

출장 보고서
2026. 07. 26. ~ 07. 30.
인도네시아 자카르타

세계여론조사협회(WAPOR) 아시아 태평양 제9차 콘퍼런스 참석 및 연구 발표

2026. 08. 21. (금)

이 용 수 경제정보분석실장
김 은 총 빅데이터분석팀 전문연구원
이 관 건 정책정보허브팀 전문연구원
정 한 나 여론분석팀 전문연구원

I. 출장 개요

1. 배경 및 목적

- ☐ 세계여론조사협회(이하 WAPOR)*가 주관하는 아시아 태평양 연례 콘퍼런스는 아시아 태평양 지역의 여론조사 전문가, 학자, 실무자들이 모여 최신 연구 동향을 공유하고 당면 과제를 논의하는 중요한 교류의 장임.

* 1947년 설립된 세계 최대 규모의 여론조사 전문 학술 단체로, 유네스코의 자문기구로 활동하며 여론조사의 국제 기준과 윤리 강령을 선도

- ☐ 경제교육·정보센터에서는 WAPOR 연례 콘퍼런스에 참석하여 경제정보분석실에서 수행하고 있는 빅데이터 분석과 설문조사를 융합한 최신의 여론 동향 분석 사례를 발표
- ☐ 또한, 본 출장에서 수집한 다양한 최신 조사 기법 및 분석 사례 등을 정리해 내부 워크숍, 브리핑 자료 등으로 공유하여 내부 직원의 역량 강화를 도모할 계획임.

2. 기간 및 장소

- ☐ 기간: 2026. 07. 26.(일) ~ 2026. 07. 30.(목)
- ☐ 장소: Universitas Islam Internasional Indonesia(UIII) (인도네시아 자카르타)

3. 출장자 및 출장 일정

- ☐ 출장자: 이용수 경제정보분석실장 외 3인

순번	이름	소속	직위
1	이용수	경제교육·정보센터 경제정보분석실	경제정보분석실장
2	김은총	경제정보분석실 빅데이터분석팀	전문연구원
3	이관건	경제정보분석실 정책정보허브팀	전문연구원
4	정한나	경제정보분석실 여론분석팀	전문연구원

□ 출장 일정

일자	시간	일정
7월 26일(일)	09:00 ~ 12:00	KDI-인천공항 이동
	14:55 ~ 20:10	인천공항-자카르타 공항 이동
	20:10 ~ 21:30	자카르타 시내로 이동
7월 27일(월)	08:30 ~ 18:00	WAPOR 등록 및 리셉션 참석(1일차)
7월 28일(화)	08:30 ~ 18:00	WAPOR 참석 및 발표(2일차)
7월 29일(수)	08:30 ~ 18:00	WAPOR 참석 및 발표(3일차)
	21:30 ~ 06:50	자카르타 공항-인천공항
7월 30일(목)	06:50	인천공항 도착
	07:00 ~ 10:00	인천공항-KDI 이동

II. 주요 사항

1. 세계여론조사협회(WAPOR) 아시아 태평양 및 제9차 콘퍼런스 소개

- 세계여론조사협회(WAPOR: World Association for Public Opinion Research)는 여론조사 및 설문조사 연구에 종사하는 연구자·전문가로 구성된 국제 학술·전문 단체임.
- 1947년 설립된 이후 70여 년 동안 공공여론 연구의 과학적 방법과 윤리 확립에 중점을 두고 활동해 왔으며, 현재 60여 개국에서 약 500명의 회원이 참여하고 있음.
- 공공여론 연구의 품질과 투명성을 위해 ‘WAPOR 전문윤리 및 실무규범 (Code of Professional Ethics and Practices)’ 을 제정·운영하고 있으며, 국제 학술지 International Journal of Public Opinion Research(IJPOR) 발간, 연례 국제학술대회 개최 등 다양한 학술 활동을 수행하고 있음.
- 북미·유럽 등 주요 지역 학회(AAPOR, ESRA/ESOMAR 등)와 협력하여 3년 주기의 국제 콘퍼런스 순환 개최 체계를 운영함으로써, 각국 학계·정부·민간 조사기관 간 교류의 장을 제공하고 있음.
- WAPOR Asia Pacific는 WAPOR의 아시아·태평양 지역 지부로, 역내 여론조사 및 설문조사 연구 협력을 담당하는 조직임.
- 2016년 7월 6일 WAPOR 이사회 승인으로 ‘WAPOR Asia’ 라는 명칭으로 공식 출범하였으며, 2022년 1월 1일부터 아시아·태평양 전체를 포괄하는 ‘WAPOR Asia Pacific’ 으로 명칭을 변경함.
- 아시아·태평양 지역에서 과학적 여론조사 방법론의 발전, 조사·데이터 공개의 자유 보장, 연구윤리 준수 등을 주요 목표로 설정하고, 각국 연구자 및 조사기관 간 네트워크 형성에 주력하고 있음.
- 2018년 대만(타이베이)에서의 첫 콘퍼런스를 시작으로 정기 연례 콘퍼런스, 웨비나, 특화 워크숍 등을 개최하며, 빅데이터·디지털 기술·AI 등 새로운 조사환경을 주제로 활발한 논의를 이어오고 있음.

- 제9차 WAPOR Asia Pacific 연례 콘퍼런스는 2026년 7월 27일~29일 인도네시아 자카르타 소재 Jakarta Public Library at Taman Ismail Marzuki, Ali Sadikin Building에서 개최되었으며, 인도네시아 국제이슬람대학교(UIN) 사회과학대학(FOSS)이 주관함.
- 이번 콘퍼런스의 공식 주제는 “빠르게 변화하는 세계 속의 여론(Public Opinion in a Fast-Changing World)”으로, 디지털 전환·소셜미디어 환경·AI·빅데이터 등 새로운 기술 여건이 여론 형성 및 여론조사 실무에 미치는 영향을 종합적으로 조망하는 데 초점을 두고 있음.
 - 21개국에서 약 110명의 연구자·실무자가 참가하였으며, 개회식은 자카르타 주지사(Pramono Anung)가 참석한 가운데 자카르타 시청(Balai Agung)에서 열려, 아시아·태평양 지역을 넘어 지역 정부 차원에서도 여론조사 연구에 대한 관심을 확인할 수 있었음.
 - 콘퍼런스 기간 중 서베이 거버넌스·표본추출 및 조사방법, 빅데이터·대안적 여론조사, AI·LLM 활용 측정, 민주주의(선거·정당성·신뢰·참여), 지정학, 종교·이념, 기후, 사회이슈(가족·젠더·건강), 미디어·디지털 사회 등 20여 개 세부 세션에서 총 90여 편의 논문이 발표됨.
 - 이번 콘퍼런스는 전통적 설문조사 기법과 빅데이터·AI 등 신기술을 어떻게 조화롭게 접목할 것인가에 대한 논의의 장으로 기능하였으며, 향후 여론조사 연구의 방법론적 정교화와 국제적 신뢰 제고에 기여할 것으로 전망함.

