根)이 1910년 12월에 무관학교 설립을 위한 독립운동자금을 모으다

> output

[1] “‘테라우치 총독암살미수사건’, ‘선천사건(宣川事件)’ 등으로도 불리며, ‘105인 사건’이라는 명칭은 제1심 재판에서 105명이 유죄 판결을 받은 데서 비롯되었다.”

[2] “ 일제가 조선을 강점한 직후에 민족의식이 높았던 황해도와 평안도 지역의 민족운동을 탄압하기 위해 테라우치 총독에 대한 암살미수사건을 날조하여 일으킨 사건이다.”

[3] “ 특히 1907년 안창호(安昌浩), 양기탁(梁起鐸), 신채호(申采浩) 등이 국권회복을 목적으로 조직한 항일비밀결사인 신민회(新民會)는 평양과 정주에 대성학교와 오산학교를 세워 민족의식을 고취시키고자 하였으며, 국외에 무관학교와 독립군 기지를 만들어 독립전쟁을 준비하려 했다.”

[4] “1910년 10월 1일 조선총독부의 초대 총독으로 부임한 테라우치 마사타케[寺內正毅]는 헌병경찰제도에 기초한 ‘무단통치(武斷統治)’를 행하는 한편, 민족의식이 높았던 황해도와 평안도 지역에 대한 대대적인 탄압을 계획하였다.”

[5] “ 그는 1910년 12월 압록강 철교 개통식에 참석하기 위해 평양(平壤), 선천(宣川), 신의주(新義州) 등을 시찰했는데, 이때 조선인들이 그를 암살하려는 계획을 추진하였으나 실패로 끝났다는 소문이 돌았다.”

WEB을 활용한 시각화

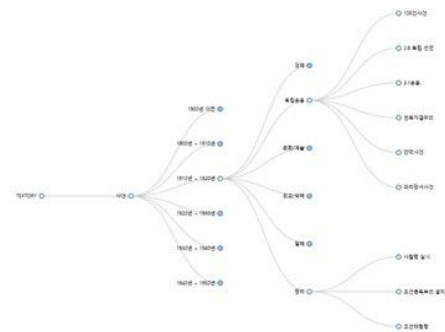
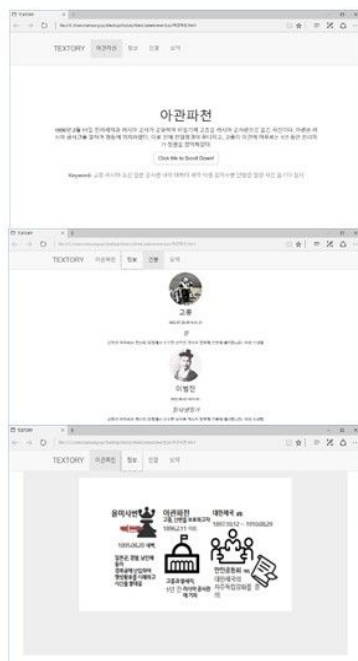
요약된 내용과 키워드



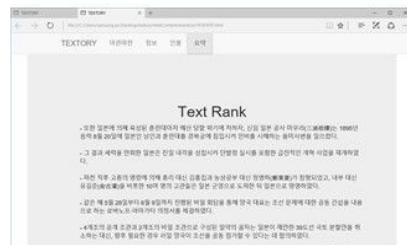
중요한 인물들!



.jpg



보고싶은 사건만!



간단한 설명



WEB에서 구현하기