



내 삶을 바꾸는 공공 빅데이터!

빅데이터 사례집



행정안전부

NIA 한국정보화진흥원

내 삶을 바꾸는 공공 빅데이터!

빅데이터 사례집

1. 누구나 안심할 수 있는, “안전한 사회”

- 1-1. 반드시 확보해야 할 골든타임, 하지만 농촌지역에서는...
- 1-2. 어둠 속 든든한 지킴이, CCTV가 있어 안심!
- 1-3. 범인 검거, 또 다른 수사의 시작!

2. 더 편리하고 건강해지는, “국민의 생활”

- 2-1. 예방해요! 유행성 질병
- 2-2. 출근길, 만원버스에 발만 동동
- 2-3. 사고 싶은 전기차, 하지만 충전은 어디서 하지?

3. 함께하면 모두가 행복해지는, “경제 활성화”

- 3-1. 사각지대 속에 방치된 이웃들의 삶
- 3-2. 한옥마을, 관광객으로 인산인해
- 3-3. 축제의 계절, 주말이 즐겁네!
- 3-4. 지난해울 이상 기온, 올해 양파 가격은 얼마?

4. 더 정확하고 신속해지는, “스마트 행정”

- 4-1. 열심히 일한 당신! 당신의 임금은 안녕하십니까?
- 4-2. 공동주택 관리비 지출내역 확인할 길 없어... 속 터지는 입주민들
- 4-3. 도로 위에는 언제 터질지 모르는 ‘지뢰’가 있다
- 4-4. 귀 기울여 듣는 시민의 목소리

1

누구나 안심할 수 있는, “안전한 사회”

- 1-1. 반드시 확보해야 할 골든타임, 하지만 농촌지역에서는…
골든타임 확보를 위한 구급차 배치·운영 최적화 모델(전라북도)
- 1-2. 어둠 속 든든한 지킴이, CCTV가 있어 안심!
CCTV 설치지역 분석 및 모델 고도화(경기도)
- 1-3. 범인 검거, 또 다른 수사의 시작!
인공지능을 활용한 과학적 민생치안강화
(임장일지 기반 유사사건 여죄분석)(경찰청·국가정보자원관리원)

1-1

반드시 확보해야 할 골든타임, 하지만 농촌지역에서는...

골든타임 확보를 위한 구급차 배치·운영 최적화 모델(전라북도)

P씨는 갑자기 추워진 날씨로 노인 화관에 다녀오는 길에 심근경색으로 인해 차가운 길 위에 쓰러지고 말았다. 119에 구조요청을 했지만 구급차는 골든타임인 5분을 넘기고서야 도착했다. 다행히 위급한 상황은 넘겼지만 앞으로 발생할 수 있는 위급 상황의 골든타임 내 신속한 대처를 위해 빅데이터 분석을 활용했다.

5분 이내 현장으로, 골든타임을 확보하라!

전북의 경우 5분 내 출동률이 54%에 불과했다. 도심지역은 출동에 평균 3km를 이동하는데, 농촌지역은 10km를 달려야 하기 때문이다. 또한 도심지역은 교통체증과 불법주차 등의 문제로 시간이 많이 소모되었다. 이에 전북소방본부는 '골든타임 확보를 위한 구급차 배치운영 최적화 모델'을 진행했다. 이 프로젝트는 대전시 '스마트 재난안전 상황관리체계'에서 시작된 것으로, 위급 상황 발생 시의 신속한 관리를 지원한다.

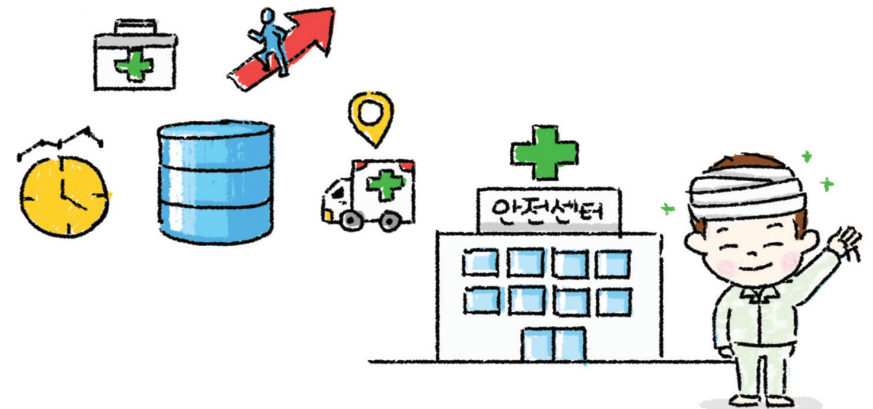


가장 적합한 시간과 장소를 빅데이터를 이용해 도출!