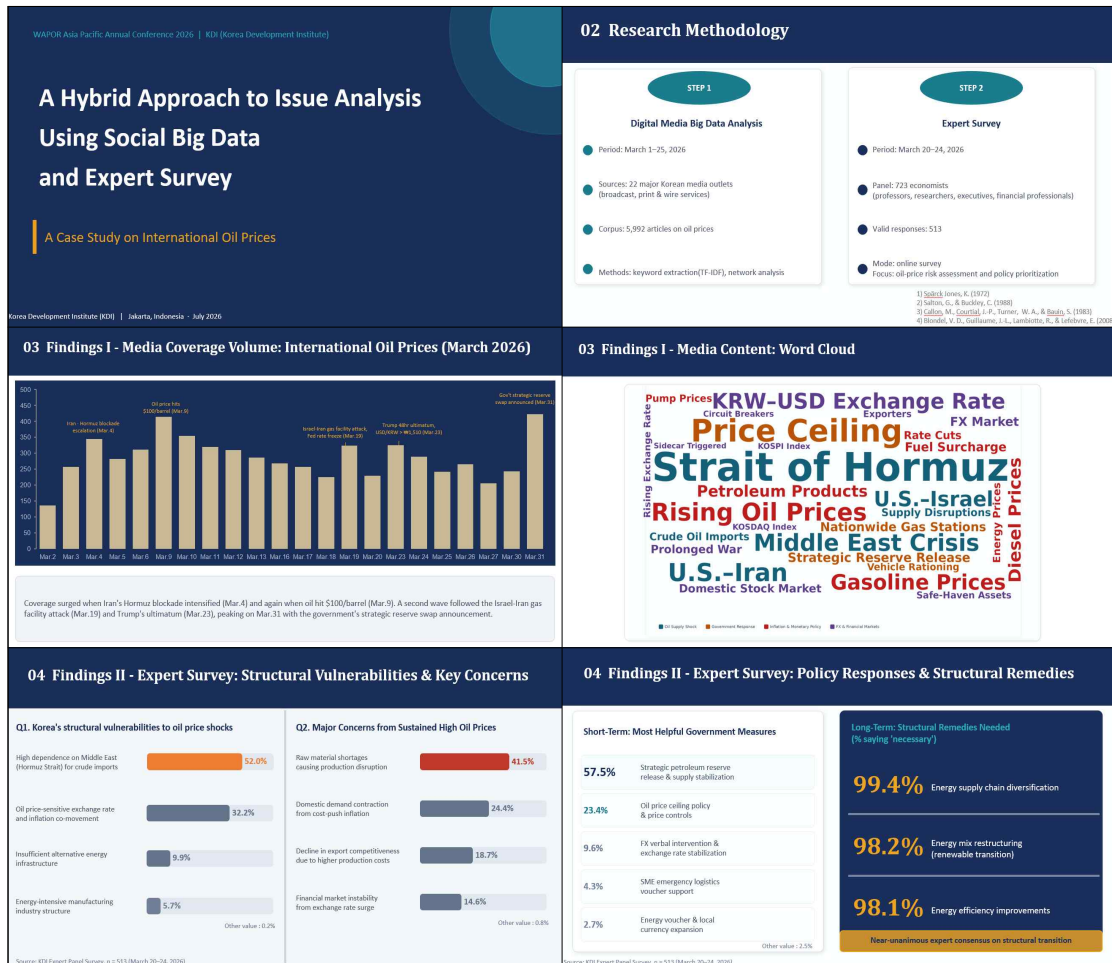
< KDI 발표 세션 현장 >



2. 발표세션: KDI 발표 내용

- KDI 경제교육·정보센터 경제정보분석실은 실의 주요 업무와 관련한 연구 결과 2개를 발표세션에서 발표함.
- KDI 첫 번째 발표(오전 세션: 빅데이터 및 대안적 여론조사, A Hybrid Approach to Economic Issue Analysis Using Social Big Data and Expert Survey: The Case of International Oil Prices)는 2026년 3월 국제 유가 급등 사태를 언론 빅데이터 분석과 KDI 경제전문가 패널조사를 결합한 하이브리드 방식으로 분석한 연구임.
- 발표 현장에 참석한 해외 연구자들은 대규모 전문가 패널 운영·응답률 관리 방식에 높은 관심을 보임.
 - KDI 여론분석팀은 자체 경제전문가 패널 보유하고 있고, 오랫동안 패널 관리를 해왔기 때문에 경제전문가 조사에서 높은 응답률을 보임.
 - 패널은 경제·부동산 등 사회과학 분야 교수, 연구원, 기업 임원 등으로 구성하며, 각 기관 홈페이지의 공개 정보와 KDI 경제교육·정보센터 여론분석팀 자체 노하우를 활용해 모집·관리함.
 - 그 외, 매월 3~4주 차 비슷한 시기에 정기조사를 실시하며, 패널 정보를 이용해 기간 중 독려 문자 및 메일을 발송하기도 함.
 - 자세한 발표 내용은 [참고 1]에 명시

< 첫 번째 발제 자료 중 일부 >



□ KDI 두 번째 발표(오후 세션: 미디어와 디지털 사회, Beyond Entertainment: Mapping the Semantic Structure of K-Content in Korean News)는 2025년 한 해 동안의 K-콘텐츠 관련 국내 언론 보도를 빅데이터 텍스트 분석으로 살펴본 연구임.

○ 발표 현장에 참석한 해외 연구자들은 방대한 언론 기사에서 특정 주제를 선별, 검증하는 방식과 분석 결과에 높은 관심을 보임.

- K-콘텐츠는 사전에 정한 주제가 아니라, 26개 언론사·5개 분야 약 188만 건의 기사에서 월별 TF-IDF 키워드로 주요 이슈를 탐지하는 과정에서 연중 지속적으로 부각된 대표 사례로 선정

- 분석 대상은 방송 · 종편 · 일간지 · 경제지 · 뉴스통신사 등 다양한 매체를 아우른 26개 언론사의 2025년 기사이며, K-콘텐츠 관련 기사는 하위 장르 · 연관 표현을 결합한 검색식으로 19,970건을 추출함.
- 형태소 분석 · TF-IDF · Louvain 군집화 등 연산은 자동화하되, 군집 해석 · 명명 등 판단이 필요한 부분은 연구자가 대표 기사를 직접 검토해 명칭이 실제 내용과 부합하는지 확인
- K-콘텐츠는 정부 주도의 산물이 아니라 시장에서 자생적으로 형성된 결과에 가까우며, 정책은 이 흐름을 촉진 · 지원하는 성격을 지님. 본 연구는 뉴스 보도에 나타난 K-콘텐츠 내용을 분석한 것으로, 정부 정책의 인과적 효과까지는 다루지 않음.
- 자세한 발표 내용은 [참고 2]에 명시

< 두 번째 발제 자료 중 일부 >

K-CONTENT : COMPUTATIONAL CONTENT ANALYSIS

Beyond Entertainment

Mapping the Semantic Structure of **K-Content** in Korean News

Kwanggun Lee • Eunhong Kim
Korea Development Institute (KDI)



04 - METHODOLOGY

Data & Methods

Data

- 19,970 online news articles from 26 major South Korean media outlets
- January 1 – December 31, 2025
- Articles containing "K-Content" and related terms

```
query = ["(K)-id*(Pop)(Os)(Ps)(Bibihaha)*(K)id*(C)content*(Korean)*(C)cotent*(C)cotent*(Industry, etc.
```

(Includes orthogonal variants of K-Content/K-pop to comprehensively capture related coverage)

