

2024-2 TEXTMINING PROJECT

Analysis of Public Needs through Text Mining

BASED ON THE NATIONAL PETITION BOARD

Team 4

21900471 Jeongin Yook

22000294 Jemin Park

22100809 Eunji Hwang



Project Overview

1. Introduction

- **What is the National Petition Board?**

- A constitutional right (Article 26) allowing citizens to petition government agencies.
- Requests include: harm relief, correction of injustice, system reforms, and law changes.

- **Petition Example**

- The Gangseo PC Room Murder Case (2018) led to the passing of the Kim Seong-su Law after 1.19 million signatures.

공개청원 보기

검색어를 입력해 주세요. 상세검색 ▾

12건 있습니다.

최신순 ▾ 참여순 ▾ 만료순

방송통신위원회 ★
청소년들의 SNS 금지 법안 제정…
저는 현재 중학교 교사이자, 중2, 초6, 초4의 세 자녀를 둔 학부모이기도 한 대한민국의 평범한 40대 여성입니다. 오늘 딥페이크, SN…
2024.11.30.~2024.12.30. D-24

방송통신위원회 ★
디지털 성범죄 관련 처벌 법 강화
요즘 딥페이크가 상당히 심각합니다 무수히 많은 피해자들이 속출하고 있고 제2의 n번방 사건과 같이 텔레그램이란 어플에서 딥페이…
의견수렴기간: 2024.11.30.~2024.12.30. D-24

의견수렴기간: 2024.11.30.~2024.12.30. D-24



Project Overview

2. Project Background and Rationale for Topic Selection

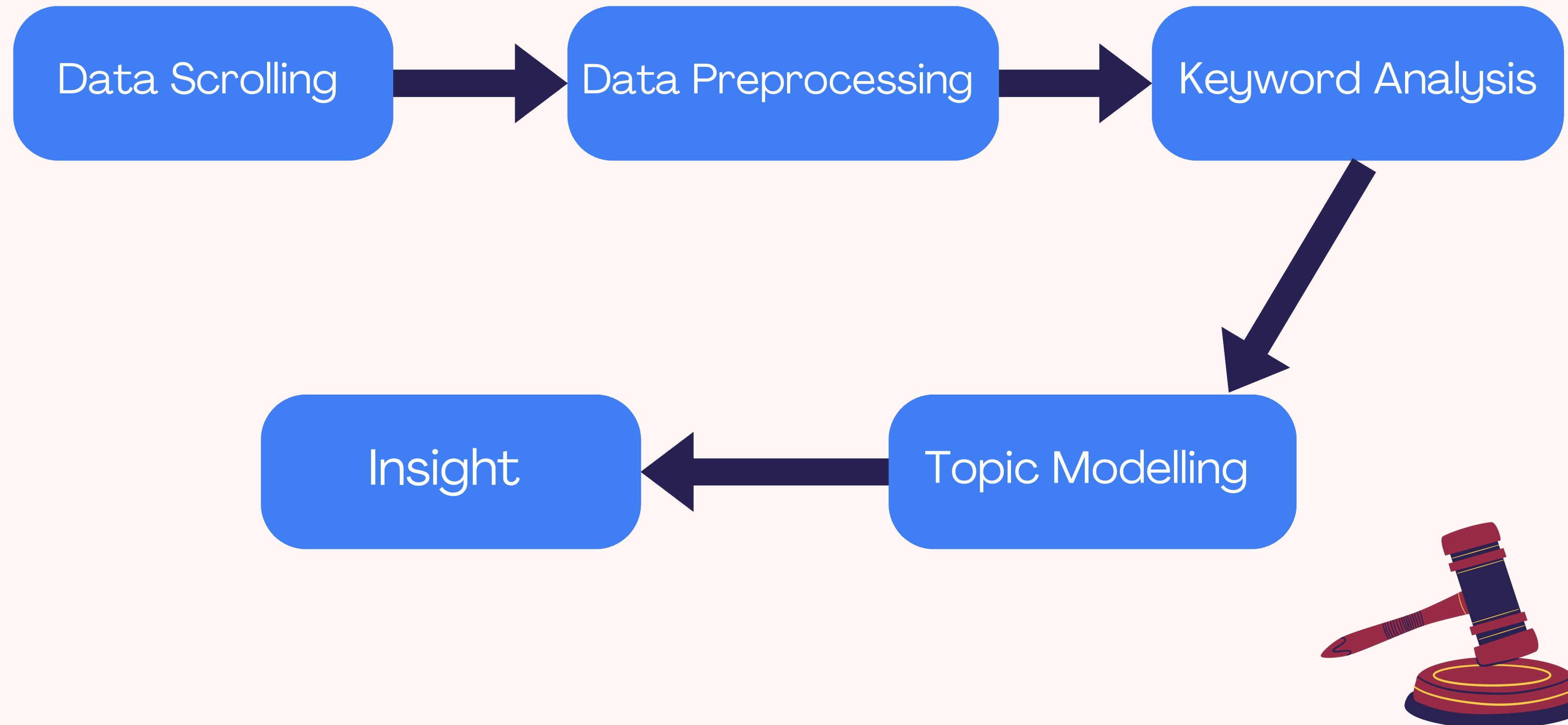
- Importance
 - Civic participation has made the Petition Board a major platform for public opinion.
- Unique Features
 - Directness: Citizens express views directly.
 - Voluntariness: Contributions are based on personal experiences.
 - Distinctiveness: Unlike news, the board captures authentic citizen voices.
- Data Analysis
 - Text mining and LDA techniques generate actionable insights from public demands.

