

# 빅데이터 분석의 새로운 방향 : 인공지능 리포트

다음소프트

2017. 11. 20

No part of this publication may be circulated, quoted, or reproduced for distribution outside the client organization without prior written approval.

**Daumsoft**  
MINING MINDS

# 데이터 분석 플랫폼



## 1 사회 트렌드 분석

- 정책 개발에 앞서 라이프스타일이나 사회의 트렌드에 대한 분석을 수행하고 방향을 설정하고자 한다
  - ✓ 트렌드에 대한 사람들의 생각의 변화가 어떻게 만들어졌는가?
  - ✓ 트렌드가 만들어진 시기는 언제이며 어떤 요인이 있었는가?
  - ✓ 다른 관심 주제와 비교할 때 어떤 차이가 존재하며 어떤 결과가 예측되는가?

## 2 신제품 & 마케팅 전략

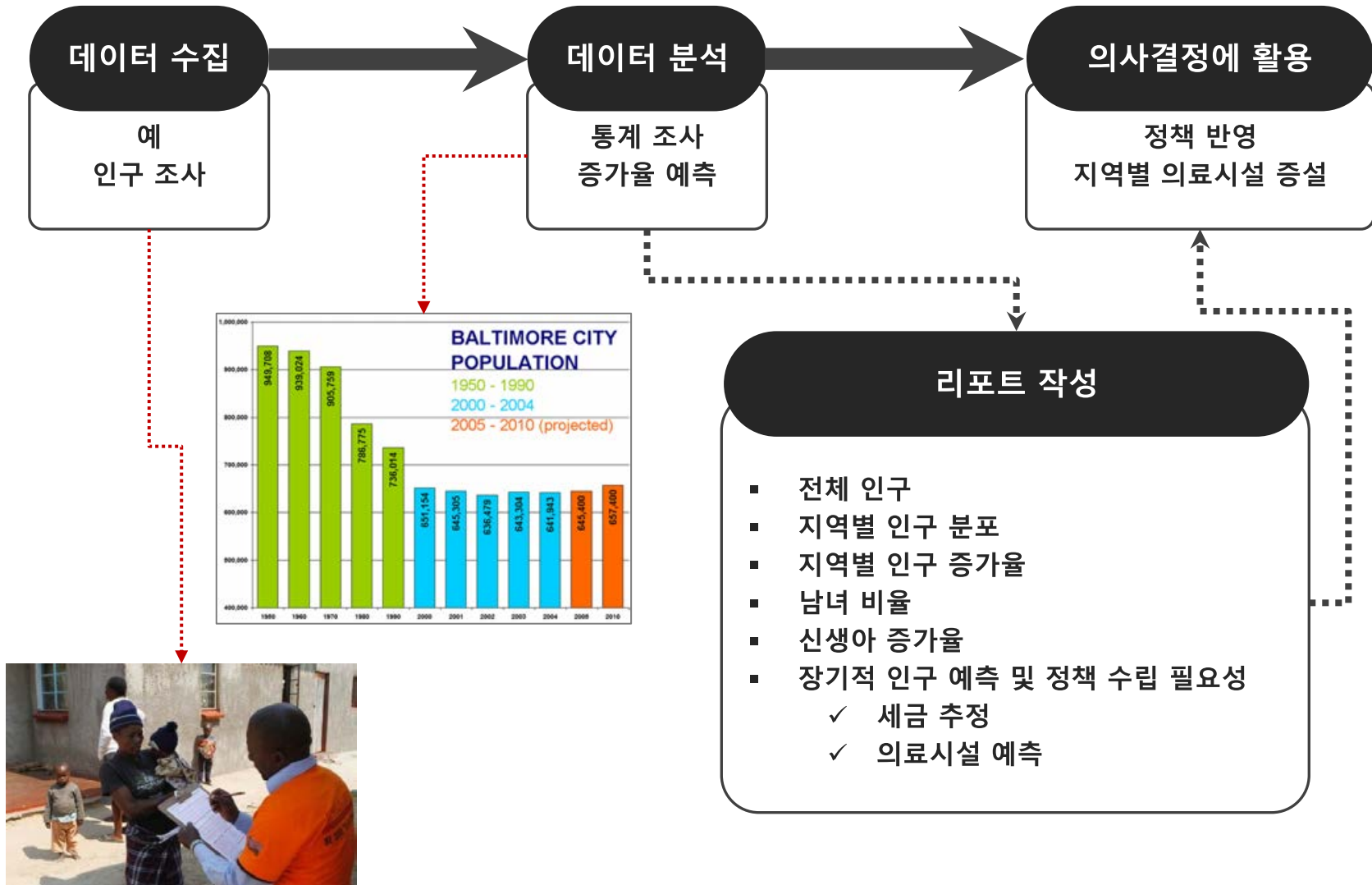
- A 전자는 기존의 아가사랑세탁기 네이밍을 변경하여, 싱글을 위한 세탁기로 리포지셔닝 하고자 한다. 싱글을 위한 제품이 되기 위한 네이밍 가이드라인은?
  - ✓ 세탁을 둘러싼 라이프 스타일 트렌드
  - ✓ 싱글의 라이프 스타일