Methods

- Monthly Media Coverage**
Article count per month (trend)
- Keyword Extraction^{[1][2]}**
TF-IDF / Bi-gram
- Co-occurrence Network^[3]**
Keyword co-occurrence
- Louvain Clustering^[4]**
Thematic cluster detection

[1] J. Leskovec et al. (2007)
[2] J. Leskovec et al. (2008)
[3] G. Adamic et al. (2005), W. A. R. R. (2005)
[4] M. E. J. Newman et al. (2004), M. E. J. Newman et al. (2004)

05 - RESULT

The Semantic Structure of Media Discourse



Cluster	Nodes
K-Content Hits	Digital Games · K-Pop Demon Hunters · Korean Cinema
Korean Culture	Korean Culture · Visiting Korea · Cultural Experience · Korean Traditions · Museum Visitors
Economic	Local Economic Growth · Korean Tourism · Tourism Industry · Tourism Products · Foreign Tourists
Cultural Content	Cultural Content · Content Industry Promotion · Enhanced Support
Content Industry	Content Production · Government Support Programs · Global Content · Intellectual Property · Content Companies · Global Market
Government Support & IP	Content Production · Government Support Programs · Global Content · Intellectual Property · Content Companies · Global Market
Global Exchange	Korea Economy · Korea Companies · Economic Cooperation · Cultural Exchange

Big data catches what we could miss.

01 - INTRODUCTION

Why study K-content, and why in the news?

Today's study: How Korean news media told the story of K-content in 2025.

Media.

Media is not public opinion. But it is the **stage** on which public issues take shape — before opinions form around them.

This study looks at that stage — the **media discourse itself**.

K-content.

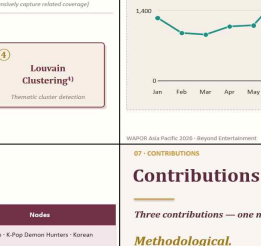
Drama, music, animation — a Korean cultural phenomenon that grew into a global one.

Globally visible from outside.

→ *But what does it look like from inside Korean news?*

02 - RESULT

Media Attention Over Time



Key Finding

Coverage remained high, peaking in September.

K-Content exceeded 1,000 articles per month throughout the year.

→ Peak in September: 2,700+ articles

→ September peak triggered by "K-Pop Demon Hunters" breaking Netflix's all-time viewership record

SOURCE: Asia Pacific 2026 - Beyond Entertainment

07 - CONTRIBUTIONS

Contributions

Three contributions — one methodological, one empirical, one conceptual

Methodological.

A scalable read of Korean news discourse.

Text mining across nearly 20,000 K-content articles from 26 major outlets — TF-IDF keyword extraction, co-occurrence networks, and Louvain community detection over a full calendar year.

Empirical.

K-content, in the news, is a multi-domain agenda.

Culture, economy, tourism, policy, and international exchange sit aside the same coverage — not as adjacent stories, but as one thematic structure.

Conceptual.

Discourse as observable structure.

Media discourse as an information environment — provides measurable representation of agenda construction.

3. 주요 세션 정리

가. 동시 운영 세션(Concurrent Session)

□ Parsing Survey Items with Large Language Models

- 설문 문항 자체를 분석 대상 데이터로 삼은 연구임. LLM을 활용해 문항을 세부 하위요인(subdimension)으로 재구성하는 AI 보조 측정 프레임워크를 제안, 대규모 공무원 연금 선택 설문에 적용한 연구
- 문항의 질문 내용과 응답 선택지를 LLM에 입력함. 경제적 제약·신념·인지/시간·결과·통제라는 5개 기준 범주 안에서 LLM이 세부 하위요인을 나눔. 각 문항이 어떤 하위요인과 얼마나 관련되는지도 가중치로 부여함. 이 가중치를 응답자별로 합산해 개인별 심리·경제적 성향 점수(메커니즘 점수)를 산출함.
- LLM이 제안한 결과는 그대로 신뢰하지 않음. 별도 표본으로 예측력을 검증하고, 하위요인 간 중복 여부도 진단함. 이 과정을 거쳐 실제로 의미 있는 구성인지 지속적으로 점검·수정함.
- 미국 16개 주 공공부문 근로자 396,948명에게 이메일로 설문을 발송함. 그 결과 7,674건(응답률 약 2.1%)을 확보함. 이 프레임워크를 적용한 결과, LLM이 만든 지표 중에서는 '근속연수(lock-in)'와 '금융이해력'이 연금 전환 수용 여부를 가장 잘 설명했고, '체감 만족도'는 금융이해력과 내용이 겹쳐 별다른 설명력을 갖지 못함.
- 연구는 LLM을 결과만 내놓는 '과정을 알 수 없는 예측 도구'가 아니라, 기존 설문 문항이 서로 다른 개념(지식·신념·제약)을 뒤섞어 놓은 지점을 찾아내고 문항을 재설계하도록 돕는, '검증 가능한 측정 보조 도구'로 활용할 수 있음을 보여줌.

□ Evaluating LLMs for Open-Ended Survey Coding and Classification in Survey Research Under a Total Survey Error Framework

- 개방형 설문 응답에 대한 LLM 자동코딩이 인간 코딩 데이터의 측정 속성을 유지하는지, 총설문오차(Total Survey Error) 틀 내에서 검증한 연구

- 영국 NatCen 사회조사의 개방형 응답(96개 코딩 범주, 총 5,231건 중 개발용 3,679건 · 평가용 1,552건)을 대상으로, 지시문 기반 LLM 코딩과 트랜스포머 문장 임베딩 기반 지도학습 분류(one-vs-rest)를 비교함.
- 정밀도 · 재현율 · F1 · 완전일치 · Jaccard 등 전통적 지표에 더해, 범주별 유병률(prevalence) 왜곡, 코딩 밀도 차이, 누락/추가 코드, 잔여범주 처리, 응답 복잡도에 따른 성능 편차 등 설문오차 관점의 지표까지 함께 분석함.
- LLM 기반 코딩은 재현율이 높은 대신 더 많은 코드를 부여하는 경향이 있으며, 일부 불일치는 단순 오류가 아니라 의미상 타당한 확장 해석으로 나타남.
- 자동코딩이 언제 여론조사 코딩에 적합한지, 자동화 방식을 어떻게 문서화해야 하는지에 대한 판단 기준으로서 총설문오차 틀의 유용성을 제시하며, 여전히 인간 검수의 필요성을 강조

□ New Frontiers in Survey Quality, Representation, and Public Interaction

- 빠르게 변화하는 정보 · 사회 환경 속에서 여론조사 자료의 품질 · 대표성 · 해석 가능성을 개선하는 방법을 다룬 4편의 논문으로 구성된 패널임.
 - ① 호주 사회태도조사(Australian Survey of Social Attitudes)의 정당 지지 편향 문제를 다루며, 표준적인 인구통계학적 사후층화만으로는 정치적 불균형이 교정되지 않을 수 있음을 보이고, 외부 선거 벤치마크를 활용한 대안적 가중치 부여 전략을 검토함.
 - ② 대만 고령층을 대상으로, 표준중국어(만다린) 대신 대만어(타이완어)를 구사 하는 AI 생성 면접원이 이해도를 높이고 무응답을 줄일 수 있는지 검증하여, AI가 비용 절감뿐 아니라 접근성 취약 집단의 포용성 제고에도 기여할 수 있음을 시사함.
 - ③ 호주의 선거 무결성 관련 대화를 소재로, 기관 출처가 명시된 챗봇과 익명 챗봇에 대한 시민들의 상호작용을 비교 · 분석하여, 태도 변화가 아닌 대화 자체의 어조 · 방식 · 내용에 초점을 맞춤.
 - ④ CEDAW(여성차별철폐협약)를 활용한 호주 설문실험을 통해, 국제 규범에 대한 준거가 젠더 평등 정책 지지와 사회 인식에 미치는 영향을 분석함.

