



인공지능에 관한 미디어 콘텐츠의 분석 -빅데이터 분석을 적용하여

**Analysis of media contents related
to artificial intelligence-by applying
big data analysis**

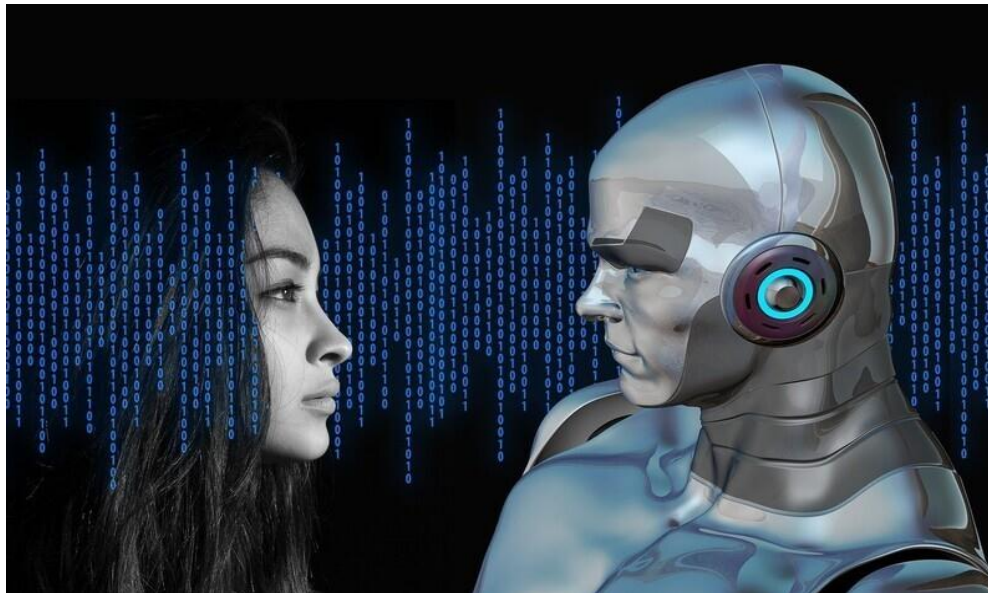
고려대학교 연구교수 허성호





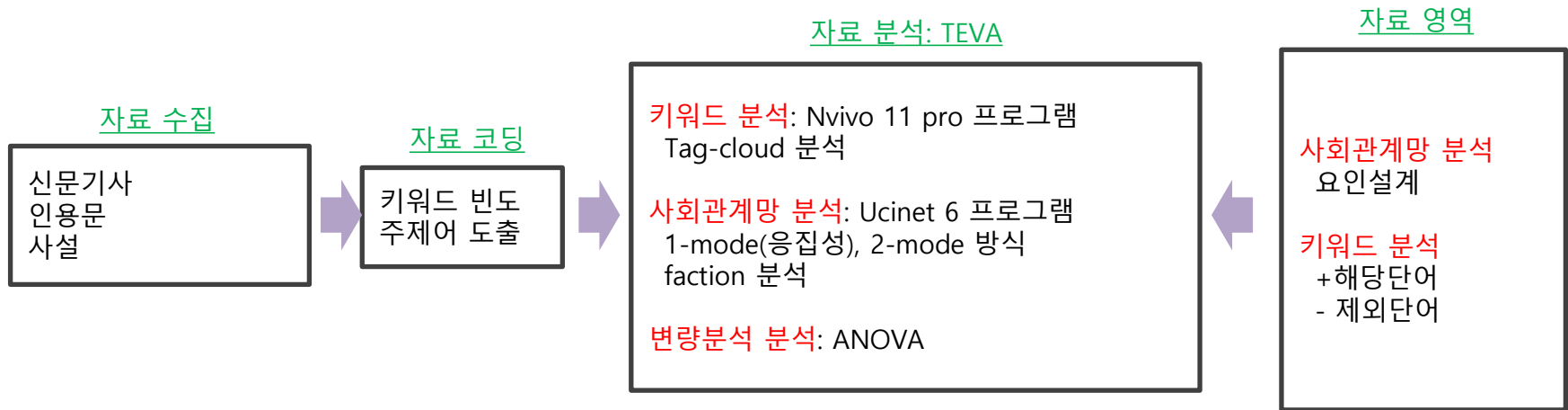
인공지능과 인간의 공존

❏ AI시대에 의심받는 인간의 가치...공존과 다양성으로 인정받아야





분석 모듈(TEVA)



