

# **TEXT MINING for BEGINNER**

Part 2. 텍스트 마이닝 실무 활용사례 알아보기

---

전병진 FINGEREDMAN (fingeredman@gmail.com)

## Part 2.

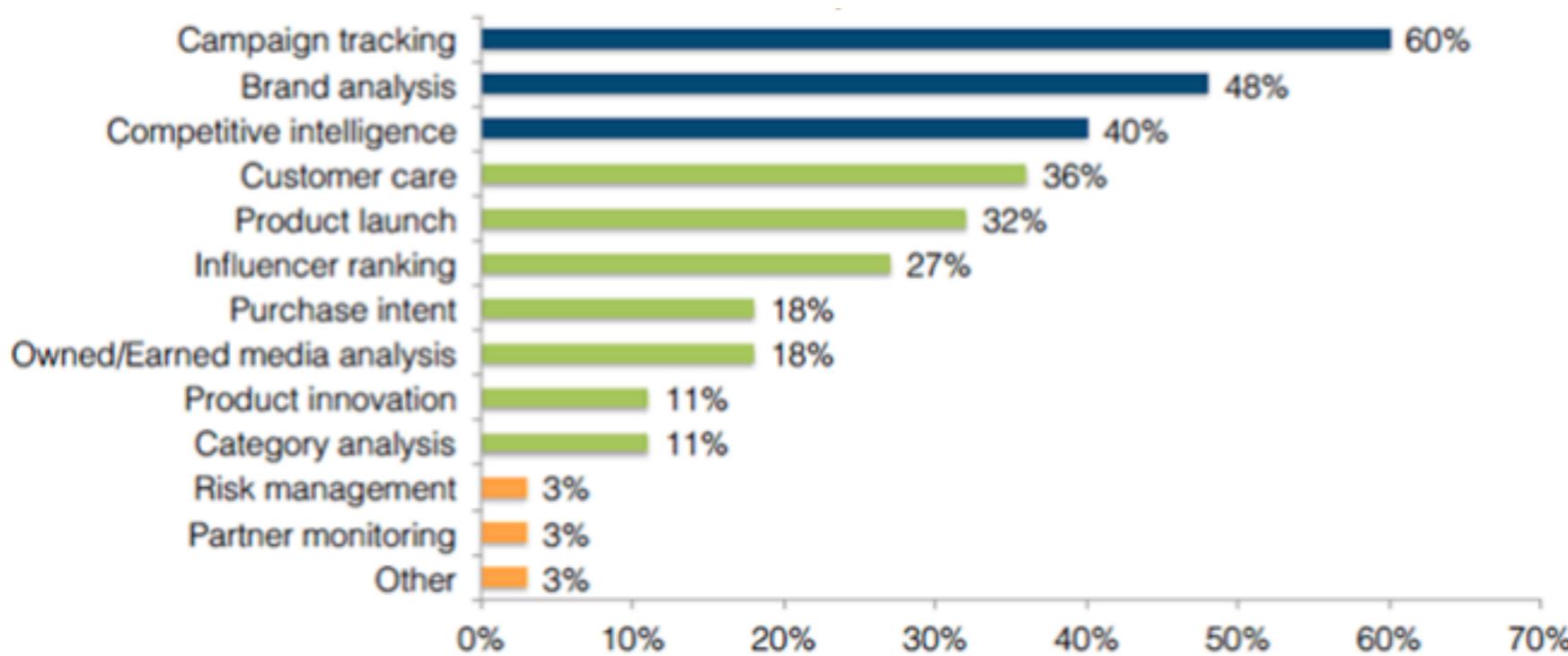
텍스트 마이닝  
실무 활용사례 알아보기

# 텍스트 마이닝 활용분야

활용분야	내용
경쟁 정보 (Competitive Intelligence)	<ul style="list-style-type: none"><li>경쟁사 정보 수집에 들어가는 작업을 효율화</li><li>시장 모니터링, 관련 자료 수집, 비교 등을 자동화</li><li>상품에 대한 고객들의 만족도조사를 효율적으로 진행</li></ul>
인적 자원 관리 (Human Resource Management)	<ul style="list-style-type: none"><li>조직 구성원들의 직무 만족도 등을 텍스트 데이터를 이용해 모니터링</li><li>이력서 관리 등 대규모 텍스트 데이터를 분석해야 하는 경우에 활용</li></ul>
고객관리 서비스 (Customer Care Service)	<ul style="list-style-type: none"><li>고객 커뮤니케이션에서 발생한 텍스트를 이용하여 고객경험을 향상시키기 위한 자료로 활용</li><li>신속한 자동 응답제공으로 콜 센터에 대한 의존도 완화 (Predictive)</li></ul>
시장분석 (Market Analysis)	<ul style="list-style-type: none"><li>고객 반응 모니터링 및 잠재고객 탐색 (Predictive)</li><li>기업의 브랜드 이미지 파악</li></ul>
리스크 관리 (Risk Management)	<ul style="list-style-type: none"><li>문서 관리, 취합, 검색 등을 도와주는 텍스트 마이닝 기반의 리스크 관리 소프트웨어 도입으로 리스크 관리 능력 향상</li></ul>
사이버 범죄 예방 (Cybercrime Prevention)	<ul style="list-style-type: none"><li>텍스트 마이닝을 이용한 범죄 예방 애플리케이션 도입 (Predictive)</li><li>개인 및 기관을 대상으로 한 인터넷 범죄 예방 (Predictive)</li></ul>
의학 (Medical)	<ul style="list-style-type: none"><li>수많은 의학 서적/논문에서 필요한 정보만 추출하거나, 새로운 정보를 찾아냄</li><li>질병과 유전자 사이의 관계를 네트워크화해서 새로운 관계를 규명</li></ul>
정책수립 (Policy Formulation)	<ul style="list-style-type: none"><li>정부 정책에 대한 대중들의 온라인 의견 분석</li><li>새로운 정책전략 수립 및 홍보 등에 활용</li></ul>

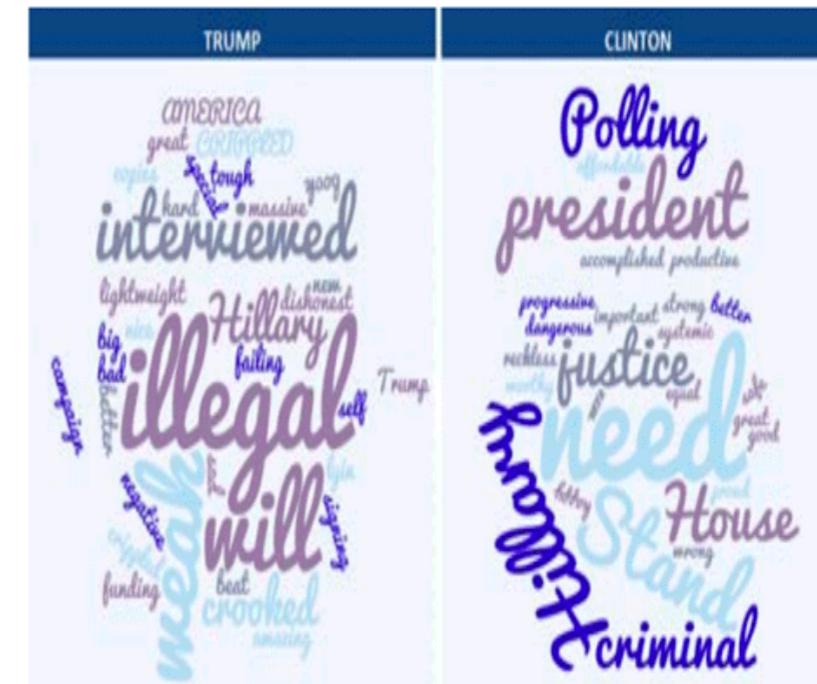
# 텍스트 마이닝 활용분야

Text Mining	Language Processing
Spam filtering	Siri (Apple) and Google Now
Document Classification	Language Understanding
Date/time event detection	Spelling Correction
Information Extraction	Statistical Language Modeling
(Web) Search engines	Website translation (Google)
Information Retrieval	Machine Translation
Watson in Jeopardy! (IBM)	"Clippy" Assistant (Microsoft)
Question Answering	Dialog System
Twitter brand monitoring	Finding similar items (Amazon)
Sentiment Analysis (Stat. NLP)	Recommender System



# 텍스트 마이닝 적용사례

## ① 2016 미국 대통령선거 분석



TRUMP	%
@CNN	16%
@FoxNews	14%
@foxandfriends	8%
@nytimes	7%
@JebBush	7%

CLINTON	%
@POTUS	36%
@billclinton	20%
@realDonaldTrump	16%
@BernieSanders	10%
@HFA	4%

TRUMP	COUNT
Terror	133
Immigration	78
Jobs	78
Taxes	53
Guns	13
Education	13
Health	8