빅데이터를 통해 도심은 시간대에 따른 이동 속도 편차를 정확히 수치화할 수 있었고 농촌은 소방서와 구급차를 반드시 배치해야 하는 취약지역을 과학적으로 선정할 수 있었다. 뿐만 아니라 내장산처럼 계절에 따라 유동인구가 급증하는 곳에는 얼마나 많은 수의 임시 구급대를 대기시켜야 할지 계산할 수 있게 되었으며 응급환자를 유형별로 분석해 지역에 맞는 구급특수장비를 준비할 수 있는 여건이 마련되었다.

빅데이터로 지키는 소중한 인명과 재산

전북소방본부는 소방활동 빅데이터를 여러 분야에 활용하고 있다. 지역과 시간대에 맞는 구급활동 대비 태세를 갖추는 한편, 한정된 자원을 응급환자, 소방, 화재관련 시설에 적절히 분배하는 것이다. 이 덕분에 응급환자의 소생률이 향상되었고 앞으로 설치될 신규 안전센터의 위치를 과학적으로 선정하여 주민들의 안전에 큰 기여를 할 것으로 기대된다. 게다가 누적 데이터를 통해 실습 매뉴얼 배포 자산을 확보했고, 데이터 분석 프로세스 표준화를 통해 표준분석모델을 전국으로 확대할 수 있는 기반을 마련할 계획이다.



1-2

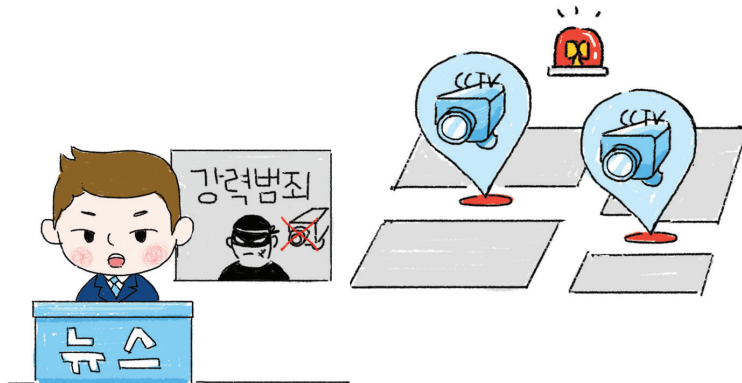
어둠 속 든든한 지킴이, CCTV가 있어 안심!

CCTV 설치지역 분석 및 모델 고도화(경기도)

K양은 여성 대상의 강력범죄 뉴스를 자주 접하면서 퇴근길 어두운 골목길이 더 무서워졌다. 골목길을 들어설 때마다 강력범죄 기사가 생각난다. 더구나 CCTV가 없던 골목길에서 수상한 사람을 만나기도 한 터라 불안감은 더 커졌다. 다른 곳으로 이사를 가지 않아도 K양의 퇴근길 안전을 지킬 수 있는 방법이 있을까?

CCTV 신규 설치 급증? 하지만 범죄예방 효과 미비!

최근 CCTV 신규 설치가 급증하는 상황 속에 현재의 CCTV는 주먹구구식 설치, 낮은 화소, 사각지대 존재, 범죄예방 효과 미비라는 문제점이 존재했다. 이에 대한 해결책으로, 경기도는 CCTV 설치지역 분석 및 모델 고도화를 통해 공공 빅데이터를 수집·통합한 후 데이터를 분석하여 우선설치지역 및 운영관리지표 등 객관적인 기준을 제공하였다.

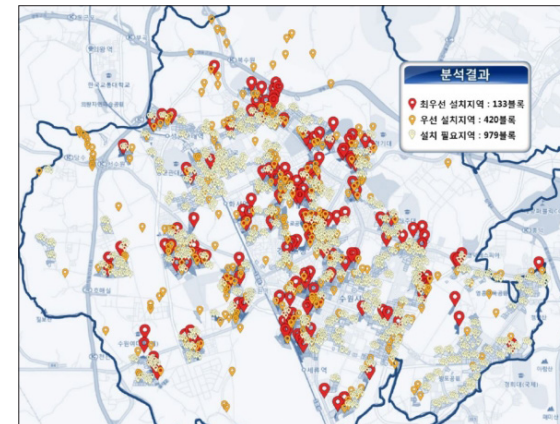


우리지역에 CCTV가 필요한 곳은?

