

NLP & AI 활용하여 한국어 중급 학습자를 위한 뉴스, 강의, 토크쇼 콘텐츠 분석 프로젝트

기말 프로젝트 발표

202192037 짬타이바오

주제 선정 배경 및 필요성

- 한국어 학습자가 실제 미디어 콘텐츠를 통해 어휘와 표현을 더 효과적으로 학습할 수 있도록 돕기 위함
- 기존 사전 앱은 단어 정의만 제공하여 맥락에 따른 이해가 부족
- 뉴스와 강의 콘텐츠는 한국어 중급자에게 실질적인 언어적 유연성을 제공할 수 있는 좋은 자료

주제 선정 배경 및 필요성

주요 뉴스 >



한겨레

“트럼프, 국무장관엔 ‘충성파’ 루비오...안보보좌관은 북-미회담 회의론자 왈츠”

1시간 전

경향신문

[속보] “트럼프, 국무장관엔 ‘대중국 매파’ 마르코 루비오 의원 발탁 예정”

2시간 전

연합뉴스


“트럼프, 국가안보보좌관에 그린베레 출신 왈츠 하원의원 발탁”(종합)

2시간 전 • 작성자: 조준형

동아일보

[속보] “트럼프, 국무장관엔 마르코 루비오 상원의원 발탁 예정”

2시간 전

 관련 콘텐츠

국제 미국 중남미

“트럼프, 국무장관엔 ‘충성파’ 루비오...
안보보좌관은 북-미회담 회의론자 왈츠”

월스트리트저널 등 보도

이본영 기자

수정 2024-11-12 11:27 등록 2024-11-12 11:20

 기사를 읽어드립니다

3:26 ▶ 🔊



마이를 볼스 미국 공화당 하원의원, AP 연합뉴스

도널드 트럼프 미국 대통령 당선자가 우크라이나 지원에 회의적이고 중국에는 강경한 입장을 지닌 마이클 왈츠 공화당 하원위원을 백악관 국가안보보좌관으로 지명할 예정이라고 월스트리트저널 등이 11일(현지시간) 보도했다.

월스트리트저널은 이에 대해 아는 관계자들을 인용해 트럼프가 외교·안보 분야에서 자신을 보좌하고 관련 기관들을 조정하는 역할을 왈츠에게 제안했다고 전했다.

왈츠는 육군 특수부대인 그린베레와 주방위군 등에서 27년간 복무하며 아프가니스탄, 중동, 아프리카에 파병된 경력을 갖고 있다. 무공을 인정받아 동성훈장 등을 받았다. 미국 국방부와 백악관에서도 근무했다. 역시 군인 출신인 왈츠의 아내도 트럼프 행정부 1기 때 백악관 보좌관을 지냈다.

주제 선정 배경 및 필요성

도널드 트럼프 미국 대통령 당선자가 우크라이나 지원에 회의적이고 중국에는 강경한 입장을 지닌 마이클 왈츠 공화당 하원의원을 백악관 국가안보보좌관으로 지명할 예정이라고 월스트리트저널 등이 11일(현지시각) 보도했다.

월스트리트저널은 이에 대해 아는 관계자들을 인용해 트럼프가 외교·안보 분야에서 자신을 보좌하고 관련 기관들을 조정하는 역할을 중책을 왈츠에게 제안했다고 전했다.

왈츠는 육군 특수부대인 그린베레와 주방위군 등에서 27년간 복무하며 아프가니스탄, 중동, 아프리카에 파병된 경력을 갖고 있다. 무공을 인정받아 동성훈장 등을 받았다. 미국 국방부와 백악관에서도 근무했다. 역시 군인 출신인 왈츠의 아내도 트럼프 행정부 1기 때 백악관 보좌관을 지냈다.