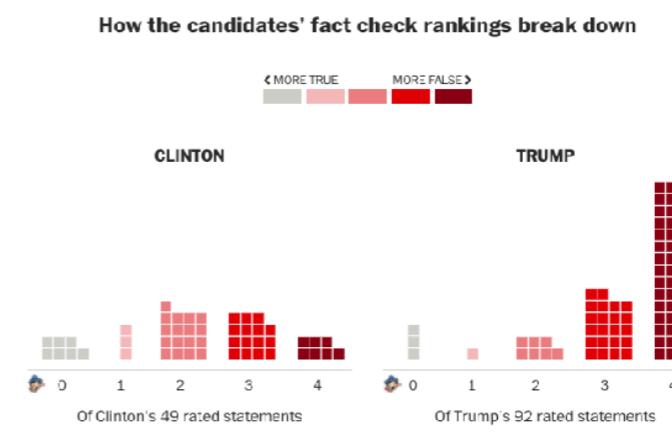
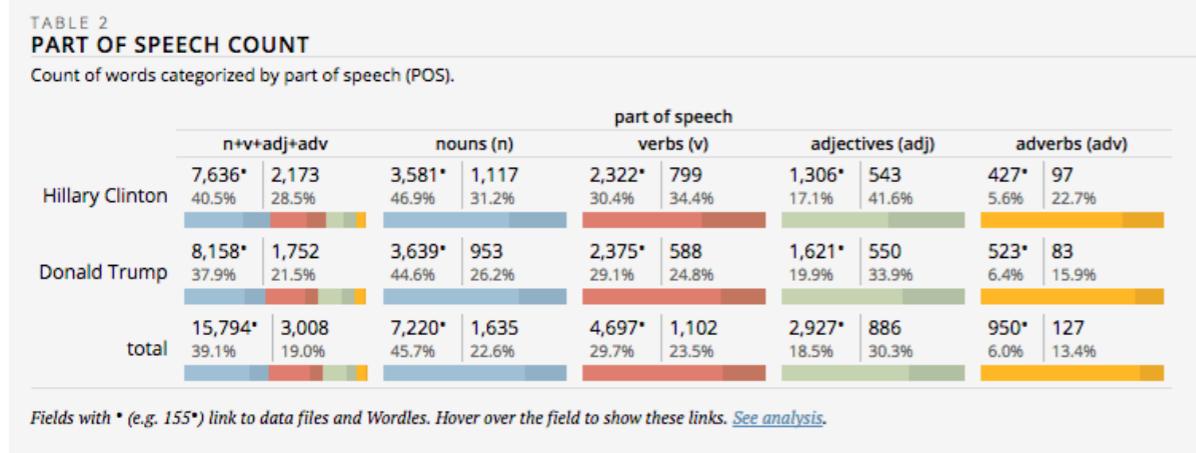
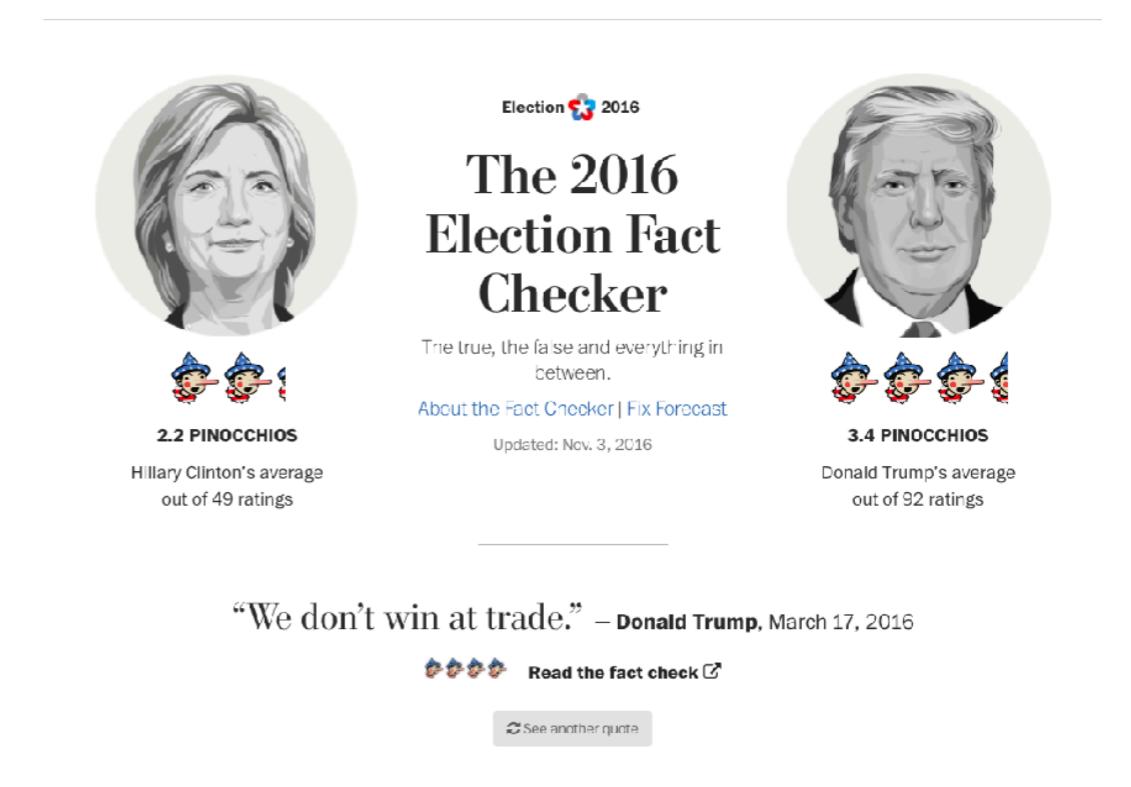
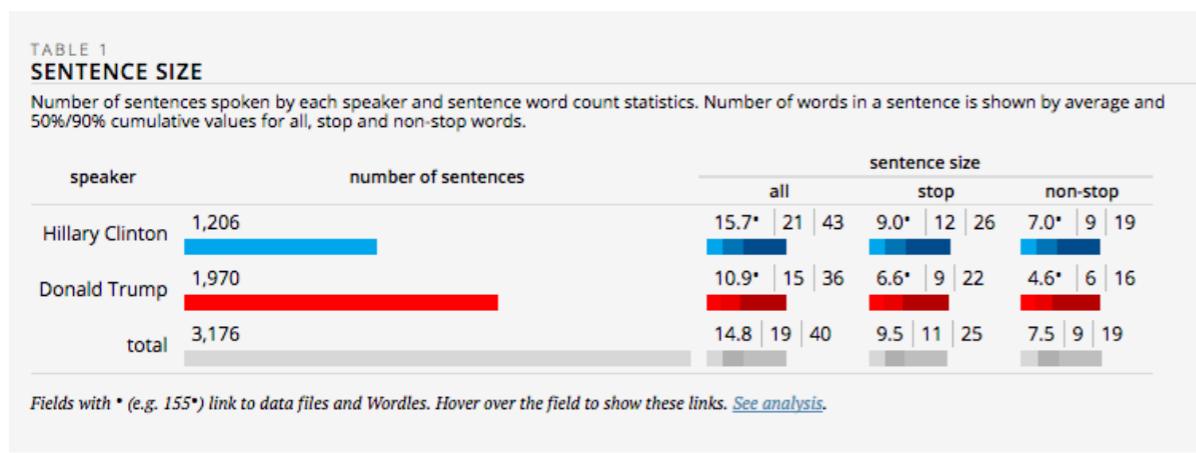
CLINTON	COUNT
Guns	250
Health	150
Taxes	83
Immigration	80
Education	53
Foreign	53
Vets	28

\*Source : 최홍규, 2016 미국 대선을 보여주는 텍스트 마이닝 분석방법들, 2017.1.9., <http://slownews.kr/60919/>.

\*\*Source : Martin Krzywinski, WORD ANALYSIS OF 2016 PRESIDENTIAL DEBATES — CLINTON VS TRUMP, 2016., <http://mkweb.bcgsc.ca/debates2016/>.

# 텍스트 마이닝 적용사례

## ① 2016 미국 대통령선거 분석



# 텍스트 마이닝 적용사례

## ② 여론조사

### 시국선언문 텍스트 분석 해보니…‘하야’ 민심 갈수록 커졌다

등록 : 2016-11-17 13:53 수정 :

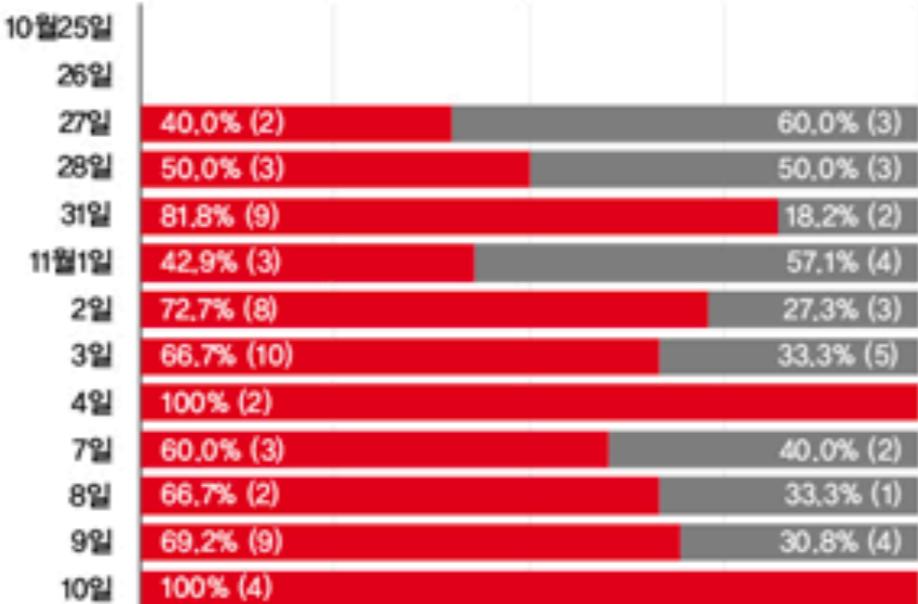


[뉴스AS] 10월25일~11월10일 나온 시국선언문 173건 분석

<한겨레>는 '최순실 게이트'에 따른 각계 민  
습니다. 본문이 없거나 짧은 선언 성격의 글

#### 갈수록 ‘탄핵’보다 ‘하야’

‘탄핵+하야’ 대비 언급 비율(언급 횟수)



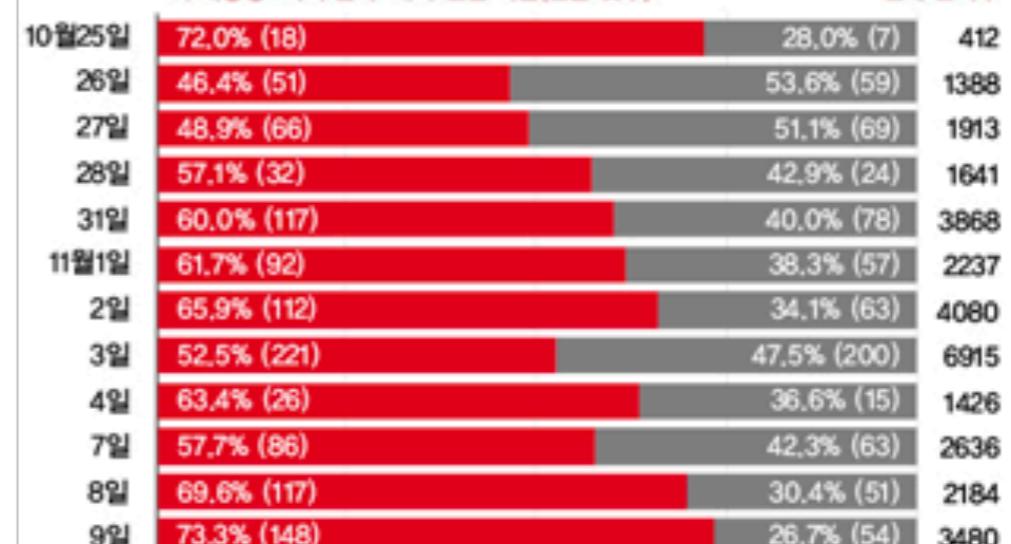
■ 하야

■ 탄핵

\*토요일, 일요일 데이터 제외

#### ‘최씨 일가’ 아닌 박 대통령

‘박 대통령+최씨 일가’ 대비 언급 비율(언급 횟수)



#### 잊혀진 ‘개헌’

전체 단어 중 ‘개헌’ 언급 비율(언급 횟수)



■ 개헌 ■ 개헌 외 모든 단어

\*토요일, 일요일 데이터 제외

# 텍스트 마이닝 적용사례

## ③ 고객성향분석 & 경제상황분석

### 보험사 빅데이터로 돈번다…‘고객성향분석’ 교보생명, 계약대출 2%→18% 증가

보험사의 고민(성장정체·규제·수익·경쟁자)을 해결해줄 빅데이터...모델링 검증후 적합성 확인 과정필요

김남희 기자 | nina@econovill.com | 승인 2015.08.20 15:21:20



#### ◇ 활용사례 ⑥ 일본 메이지야스다생명 : 지급심사 자동화

메이지야스다생명은 80만 어휘로 구성된 자사 고유 보험용어 사전과 3500개의 지급관련 키워드로 구성된 '의료사전'을 구축했다. 보험금 청구접수가 들어오면 진단명, 급부금 등 보험금 규모가 자동산정이 되도록 해 지급심사 과정을 대폭 줄였다. 그 결과 2일내 지급처리 비율이 50%에서 80%로 30%p 향상된 가운데 지급누락도 사전에 예방됐다. 진단서 텍스트 마이닝을 통한 지급심사 효율을 개선한 경우다.