Project Overview

3. Project Purpose

- Identify Key Issues
 - Utilize text mining techniques to analyze the public's concerns and demands
- Extract Keywords and Topics
 - Keyword Analysis: Use Term Frequency (TF) and N-gram analysis to extract key keywords
 - Topic Modeling: Apply LDA to identify and visualize major topics and related issues
- Suggest Policy Implications
 - Based on the analysis results, propose data-driven policy recommendations

Work Flow



Data

- Source: 2,953 public petitions on Cheongwon 24
(<https://www.cheongwon.go.kr/portal/petition/open/view>)
- Date: January 2023 – November 2024
- Collection Method: HTML documents of the petition board were extracted using the Rvest library and saved as RData.
- Data Content: 2,953 observations / 5 variables: "page_id", "title", "text", "label", "date"



Data Crawling

Refer to the petition bulletin HTML Tag with Rvest Library

- Using the rvest library, petition board entries are scraped through crawling.
- Key data such as date, department, title, and content are extracted based on the tags containing the text information.



The screenshot shows a petition bulletin board from Chongwon24. The main page displays a search bar and a list of petitions. A specific petition entry for '중증장애인 기준 개정 요구' (Amendment Request for Persons with Severe Disabilities) is highlighted. The DevTools Elements tab is open over the petition details, showing the HTML structure of the petition card. The selected element is a list item (``) containing a label for a favorite petition (`<label class="favorite" data-id="N" title="관심청원 등록">...`) and a link to the petition detail view (`...`). The petition title is shown in a blue box (`span.subject`).

```
<li><label class="favorite" data-id="N" title="관심청원 등록">...</label>
<a href="#" onclick="fn_detailView('PRI6be0361b9cc34b22ada259fa6cbebfd3')>고용노동부</a>
<span class="subject" title="중증장애인 기준 개정 요구"> 중증장애인 기준 개정 요구 </span> == $0
<span class="text" style="word-break: break-all;">...</span>
<span class="date">...</span>
<span class="deemical red">종료</span>
</a>
</li>
```

The CSS styles for the petition card are also visible in the Styles tab of DevTools:

```
.open_box .list_card a .subject {
    display: block;
    font-size: 24px;
    font-weight: 600;
    line-height: 1.2em;
    letter-spacing: -0.5px;
    color: #111;
    white-space: nowrap;
    text-overflow: ellipsis;
    overflow: hidden;
}
```

Data Crawling

- Texts received through crawling
- The extracted data is organized into a structured dataset by converting it into a data frame and saved in R.Data format.

```
save(petition_data, file="petition_data.RData")
```

```
load("~/Downloads/petition_data.RData")
```

```
> petition_data %>% select(-text) %>% head
```

page_id	title	label
1	대한민국 피겨 스케이팅 국가대표 남자 선수들의 선수 활동 기간을 보장하는 제도 도입이 조속히 필요한 이유 문화체육관광부	
2	딥페이크 사건 관련, 성범죄를 야기시키는 불법 유해 성인사이트, 성매매 사이트에 대한 처벌 수위를 강화주세요. 방송통신위원회	
3	청소년들의 SNS 금지 법안 제정에 관한 요구 방송통신위원회	
4	디지털 성범죄 관련 처벌 법 강화 방송통신위원회	
5	금속성 분진 관련 안전 조치 강화 및 관리미흡에 대한 형사처벌 규정 제정 청원 고용노동부	
6	숙박시설 CCTV 설치 의무화 요청 보건복지부	
date		
1	2024.12.02	
2	2024.11.30	
3	2024.11.30	
4	2024.11.30	
5	2024.11.30	
6	2024.11.30	

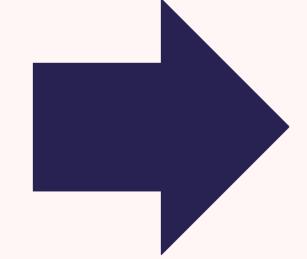


Text

Preprocessing

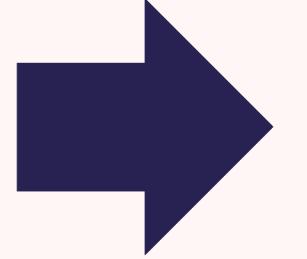
```
Top Text Keywords:  
# A tibble: 10 × 2  
  keywords     n  
  <chr>    <int>  
1 있      315  
2 도      286  
3 것      272  
4 은      269  
5 과      147  
6 저      139  
7 되      135  
8 로      134  
9 주      121  
0 수      113
```

Remove
1Letter



```
Top Text Keywords:  
# A tibble: 10 × 2  
  keywords     n  
  <chr>    <int>  
1 으로      230  
2 소음      69  
3 플라스틱  66  
4 환경      63  
5 사람      62  
6 사용      44  
7 생각      41  
8 쓰레기     41  
9 경우      40  
10 따르     40
```