1. 도널드 트럼프 미국 대통령 당선자가 우크라이나 지원에 회의적이고 중국에는 강경한 입장을 지닌 마이클 왈츠 공화당 하원의원을 백악관 국가안보보좌관으로 지명할 예정이라고 월스트리트저널 등이 11일(현지시각) 보도했다.
2. 월스트리트저널은 이에 대해 아는 관계자들을 인용해 트럼프가 외교·안보 분야에서 자신을 보좌하고 관련 기관들을 조정하는 역할을 중책을 왈츠에게 제안했다고 전했다.
3. 왈츠는 육군 특수부대인 그린베레와 주방위군 등에서 27년간 복무하며 아프가니스탄, 중동, 아프리카에 파병된 경력을 갖고 있다.
4. 무공을 인정받아 동성훈장 등을 받았다.
5. 미국 국방부와 백악관에서도 근무했다.
6. 역시 군인 출신인 왈츠의 아내도 트럼프 행정부 1기 때 백악관 보좌관을 지냈다.

주제 선정 배경 및 필요성

1. 도널드 - 唐纳德 (Tángnàdé) - Donald
2. 트럼프 - 特朗普 (Tèlǎngpǔ) - Trump
3. 미국 - 美国 (Měiguó) - United States
4. 대통령 - 总统 (Zǒngtǒng) - President
5. 당선자 - 当选人 (Dāngxuǎnrén) - President-elect
6. 우크라이나 - 乌克兰 (Wūkèlán) - Ukraine
7. 지원 - 支持 (Zhīchí) - Support
8. 회의적이다 - 怀疑的 (Huáiyí de) - Skeptical
9. 중국 - 中国 (Zhōngguó) - China
10. 강경한 - 强硬的 (Qiángyìng de) - Firm
11. 입장 - 立场 (Lìchǎng) - Position
12. 마이클 - 迈克尔 (Màikè'ěr) - Michael
13. 왈츠 - 沃尔兹 (Wò'ěrzi) - Waltz
14. 공화당 - 共和党 (Gònghédǎng) - Republican Party
15. 하원의원 - 众议员 (Zhòngyìyuán) - Congressman
16. 백악관 - 白宫 (Báigōng) - White House
17. 국가안보보좌관 - 国家安全顾问 (Guójiā Ānquán Gùwèn) - National Security Advisor
18. 지명하다 - 指定 (Zhīdìng) - Nominate
19. 예정 - 预定 (Yùdìng) - Expected
20. 월스트리트저널 - 华尔街日报 (Huá'ěrjiē Rìbào) - Wall Street Journal
21. 보도하다 - 报道 (Bàodào) - Report
22. 관계자 - 相关人士 (Xiāngguān rénshì) - Source
23. 인용하다 - 引用 (Yīnyòng) - Cite
24. 외교 - 外交 (Wàijiāo) - Diplomacy
25. 안보 - 安保 (Ānbǎo) - Security
26. 분야 - 领域 (Lǐngyù) - Field
27. 보좌하다 - 辅佐 (Fǔzuǒ) - Assist
28. 기관 - 机构 (Jīgòu) - Agency
29. 조정하다 - 协调 (Xiétiáo) - Coordinate
30. 역할 - 角色 (Juésè) - Role
31. 중책 - 重责 (Zhòngzé) - Responsibility
32. 제안하다 - 提议 (Tíyì) - Propose
33. 육군 - 陆军 (Lùjūn) - Army
34. 특수부대 - 特种部队 (Tèzhōng bùduì) - Special Forces
35. 그린베레 - 绿色贝雷帽部队 (Lǜsè bèiléimào bùduì) - Green Berets
36. 주방위군 - 州国民警卫队 (Zhōu guómín jǐngwèiduì) - National Guard
37. 27년간 - 27年间 (27 Niánjiān) - For 27 Years
38. 복무하다 - 服役 (Fúyì) - Serve
39. 아프가니스탄 - 阿富汗 (Āfúhàn) - Afghanistan
40. 중동 - 中东 (Zhōngdōng) - Middle East
41. 아프리카 - 非洲 (Fēizhōu) - Africa
42. 파병되다 - 派兵 (Pàibīng) - Deploy
43. 경력 - 经历 (Jīnglì) - Experience
44. 무공 - 勋功 (Xūngōng) - Military Merit
45. 인정받다 - 认可 (Rènkě) - Recognized
46. 동성훈장 - 铜星勋章 (Tóngxīng xūnzhāng) - Bronze Star Medal
47. 국방부 - 国防部 (Guófángbù) - Department of Defense
48. 근무하다 - 工作 (Gōngzuò) - Work
49. 군인 - 军人 (Jūnrén) - Soldier
50. 출신 - 出身 (Chūshēn) - Background
51. 아내 - 妻子 (Qīzi) - Wife
52. 행정부 - 政府 (Zhèngfǔ) - Administration
53. 1기 - 第一期 (Dì yī qī) - First Term
54. 보좌관 - 助理 (Zhùlǐ) - Advisor
55. 지내다 - 担任 (Dānrèn) - Serve