경기도는 지난 2014년 범죄 데이터를 연계한 빅데이터 분석을 통해 수원시의 CCTV 사각지대와 우선설치 지역을 선정했다. 또한 경기 지방 경찰청과의 업무양해각서 체결을 통해 범죄 발생현황 정보를 접목하여 범죄 다발지역에 대한 위험지수와 예측지수를 개발했다. 여기에 범죄 발생에 영향을 미치는 유동인구 분포지수와 기존 CCTV 현황을 바탕으로 CCTV 우선 설치 지역을 도출하여 방법효과를 극대화하였다.

CCTV 설치로 안전해진 우리동네

수원시는 979개의 CCTV 우선 설치 지역 블록을 도출하고 그 중 CCTV 최우선 설치 지역 133개 블록, 우선 설치 지역 420개 블록을 선정하였다. 경기도는 이를 토대로 수원시 126곳에 CCTV를 설치했으며 계속 늘려갈 예정이다. 이에 따라 절도, 강도 등 범죄 발생률이 낮아질 것이며, CCTV 관련 민원 또한 낮아질 것으로 예상하고 있다. 경기도는 분석 모델을 표준화하고 확대하여 31개 시·군으로 확대 적용할 계획이다. 경기도의 보안 취약지역이 개선되고 범죄율이 감소되어 시민이 안전한 생활을 할 수 있기를 기대해 본다.



CCTV 우선
설치지역 분석결과

1-3

범인 검거, 또 다른 수사의 시작!

인공지능을 활용한 과학적 민생치안강화(임장일지 기반 유사사건 여죄분석)
(경찰청·국가정보자원관리원)

빈집털이범을 현장에서 붙잡은 C경사는 범인의 여죄에 대한 의심이 들었다. 자물쇠를 열고 들어가 귀중품을 훔친 범인은 분명히 여러 차례의 범행을 저질렀음이 틀림없었기 때문이다. C경사의 심증을 확신으로 바꿔줄 증거를 찾기 위해 경찰청이 빅데이터를 활용하여 민생치안강화 프로젝트를 추진하였다.

수많은 구슬을 어떻게 꿰어야 할 것인가

사건 현장을 기록한 임장일지는 사건 해결을 위한 기초 자료인데, 그동안 누적된 임장일지가 너무 많아 활용에 적잖은 시간이 걸린다는 문제가 있었다. 특히 미제사건으로 남은 범죄의 경우, 피의자의 여죄를 입증하기 위해 임장일지를 들춰보는 데에 상당한 시간과 인력이 필요했다. 행정안전부와 경찰청은 이에 “인공지능을 활용한 과학적 민생치안강화 프로젝트(임장일지 기반 유사사건 여죄분석)”를 수행했다.

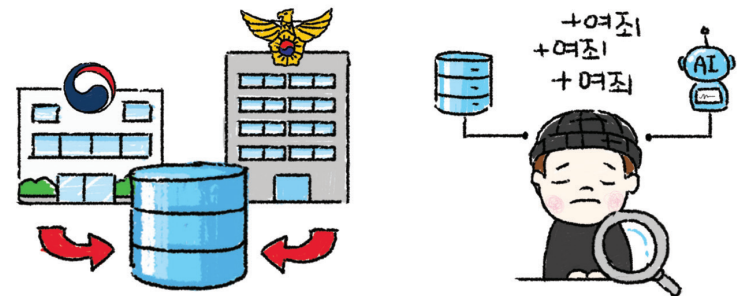


숨어 있는 범죄는 데이터로 찾아낸다

국가정보자원관리원과 경찰청은 데이터와 분류·분석기법을 통해 협업을 시작했다. 국가정보자원관리원은 경찰청이 제공한 최근 2년의 임장일지 데이터를 활용해 최적의 여죄 추적 모델을 구현했다. 그리고 기존의 알고리즘(TF-IDF)뿐 아니라 구글 최신기술(Doc2Vec) 등 4개의 알고리즘을 적용시켜 테스트를 시행했다. 또한 약 1만 건의 동의어와 약 7백 건의 불용어 사전을 자체 개발해 최적화에 성공했다.