Add
Stopwords



```
# add stopwords  
stopwords <- c(  
  "그리고", "하지만", "또한", "정말",  
  "위해", "하여", "있다", "없다",  
  "으로", "에서", "하고", "대하",  
  "위하", "아니", "따르", "에게",  
  "부터", "까지", "관하", "인하",  
  "못하", '하는'
```

```
Top Text Keywords:  
# A tibble: 10 × 2  
  keywords     n  
  <chr>    <int>  
1 소음      69  
2 플라스틱  66  
3 환경      63  
4 사람      62  
5 사용      44  
6 쓰레기     41  
7 빨대      38  
8 발생      33  
9 기준      32  
10 때문     32
```

```
stopwords <- unique(c(stopwords,  
  intToUtf8(44032:55203, multiple = TRUE),  
  letters, LETTERS,  
  as.character(0:9)))
```

Text Preprocessing

library(KoNLP)

```
$`대통령님,`  
[1] "대통령님/NC+,/SP"
```

```
$문화체육관광부  
[1] "문화체육관광부/NC"
```

```
$유인촌  
[1] "유인촌/NC"
```

```
$`장관님,`  
[1] "장관님/NC+,/SP"
```

```
$대한빙상경기연맹  
[1] "대한빙상경기연맹/NC"
```

```
$윤홍근  
[1] "윤홍근/NC"
```

```
SimplePos22(petition_data$text[1])
```

- **Korean Language Specialization**

- The Korean language combines particles, endings, and affixes, making morphological analysis challenging.
- KoNLP is specifically designed to handle these complexities.

- **Morphological Parsing and POS Tagging**

- KoNLP splits input text into morphological units and provides Part-Of-Speech (POS) tags.
- Common POS categories include nouns (N), verbs (V), adjectives (A), particles (J), and more.

Text Preprocessing

- Morphological analysis is performed on the content of petition posts using the KONLP library.
- During the text preprocessing phase, Korean text is processed to extract and return words in the form of nouns.
- For example, the word '청원합니다' is divided into '청원'(noun) and '합니다'(verb). Meaningful terms are primarily composed of nouns.

```
# Keyword analysis function
analyze_keywords <- function(label, petition_data, stopwords) {
  # Filter data by label
  filtered_data <- petition_data %>%
    filter(label == !!label)

  # Function to extract keywords
  extract_keywords <- function(text) {
    clean_text <- str_replace_all(text, "[^가-힣\\s]", "") %>% str_squish() # Remove special characters
    stems <- stem_text(clean_text) # Extract stems
    nouns <- extract_nouns(stems) # Extract nouns only
    keywords <- nouns[!nouns %in% stopwords] # Remove stopwords
    keywords <- keywords[nchar(keywords) > 1] # Remove single-character words
    return(keywords)
  }

  # Extract Title keywords
  title_keywords <- filtered_data %>%
    mutate(keywords = sapply(title, extract_keywords)) %>%
    unnest(keywords) %>%
    count(keywords, sort = TRUE)

  # Extract Text keywords
  text_keywords <- filtered_data %>%
    mutate(keywords = sapply(text, extract_keywords)) %>%
    unnest(keywords) %>%
    count(keywords, sort = TRUE)

  # Return results
  list(label = label, title_keywords = title_keywords, text_keywords = text_keywords)
}
```

Text Preprocessing

```
# 데이터를 로드
load("~/Downloads/petition_data.RData")

# 상위 10개 레이블 추출
top_labels <- petition_data %>%
  count(label, sort = TRUE) %>%
  slice_head(n = 10) %>%
  pull(label)

# 각 레이블에 대해 키워드 분석 실행
all_results <- lapply(top_labels, analyze_keywords, petition_data = petition
```

- Through the above process, preprocessed text values are analyzed based on frequency to identify the most commonly appearing words for each department.
 - Using a word cloud, it becomes possible to visualize the keywords that frequently appear for each department.