목표 결과물

- PDF 및 Excel 출력: 분석된 단어 목록, 동의어 및 반의어, 유사 문장과 관련 기사 링크 포함
- 사용자 친화적인 인터페이스: URL 입력 시 자동으로 PDF와 Excel 파일의 생성
- 자동화된 콘텐츠 추천 시스템: 원문에서 추출된 핵심 키워드와 유사한 주제를 가진 추가 콘텐츠 추천

추진 전략

- 기술 선정: Python (Flask), Selenium, ~~transformers~~ 라이브러리로 NLP 모델 구현
- 모듈별 개발:
- ~~텍스트 파싱 및 URL 분류~~
- NLP 처리를 통한 단어 추출 및 유사 문장 생성
- ~~PDF 및 Excel 파일 생성 -> CSV FILE~~
- 추천 시스템 및 결과 출력
- 모델 학습: BERT 기반 모델을 사용하여 한국어 자연어 처리 강화

추진 일정

- 1주차: 주제 선정 및 기획, 기술 조사 및 라이브러리 설치
- 2-3주차: URL에서 텍스트 추출 모듈과 NLP 모듈 개발
- 4-5주차: PDF 및 Excel CSV 생성 모듈 개발, 콘텐츠 추천 시스템 구축
- 6주차: 최종 테스트 및 디버깅, 프로젝트 발표 자료 준비

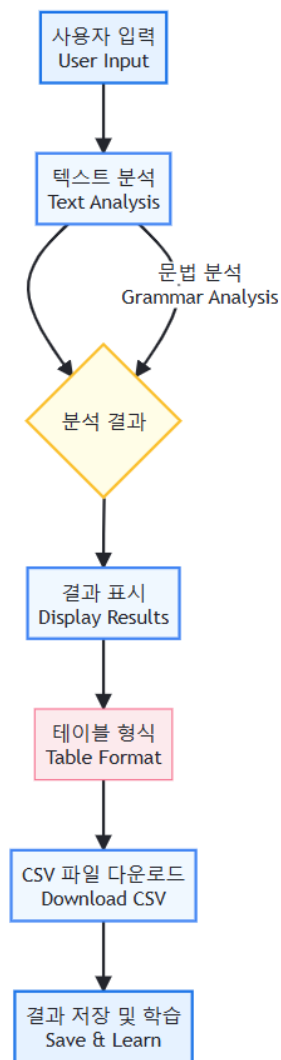
참고 자료

- KONLPy 라이브러리 문서: <https://konlpyko.readthedocs.io/>
- Transformers 라이브러리: <https://huggingface.co/transformers/>
- Flask Framework: <https://flask.palletsprojects.com/>
- API : Claude 3.5 Haiku, Gemini
- Streamlit : A faster way to build and share data apps