신속한 추가 수사로 강화되는 민생치안

임장일지를 바탕으로 구현된 분석 모델은, 부산지방경찰청의 범죄 피의자 여죄추적에 활용되어 3건의 추가 여죄 입증에 성공했다. 부산 뿐 아니라 타 지역의 침입·절도사건 등 6건의 유사사건에서도 활용되는 중이다. 경찰청은 앞으로도 빅데이터 및 인공지능을 적극적으로 활용하는 한편, 더 많은 유관기관과 협력해 의미 있는 데이터를 창출하고 그 데이터를 분류·분석하기 위해 다양한 분석도구를 이용하여 딥러닝을 통한 인공지능 고도화도 진행할 계획이다.



2

더 편리하고 건강해지는, “국민의 생활”

2-1. 예방해요! 유행성 질병

국민건강알람서비스(건강보험공단)

2-2. 출근길, 만원버스에 발만 동동

빅데이터 분석을 통한 시내버스 효율적 운영(광주광역시)

2-3. 사고 싶은 전기차, 하지만 충전은 어디서 하지?

전기차 충전 인프라 설치 입지선정(대구광역시)

2-1

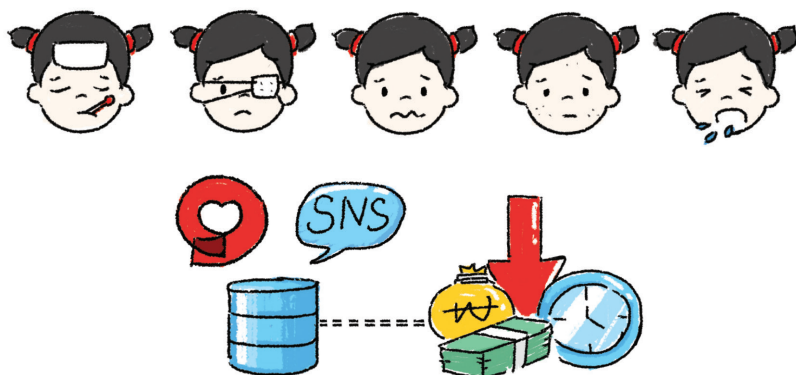
예방해요! 유행성 질병

국민건강알람서비스(건강보험공단)

평소 잔병치레가 심한 초등학교 딸을 둔 엄마 L씨는 이번 환절기도 걱정이다. 딸아이가 감기에 걸리면 며칠씩 병원을 다녀야 하기 때문이다. 아이에게 항상 외출 후 집에 오면 손을 씻으라고 말하지만, 아이가 커갈수록 엄마 말을 듣질 않는다. 아이가 스스로 관심을 갖고 감기를 예방할 수 있도록 할 방법이 있을까?

빅데이터로 질병을 예측하라

건강보험공단은 빅데이터 분석을 기반으로 2013년부터 '국민건강알람서비스'(<http://forecast.nhis.or.kr>)를 통해 5개 질병(감기, 눈병, 식중독, 천식, 피부염)에 대해서 4단계(위험, 경고, 주의, 관심)로 위험 정도를 제공하고 있다. 건강보험공단 국민건강정보DB와 식약처·기상청·환경부의 식중독·기상·환경 자료를 연계하고 소셜정보를 융합하여 진료 동향과 질병 관련 트위터·블로그 등의 정보를 확인할 수 있도록 하였다.