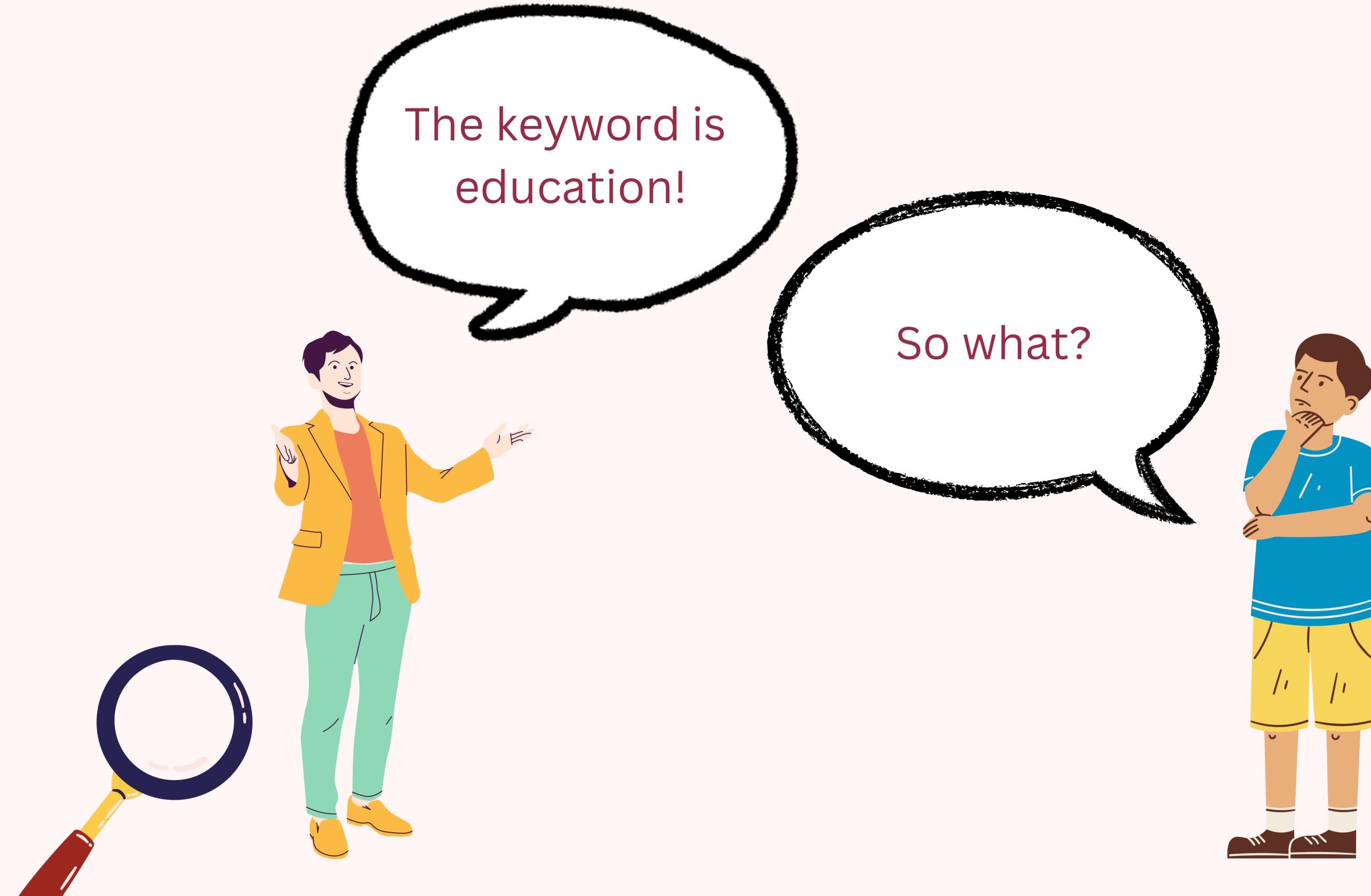
Keyword Analysis

- 3 keywords that appeared the most in each department

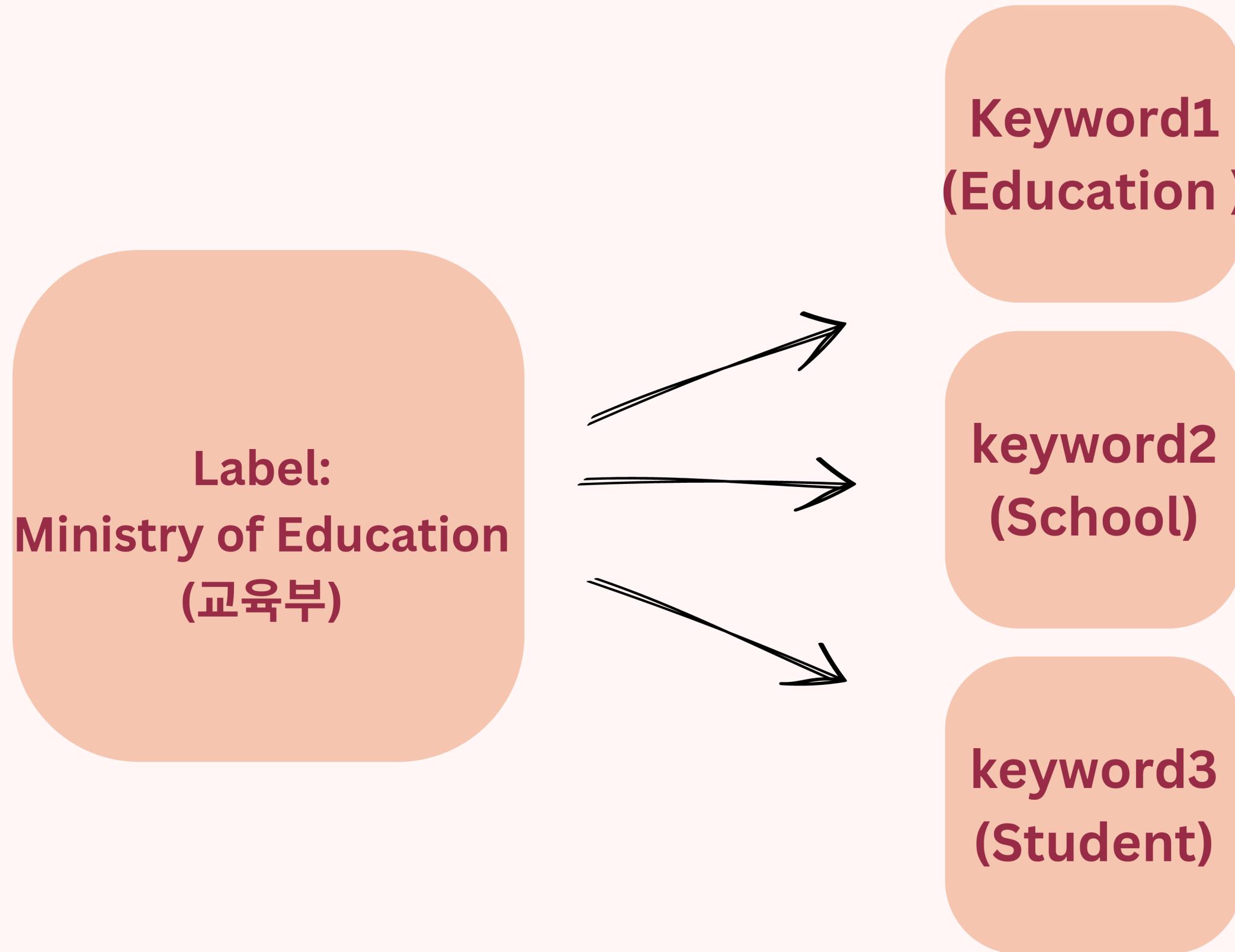
#	A tibble: 30 × 3	label	keyword	frequency
		<chr>	<chr>	<int>
1	경찰청 경찰	경찰청	경찰	205
2	경찰청 차량	경찰청	차량	159
3	경찰청 보행자	경찰청	보행자	135
4	고용노동부 근로자	고용노동부	근로자	221
5	고용노동부 교육	고용노동부	교육	197
6	고용노동부 육아휴직	고용노동부	육아휴직	176
7	교육부 교육	교육부	교육	347
8	교육부 학교	교육부	학교	257
9	교육부 교사	교육부	교사	157
10	국토교통부 개정	국토교통부	개정	103
11	국토교통부 정부	국토교통부	정부	84
12	국토교통부 설치	국토교통부	설치	79
13	법무부 범죄	법무부	범죄	485
14	법무부 개정	법무부	개정	355
15	법무부 처벌	법무부	처벌	306
16	보건복지부 지원	보건복지부	지원	272
17	보건복지부 치료	보건복지부	치료	192
18	보건복지부 국민연금	보건복지부	국민연금	175
19	서울특별시 지하철	서울특별시	지하철	53
20	서울특별시 장애인콜택시	서울특별시	장애인콜택시	45
21	서울특별시 중증장애인	서울특별시	중증장애인	44
22	인천광역시 사업	인천광역시	사업	663
23	인천광역시 원당태리	인천광역시	원당태리	566
24	인천광역시 시설	인천광역시	시설	314
25	행정안전부 처리	행정안전부	처리	161
26	행정안전부 법률	행정안전부	법률	142
27	행정안전부 공무원	행정안전부	공무원	126
28	환경부 소음	환경부	소음	69
29	환경부 플라스틱	환경부	플라스틱	66
30	환경부 환경	환경부	환경	63