### 주식 커뮤니티보면 코스피지수도 보인다

조선비즈 | 이종현 기자

입력 2018.12.10 10:19 | 수정 2018.12.10 12:08

#### 주식 커뮤니티 게시글 18만여개 키워드 분석

10월 들어 코스피지수가 급락하자 온라인 주식 커뮤니티에 올라오는 글에서 '수익'이라는 키워드가 빠르게 줄어든 것으로 조사됐다. 대신에 '장외' '비상장' 같은 키워드는 등장하는 빈도 수가 늘었다. 코스피지수가 급락하자 투자자들이 당장의 수익보다 새로운 투자처를 찾는데 골몰했기 때문으로 보인다.

7월		8월		9월		10월		11월	
주식	1,662	주식	1,711	주식	3,399	주식	4,931	주식	10,298
투자	1,220	투자	1,469	투자	2,392	투자	3,451	투자	3,774
종목	942	수익	1,041	수익	2,008	주식시장	2,527	장외	3,353
주식시장	867	주식시장	954	주가	1,548	시장	2,398	거래	3,074
수익	740	종목	927	종목	1,459	미국	1,898	비상장	2,814
시장	663	시장	865	주식시장	1,360	종목	1,872	시장	2,575
매수	563	상승	669	시장	1,337	수익	1,709	주식시장	2,332
상승	527	주가	602	부동산	1,141	주가	1,667	기업	2,309
기업	483	미국	530	상승	1,083	상승	1,472	종목	2,092
주가	482	부동산	501	정보	1,078	금리	1,464	수익	1,976
미국	424	투자자	494	매매	1,069	기업	1,441	정보	1,757
정보	410	가격	477	미국	873	부동산	1,395	미국	1,688

최근 5개월 동안 주식 커뮤니티에 많이 등장한 키워드. /몬데이터 제공

\*Source : 김남희(이코노믹리뷰), 보험사 빅데이터로 돈번다…‘고객성향분석’ 교보생명, 계약대출 2%→18% 증가, 2015.8.20., <http://www.econovill.com/news/articleView.html?idxno=258698>.

\*\*Source : 이종현(조선비즈), 주식 커뮤니티보면 코스피지수도 보인다, 2018.12.10., [http://biz.chosun.com/site/data/html\\_dir/2018/12/10/2018121000888.html](http://biz.chosun.com/site/data/html_dir/2018/12/10/2018121000888.html).

# 텍스트 마이닝 적용사례

## ④ 사이버 수사 & 디지털 포렌식

### Identifying and Profiling Key Sellers in Cyber Carding Community: AZSecure Text Mining System

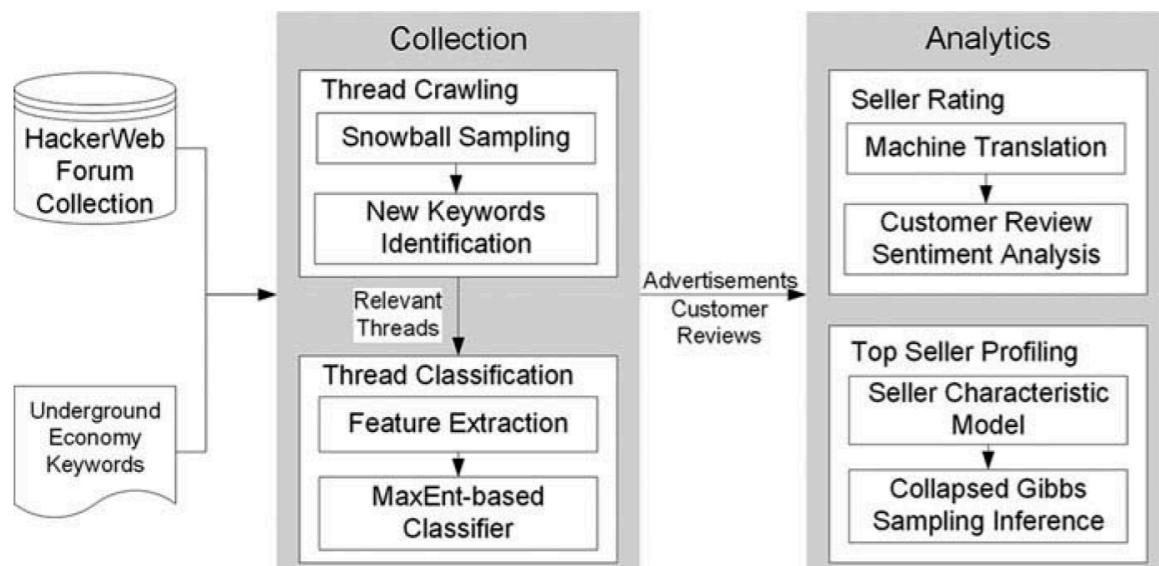


Figure 2. The AZSecure Text Mining Research System

"자살자 휴대전화 분석해 원인 파악 가능…예방에도 기여"

송고시간 | 2017/03/07 07:08

f t , G+ BAND blog | ☰ + -

경남경찰청 강민규 경사 '디지털포렌식' 석사 논문서 주장

(창원=연합뉴스) 김선경 기자 = 자살자가 생전 휴대전화에 남긴 기록을 통해 자살 원인을 분석한 연구 결과가 나왔다.

경남지방경찰청 사이버범죄수사대 소속 강민규 경사는 최근 발표한 연세대학교 정보대학원 디지털포렌식(Digital Forensics·디지털 감식) 전공 석사 논문 '텍스트 마이닝 기법을 활용한 자살 데이터 분석'에서 이런 내용을 밝혔다.

텍스트 마이닝(text-mining)이란 구조화되지 않은 대규모 텍스트를 통계적으로 분석해 새 지식을 도출해내는 기법을 의미한다.



\*Source : Li et al., (2016). Identifying and profiling key sellers in cyber carding community: AZSecure text mining system. Journal of Management Information Systems, 33(4), 1059-1086.

\*\*Source : 김선경(연합뉴스), "자살자 휴대전화 분석해 원인 파악 가능…예방에도 기여", 2017.3.7., [http://biz.chosun.com/site/data/html\\_dir/2018/12/10/2018121000888.html/](http://biz.chosun.com/site/data/html_dir/2018/12/10/2018121000888.html/).

# 텍스트 마이닝 적용사례

## ④ 사이버 수사 & 디지털 포렌식

'IBK 보이스피싱 탐지 앱' 핵심은.."텍스트마이닝으로 객관적 확률 판단"

업계 최초 앱 개발한 이봉기 기업은행 과장 인터뷰  
통화내용 실시간 분석 통해 확률 높으면 '주의경보'  
9일 企銀-금감원-NIA 업무협약..내년 1월 출시목표

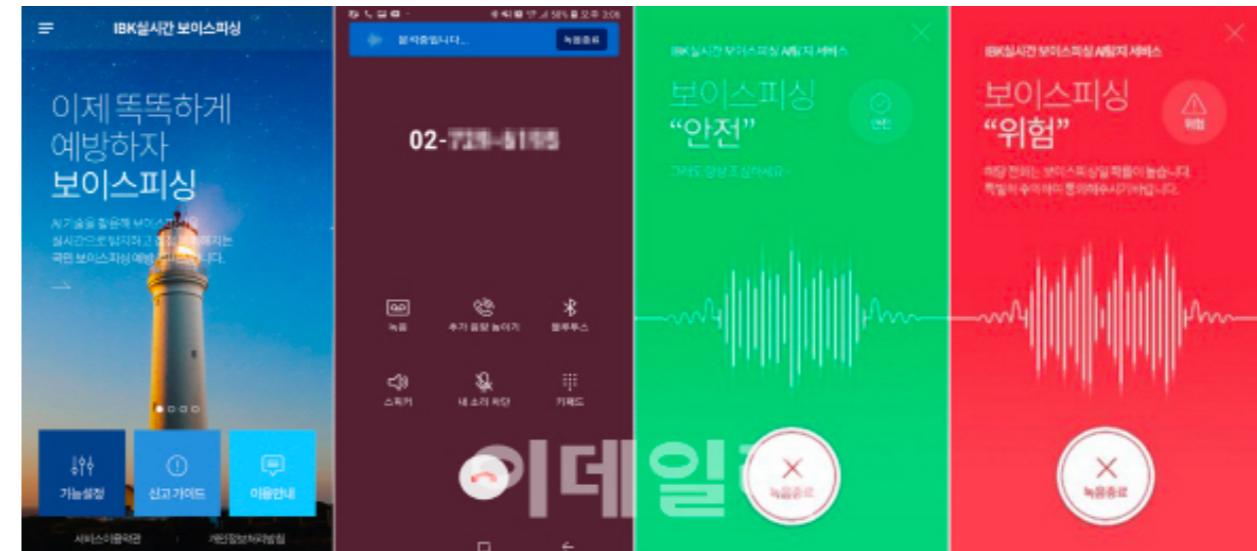
등록 2018-11-09 오후 2:46:00  
수정 2018-11-09 오후 4:14:23



김범준 기자



지난 7일 서울 중구 BK기업은행 을지로본점에서 이봉기 기업은행 IT정보부 과장이 개발 중인 'BK 실시간 보이스피싱 인공지능 탐지 앱(창)'에 대해 이데일리와 인터뷰를 하고 있다.



[이데일리 김범준 기자] “보이스피싱에 사용되는 단어들과 문맥은 어찌보면 뻔하고 정형화돼 있는데 심리적으로 말리면 객관성을 잃게 돼요. 이때 컴퓨터가 텍스트 정보를 객관적으로 판단해 ‘조심하라’고 한마디만 해주면 정신을 차릴 수 있어요.”