- 4편 모두를 관통하는 결론은, 현대 여론조사 연구가 표본추출 · 소통 · 상호작용 · 프레이밍 문제에 동시에 대응해야 한다는 점임.

□ Social Change and Scientific Contributions in Asia: Evidence from the International Social Survey Programme

- 국제사회조사프로그램(ISSP)의 국가 간 · 시계열 자료를 활용해, 아시아 · 태평양 지역의 사회변동을 종교 · 불평등 · 사회적 상호작용의 관점에서 조명한 4편의 연구로 구성된 패널임.
- ISSP 종교 모듈을 활용해 필리핀 · 인도네시아 · 태국 · 대만 · 일본의 종교적 신념 · 실천 · 공동체 지원이 청년 행복도와 맺는 관계를 국가 간 비교 분석함.
- ISSP와 Pew 자료를 비교해 인도의 종교 간 · 공동체 간 태도를 심층 분석함.
- ISSP 사회불평등 모듈을 활용해, 공공교통이 일상적 계층 간 접촉(cross-class encounter)을 촉진하는 반면 디지털 플랫폼 경제의 확산은 사회적 회피를 통해 접촉 기회를 축소시킬 수 있음을 보임.
- ISSP 기반 학술 성과를 모듈 · 변수 · 국가 · 언어별로 연결한 개방형 검색 가능한 데이터베이스 구축 프로젝트를 소개함.
- 조화된 설문 설계와 실질적 주제 폭을 결합한 ISSP가 아시아 · 태평양 지역의 종교 · 불평등 · 사회변동에 관한 비교연구를 뒷받침하는 핵심 자원임을 강조

□ Probing LLM Political Knowledge: Encoding Artifacts in Historical Roll-Call Data

- 대규모 언어모델이 정치 영역에서 갖는 시간적 편향(temporal bias)을 미국 의회 표결 기록을 통해 검증한 연구
- 1984년 미국 의회 표결 자료를 바탕으로, 16개 쟁점별 표결 내용을 반복 제시해 LLM(GPT-5-4, Claude Sonnet 4.6 등)에게 의원의 정당 소속(공화 · 민주)을 제로샷 분류하도록 함.
- 두 모델 모두 실제 1984년 투표 패턴 기준으로는 무작위 추정(50%) 수준에 크게 못 미치는 정확도를 보였으나, 모델 간 예측은 서로 상당히 일치함.

- 정당별 쟁점 연관성이 변화·소멸된 영역(냉전기 외교정책, 의료규제 등)에서 오분류가 두드러진 반면, 오늘날까지 이념적 정당 구도가 유지되는 쟁점에서는 상대적으로 정확도가 높음.
- 이는 LLM이 현재의 정당 정치 구도를 과거 표결 기록에 그대로 투영하는 '현재 편향'을 내재하고 있음을 시사.

□ Measuring Social Perceptions of Everyday Discrimination: A Korean Vignette Survey and Initial LLM Comparison

- 일상적 차별에 대한 사회적 인식을 한국인 대상 비네트(vignette) 설문과 LLM 판단을 비교해 측정한 연구
- 성인 2,270명을 대상으로, 성별·성적지향, 학력, 가족형태, 인종·국적, 직업, 연령, 지역, 장애, 건강, 외모 등을 아우르는 30개의 일상적 상황 비네트를 제시하고, 각 상황이 얼마나 차별적으로 받아들여지는지를 6점 척도로 평가하게 함.
- 동일한 비네트에 대해 GPT-5.5 Pro, Claude Opus 4.8, Gemini 3.5 Thinking의 판단과 인간 응답을 비교한 결과, 인간 응답 평균(4.09)보다 세 모델 모두 더 높은 평균(GPT 4.93, Claude 4.50, Gemini 4.53)을 보였으며, 척도의 하한값을 거의 사용하지 않음(GPT-5.5 Pro는 4점 미만 응답이 전무).
- 이러한 '모호성 압축(ambiguity compression)' 현상은 LLM이 사회적으로 편차가 있는 판단을 지나치게 확정적인 도덕적 평가로 수렴시켜, 집단 간 인식 차이를 가리고 합의 수준을 과장할 수 있음을 시사하며, 측정 대상이 사회적으로 모호할 때 단순 평균 일치도만으로는 LLM의 타당성을 판단하기 부족함을 지적

< 학회 발표 및 토론 현장 >



Ⅲ. 시사점

- AI(LLM)를 활용한 조사·분석 방법론이 국제적으로 빠르게 확산되고 있으며, 동시에 편향 재생산·판단 과잉확정 등 방법론적 한계도 함께 지적되고 있어, 여론조사·빅데이터 분석 업무에 AI를 적용할 때 결과 해석에 신중을 기해야 한다는 점을 재확인함.
 - 이번 콘퍼런스는 “빠르게 변화하는 세계 속의 여론(Public Opinion in a Fast-Changing World)”을 주제로, LLM을 활용한 설문 문항 구성·개방형 응답 코딩·문화 차원 측정 등 AI의 조사방법론 적용 사례가 다수 발표됨.
 - LLM이 문화적 고정관념이나 정치적 편향을 답습·재생산할 위험, 인간 응답에 비해 판단을 지나치게 확정적으로 수렴시키는 경향 등 AI 활용에 따른 방법론적 한계에 대한 비판적 검토도 다수 이루어져, 조사·분석 역량 강화 방안뿐 아니라 AI 기반 분석 결과를 해석·활용할 때 유의해야 할 지점에 대해서도 폭넓게 확인할 수 있었다는 점에서 의의를 가짐.
- 또한 확률표본조사와 비확률(옵트인) 온라인조사 간 정확도 격차, 적정 표본 크기, 무응답 최소화 기법 등 전통적 조사방법론에 대한 최신 실증 연구도 다수 공유됐으며, 참고할 만한 시사점을 확인함.
 - 일본·태국·러시아 사례 연구에서 전화(RDD)·대면조사와 온라인조사 간 응답 격차가 문항 유형(태도·평가 문항일수록 격차 확대)과 표본 크기에 따라 체계적으로 달라진다는 점, 후속탐침(follow-up probing) 등 응답 유도 기법으로 금융·소득 문항 무응답을 상당폭 낮출 수 있다는 점을 확인
 - 이는 설문 운영 시 문항 설계·무응답 관리 방식을 점검하는 데 참고할 만한 시사점을 제공함.
- 이번 아시아 태평양 콘퍼런스는 여론조사·빅데이터 분석에 대한 글로벌 전문가의 최신 연구 결과를 청취하고, 향후 분석 방법론에 새로운 조사 트렌드와 데이터 활용 방안을 파악할 수 있는 장이었음.
 - 해당 콘퍼런스에는 21개국 약 110명의 글로벌 민·관·학계 여론조사 전문가

가 참여했으며, AI 활용 조사방법론부터 미디어 담론·사회변동 연구까지 급변하는 여론조사 동향에 대한 다각적인 논의와 의견 교류가 이루어짐.