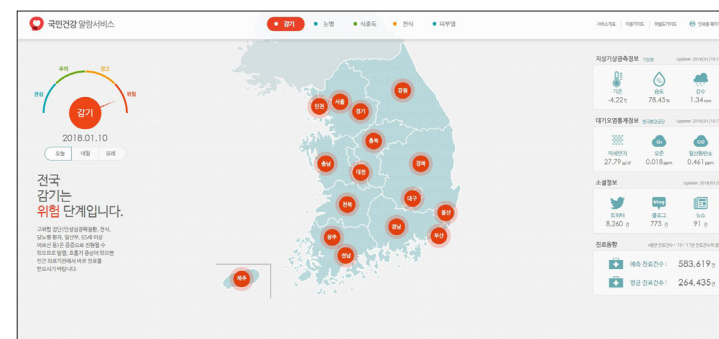


SNS로 보는 건강 위험 신호

'국민건강알람서비스'는 건강보험공단의 진료정보와 소셜정보를 융합하여 질병발생 예측 모델을 구축함으로써 국민 의료비 절감과 건강 증진을 기대하고 있다. 또한, 감기가 유행한다고 하면, SNS상에 감기에 대한 이야기가 많이 있을 것이라는 전제를 바탕으로 SNS 등 비정형데이터를 활용했다. SNS 속 텍스트의 빈도와 과거 진료 건수를 확인한 결과 날짜에 따라 등락폭이 비슷한 그래프를 그렸다. 이것은 소셜정보가 진료 동향과 융합해서 사용될 수 있음을 검증한 것이다.

믿을 가는 국민건강알람서비스

앞으로 건강보험공단은 기상 자료, 대기오염 자료 등의 환경측정 자료와 뉴스미디어 등의 데이터 수집 등 채널을 다양화하여 알람의 정확도와 신뢰도를 높일 계획이다. 아울러 알람 대상 질병을 확대해 나감으로써 국민건강 지킴이가 되겠다는 각오를 다지고 있다.



국민건강알람서비스

2-2

출근길, 만원버스에 발만 동동

빅데이터 분석을 통한 시내버스 효율적 운영(광주광역시)

광주시에 거주하는 직장인 P씨는 아침마다 벌어지는 버스 쟁탈전이 두렵다. 승객들이 몰려들어 버스에 타지 못하는 상황이 매일 발생하기 때문이다. 문제를 해결하기 위해 광주시는 이용객이 많은 시간대의 배차를 늘리는 등 버스를 탄력적으로 운영하여 주민의 삶의 질을 높이기 위해 빅데이터 분석을 활용했다.

교통 약자를 위한 작은 배려의 시작!

행정안전부는 지난 2014년 광주시와 함께 ‘광주지역 시내버스 효율적 운영방안 마련’ 사업을 진행했다. 빅데이터 분석을 통해 불편사항을 없애고 시민의 편의성을 제고하여 편리하게 시내버스를 이용할 수 있도록 한 것이다. 시는 교통카드 이력, 주거·유동인구 데이터 등을 활용해 시민의 거주 위치와 버스, 지하철 정류장의 위치데이터를 분석하였으며, 시내버스의 효율적 운영을 위한 과학적인 분석 결과를 도출하였다.

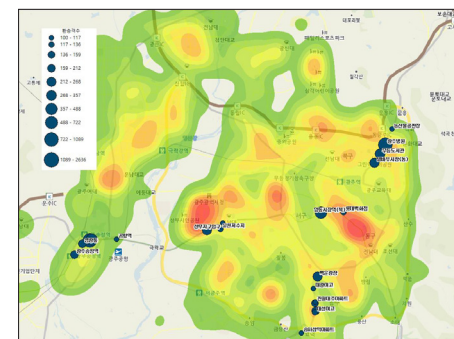


빅데이터 분석으로 시민의 발, 버스를 더욱 편리하게

광주시는 전 지역을 작게 나눠 버스 정류장 수가 부족한 지역 중 10개 지역을 대중교통 사각지대로 선정하여 교통약자가 많이 이용하는 편의시설 위치를 분석해 16개 버스 노선을 선정했다. 또 분석 결과에 따라 교통약자가 많이 이용하는 노선에 저상버스 40대를 추가 도입했다. 이와 함께 교통카드 데이터를 분석해 출퇴근 시간대 이용자가 많은 노선에 배차간격을 조정해 시민들의 불편을 대폭 줄일 수 있었다.

데이터에 기반 한 시내버스 수요 분석

한편 행정안전부는 전주시를 대상으로 2015년 교통 분야 빅데이터 분석 사업을 추가로 진행했다. 시내버스 정류장·노선 및 유동인구 데이터의 특성에 맞추어 지도를 이용한 시각화 결과를 활용하여, 신규 노선 개발과 기존 노선 조정, 탄력배차제 도입을 위한 발판을 마련하고, 관광객을 위한 노선도 개발했다. 또한, 실제 시내버스 탑승 데이터를 바탕으로 수혜를 받는 고객 수 등을 집계할 수 있는 시뮬레이션을 개발하여 환승 할인 정책 변경 시 근거자료로 활용할 수 있게 되었다. 이를 통해 대중교통 이용의 편의성이 제고되어 시민 만족도가 높아졌다. 더불어, 지자체의 관광 활성화에도 기여하여 지역경제도 살아나게 되었다.