# 텍스트 마이닝 적용사례

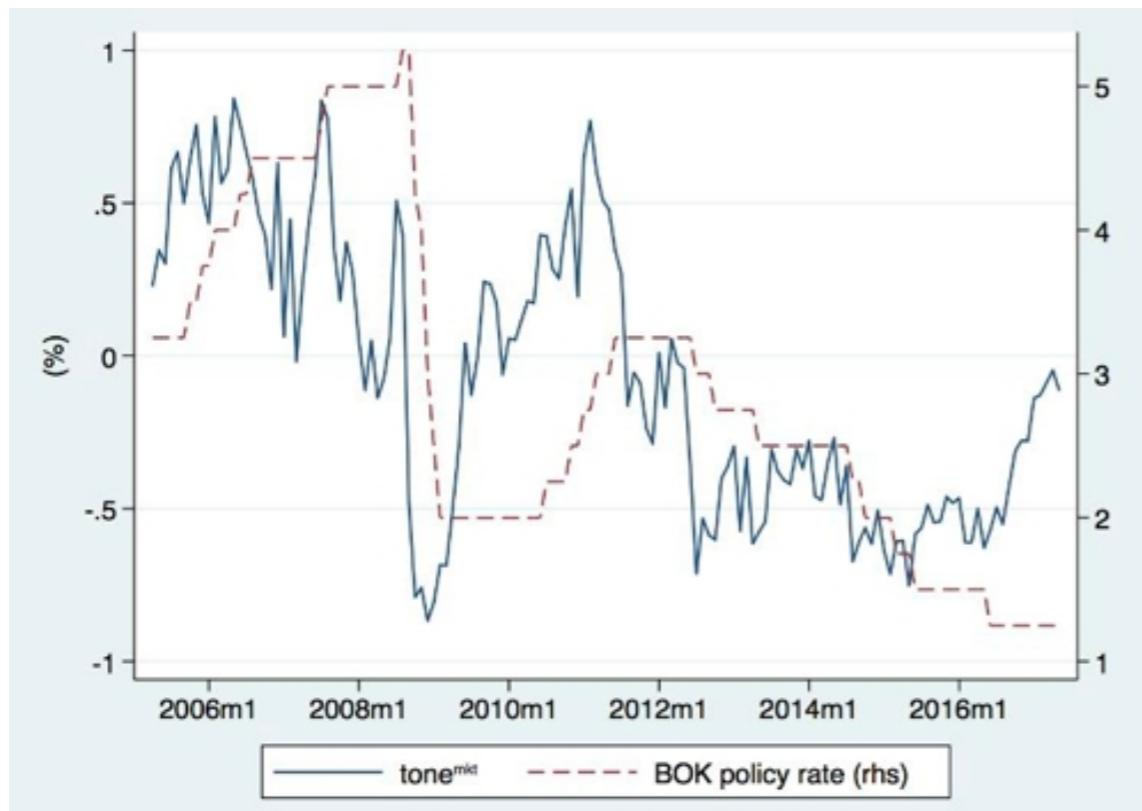
## ⑤ 경제 트렌드 예측

### 한은 "텍스트마이닝 활용해 의사록 분석...기준금리 예측 가능"

◇ 전소영 기자 | ◇ 승인 2019.01.06 12:00 | ◇ 댓글 0

(서울=연합인포맥스) 전소영 기자 = 한국은행 금융통화위원회 의사록을 텍스트마이닝(text mining) 기법으로 활용하면 기준금리의 방향을 예측할 수 있다는 결과가 나왔다.

김수현 한은 경제연구원 국제경제연구실 부연구위원은 6일 '텍스트 마이닝을 활용한 금융통화위원회 의사록 분석' 보고서를 통해 이같이 밝혔다.

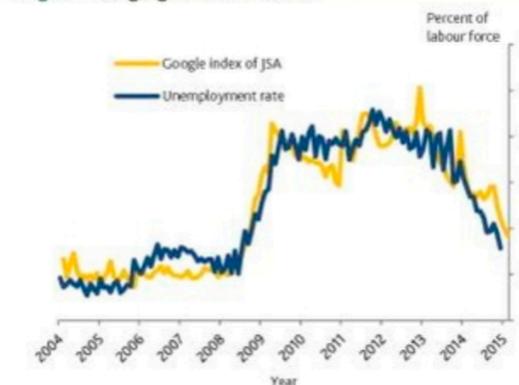


### 한은, SNS·댓글·기사 수집해 경제 분석한다

조선비즈 | 조귀동 기자

입력 2017.03.02 08:46

Figure 1 Googling the labour market



한국은행이 페이스북이나 트위터 등 소셜네트워크서비스(SNS), 네이버 등 인터넷 웹사이트에 올라온 댓글, 경제 관련 언론 기사 등을 분석해 경제 상황 해석에 활용한다.

가령 '물가가 너무 많이 올랐다'거나 '일자리 잃는 사람이 너무 많아 장사가 힘들다'는 식의 인터넷 게시물이나 언론 기사를 모아서 물가 상승률, 실업률, 각 경제 주체들의 심리 등을 파악하는 게 활용한다는 것이다. 이른바 '텍스트 마이닝(text mining: 많은 텍스트를 '채굴'해 쓸만한 데이터를 만들어낸다는 의미)'을 경제 조사와 통화 정책 수립에 반영하겠다는 얘기다.

# 텍스트 마이닝 적용사례

## ⑥ 금융사 빅데이터 플랫폼 구축

세계 최초 디지털 콘택트센터 구축한다...KB금융 스타링크  
내달 출범

발행일 : 2018.09.17



[AD] 4월 25일 차세대 디스플레이 기술응용 세미나 개최

| 내달 '스타링크' 본격 상용화...민원처리 업무서 역할 재정립



<KB금융이 아날로그 콜센터를 빅데이터 기반 전문 상담센터로 업그레이드한다. 17일 서울 마포구 KB국민은행 스타링크 직원이 상담내역을 빅데이터화한 정보로 고객과 상담하고 있다. 김동욱기자 gphoto@etnews.com>

은행 창구를 가지 않아도 콘택트센터(일명 콜센터)를 통해 예금, 대출, 자산 관리가 가능해진다. 약 7~8단계에 이르는 자동응답전화(ARS) 기계음을 모두 없애고, 고객이 상담한 음성 내역은 텍스트로 전환돼 빅데이터로 쌓인다. 본부와 은행 창구가 이를 활용, 고객에게 먼저 서비스를 제시하는 '디지털 오퍼' 시대도 열린다. 텔레마케팅 전담 조직도 없어진다.

구분	내용
“ 콜센 터 디지 털화 ”	△인프라 구축(지능형 전화 연결, 콘텐츠 발송, 新상담관리 시스템 등) △빅데이터 분석 기반 마련(STT/TA 고도화) △채팅 상담 강화 : 다중(1명의 직원이 최대 3명 고객 응대), 멀티(유선+채팅) 상담 운영 △디지털 전문 인력 운영, 디지털화 업무(RPA) 확대
“ 스타 링크 고 도화 ”	△스타링크 브랜드화 △고객경험관리 프로세스, 비대면 고객관리체계 수립 △스타링크 Plus+ 서비스 시행 : 관리 사각지대 잠재 우수 고객 발굴 및 관리 △상담 직원 캐어 프로 그램 구축
“ 비대 면 영업 채널화 ”	△비대면 금융 마케팅 허브 시스템 구축 △비대면 대출 운영 집중화 △고객분석(STT/TA, Intelligent Routing 활용) 기반 개인화 마케팅 △비대면 고객 전담 관리 강화
“ 채널 간 연계 강화 ”	△챗봇 운영(4119/대고객) △그룹 계열사 콜센터 간 콘트롤타워 구축 △전문 상담 인력 양성 및 확보

# 텍스트 마이닝 적용사례

## ⑦ 인공지능 (AI) 비서/챗봇

한경BUSINESS

“이제는 ‘AI 퍼스트’ 시대”...가전에서 차량까지 필수 탑재

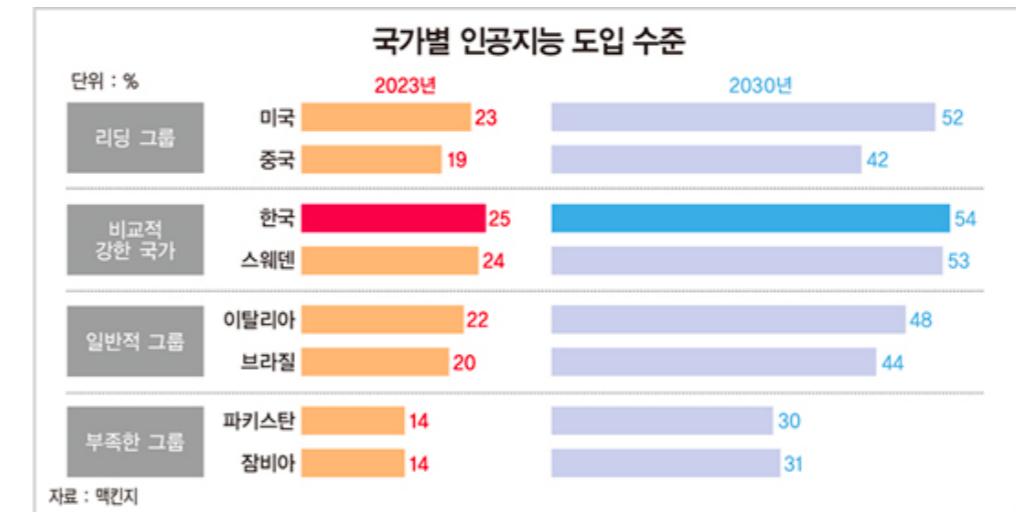
기사입력 2018-11-20 15:25 최종수정 2018-11-20 18:06 기사원문 스크랩 본문듣기 설정



[커버스토리-'미리 보는 2019년 IT 시장 빅 이슈 5']-시각 대신하는 ‘비전 AI’ 고도화 예상, “AI가 모든 산업 재정의할 것”



존 크라프치 웨이모 CEO가 지난 5월 열린 구글 IO 컨퍼런스에서 발표하고 있다.(연합뉴스)



순다 피차이 구글 CEO는 미국 캘리포니아 주에서 지난 8월 열린 구글 연례 개발자 회의 ‘구글 IO’에서 구글의 새로워진 AI 서비스들을 선보였다.

구글 어시스턴트에 AI를 활용해 더 많은 목소리가 추가됐고 구글 어시스턴트가 직접 가게에 전화를 걸어 예약할 수 있게 됐다. 구글은 이용자들이 구글 어시스턴트와 자연스러운 대화가 가능하도록 기술을 발전시켰기 때문이다.

피차이 CEO는 이를 “자연어 처리, 딥러닝, 텍스트 투 스피치 등이 모두 결합된 기술”이라며 이 기술을 ‘구글 듀플렉스(Google Duplex)’라고 소개했다. 또 e메일을 작성할 때 문장을 추천해 주는 ‘스마트 작성(Smart Compose)’ 기능을 선보였다.

# 텍스트 마이닝 적용사례

## ⑦ 인공지능 (AI) 비서/챗봇

인터넷 | 김종윤 스캐터랩 "강아지만큼 친근한 AI 챗봇 만든다"

"개발중인 챗봇 '핑퐁'은 감정까지 표현"

김민선 기자 | 입력 : 2018.05.29.18:12 | 수정 : 2018.05.31.08:23

0

[Asia Tech Summit 2018] AI, 이제는 플랫폼이다 | 12월 12일, 포시즌스 호텔 서울 |  
[ZDNet Korea Webinar] 마이크로소프트, HPE 등 최신 IT 트렌드를 확인해보세요!

"인공지능을 단순한 기계로 생각하지 않는다. 좀 더 살아있는 것 같은 느낌이 든다. 그렇기 때문에 기계나 컴퓨터와는 근본적으로 다른 특성이 있는 것 같다. 우리 경쟁자는 강아지다. 인공지능이 탑재된 제품과의 관계나 감정이 반려동물과 아닐까?"