Keyword Analysis



N-gram Analysis



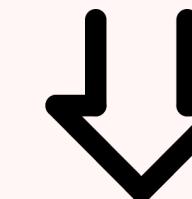
N-gram Analysis

	label	keyword	ngrams	n
	<chr>	<chr>	<chr>	<int>
1	교육부 교육	생존 수영 교육		19
2	교육부 교육	교육 지원 사업		8
3	교육부 교육	수영 교육 실시		7
4	교육부 교육	자기 맞춤 교육과정		7
5	교육부 교육	교육 품질 향상		6
6	교육부 교육	교사 교육 품질		5
7	교육부 교육	교육 급여 지급		5
8	교육부 교육	선행학습 사교육 학생		5
9	교육부 교육	우리나라 교육 백년대계		5
10	교육부 교육	맞춤 교육과정 운영		4

	label	keyword	ngrams	n
	<chr>	<chr>	<chr>	<int>
1	고용노동부 육아휴직 아빠 육아휴직 보너스			10
2	고용노동부 육아휴직 근로자 육아휴직 신청			5
3	고용노동부 육아휴직 사업 주의 육아휴직			5
4	고용노동부 육아휴직 개월 육아휴직 사용			4
5	고용노동부 육아휴직 올해 아빠 육아휴직			4
6	고용노동부 육아휴직 육아휴직 급여 신청			4
7	고용노동부 육아휴직 육아휴직 산재 휴직			4
8	고용노동부 육아휴직 육아휴직 허용 의무			4
9	고용노동부 육아휴직 기간 근로자 육아휴직			3
10	고용노동부 육아휴직 육아휴직 급여 수령			3

1. People want to teach children to swim for survival.
2. People want to raise the level of customized education

1. Workers are dissatisfied with parental leave.
2. There are many opinions about improving fathers' parental leave.



The petitioners must be parents!

N-gram Analysis

Google

교육부 수영

전체 이미지 동영상 뉴스 쇼핑 지도 도서 : 더보기

도구

◐ 중도일보
[교육부 생존수영 전 학년 확대 계획, 4년째 '지지부진'](#)
19일 교육부 등에 따르면 생존수영 전 학년 확대를 위한 교육부 차원의 지원은 현재 전무하다. 앞서 교육부가 2018년 발표한 '제2차 학교체육진흥 기본...'
2024. 3. 19.

◐ 조선일보
[초등학교 '생존 수영' 의무화됐더니... 물엔 안 들어가고 참관 수업](#)
세월호 참사 이후 교육부는 초등학생을 대상으로 '생존 수영' 수업을 의무화했다. 세월호 침몰 같은 위기 상황에서 학생들이 최소한의 대응을 할 수...
2024. 4. 16.

◐ 매일신문
[문경중에 최신수영장 들어서... 교육부 학교복합시설 공모사업 선정](#)
국비 44억 확보 25m 6레인 규모.. 2026년말 완공계획. 문경시가 운영 ... 경북 문경중학교 운동장에 학생과 지역주민이 함께 이용할 수 있는 최신 수영장이...
2024. 8. 13.