## 3 홍보/마케팅 효과 분석

- 6개월간 진행된 C 항공사의 마케팅 활동의 결과가 어느 정도 효과가 있었는지 분석해보고자 한다.
  - ✓ 블로그, 트위터, 페이스북, 유튜브 등 다양한 채널을 통해 마케팅 활동의 효과를 정량적으로 분석해본다
  - ✓ 마케팅의 결과로 나타난 사람들의 인식의 변화는?

## 4 여론 분석

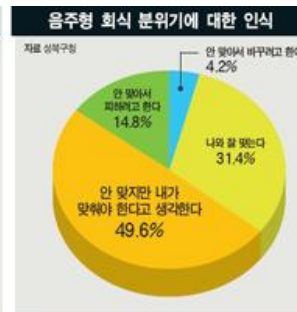
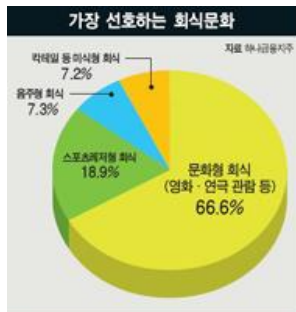
- 새 정부가 들어서고 나면 다양한 사회적/정책적 이슈가 발생하게 되는데 이와 관련된 사람들의 여론은 어떠한지 분석해보고자 한다
  - ✓ 일반 국민들은 어떤 사회/정책 이슈에 관심을 가지고 있는가?
  - ✓ 국민들이 정책 이슈에 대해 어떤 생각을 가지고 있으며 무엇이 문제인가?



- 일반적으로 데이터의 수집을 위해 전수조사나 설문 방법이 이용
- 국민이나 고객의 의견을 듣기 위해서는 샘플링을 통한 설문이 많이 이용되나 질문이 제한적이고 자발적 의견이 반영되지 않는 경우가 많음