김종윤 스캐터랩 대표는 이상적인 인공지모습이라고 강조했다. 그가 2011년 회시 키워드를 유난히 강조한 것도 그 때문이

실제로 그가 내놓은 서비스들은 거의 대

핑퐁 월 그런걸루 스트레스 받아

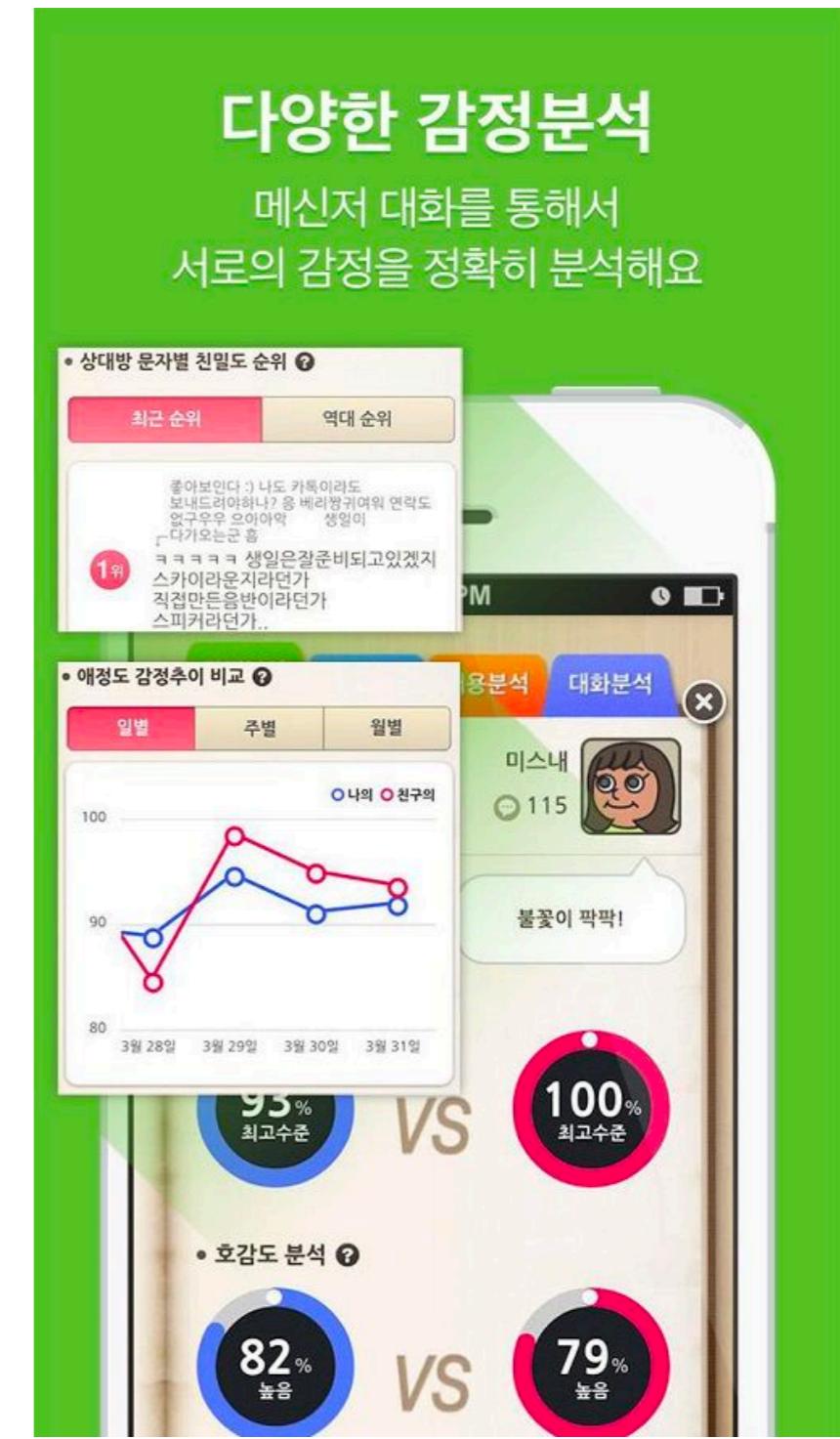
2013년 2월 처음 내놓은 '텍스트앳'은 키저 앱 비트원에서 오간 대화를 통해 연애을 파악하는 '연애의과학'까지 모두 대화

시리 (애플) 흥미로운 질문이네요

빅스비 (삼성) 인터넷에는 이렇게 나와있네요 : 검색

누구 (SKT) 원하시는 답변을 찾지 못했습니다 : 처리불가

클로바 (네이버) '일하는 거 너무 스트레스야'에 대한 내용을 찾아봤어요 : 검색



# 텍스트 마이닝 적용사례

## ⑦ 인공지능 (AI) 비서/챗봇

### [4차산업혁명] 마인즈랩, KEB하나은행 '하이뱅킹' 구축…AI 신기술 총망라

3D AI 금융비서와 음성, 텍스트로 대화  
하이뱅킹 2차 프로젝트 진행

총정민 기자 | 승인 2018.11.17 13:19 | 댓글 0



이번에 KEB하나은행과 진행한 하이뱅킹 2차 프로젝트를 통해 마인즈랩은 고객이 3D 인공지능 금융 비서와 음성 및 텍스트로 대화하며 계좌이체, 상품 가입 및 추천, 환전 및 해외 송금 등을 간편하게 이용할 수 있도록 서비스에 적용된 자연어 및 대화 처리 관련 AI 신기술을 도입했다.

또한 시각지능 프레임워크를 도입한 하이렌즈 기능을 통해 외화 지폐 인식, 고지서 자동 납부 등이 가능해지면서 AI 뱅킹 서비스 인터페이스가 시각 영역까지 확장됐다.

마인즈랩은 KEB하나은행과의 이번 프로젝트에 최신 딥러닝 알고리즘을 기반으로 한 마인즈랩의 AI 플랫폼 마음에이아이(maum.ai)의 다양한 AI 엔진을 통합적으로 적용했다. 이를 통해 AI의 인지 채널을 텍스트에서 음성, 시각으로 확대하고 대화 이해 수준을 한층 높여 동종 업계 최고 수준의 AI 비서 서비스를 구현했다.

HAI 서비스는  
손님이 '문자' '음성' '이미지촬영'을 통해  
요청한 의도를 인공지능 금융비서 HAI가  
분석하여 각종 금융업무를 처리해주는  
금융비서 서비스입니다.



# 텍스트 마이닝 적용사례

## ⑧ 기타 금융사 적용사례

### 통합 보험 앱 '위플랫' 출시… "보험약관에 집중"

지난달 데모버전 출시, 9월 중 정식버전 오픈… 병원 콘텐츠 유기적 활용 가능

배승희 기자 승인 2018.08.20 11:13 댓글 0

[이뉴스투데이 배승희 기자] 숨은 보험금 찾기에 대한 국민 관심이 높다. 모바일 앱 '위플랫'에서도 잠자고 있는 보험금을 찾을 수 있다.

지난해 10월말 기준 소비자의 숨은 보험금은 약 7조4000억원이다. 6월 말 기준, 정부의 '내 보험 찾아줌(ZOOM)'을 이용해 2조1000억원을 찾아간 것으로 나타났다.

민간의료보험은 우리나라 10가구 중 9가구가 가입하고 있을 정도로 대중화됐다. 그러나 정작 필요할 때 보험금을 지급받지 못하는 경우도 허다하다. 실제 금융감독원의 '2017년 금융민원 및 금융상담 동향' 자료에 따르면, 보험관련 민원건수가 생명보험 1만8101건, 손해보험 2만9641건인 가운데, '보험금 산정 및 지급 문제'로 인한 민원이 각각 18.7%, 46%를 차지했다. 제대로 된 보험정보의 필요성이 커지는 이유 중 하나다.



대부분 놓치기 쉬운 수술비, 입원비와 같은 정액보험 특약사항도 분석하여 알려준다. 이렇게 찾은 보험금은 해당 보험사로 바로 청구를 할 수 있다. 여러 보험사로 동시에 청구도 가능하다. 국내 보험관련 앱 중에서는 유일한 서비스다.

# 텍스트 마이닝 적용사례

## ⑧ 기타 금융사 적용사례

### [금융과데이터만남②] SNS신용등급

박형재 기자 news34567@the-pr.co.kr | 승인 2018.07.26 17:25 | 댓글 0

| 댓글이 대출금리 결정...페북·인스타 대화로 고객 행동 분석, 개인화 마케팅 연결



[더피알=박형재 기자] SNS 맞춤법을 자주 틀리면 신용대출 등급이 낮아지고, 주기적으로 택배를 시키면 반대로 등급이 올라간다. 농담 같지만, 실제로 외국 기업이 정해놓은 기준이다. 금융사들이 은행이나 카드 거래 같은 공식 데이터 외에 SNS를 비롯한 비정형 데이터에 주목하고 있다.

독일의 핀테크 기업 크레디테크(Kreditech)는 대출심사를 할 때 기존 은행거래정보 외에 페이스북, 이베이, 아마존에서의 행동 패턴까지 반영한다. SNS에서 맞춤법을 틀리는지, 대출 약관을 꼼꼼히 읽었는지, 주기적으로 온라인 쇼핑을 했는지 등이다. 맞춤법을 틀리지 않는 대출자는 상대적으로 덜 연체하는 특성을 갖고 있으며, 주기적으로 택배기사가 방문한 기록이 있으면 일정한 수입이 있다고 보고 신용평가 점수가 올라간다.

미국의 렌도(Lenddo)는 SNS와 통신사 데이터 등을 신용평가에 활용한다. SNS 친구 중 연체자가 있거나, '사고' '실직' 같은 부정적인 단어가 많이 나오면 신용점수가 낮아진다. 또한 통화량 분석을 통해 '주간 통화량이 심야 통화량보다 많은 사람이 대출 상환률이 높다', '매일 비슷한 시간에 통화·검색·이메일 확인을 하는 사람이 불규칙한 사람보다 신용도가 높다'는 결과값을 도출했다.

페이스북의 전 세계 가입자 수는 22억명에 달한다. 인스타그램은 7억명 가량이다. 금융사들이 SNS를 비롯한 비정형 데이터에 주목하는 흐름은 어쩌면 당연해 보인다. 기존 신용평가모델에서는 고객의 대출액 규모나 연체율 등 금융 거래 이력을 기반으로 고객의 신용도를 평가했으나, 이제는 빅데이터를 이용해 고객들의 SNS, 소비행태 등 비재무적 요소까지 함께 고려하는 경우가 많아지고 있다.