- KDI 발표 세션에서도 참가 연구자들과 방법론에 대한 의견을 나누었으며, 이를 계기로 관련 분야 전문가와의 교류와 소통을 지속·확대해나갈 필요

[참고 1] A Hybrid Approach to Economic Issue Analysis Using Social Big Data and Expert Survey: The Case of International Oil Prices

Abstract

This study analyzes the March 2026 international oil price shock in Korea with a hybrid approach that combines social big data analysis and an expert survey. In Stage 1, 5,992 news articles published by 22 Korean outlets in March 2026 were analyzed with TF-IDF keyword extraction, co-word network analysis, and Louvain community detection to identify the key issues in oil-related press coverage. In Stage 2, a survey of the KDI economic expert panel (n = 513) was designed around the detected issues to assess Korea's structural vulnerabilities, the expected economic impact of the shock, and policy priorities. Media coverage organized into four themes—the oil supply shock, government response, inflation and monetary policy, and financial markets—while experts identified dependence on Middle East crude imports as the core vulnerability (52.0%) and production disruption as the top concern (41.5%), rating strategic reserve releases the most helpful short-term measure (57.5%) and expressing near-unanimous agreement that the three proposed long-term measures were necessary (98-99%). The findings show that media analysis and expert surveys can be integrated into a timely, complementary procedure for economic issue analysis.

Keywords: text mining, network analysis, expert panel survey, expert opinion, oil prices

1. Introduction

International oil prices are among the most consequential external variables for the Korean economy, shaping consumer prices, industrial production costs, the trade balance, and consumer sentiment. In March 2026, geopolitical risk in the Middle East—including the blockade of the Strait of Hormuz—disrupted global crude supply and pushed oil prices to the top of Korea's economic agenda, with the shock quickly spreading to the foreign exchange and capital markets. Because oil prices touch the Korean economy through so many channels, understanding how such a shock unfolds, and how it is understood, is a precondition for drawing the right policy implications.

To that end, this study takes a hybrid approach that combines text analysis of digital media big data with a survey of the KDI economic expert panel, diagnosing the causes and ripple effects of the Middle East shock on the Korean economy and identifying the policy tasks it raises. Rather than simply tracking market indicators, the study looks at two things at once: the issue agenda that emerged in media coverage, and the structural remedies that experts identify beneath it.

This design reflects two considerations. The first consideration was timeliness. Many conventional economic indicators are published with a lag and may therefore be unavailable during the early stages of a rapidly developing shock. Media data can support timely issue detection during such periods by revealing which concerns and policy questions are receiving immediate attention. To capture the issue as it unfolded, the study therefore applied text analysis to news coverage of international oil prices. The second consideration was interpretation. Media data reveal which issues are being discussed and how they are framed, but their economic meaning does not speak for itself. To interpret these signals—and to confirm how experts perceive the economic impact of the oil price shock and the priorities for response—the study fielded a structured survey of economic experts, designed around the issues that the media analysis had surfaced.

Using the March 2026 oil price shock as a case, the study pursues four objectives:

(a) examine how the causes and economic repercussions of the shock were diagnosed by experts; (b) to identify the key sub-issues in oil-related press coverage through media text analysis; (c) to assess the economic impact, policy importance, and response priorities through an expert survey; and (d) to synthesize the two analyses to draw economic and policy implications. By linking online text analysis with survey research, the study also aims to provide baseline evidence that supports policy responses and the decision-making of economic agents.

The remainder of the paper proceeds as follows. The next section describes the two-stage research design and data. The following section presents the results of the media analysis, the expert survey, and their integration. The final section concludes with the study's findings, methodological significance, and limitations.

2. Methodology

Research Design

The study employs a sequential two-stage design in which each stage addresses a different informational limitation. Stage 1 applies text mining to media coverage to detect, in near real time, the key issues surrounding international oil prices and the structure of their coverage. Stage 2 administers an expert survey designed around the issues detected in Stage 1, translating the salient themes of public discourse into questions on economic impact and policy response. The survey window (March 20–24, 2026) is nested within the media analysis window (March 1–25, 2026), so the two instruments observe the same phase of the shock through different lenses.

Stage 1: Media Text Analysis

The media corpus consisted of 5,992 news articles published between March 1 and March 25, 2026 by 22 Korean outlets. For descriptive purposes, the daily coverage-volume series was extended through March 31 to capture subsequent developments in the issue. Articles were retrieved when they mentioned terms related to international oil prices, including crude benchmarks prices (WTI, Dubai, Brent), the Strait of Hormuz, and Middle East energy supply. After standard preprocessing (tokenization, part-of-speech filtering, and stop-word removal), keywords were weighted using the TF-IDF scheme (Salton & Buckley, 1988; Spärck Jones, 1972) to identify the most salient terms. A co-word network was then constructed from keyword co-occurrence within articles, following the co-word analysis tradition (Callon et al., 1983), and thematic clusters were identified with the Louvain community detection algorithm (Blondel et al., 2008).

Stage 2: Expert Survey

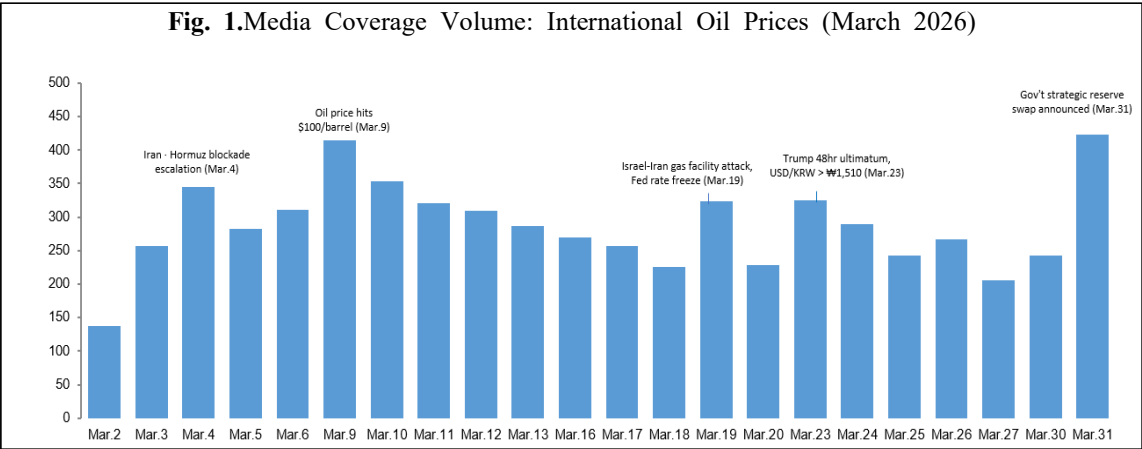
The survey was administered online to the KDI economic expert panel of 723 members—professors, researchers, corporate executives, and financial professionals—between March 20 and 24, 2026; 513 experts responded (a 71.0% response rate). The questionnaire was built on the Stage 1 results, with response options drawn from the issue categories that the media analysis had surfaced. It comprised two blocks, each with two questions. The first block addressed the impact of rising international oil prices: Korea's structural vulnerabilities to oil price shocks, and the greatest concern should elevated prices persist. The second block addressed countermeasures: an evaluation of which recent government responses are most helpful in the short term, and the necessity of fundamental measures to mitigate oil price volatility over the longer term. The first three questions were single-response items, whereas the final question asked respondents to assess each of three long-term measures separately; the reported

percentages represent the share rating each measure as necessary.