응집성: K-core 방식의 군집화 분석 결과를 도출하며, 공통성을 추출하는데 용이하다.

집합성: Tabu search 방식의 군집화 분석 결과를 도출하며, 이질성을 중심으로 추출하는데 용이하다.

2-mode: 행열에서 행은 평가대상, 열은 평가자로 구분되는 매트릭스 구조를 의미한다.

faction: Tabu search의 집합성 원리를 이용한 군집 방법이다.

fitness: 1-mode 변환 시 대각선 블록의 오류 수와 대각선 블록을 벗어난 오류 수를 표준화하여 계산한다. 일반적으로 집단을 많이 나눌수록 대각선 블록의 오류수가 상대적으로 적게 증가하기 때문에 fitness는 줄어든다.

응집성로직: Maximize $M = \sum_s \left[\frac{l_s}{L} - \left(\frac{d_s}{2L} \right)^2 \right]$

집합성로직: Minimize $z = \sum_{j \in J} f_j w_j + \sum_{i \in I} \sum_{j \in J} c_{ij} x_{ij}$



연구대상

❑ Covid 19 시기

- 전(2017.01~2019.12)
- 후(2020.01~2022.12)

❑ 언론사 유형

- 신문사: 중앙일보, 조선일보, 동아일보, 국민일보, 경향신문
- 방송사: KBS, MBC, SBS, YTN, OBS

	Covid 19 전	Covid 19 후
신문사	8457 건	7790 건
방송사	2677 건	3882 건

❑ 총 22804 건의 기사글 분석



워드 클라우드분석 결과

반도체
5G기술
딥러닝
머신러닝
자율주행
모빌리티
헬스케어
가상현실
클라우드
네트워크
사물인터넷
빅데이터
ict
핀테크
블록체인

전-신문사

전-방송사

후-신문사

후-방송사

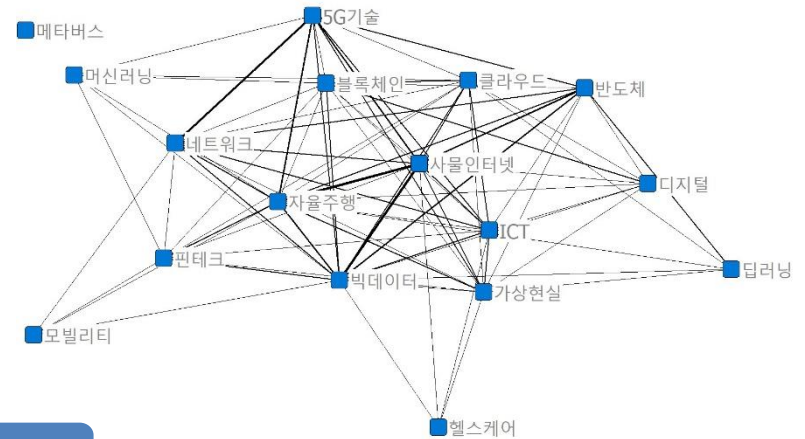
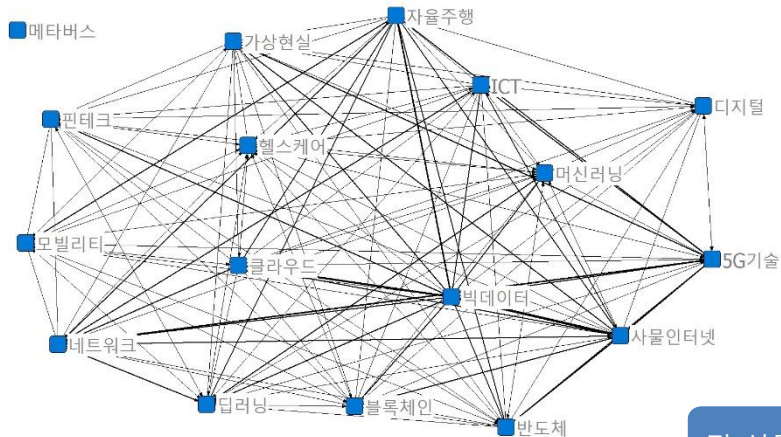
네트워크
5G기술
머신러닝
클라우드
사물인터넷
모빌리티
핀테크
헬스케어
자율주행
블록체인
가상현실
반도체
딥러닝
빅데이터
ict