## 설문지

- 귀하께서 현재 사용중인 휴대폰은 무엇입니까?  
 ① 011/017(SK)                      ② 016/018(KTF)                      ③ 사용안함 )  
 ④ 019(LG)                            ⑤ 010
- 귀하께서 가장 선호하는 휴대폰 브랜드는 무엇입니까?  
 ① 삼성 (애니콜)                      ② 스키야(SKY)                      ③ 기어K )  
 ④ LG                                      ⑤ 모토로라
- 귀하께서 휴대폰을 구입하신 동기는 무엇입니까?  
 ① 꼭 필요해서                      ② 주위에서 다 가지고 다니니까                      ③ 휴대폰의 가격 하락으로 인해                      ④ 선물 사온품종으로                      ⑤ 기어K )
- 귀하께서 휴대폰 구입할 때 가장 고려한 요인은 무엇입니까?  
 ① 디자인                      ② 가격                      ③ 광고                      ④ 성능                      ⑤ 인기도
- 귀하께서 구입하신 휴대폰의가격대는 어떠한가?  
 ① 10만원~20만원                      ② 20만원~30만원                      ③ 30만원~40만원                      ④ 40만원~50만원                      ⑤ 50만원 이상
- 귀하께서 가입하신 통신업체를 선택하신 이유는 무엇입니까?  
 ① 요금제 저렴해서                      ② 통화품질이 좋아서                      ③ 업체에 대한 신뢰도                      ④ 주위의 추천                      ⑤ 기어K )
- 귀하께서 휴대폰을 사용하신 기간은?  
 ① 3개월 이내                      ② 6개월 이상                      ③ 1년 이상                      ④ 2년 이상                      ⑤ 3년 이상
- 귀하께서 휴대폰을 어떤매체를 통해서 접하게 되었는가?  
 ① TV/CF                      ② 라디오                      ③ 신문 잡지                      ④ 인터넷                      ⑤ 기어K )
- 귀하께서 사용중인 휴대폰의 한달 사용료는?  
 ① 1~3만원                      ② 3~5만원                      ③ 5~7만원                      ④ 7~10만원                      ⑤ 10만원 이상



### 직장인 음주실태 (남자/기혼)

음주관련 항목	비율(%)
1회 음주시 적정음주량(3~4잔)을 초과해 마시는 비율	83.2
주 1회 이상 과음(소주 1병 또는 맥주 4병 이상) 하는 비율	45.5
소주 한병, 주 2~3회 이상 비율	33.1
1년 간 한번이라도 술이 덜 잔 상태로 출근한 경험	67.6
술 마신 다음 날 업무집중 안 된 경험	65.1
과음으로 다음 날 근무 지반한 경험	30.9
술 마신 것 때문에 직장에 손해를 입힌 경험	24.5

자료: 서울시 직장인 음주행태 및 사업장 질주환경 실태조사(2005)

비정형 데이터에는 많은 사람들의 의견이 여과없이 담겨 있으며 내용의 제한이 없으나 분석이 어렵고 대용량 분석이 필요함



소셜미디어



증권사 보고서

【국세청원후사무처리규정 제13호 제5지】

### 민원서류 위임장 (대리인 신청용)

원	표	주	주
위임하는 사람 (납세자)	장소 (납세지)	주인등록번호	-
연번	사실장	사업자등록번호	-
연번	납세지	업	행도문

위임하는 사람 (납세자) 연월일 서명 또는 (인)

【별첨인 위임하는 경우에는 별첨인강(사용인강) 가능을 날인하여야 합니다.  
 별첨인강(사용인강) 제1호에 의한 "수집·이용 동의(신청인(본인))"  
 ○ 수집·이용목적(송명합금, 사후관리 등) ○ 보유·이용기간(5년)  
 ○ 수집대상·고유성상(별첨) ○ 주인등록번호·사업자등록번호·사업장번호  
 ※ 상태변화에 대해 통지함 □ 동의하지 않음 □  
 ○ 통지를 거부할 권한이 있으며, 동의 거부에 따라 불이익(송명합금 불과)이 있을 수 있음

성명	성명	주인등록번호	-
납세자대표	또는(인)	대리인	-

민원



뉴스기사

### 피기항습심 진흙층 전주시법병원 2014.4.5.329

만결요지

- 주치의의 진단이 일반적인 의료기준에 미흡하다고 볼 수 있는 객관적인 사정들이 나타나 있다면 그 진단 사실만으로 악관예 정한 보험사고가 발생했다고 할 수 없다.