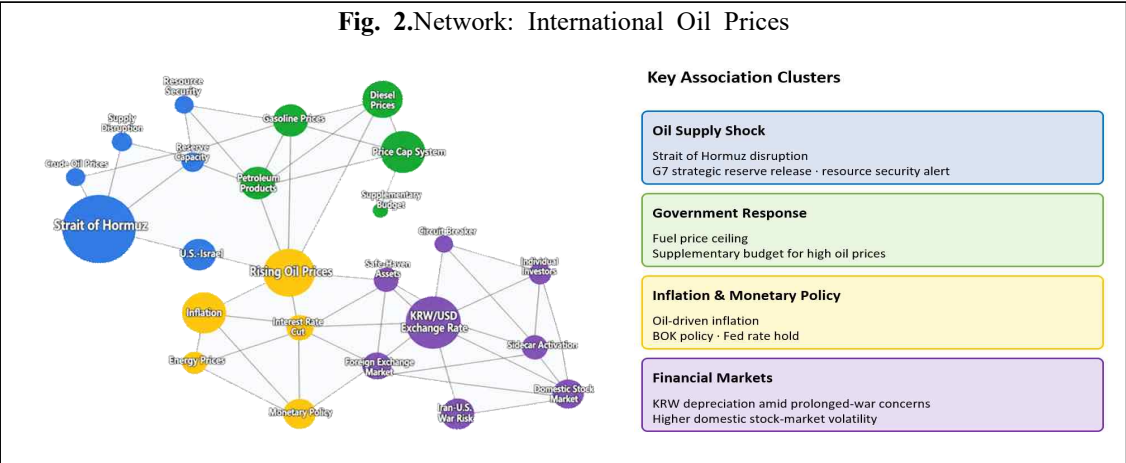
3. Results

3.1 Media Coverage of the Oil Price Issue

The text analysis shows that ‘international oil prices’ was the most prominent economic keyword in the Korean press during March 2026, and coverage volume closely tracked the unfolding of the crisis. Reporting surged when Iran escalated the Hormuz blockade (March 4) and again when crude prices reached \$100 per barrel (March 9). A second wave followed the attack on gas facilities amid the Israel–Iran confrontation and the Federal Reserve’s decision to hold rates (March 19), and the United States’ 48-hour ultimatum, during which the won–dollar exchange rate surpassed

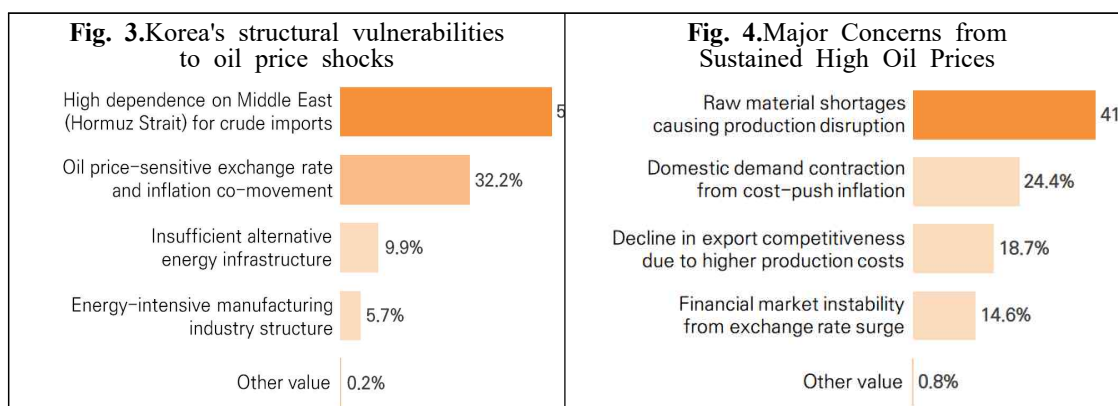


The co-word network and Louvain clustering organize this coverage into four association clusters: (a) the oil supply shock, centered on the disruption of the Strait of Hormuz, G7 strategic reserve releases, and resource-security alerts; (b) government response, including the fuel price ceiling and a supplementary budget for high oil prices; (c) inflation and monetary policy, linking oil-driven inflation to the Bank of Korea’s policy stance and the Federal Reserve’s rate hold; and (d) financial markets, where won depreciation amid prolonged-war concerns accompanied higher domestic stock-market volatility. The shock thus registered in the press not as a single story but as a connected system of energy, fiscal, monetary, and capital-market anxieties.

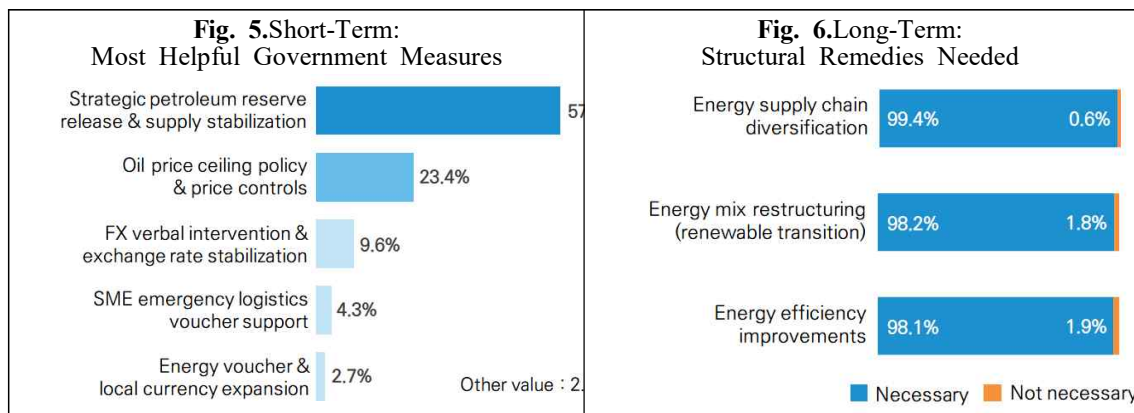


3.2 Expert survey

The survey translates these themes into economic judgment. Among responding members of the KDI economic expert panel, 52.0% identified the country's high dependence on Middle East crude imports passing through the Strait of Hormuz as Korea's core structural vulnerability to oil price shocks, followed by the oil-price-sensitive co-movement of the exchange rate and inflation (32.2%), insufficient alternative-energy infrastructure (9.9%), and the energy-intensive structure of Korean manufacturing (5.7%). Asked about the greatest concern if elevated oil prices persist, experts most frequently cited disruption of corporate production due to raw-material supply instability (41.5%), followed by a contraction of domestic demand driven by cost-push inflation (24.4%), weaker export competitiveness due to higher production costs (18.7%), and financial-market instability stemming from exchange-rate surges (14.6%).