자율주행
모빌리티
가상현실
딥러닝
클라우드
머신러닝
ict
핀테크
5G기술
블록체인
반도체
디지털
네트워크
메타버스
사물인터넷
빅데이터

모빌리티
자율주행
머신러닝
클라우드
가상현실
디지털
네트워크
ict
핀테크
블록체인
딥러닝
5G기술
메타버스
헬스케어
사물인터넷
반도체
빅데이터



연관어 분석 결과

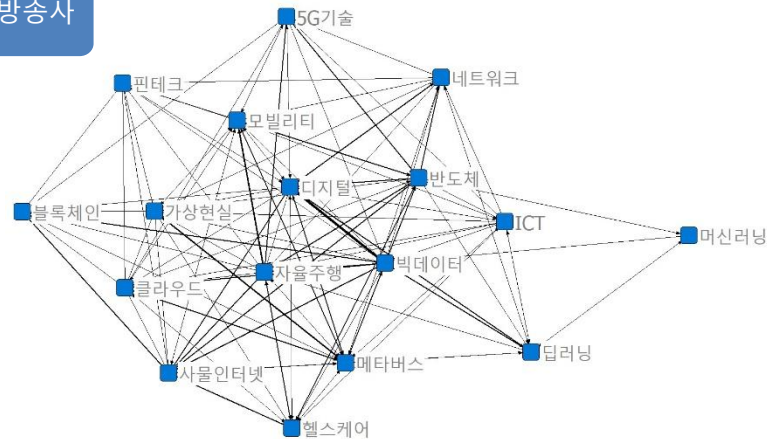
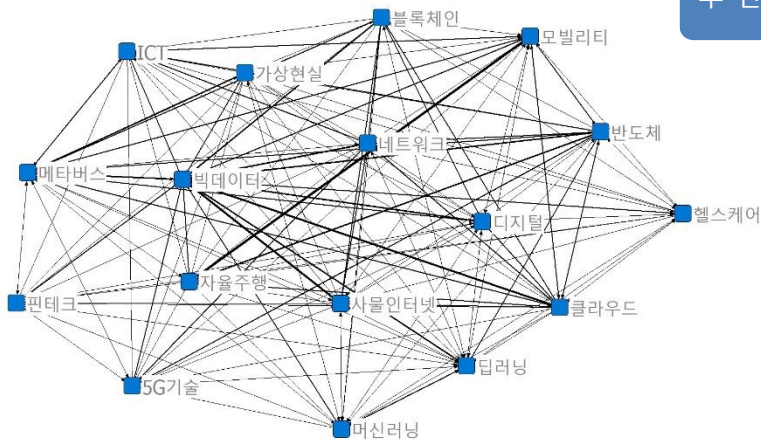