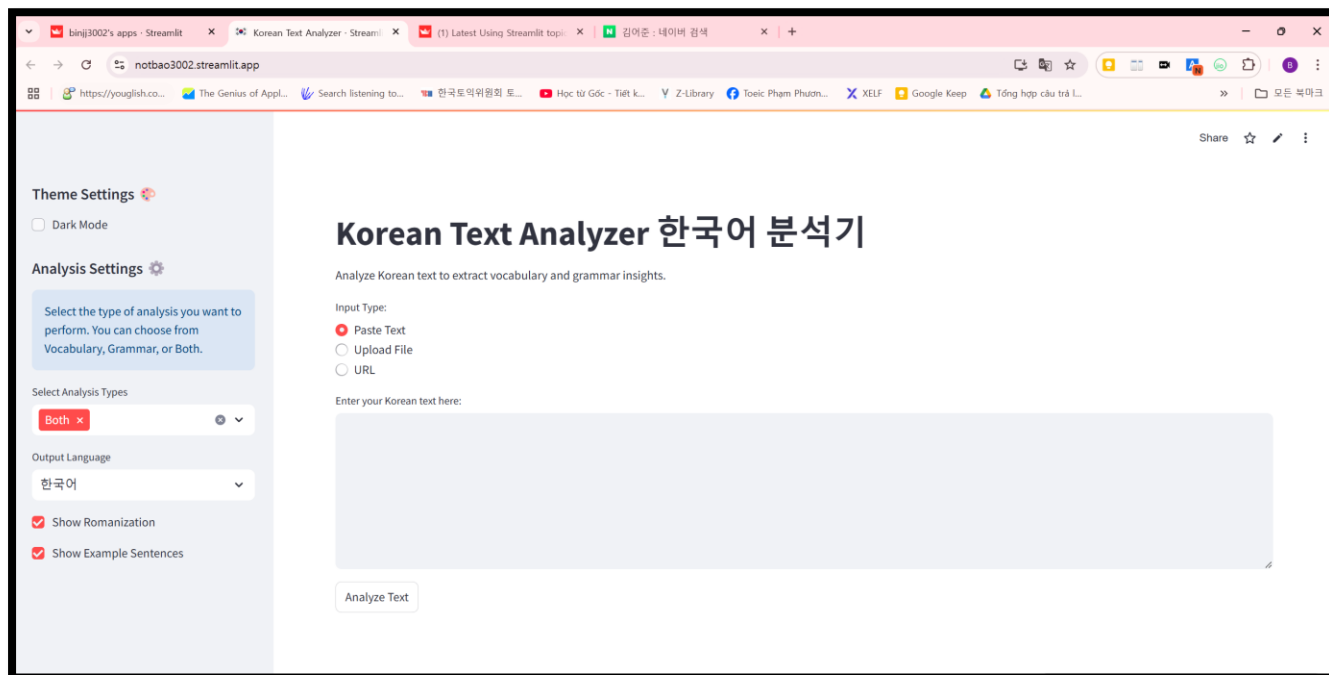
기말 보고서

The screenshot shows a web application titled "Korean Text Analyzer 한국어 분석기". On the left is a sidebar with "Theme Settings" (Dark Mode toggle) and "Analysis Settings" (Input Type: Paste Text, Upload File, or URL; Select Analysis Types: Both, Vocabulary, Grammar, or Both; Output Language: 한국어; Show Romanization; Show Example Sentences). The main area has a heading "Korean Text Analyzer 한국어 분석기" and a subtitle "Analyze Korean text to extract vocabulary and grammar insights." Below this are input options (Paste Text, Upload File, URL), a text input field "Enter your Korean text here:", and an "Analyze Text" button.

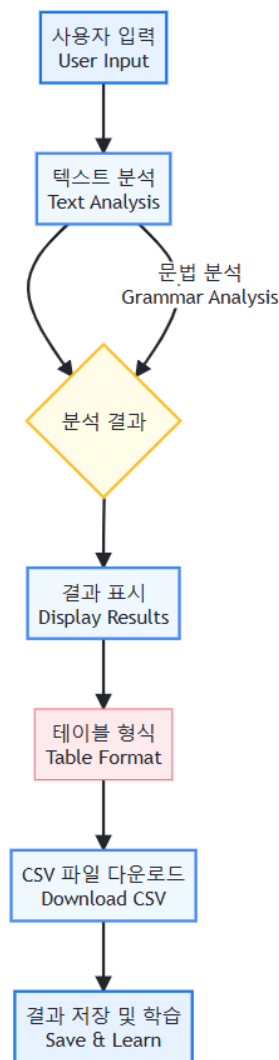
Pipeline



기말 보고서



Deploy



시연



기말 보고서

1. API 불러오는데 제일 중요한 건 TOKEN 관리
2. 추가로 단어를 쉽게 설명해주는 기능 추가
3. PROMT 개선하여 더 좋은 성능
느낀 점