Regarding recent government responses, releases from strategic petroleum reserves and other supply-stabilization measures was most frequently selected as the most helpful short-term response (57.5%), well ahead of price controls such as the oil price ceiling (23.4%); foreign-exchange market stabilization measures (9.6%), emergency logistics support for small and medium-sized enterprises (4.3%), and energy vouchers with local-currency programs (2.7%) drew limited support. Most strikingly, when asked about fundamental remedies for oil price volatility, experts were nearly unanimous: 99.4% judged the diversification of energy supply chains necessary, 98.2% supported restructuring the energy mix, and 98.1% called for improvements in energy efficiency.



3.3 Reading the Two Analyses Together

Placed side by side, the two analyses describe the same episode at different levels. The themes that organized press coverage—supply disruption, foreign exchange and financial market effects, inflation and monetary policy, and government response—reappear in the survey results as the objects of expert evaluation: production disruption and inflation-driven demand contraction were the most frequently cited concerns, and experts favored supply stabilization in the short term while calling for fundamental measures such as supply chain diversification over the longer term. In this reading, the media analysis indicates what was at issue during the shock, and the expert survey indicates how those issues were weighed and what responses were preferred. The study does not claim that media salience validates expert judgment, or the reverse; the contribution of the juxtaposition is descriptive—a fuller account of the episode than either source provides alone.

4. Conclusion

This study analyzed the March 2026 international oil price shock using a hybrid approach that combines social big data analysis with an expert survey. The media text analysis identified key sub-issues and coverage trends in oil-related reporting in near real time; the expert survey captured panel members' assessments of the shock's economic impact and policy priorities; and, taken together, the two analyses provided complementary perspectives on the oil price issue.

The study's contribution is primarily methodological: it illustrates a sequential procedure in which media analysis supports timely issue detection when conventional indicators are not yet available, and an expert survey assesses the economic significance and policy priority of the issues identified. Applied to the March 2026 oil price shock, the approach linked patterns of media coverage with expert assessments of economic risks and policy responses. The procedure may also be applicable to other rapidly evolving economic and social issues, including exchange rates, supply chains, and financial stability.

Three boundaries of the analysis should be noted. First, the media analysis examines the content and structure of coverage; it does not measure the magnitude of the economic shock or identify causal effects. Second, the survey captures expert perceptions at a single point in the shock, not their evolution over time. Third, the corpus is limited to press articles, so social media and online communities lie outside the scope of the analysis. These boundaries define a natural agenda for future research: linking text-based indicators to market and macroeconomic data to quantify magnitudes, fielding repeated survey waves to track how expert perceptions evolve, and extending the corpus to social media discourse.

References

- Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10), P10008. <https://doi.org/10.1088/1742-5468/2008/10/P10008>
- Callon, M., Courtial, J.-P., Turner, W. A., & Bauin, S. (1983). From translations to problematic networks: An introduction to co-word analysis. *Social Science Information*, 22(2), 191–235. <https://doi.org/10.1177/053901883022002003>
- Salton, G., & Buckley, C. (1988). Term-weighting approaches in automatic text retrieval.

Information Processing & Management, 24(5), 513–523.
[https://doi.org/10.1016/0306-4573\(88\)90021-0](https://doi.org/10.1016/0306-4573(88)90021-0)

Spärck Jones, K. (1972). A statistical interpretation of term specificity and its application in retrieval. *Journal of Documentation*, 28(1), 11–21.
<https://doi.org/10.1108/eb026526>

[참고 2] Beyond Entertainment: Mapping the Semantic Structure of K-Content in Korean News

Abstract

K-content—spanning drama, music, and animation—is widely recognized as a global cultural phenomenon. Yet how Korean news media themselves represent K-content, and around which themes, remains a separate empirical question. This study analyzes 19,970 online news articles from 26 major South Korean outlets published in 2025, applying TF-IDF keyword extraction, co-occurrence network analysis, and Louvain community detection to map the semantic structure of K-content coverage. Media attention was sustained across the year and peaked in September, coinciding with the record-breaking global success of the animated film K-Pop Demon Hunters. Beyond entertainment, coverage connected K-content to the economy, tourism, government policy, intellectual property, and international exchange, forming a multi-domain thematic structure. We read media discourse as the information environment in which public opinion later develops, and treat computational text analysis as a scalable way to trace the shape of issue discourse.

1. Introduction: Media Discourse as an Information Environment

Public opinion does not arrive fully formed. Before people form views, issues are introduced, interpreted, and circulated through the media. WAPOR is centrally concerned with how public opinion is measured, interpreted, and shaped; this study focuses on an earlier stage in that sequence—the media discourse through which public issues first become visible.

It is useful to distinguish four stages. Issues emerge as events, incidents, and moments. Media discourse constructs and circulates meanings around them. Public discourse follows as conversation among audiences. Public opinion is the aggregated stances and attitudes that result. This study looks at the second stage. It does not measure public opinion or attitudes; it examines the information environment in which opinion later develops. Two routes lead to public opinion: survey research measures it directly, while the media supply the agenda and vocabulary from which public discourse, and eventually opinion, are built. Understanding that information environment is therefore complementary to opinion measurement rather than a substitute for it.

K-content is one of the most globally visible cultural phenomena associated with Korea. Drama, music, and animation that began as domestic entertainment have grown into a worldwide presence. That visibility, however, is largely a view from the outside. How Korean news media themselves tell the story of K-content—what themes they attach to it, and how those themes connect—is a separate question. This study does not

ask whether K-content is popular. It asks how Korean news media constructed the meaning of K-content throughout 2025.

This positioning is grounded in agenda-setting theory (McCombs & Shaw, 1972), which holds that the volume and emphasis of media coverage shape what the public regards as important. We follow this lineage: the same theoretical premise, examined through a new measurement structure made possible by big-data text analysis. Our aim is not to extend agenda-setting theory but to demonstrate a scalable, big-data approach to issue detection, using K-content as a representative case.

2. Research Questions

The study asks a single overarching question: across a year of Korean news, how does the K-content discourse take shape? We address it through three sub-questions: (RQ1) When did media attention rise and peak? (RQ2) What themes did coverage bring together around K-content? (RQ3) How were those themes connected to one another?

3. Data and Methods

The corpus comprises 19,970 online news articles published between January 1 and December 31, 2025, drawn from 26 major South Korean news outlets. Articles were retrieved using a query capturing K-content and its orthographic variants, including "K-pop," "Hallyu," "Korean content," and "content industry," so as to comprehensively capture related coverage.

Because Korean is an agglutinative language in which meaning attaches to word stems and particles, the articles were first processed through morphological analysis, and candidate terms were formed as bigrams and trigrams so that multi-word expressions—rather than isolated tokens—served as the units of analysis.

The analysis then proceeds in four steps. First, we measured monthly article volume to trace the temporal shape of attention. Second, we extracted salient keywords using TF-IDF weighting (Spärck Jones, 1972; Salton & Buckley, 1988) over the bigram and trigram terms, a weighting that raises the score of expressions distinctive to K-content coverage while discounting terms common across all news. Third, we built a co-occurrence network in which keywords appearing together within articles are linked (Callon et al., 1983). Fourth, we partitioned that network into thematic communities using the Louvain algorithm (Blondel et al., 2008), which identifies densely connected groups of keywords.