전-신문사

전-방송사

후-신문사

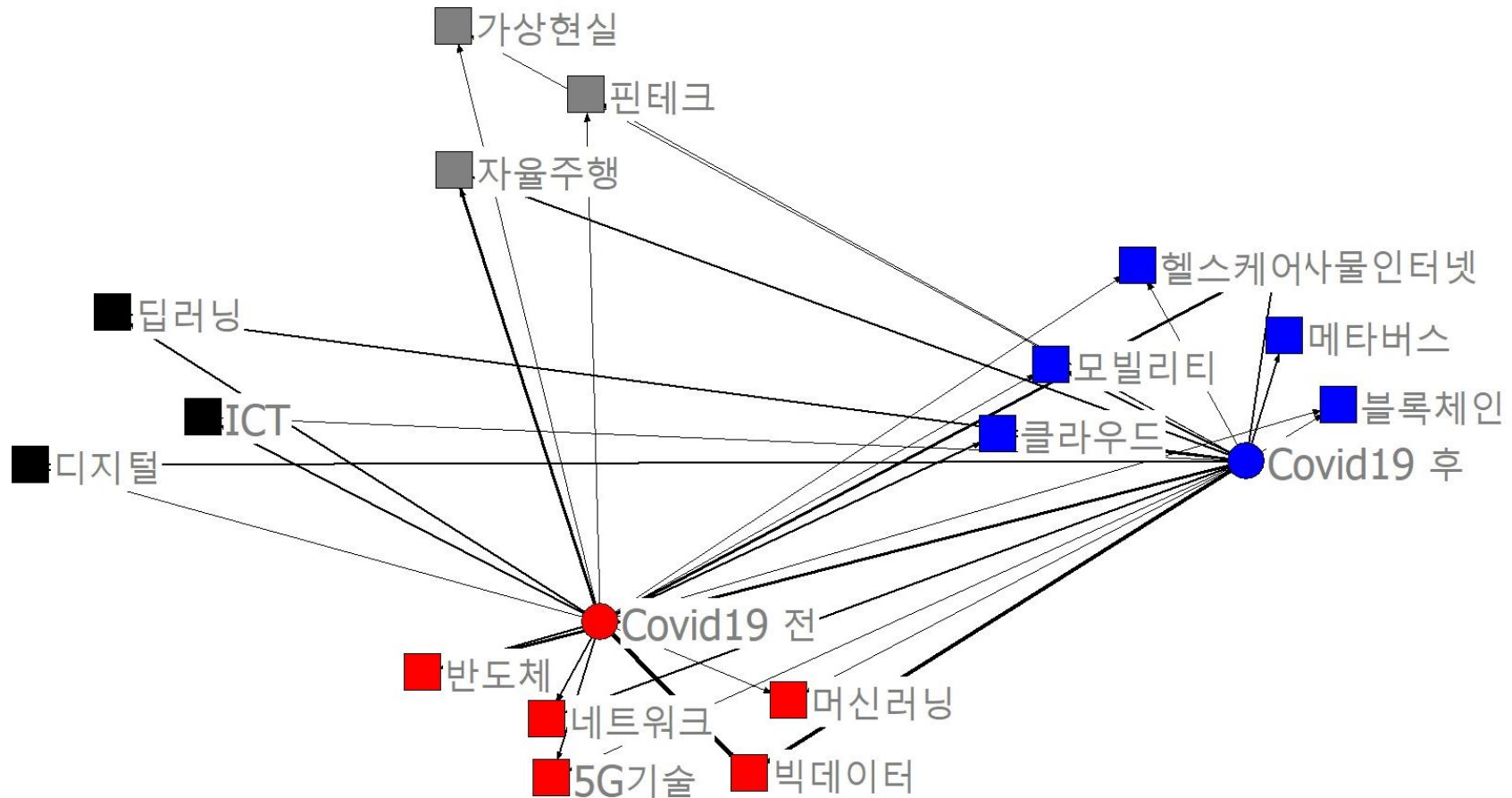
후-방송사



구조방정식

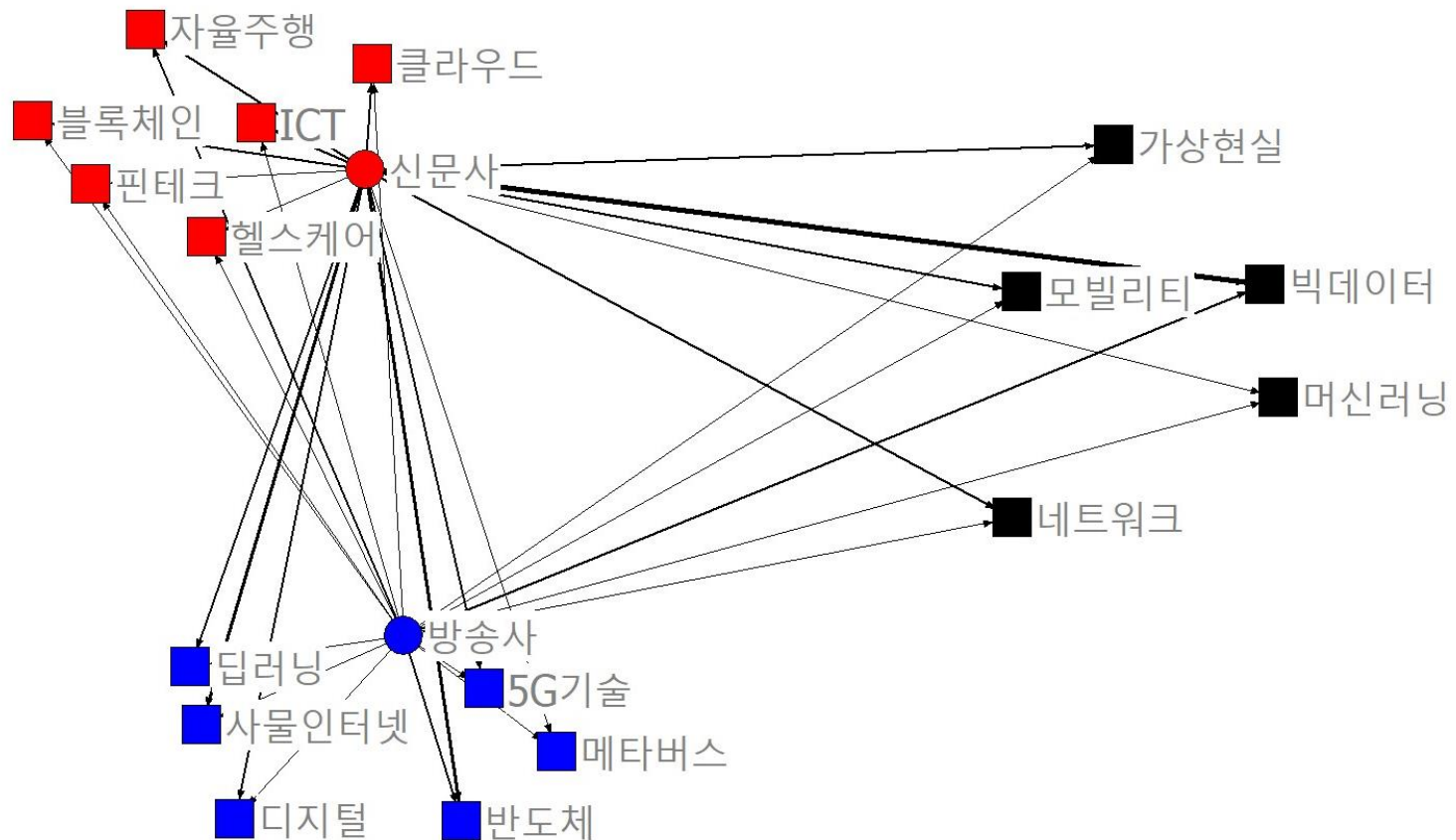


군집분석 결과-시기별



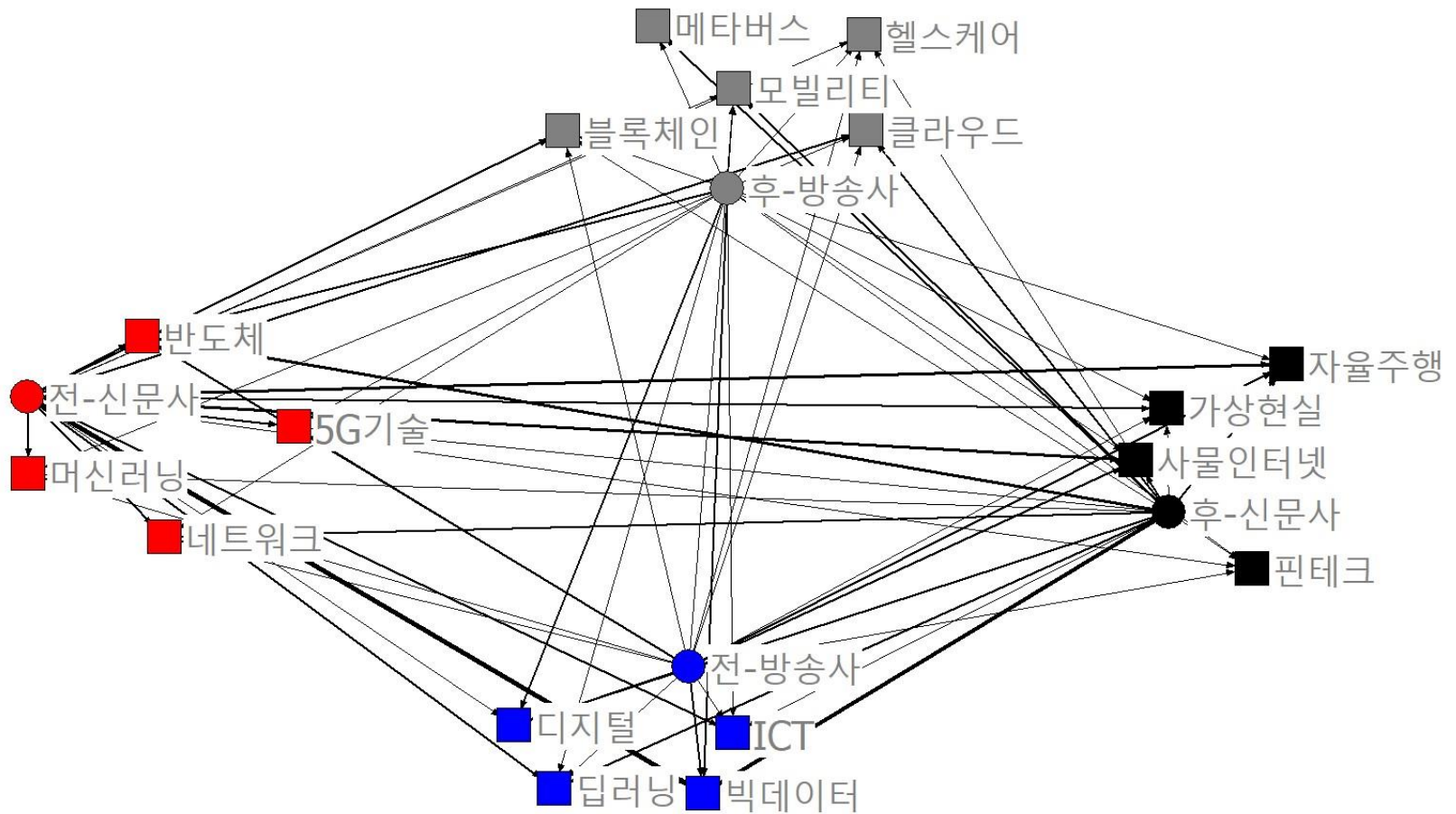


군집분석 결과-언론사별





군집분석 결과-교차





요인분석

	1	2	3	4	5	6	7
5G기술	0.725	0.082	0.008	0.027	-0.062	-0.011	-0.056