Google 육아휴직

전체 이미지 동영상 뉴스 쇼핑 지도 도서 : 더보기

도구

◐ 한겨레
[6.8% 남성 육아휴직률, 정부 "2030년 70%까지 올리겠다"](#)
정부가 현재 한자릿수에 불과한 남성의 육아휴직 사용률을 2030년까지 70%까지 끌어올리겠다고 밝혔다.
대통령 직속 저출산 고령사회위원회(저고위)는 3일 제6차...
2일 전

◐ 경향신문
[전북, '남성 육아휴직' 장려금 준다... 3개월간 90만원 지원](#)
0.78명. 2023년 전북의 합계출생률이다. 전국 9개 광역시도 중 두 번째로 낮은 기록이다. 인구수는 174만명으로 10년 새 13만명 가까이 줄었다.
22시간 전

◐ 뉴스핌
[정부, '자동육아휴직제' 도입 사실상 백지화... 출산·육아휴직 통합 신청 허용](#)
엄마 독박 유아 가능성... 남녀 형평성도 안 맞아 "근로자가 기간을 정해 육아휴직 가는 게 효율적".
[세종=뉴스핌] 정성훈 기자 = 고용노동부가 출산...
16시간 전

◐ 디지털타임스
['男 육아휴직 70%·女 경력단절 10%' 쌍끌이로 출생 가속도](#)
정부가 남성의 육아휴직 사용률을 2030년까지 70%까지 끌어올리겠다는 목표를 세웠다. 같은 기간 30~44세 여성의 경력 단절 비율은 현 수준의 절반...
2일 전

disagreement with the opinion of the petition

agreement with the opinion of the petition

N-gram Analysis

label	keyword	ngrams	n
			<int>
1	법무부 범죄	중대 범죄 사건	10
2	법무부 범죄	범죄 장기 미만	7
3	법무부 범죄	특정 중대 범죄사건	6
4	법무부 범죄	공무원 직무상 범죄	5
5	법무부 범죄	직무상 범죄 가중	5
6	법무부 범죄	금고 범죄 장기	4
7	법무부 범죄	대상 강력 범죄	4
8	법무부 범죄	대한 민국 중범죄	4
9	법무부 범죄	범죄자 인권 중요	4
10	법무부 범죄	징역 금고 범죄	4

Highlighting various criminal issues

- major criminal cases
- crimes in the line of duty



label	keyword	ngrams	n
			<int>
1	보건복지부 지원	은둔 외톨이 지원	11
2	보건복지부 지원	활동 지원서 비스	9
3	보건복지부 지원	지원 은둔 외톨이	7
4	보건복지부 지원	아동 보호 지원	4
5	보건복지부 지원	보호 지원 법률	3
6	보건복지부 지원	생활 지원 직접	3
7	보건복지부 지원	일상 생활 지원	3
8	보건복지부 지원	폐지 시술 지원	3
9	보건복지부 지원	고위험 산모 지원	2
10	보건복지부 지원	난방지원금받고 중증장애인 난방	2

Emphasize social welfare policies

- Support for reclusive individuals
- Child protection support
- Living assistance

N-gram Analysis

label	keyword ngrams		
	<chr>	<chr> <chr>	
1 행정안전부 공무원	지방공무원 적극 행정	13	
2 행정안전부 공무원	공무원 위법 부당	5	
3 행정안전부 공무원	구제 공무원 위법	5	
4 행정안전부 공무원	기관 구제 공무원	5	
5 행정안전부 공무원	자서전 청렴등이공무원이 이유교보문고	5	
6 행정안전부 공무원	장관 자서전 청렴등이공무원이	5	
7 행정안전부 공무원	청렴등이공무원이 이유교보문고 질문	5	
8 행정안전부 공무원	공무원 오해 근무	4	
9 행정안전부 공무원	공무원 징계 예방	4	
10 행정안전부 공무원	민원 법률 국가공무원법	4	

Administrative management issues

- Civil service administration
- Civil service procedures and duties
- Civil servant burdens, etc.

label	keyword ngrams		n `label == "경찰청"``
	<chr> <chr> <chr>		<int> <lgl>
1 경찰청 차량	횡단보도 차량 일시정지	9	TRUE
2 경찰청 차량	차량 일시정지 의무	8	TRUE
3 경찰청 차량	운전자 횡단보도 차량	4	TRUE
4 경찰청 차량	감시 차량 시동	2	TRUE
5 경찰청 차량	마지막 차량 시동	2	TRUE
6 경찰청 차량	상대 차량 우회전	2	TRUE
7 경찰청 차량	세자녀 차량 스티커	2	TRUE
8 경찰청 차량	신호등 횡단보도 차량	2	TRUE
9 경찰청 차량	유턴을 시도 차량	2	TRUE
10 경찰청 차량	인승 차량 최소	2	TRUE

Safety issues between crosswalks and vehicles

- Temporary stop for vehicles at crosswalks
- Obligation for vehicles to stop
- Traffic-related policies