대법원 2014.06.12. 판결 2013다208661 보험금 사건

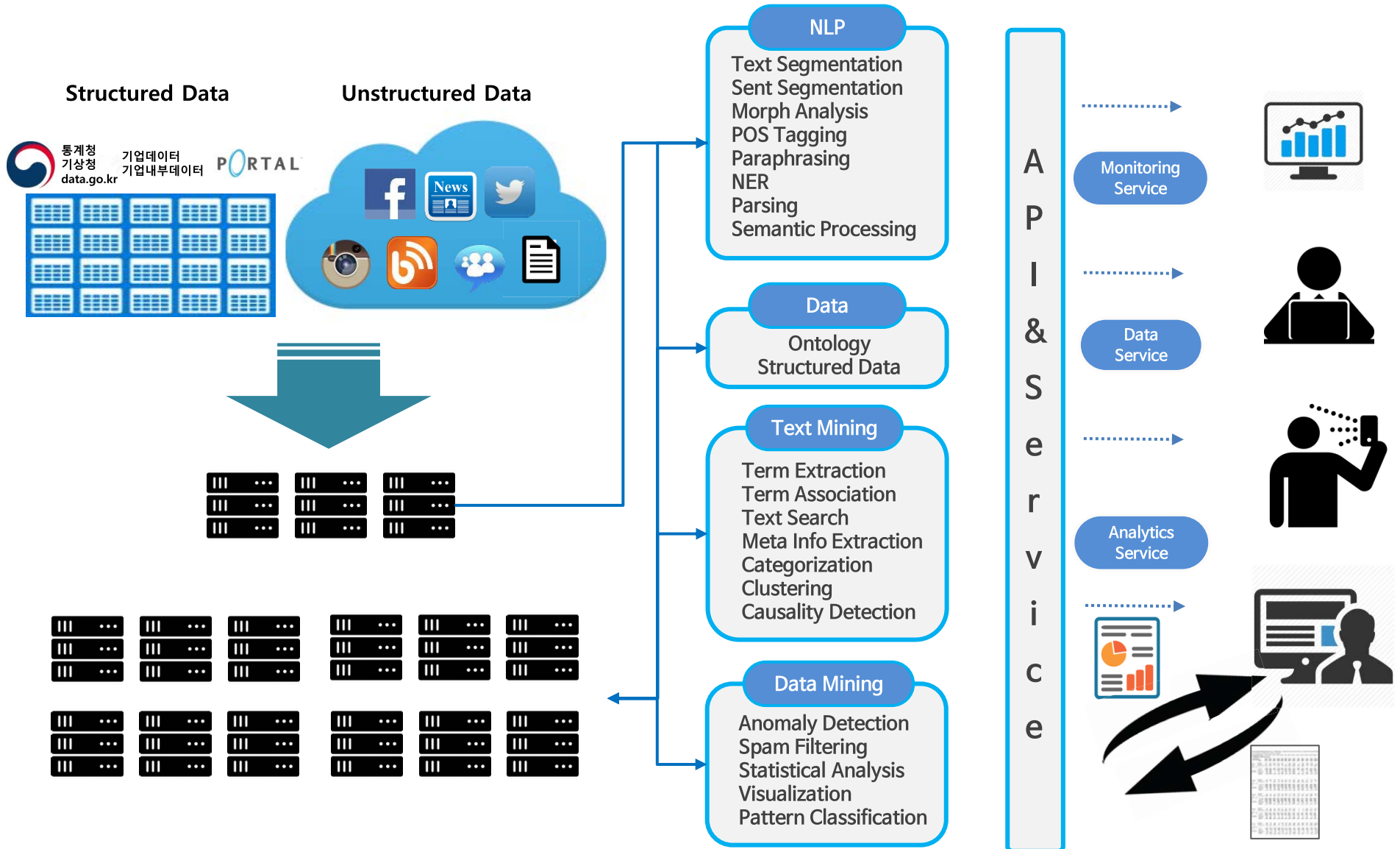
※ 현재 피기항습심 진흙층 전주시법병원 2014.4.5.329

만결내용

의사가 일정한 검사를 거쳐 상세불명의 협심증으로 진단한 경우에 이는 의사권 특별배려 규정에서 정한 보험사고의 근거자료가 될 수 있다고 할 것이지만, 그 진단의 기초가 된 병력 및 심전도 등의 객관적인 검사 결과가 충분하지 아니하거나 그러한 검사 결과 등에 기초한 진단이 일반적인 