네트워크	0.717	-0.022	-0.044	-0.091	0.125	0.027	-0.022
ICT	0.370	-0.001	0.191	0.147	-0.091	-0.044	0.103
모빌리티	-0.045	0.756	-0.121	0.055	0.029	-0.048	0.054
자율주행	0.103	0.755	0.073	-0.033	-0.039	0.075	-0.031
사물인터넷	0.146	0.036	0.626	0.031	-0.064	-0.093	-0.033
빅데이터	-0.037	-0.039	0.598	-0.068	0.147	0.100	0.151
클라우드	-0.002	-0.034	0.480	0.058	0.073	0.052	-0.052
메타버스	-0.067	0.008	-0.083	0.748	0.183	-0.040	0.045
가상현실	0.108	0.013	0.131	0.714	-0.141	0.046	-0.048
블록체인	0.098	0.006	0.100	0.097	0.681	-0.092	0.081
핀테크	-0.067	0.125	0.170	-0.099	0.593	-0.035	-0.199
반도체	0.055	0.123	0.072	-0.013	-0.334	-0.143	-0.097
딥러닝	0.031	-0.009	-0.016	-0.010	-0.064	0.724	0.045
머신러닝	-0.047	0.035	0.074	0.013	0.072	0.698	-0.061
디지털	0.080	-0.056	-0.107	0.018	0.155	0.060	0.830
헬스케어	-0.117	0.128	0.280	-0.030	-0.187	-0.113	0.480

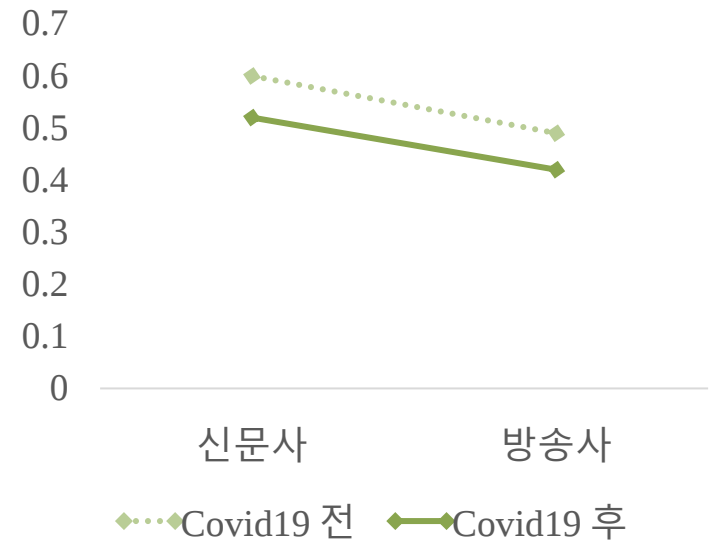


변량분석-전체

변수명	제곱합	자유도	제곱 평균	F값
시기별	22.75	1	22.75	35.44**
언론별	50.34	1	50.34	78.41**
시기 × 언론유형	0.10	1	0.10	0.15