The four steps are complementary. Monthly volume shows how much attention the issue drew; TF-IDF shows which terms defined it; the co-occurrence network shows how those terms related; and Louvain detection groups the relations into interpretable themes. TF-IDF alone identifies prominent words but not the structure connecting them, which is what the network and community-detection stages recover. Together they move from how much coverage appeared, to what it contained, to how its themes were organized.

4. Results

4.1 Media Attention Over Time

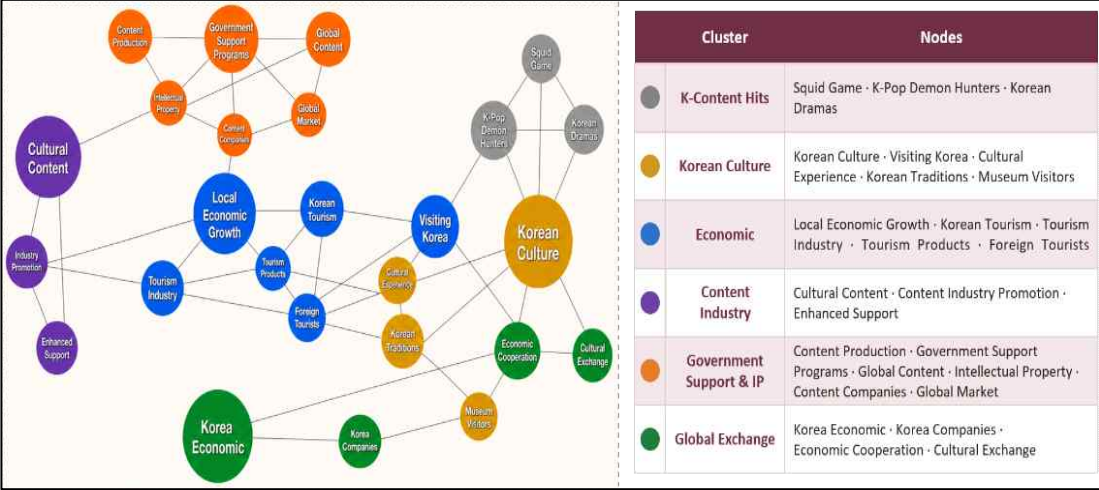
Coverage of K-content was sustained throughout 2025, exceeding 1,000 articles in every month of the year (Figure 1). Attention was not evenly distributed: it rose gradually across the year, peaked in September at more than 2,700 articles, and tapered off in the following months. This September concentration coincided with the animated film *K-Pop Demon Hunters* breaking Netflix's all-time viewership record—an

4.3 The Semantic Structure of Media Discourse

Louvain community detection organized the co-occurrence network into six thematic clusters (Figure 3). The first, K-content hits, gathered the breakout titles carrying the Korean Wave—Squid Game, K-Pop Demon Hunters, and Korean dramas. A second cluster, Korean culture, linked K-content to visiting Korea, cultural experiences, traditional culture, and museum attendance. A third, economic, connected K-content to local economic growth, tourism, and foreign visitors. A fourth, content industry, concerned cultural content, industry promotion, and enhanced support. A fifth, government support and intellectual property, brought together content production, government programs, intellectual property, and market expansion. A sixth, global exchange, tied K-content to economic cooperation and cultural exchange between countries.

These clusters were not abstract groupings. The Korean culture cluster tracked rising interest in traditional Korean culture and heritage tourism, including record museum attendance during the year. The economic cluster reflected the tie between K-content and regional tourism, drawing foreign visitors and supporting local economies. The content industry and government-support-and-IP clusters followed policy moves to strengthen the sector—administrative reorganization, expanded support programs, content production, and intellectual property, including webtoons. The global exchange cluster reflected coverage in which K-content surfaced in diplomatic settings, from trilateral foreign-ministerial meetings to the APEC summit hosted in Korea. Read together, these clusters show K-content situated not as a single topic but as a connected, multi-domain structure spanning culture, economy, tourism, policy, and diplomacy.

Figure 3.Co-occurrence network of K-Content keywords with Louvain community detection identifying six thematic clusters, 2025.



5. Discussion

Three observations follow from the structure of the coverage itself. First, the K-content discourse reached beyond entertainment: economy, tourism, and national image appeared inside the same articles, not as separate stories. Second, coverage repeatedly linked cultural products to national competitiveness—K-content appeared tied to Korea's global presence and a soft-power narrative. Third, government policy and industry were

integral to the discourse rather than peripheral to it: K-content was discussed alongside support programs, content production, intellectual property, and market expansion.

What computational analysis adds here is scale and structure. Reading nearly 20,000 articles by hand could surface these themes anecdotally; co-occurrence and community detection make the connections among them visible and reproducible. In the terms of agenda-setting theory, the analysis renders the composition of a media agenda observable—not only which issue drew attention, but the thematic architecture through which that attention was organized.

This study examines the information environment in which public opinion later develops, rather than measuring public opinion itself. The K-content case shows that environment concretely: a subject an audience might first meet as entertainment already appeared, in the coverage, alongside the economy, tourism, and national image.

6. Contributions

The study offers three contributions. Methodologically, it demonstrates a scalable read of Korean news discourse—TF-IDF keyword extraction, co-occurrence networks, and Louvain community detection applied across nearly 20,000 articles over a full calendar year. Empirically, it shows that K-content, in the news, is a multi-domain agenda: culture, economy, tourism, policy, and international exchange sat within a single thematic structure rather than as adjacent stories. Conceptually, it treats media discourse as an information environment that can be observed directly, providing a measurable representation of how an agenda is constructed.

7. Limitations

Four boundaries define what these findings can and cannot support. First, this is discourse, not opinion: the study examines how K-content is represented in coverage, not how citizens or global audiences regard it. Second, it is a slice of the press—26 Korean outlets over a single year—and does not generalize to other media platforms, other nations, or longer-term trends. Third, co-occurrence reveals which terms cluster together but does not, by itself, recover tone, intent, or causal direction. Fourth, Louvain detection identifies clusters, but naming them "culture," "economy," or "policy" reflects researcher judgment rather than an objective measurement. To limit this subjectivity, cluster labels were independently reviewed by the researchers and finalized through consensus.

8. Conclusion

Our focus was the stage before opinion. Across a year of Korean news, K-content coverage was not limited to entertainment; it ran together with economy, tourism, policy, and international exchange—a media discourse seen from the inside, at scale. Computational text analysis offers a scalable way to trace the shape of issue discourse, and with it a clearer view of the information environment in which public issues become visible and public opinion later forms.

References

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10), P10008.

Callon, M., Courtial, J.-P., Turner, W. A., & Bauin, S. (1983). From translations to problematic networks: An introduction to co-word analysis. *Social Science Information*, 22(2),191–235.

McCombs, M. E., & Shaw, D. L. (1972). The agenda-setting function of mass media.*Public Opinion Quarterly*, 36(2), 176–187.

Salton, G., & Buckley, C. (1988). Term-weighting approaches in automatic text retrieval.*Information Processing & Management*, 24(5), 513–523.

Spärck Jones, K. (1972). A statistical interpretation of term specificity and its application in retrieval. *Journal of Documentation*, 28(1),11–21.