의료기준에 미흡하다고 볼 수 있는 객관적인 사정들이 나타나 있다면 그 진단 사실만으로 악관예 정한 보험사고가 발생하였다고 할 수 없다.

법, 판례

# 트렌드맵 & 소셜메트릭스



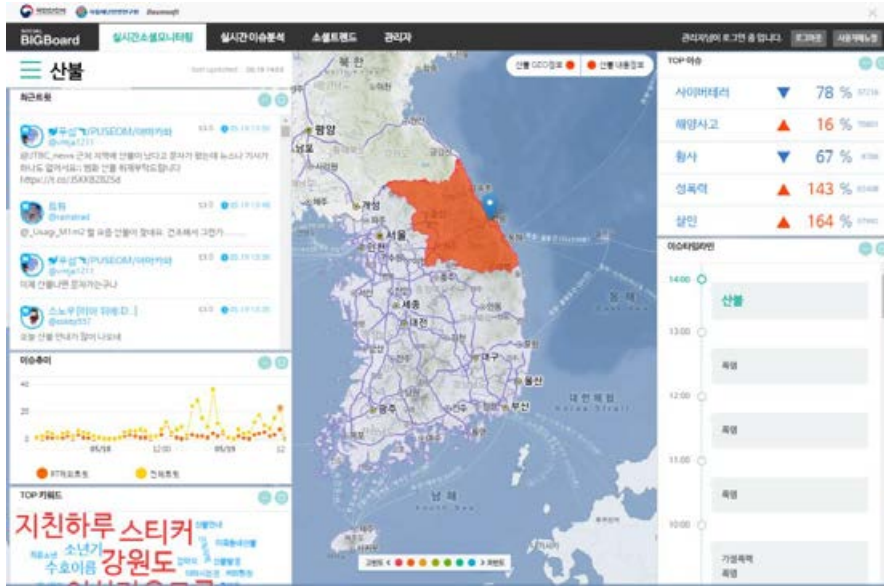
# 인공지능 리포트





# 대시보드 or 리포트?

명확한 목적하에 전문가들이 직접 활용하거나 발생 이벤트를 모니터링하기 위해서는 대시보드가 필요하나 비교/분석 등 포괄적 분석 통한 인사이트 발굴을 위해서는 리포트가 필요