# 실무 적용사례

Case 1:

네이처메이드 수입영양제

브랜드 포지션 전략 수립

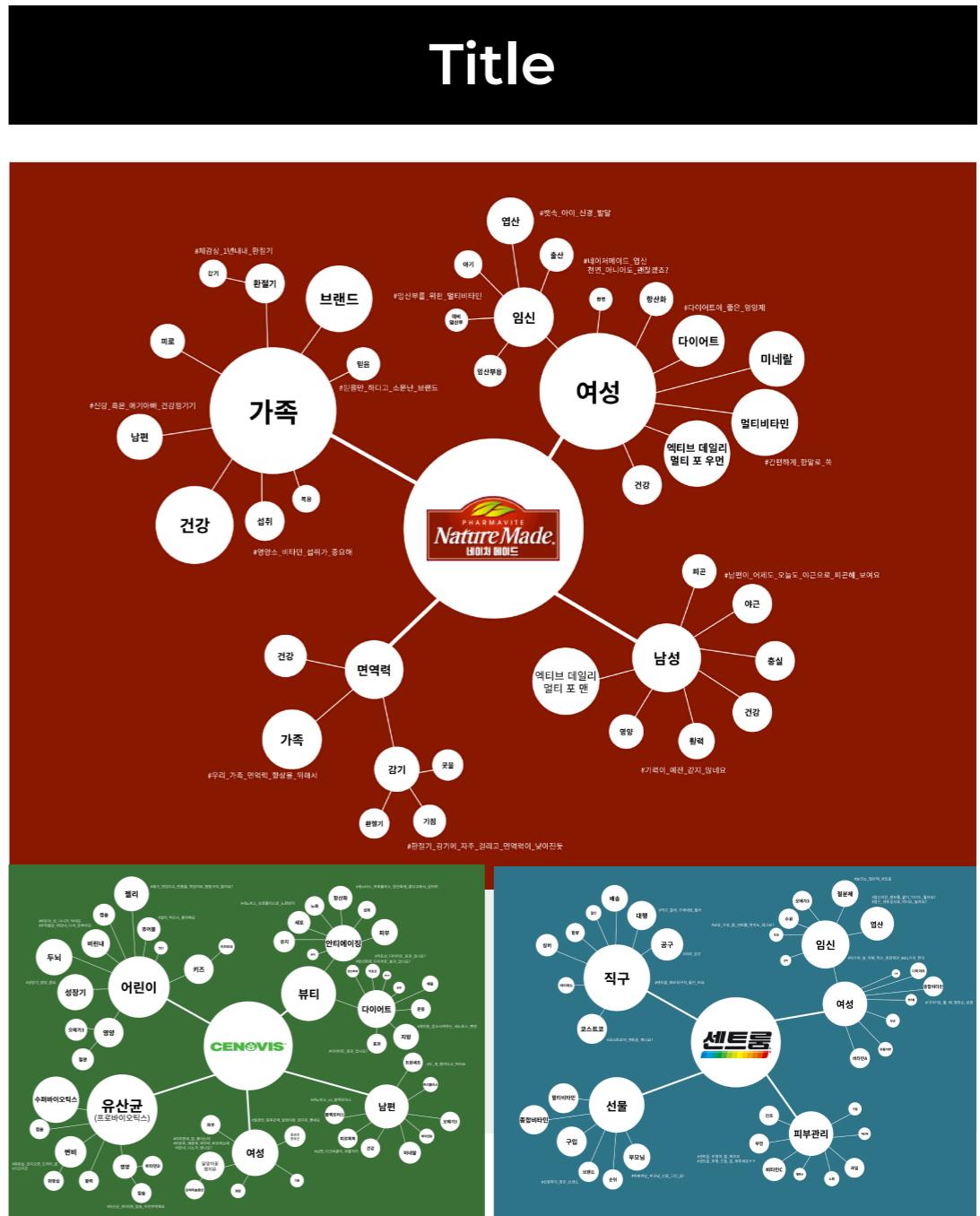
# 텍스트 마이닝 실무 적용사례

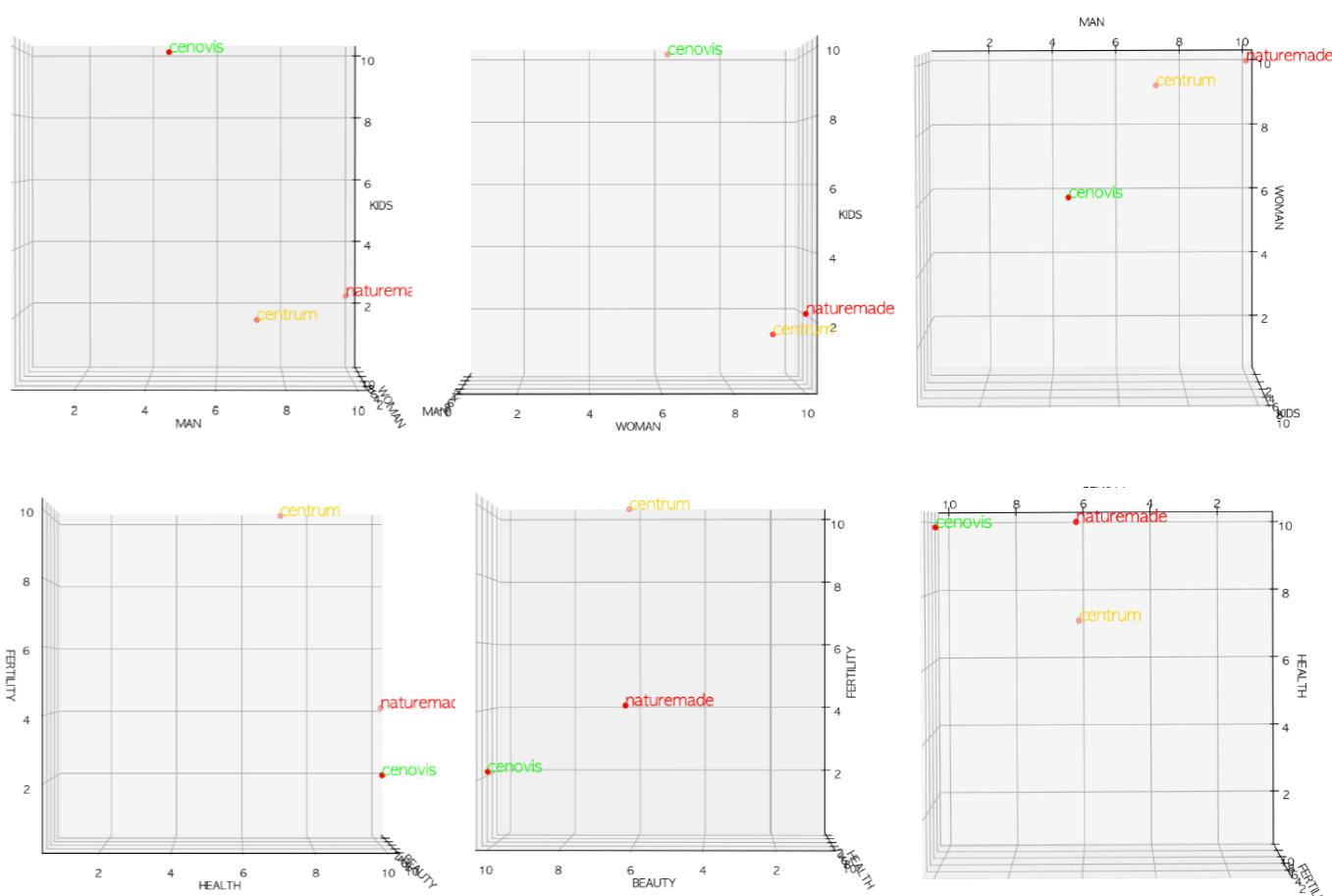
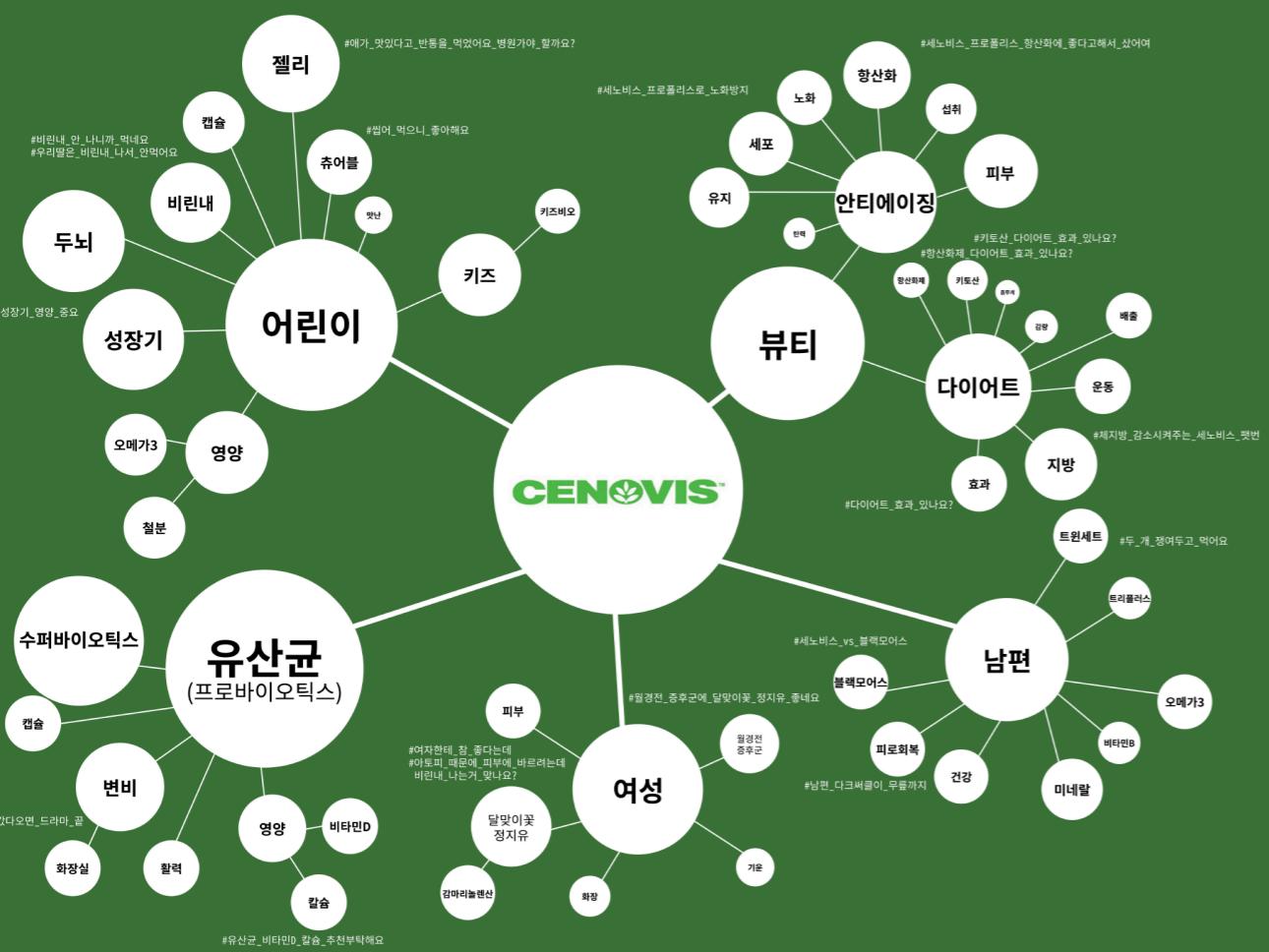
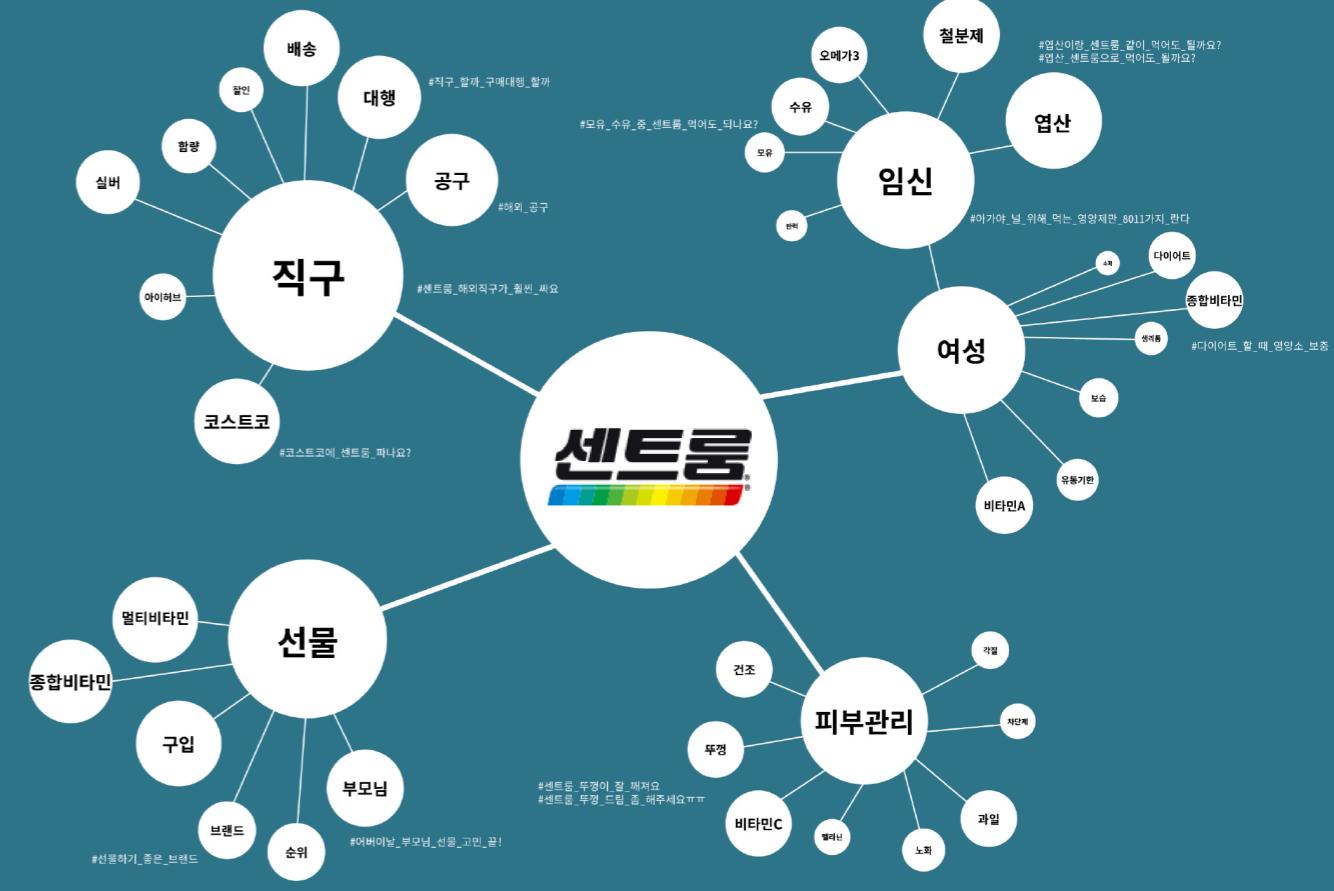
## <네이처메이드> 브랜드 포지션 전략 수립