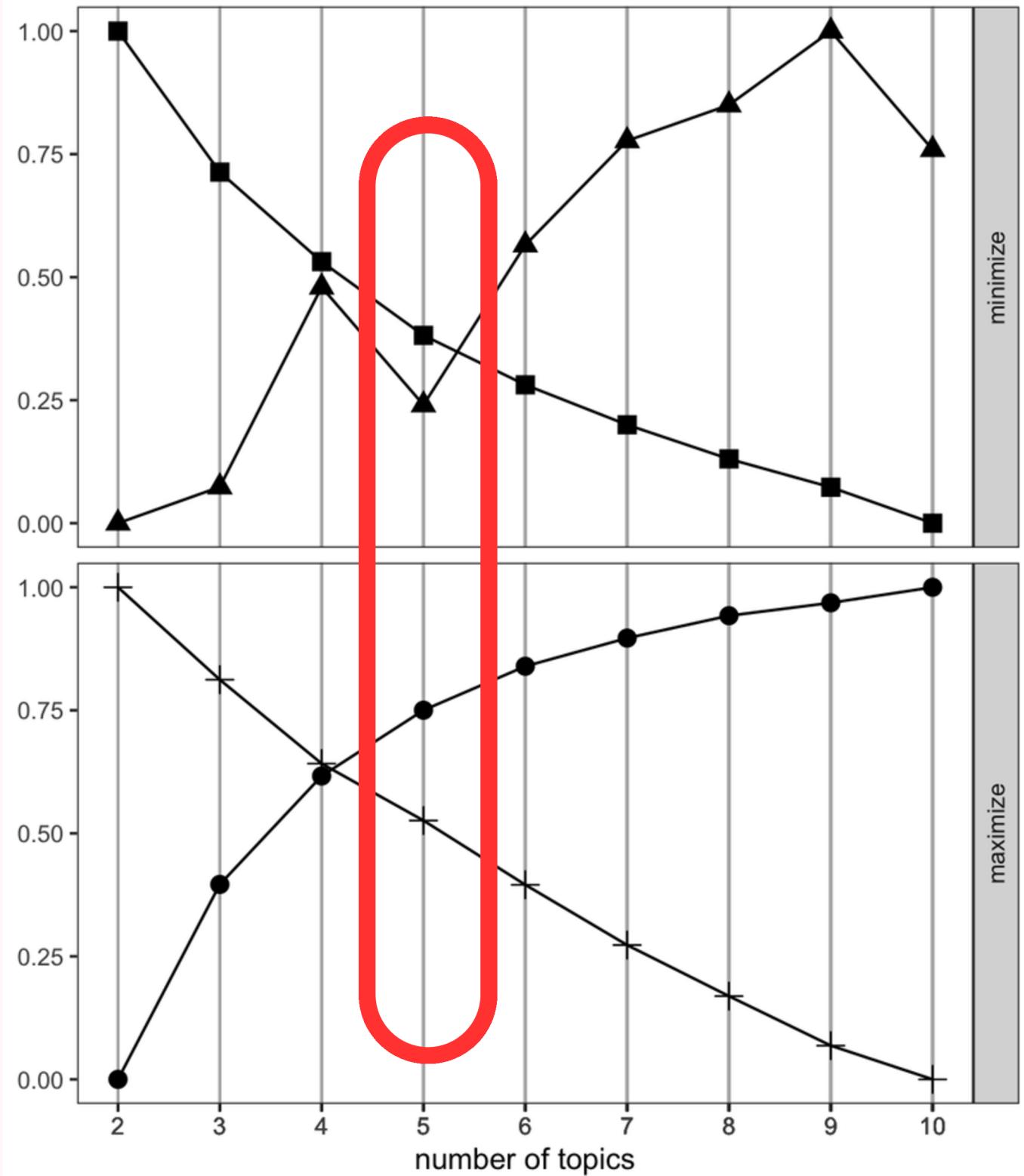
Topic Modeling - LDA

LDA is an unsupervised learning algorithm for topic modeling that identifies hidden topics within a collection of text. It assumes that each document is a mixture of multiple topics, and each topic consists of a specific distribution of words.

Process

1. Convert data into DTM
2. Determining the optimal K
 - a. "Griffths2004", "Deveaud2014", "CaoJuan2009", "Arun2010"
3. Run LDA

Topic Modeling - LDA



```
# 5. Use lda tuning to find the best number of topics
result <- FindTopicsNumber(
  dtm,
  topics = seq(2, 10, by = 1), # 테스트 할 k 값 (2~10)
  metrics = c("Griffiths2004", "CaoJuan2009", "Arun2010", "Deveaud2014"),
  method = "Gibbs",          # LDA 학습 방법
  control = list(seed = 1234), # 랜덤 시드 고정
  mc.cores = 2L,             # 병렬처리 코어 수
  verbose = TRUE)            # 진행상황 출력

# Main Analysis

# Assume `petition_data$title` contains the title text from your dataset
# Preprocess the titles
preprocessed_tokens <- preprocess_titles(petition_data$title, stopwords)

# Create a Document-Term Matrix (DTM) from the preprocessed tokens
dtm <- create_corpus(preprocessed_tokens)

# Perform LDA to find 5 topics
lda_results <- perform_lda(dtm, k = 5)

# Output the top 10 terms for each topic
print(lda_results$topics)
```

Topic Modeling

```
> print(lda_results$topics)
```

	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5
[1,]	"폐지"	"처벌"	"제정"	"폐지"	"지원"
[2,]	"변경"	"제정"	"국민연금"	"설치"	"설치"
[3,]	"규제"	"규정"	"폐지"	"규제"	"고가도로"
[4,]	"처벌"	"원당태리"	"금지"	"규정"	"폐지"
[5,]	"고가도로"	"규제"	"교육"	"국민연금"	"법률"
[6,]	"법률"	"불법"	"자동차"	"외국"	"원당태리"
[7,]	"원당태리"	"청소년"	"장애인"	"고가도로"	"자동차"
[8,]	"지원"	"유현사거리"	"지하철"	"법률"	"신설"
[9,]	"규정"	"시행"	"청소년"	"음주운전"	"제정"
[10,]	"국민연금"	"요청"	"도입"	"장애인"	"금지"

Total 86 cases!!!