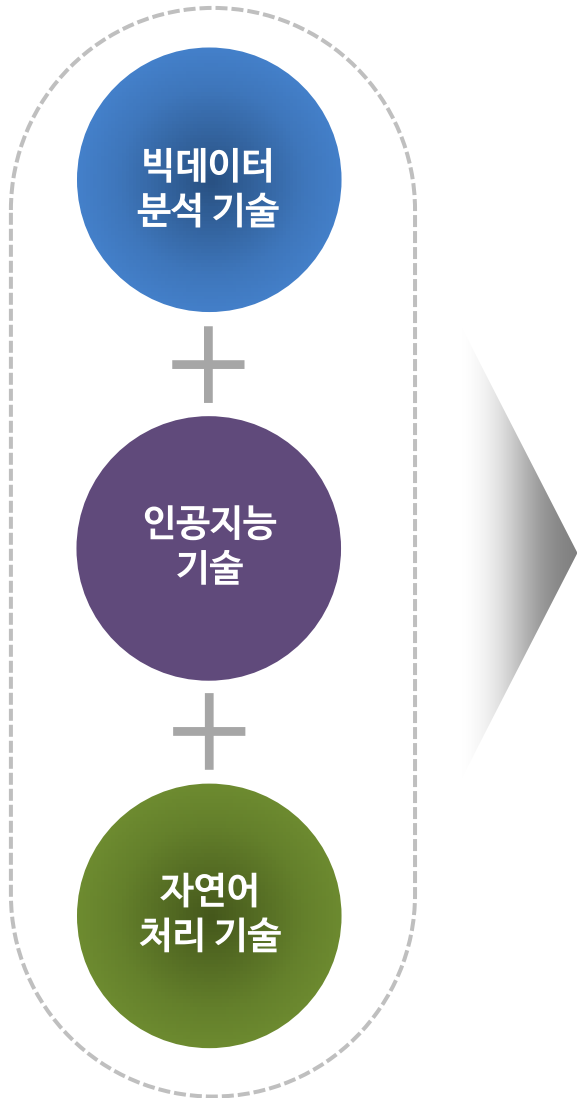
산불의 추이  
작년 동기 대비 산불 관련 메시지의 증감  
어느 지역에서 산불 관련 메시지가 늘고 있는지  
산불과 관련하여 함께 늘고 있는 재해는?  
이러한 사실들을 바탕으로 어떤 정책적 결정을 해야 하는지?



인사이트를 담은  
리포트

현재 상태에서 발생한 이벤트에 대응하는 목적으로 적합  
발생하는 현상 하나하나를 모니터링

개별 데이터보다는 비교/분석을 통해  
전체적 관점에서의 인사이트가 필요



## 인공지능 리포트

01

반복적이고 시간이 많이 걸리는 검색, 계산, 분석 작업을 대신 수행

02

전문지식이 필요한 통계 분석이나 기계학습을 이용한 패턴 인식을 자동으로 수행

03

단순한 '표'나 '그림' 뿐 아니라 사람에게 친숙한 '글'로 표현된 해석을 포함하여 이해를 도움

04

분석 결과에서 가치있는 정보만을 선별하여 제공하여, 신속하고 효율적인 인사이트 발굴을 지원

# 인공지능 리포트 작성 과정



사드에 대한 여론분석 보고서를 작성해줘



1. '사드'에 대한 사람들의 관심의 변화



사람들의 이슈 탐지 패턴 학습

이슈와 관련된 문서 검색 및 요약

'사드'와 관련된 이슈가 크게 발생한 날짜는 2016년 3월이며 '경북 성주군', '배치', '결정' 등이 주요 이유로 나타났습니다. 경북 성주군 성산리 일대가 고고도 미사일 방어체계(사드·THAAD) 배치 지역으로 사실상 결정됐다는 소식이 전해진 12일, 이 지역 민심은 벌집 쑀신 듯 들쭉였다.

'사드'에 대한 여론분석 보고서

여론 분석 보고서에 대한 분석 요소 결정

1. '사드'에 대한 사람들의 관심의 변화

2. '사드'의 여론과 가장 비슷한 패턴을 보였던 이슈 검색 및 향후 여론 변화 예측

3. '사드'와 관련해서 가장 두드러지게 나타나는 연관어를 바탕으로 연관 문서 정리 및 요약

## 인공지능 기반기술



## 01

### 인공지능 기반 자동 리포트 작성 시스템

- 분석 유형 및 요소들에 대한 **능동적** 결정
- **기계 학습**을 이용한 패턴 인식 및 분류
- 통계 분석을 이용한 **가설 수립 및 검증**
- **언어생성 및 로봇 저널리즘** 기술을 기반으로 모든 분석 결과를 인간에게 익숙한 '글'과 '그림'로 표현
- **인과 관계** (Causal Relationship) 분석 및 **추론** (Inference)