- ▶ 브랜드 포지셔닝을 위한 경쟁사 고객 관심사 파악
- ▶ 경쟁 영양제 업체(네이처메이드, 센트룸, 세노비스)와의 차별성을 파악하고 제품 포지셔닝을 위한 온라인 고객 데이터 분석
- ▶ 소셜미디어 텍스트 분석결과를 기반으로 브랜드 포지셔닝에 근거 마련

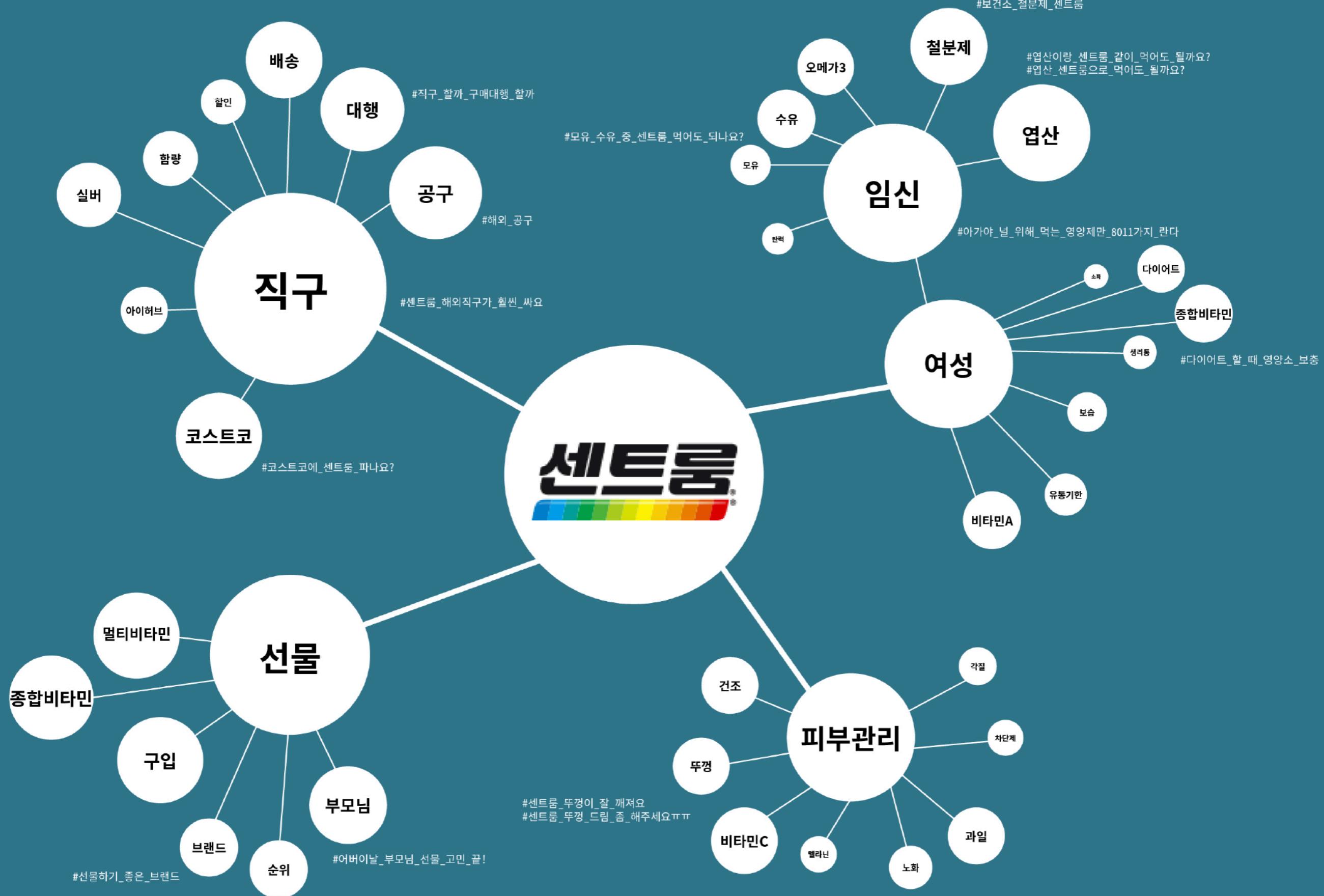
## 분석내용

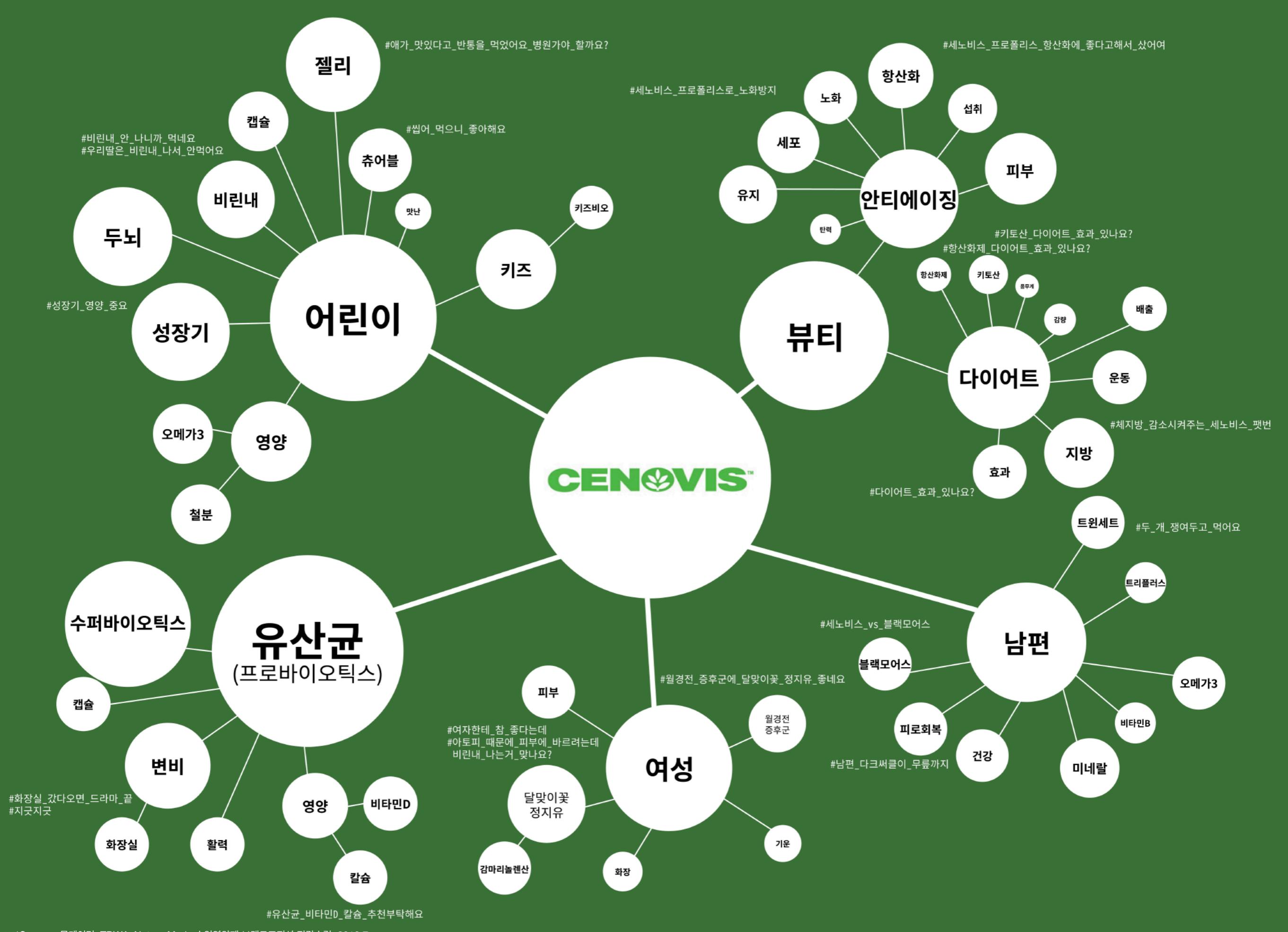
- ▶ Twitter, 블로그, 카페 등 총 14만여건의 텍스트 데이터 수집
- ▶ 수집소스의 인구통계학적 특성을 활용해 수입 영양제 주 고객층인 30-50대 육아맘들을 대상으로 타겟팅
- ▶ 클러스터링, 연관어 분석, 키워드 네트워크 분석 수행
- ▶ 소셜미디어 기반 브랜드 포지셔닝 맵 도출