** $p < 0.01$

AI연관어전체

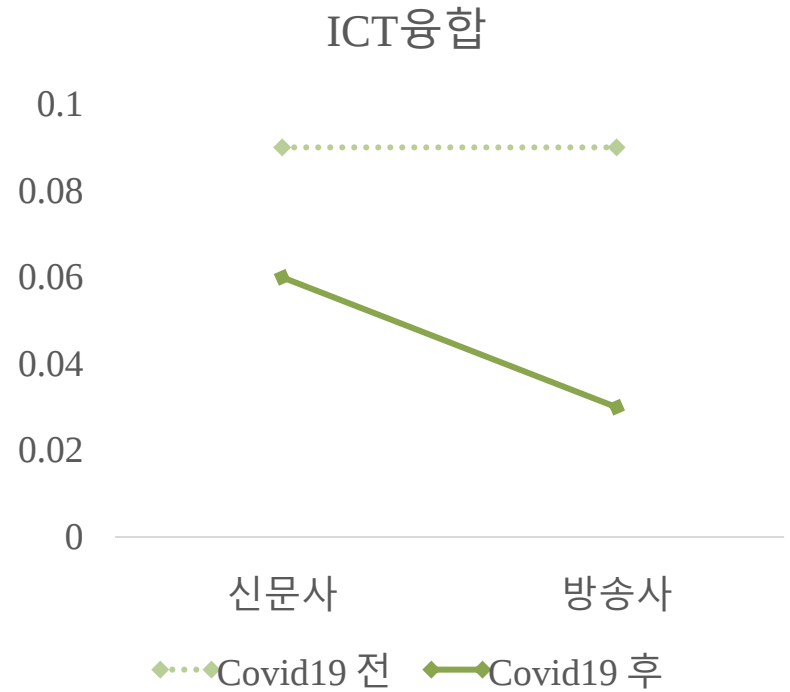




변량분석-ICT융합

변수명	제곱합	자유도	제곱 평균	F값
시기별	9.66	1	9.66	115.78**
언론별	1.29	1	1.29	15.44**
시기 × 언론유형	0.76	1	0.76	9.05**

** $p < 0.01$

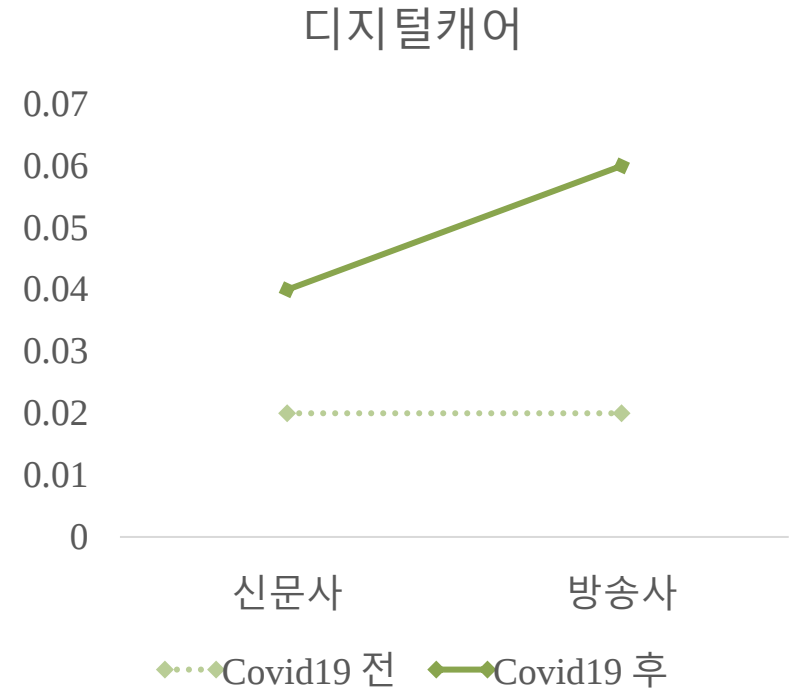




변량분석-디지털케어

변수명	제곱합	자유도	제곱평균	F값
시기별	3.56	1	3.56	107.75**
언론별	0.30	1	0.30	9.15**
시기 × 언론유형	0.69	1	0.69	20.95**

** $p < 0.01$





논의

- ❖ 전반적으로 AI 관계망에 “빅데이터” 개념이 가장 크게 연관되어 출현하는 것으로 나타났다.
- ❖ 연관어 분석결과에서는 “메타버스”는 Covid19 전 시기에는 전혀 나타나지 않았다.
- ❖ 신문사는 “보안”관련 개념이 상대적으로 우세하게 나타나고 있었다.
- ❖ [ICT융합] 개념들은 Covid19 이후 방송사에서 가장 낮게 출현하는 것으로 나타났다.
- ❖ [디지털케어] 개념들은 Covid19 이후 방송사에서 가장 높게 출현하는 것으로 나타났다.
- ❖ 전반적으로 Covid19 이후 방송사의 보도 주제가 상대적으로 크게 변화하는 것으로 나타났다.



감사합니다.