인천광역시 ☆
원당태리 3번도로 고가도로 결…
안녕하세요. AA10-1 검단 e편한세상 웰카운티 입주민입니다. 검단신도시 내에 교통체증 완화를 위하여 생존권을 위협하는 행정을 일삼…
의견수렴기간: 2024.01.18.~2024.02.16. 종료

인천광역시 ☆
원당 태리 3번도로 고가 도로 절…
안녕하세요. AA10-1 검단 e편한세상 웰카운티 입주민입니다. 검단신도시 내에 원당~태리 3번 입체화시설을 고가도로로 건설하려는 …
의견수렴기간: 2024.01.18.~2024.02.16. 종료

인천광역시 ☆
원당태리 3번 고가도로 결사반대
안녕하세요. AA10-1 검단 e편한세상 웰카운티 입주민입니다. 검단신도시 내에 교통체증 완화를 위하여 생존권을 위협하는 행정을 일삼…
의견수렴기간: 2024.01.16.~2024.02.14. 종료

인천광역시 ☆
검단드림로 유현사거리 3번도로…
안녕하세요. 저는 AA10-1 인천광역시 서구 검단 e편한세상 웰카운티 입주민입니다. 유현사거리 3번 고가도로 신설에 관련하여 도로 …
의견수렴기간: 2024.01.16.~2024.02.14. 종료

인천광역시 ☆
검단드림로 유현사거리 3번도로…
안녕하세요. 저는 AA10-1 인천광역시 서구 검단 e편한세상 웰카운티 입주민입니다. 유현사거리 3번 고가도로 신설에 관련하여 도로 …
의견수렴기간: 2024.01.16.~2024.02.14. 종료

인천광역시 ☆
원당태리 3번 도로 지하화 요청
인천 서구 주민으로서 청원 하고자 하는게 있어 글을 남깁니다. 인천 검단신도시에 고가도로 설치가 논의 되고 있는걸로 들었습니다…
의견수렴기간: 2024.01.16.~2024.02.14. 종료

Deleting "원당태리 3번도로 고가도로 반대 청원"

```
> print(lda_results$topics)
```

	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5
[1,]	"폐지"	"처벌"	"폐지"	"폐지"	"설치"
[2,]	"규제"	"제정"	"설치"	"청소년"	"처벌"
[3,]	"법률"	"규정"	"장애인"	"지원"	"음주운전"
[4,]	"장애인"	"부활"	"규제"	"전면"	"법률"
[5,]	"요청"	"금지"	"불법"	"설치"	"제정"
[6,]	"국민연금"	"법률"	"반대"	"교육"	"규제"
[7,]	"전면"	"외국"	"교육"	"변경"	"국민연금"
[8,]	"처벌"	"가해자"	"해외직구"	"사용"	"금지"
[9,]	"변경"	"도입"	"방지"	"제정"	"변경"
[10,]	"규정"	"다자녀"	"육아휴직"	"도입"	"장애인"

LDA Analysis

ChatGPT Answer

Topic 1 : "폐지", "규제", "법률", "장애인", "요청", "국민연금", "전면", "처벌", "변경", "규정"

Public System Reforms and Legal Regulation Amendments on National Pensions and Disability

Topic 2 : "처벌", "제정", "규정", "부활", "금지", "법률", "외국", "가해자", "다자녀", "도입"

Crime Punishment and New Legal Enactments on Foreigners and Multi-child Families

Topic 3 : "폐지", "설치", "장애인", "규제", "불법", "반대", "교육", "해외직구", "방지", "육아휴직"

Consumer Protection and Infrastructure Development

Topic 4 : "폐지", "청소년", "지원", "전면", "설치", "교육", "변경", "사용", "제정", "도입"

Youth Protection and Educational Policy Improvements

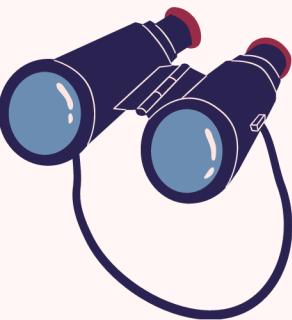
Topic 5 : "설치", "처벌", "음주운전", "법률", "제정", "규제", "국민연금", "금지", "변경", "장애인"

Social Issues and Transportation Safety Legislation



Conclusion & Insight

- The results of the n-gram keyword analysis and the LDA analysis showed some differences. This is likely due to the petitions being highly specific and containing many detailed topics.
- As a result, it was challenging to fully grasp the concrete content of the petitions through LDA analysis alone.
- However, the keyword analysis allowed for a better understanding of the specific content of the petitions, revealing that the concerns and needs of the public often differ from socially trending issues. (횡단보도 보행자 보호, 은둔형 외톨이 지원, 육아휴직 확대)



Thank you