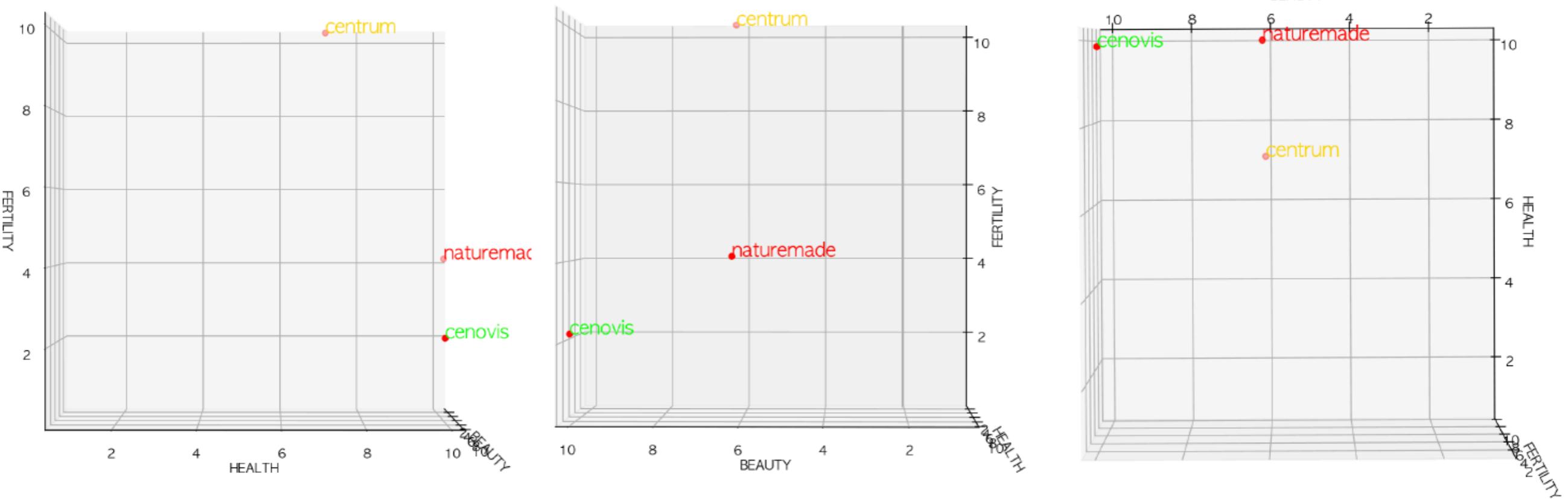
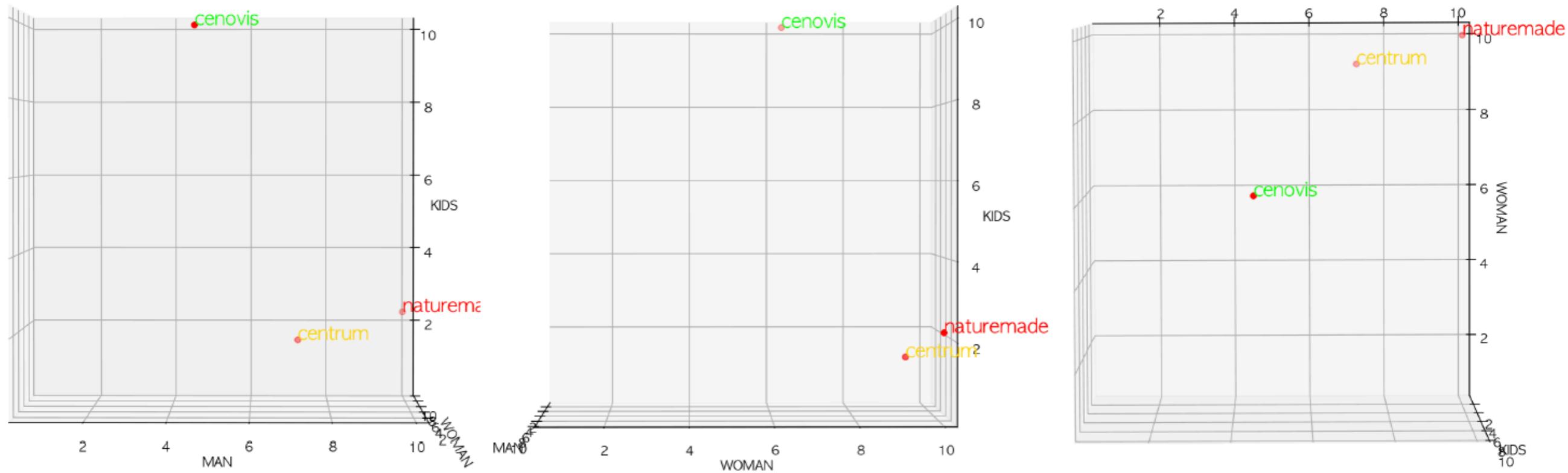












# 실무 적용사례

Case 2:

COSMOPOLITAN x SK2

2030 여성의 속마음

# 텍스트 마이닝 실무 적용사례

## <내나이가 어때서 - 서른의 속마음>

- ▶ 2017년 SK-II와 패션잡지 COSMOPOLITAN이 공동으로 진행한 “내나이가 어때서 - 서른의 속마음” 프로젝트
- ▶ 20-30대 여성들의 솔직한 속마음과 고민을 알아보기 위한 소셜미디어 텍스트 분석
- ▶ COSMOPOLITAN 2017년 7월호 게재

## 분석내용

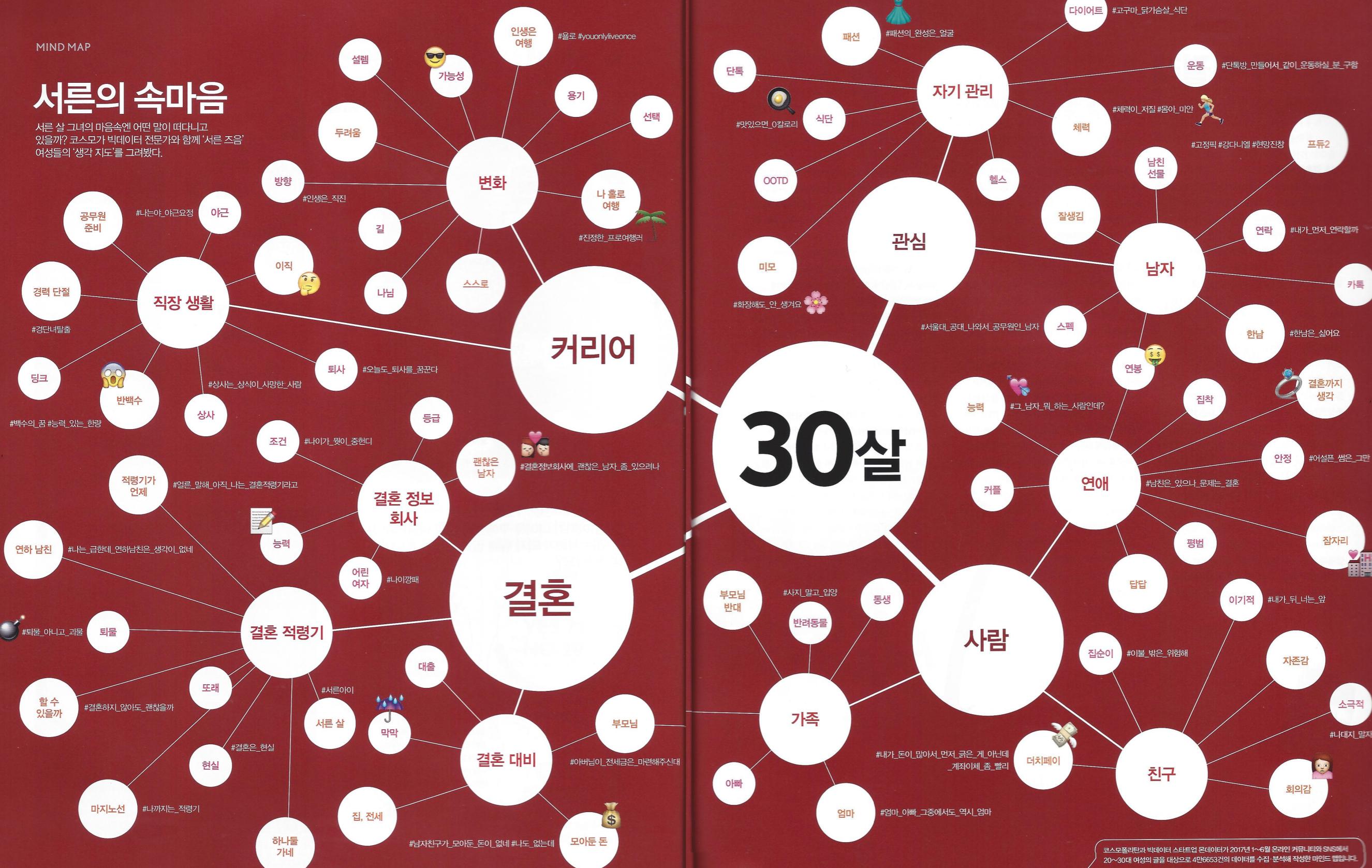
- ▶ Twitter, Instagram, 카페, 온라인 커뮤니티 등 총 4만여건의 20-30대 여성들이 작성한 텍스트 데이터 수집
- ▶ 수집소스의 인구통계학적 특성을 활용해 20-30대 여성 타겟팅
- ▶ 클러스터링, 연관어 분석, 키워드 네트워크 분석 수행



## MIND MAP

# 서른의 속마음

서른 살 그녀의 마음속에 어떤 말이 떠다니고 있을까? 코스모가 빅데이터 전문가와 함께 '서른 즈음' 여성들의 '생각 지도'를 그려봤다.



코스모폴리탄과 빅데이터 스타트업 몬데이터가 2017년 1~6월 온라인 커뮤니티와 SNS에서 20~30대 여성의 글을 대상으로 4만6653건의 데이터를 수집 분석해 작성한 마인드 맵입니다.

# 실무 적용사례

Case 3:

**뜻밖의 텍스트 마이닝 - 네오플**



# 뜻밖의 텍스트마이닝 유저동향분석에서 인게임데이터까지

네오플 데이터분석팀

김대영 [emotionalcode@gmail.com](mailto:emotionalcode@gmail.com)

NEXON COMPANY

**E.O.D**