## 02

### 빅데이터 활용

- 소셜 빅데이터, 기업 내부 데이터 등 빅데이터 분석 플랫폼 활용
- 통계청, 기상청 등 공공데이터 및 매출 등 내부 정형데이터 이용 가능
- 빅데이터에 대한 **반복적 검색과 연산**

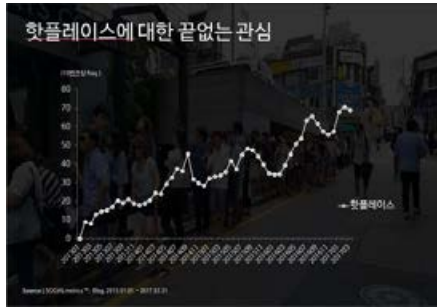
## 03

### 사람과 기계의 협업

- 리포팅 로봇은 빅데이터에 대한 검색 및 반복적 분석 실행
- 또한 많은 일반 리포트 작성을 위한 지식 보유
- **인간의 직관과 통찰**을 이용해 수정 및 가치가 더해진 포괄적 리포트 작성 가능

# 리포트 생성 방법

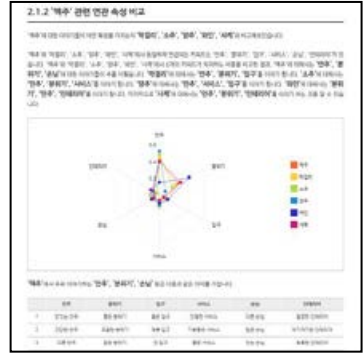
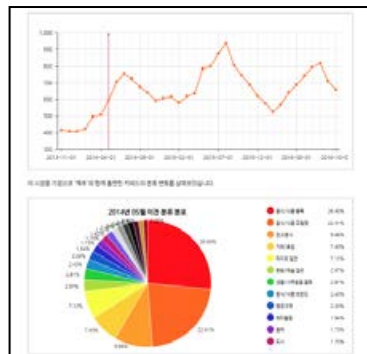
## Report by Human Expert



## Report Assistant by AI



## Auto Report by AI



# 인공지능 리포트 전문비서

저장

분석

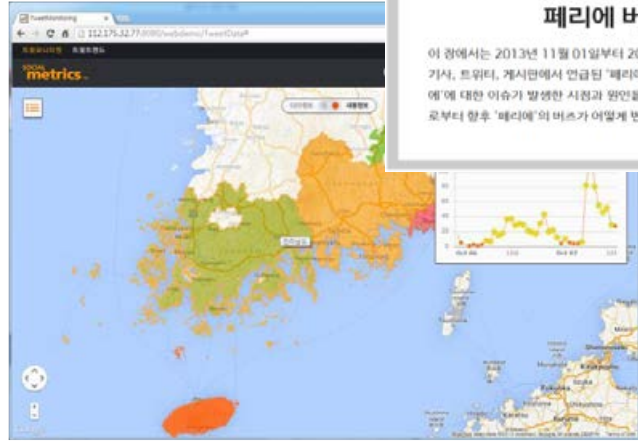
서비스

블로그   트위터   커뮤니티   인스타그램   뉴스   오픈데이터   정형데이터



**페리에 버즈 추이와 이슈**

이 장에서는 2013년 11월 01일부터 2016년 10월 31일까지 3년의 기간 동안 블로그, 신문 기사, 트위터, 게시판에서 언급된 '페리에'의 버즈 추이를 살펴봅니다. 또한 3년 사이에 '페리에'에 대한 이슈가 발생한 시점과 원인을 분석합니다. 그리고 3년 동안 '페리에'의 버즈 추이로부터 향후 '페리에'의 버즈가 어떻게 변화할 것인지를 예측합니다.





**DEMO**



# End of Document

**Daumsoft**  
MINING MINDS

No part of this publication may be circulated, quoted, or reproduced for distribution outside the client organization without prior written approval.