교통 취약 지역 분석

2-3

사고 싶은 전기차, 하지만 충전은 어디서 하지?

전기차 충전 인프라 설치 입지선정(대구광역시)

오래된 자동차를 바꿀 때가 된 K씨는 요즘 전기차에 대한 관심이 부쩍 높아졌다. 하지만 일반 주유소에 비해 현저히 적은 수의 충전소 때문에 K씨는 전기차를 구입하겠다는 결정을 선뜻 내리지 못하고 있다. 이를 위해 대구시에서는 빅데이터 기반으로 최적의 위치를 찾아 전기차 충전 인프라의 접근성을 높였다.

쉽게 접근할 수 있는 최적의 위치를 찾아라

대구광역시에서는 환경보존과 차세대 성장 산업 육성이라는 두 마리 토끼를 잡기 위해 전기차 보급을 위한 현황 조사를 시작했다. 현재의 전기차는 한 번 충전에 평균 150~250km를 주행하여 일반 차량보다 충전소를 찾는 일이 더 잦을 수밖에 없다. 그러나 현재 충전소의 숫자는 전기차의 보급 속도를 따라가지 못하고 있다. 충전소의 절대적인 숫자가 적은 것도 문제지만, 2017년 4월~8월의 조사에 따르면, 총 93개 충전소 중 39개 충전소는 사용되지 않고 있다는 사실이 드러났다. 입지선정이 잘못되었다는 뜻이다.

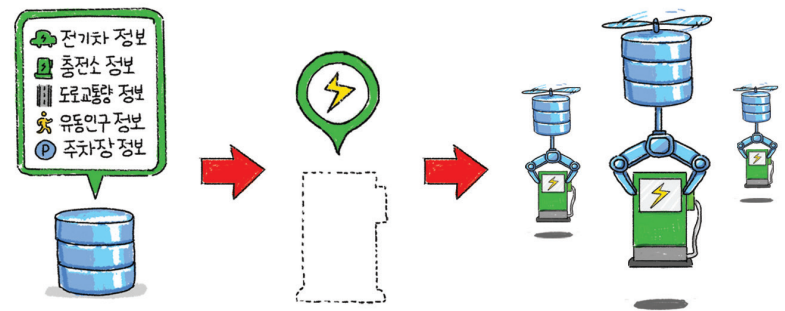


멈추지 않고 달릴 수 있는 환경, 데이터 기반으로 구축하다

대구광역시는 우선 데이터를 분석해 이를 모델화하였다. 대구광역시시의 현황과 전기차 인프라에 대한 점검 역시 철저히 진행했으며 이를 선진 사례와 비교 분석하였다. 전기차 관련 담당자의 요구사항 역시 정확하게 반영하여 과학적 입지선정을 통한 업무 효율화를 목표로 했다. 시간에 따라 변화하는 데이터나 가설적 실험 등 통계적 예측에 이용되는 회귀분석을 통해 전기차 사용량을 예측하는 한편 주변 환경 요인 분석을 통해 충전소 운영에 이를 적용하기로 했다.

데이터가 만들어내는 깨끗한 미래

대구광역시는 과학적 입지선정과 정확한 수요예측을 바탕으로 충전소의 최종 입지선정모델을 시각화했다. 지도를 셀 단위로 쪼개 지정된 곳의 충전소로서의 입지 조건이 어느 정도인지 수치화하는 데에 성공한 것이다. 대구광역시는 이 수치로 여러 차례의 검증작업을 거쳐 최적의 충전소를 선정하는 작업에 착수할 수 있었다. 이를 통해 궁극적으로는 지자체의 예산수립과 자원관리에 효율성을 담보할 수 있게 되었다.



3

함께하면 모두가 행복해지는, “경제 활성화”

3-1. 사각지대 속에 방치된 이웃들의 삶

잠재적 사회취약계층 일자리 창출 및 자립지원(남양주시, 국민연금공단)

3-2. 한옥마을, 관광객으로 인산인해

전주 한옥마을 관광분석을 통한 경제활성화(전주시)

3-3. 축제의 계절, 주말이 즐겁네!

빅데이터 활용 문화·관광 축제 분석(한국관광공사)

3-4. 지난 겨울 이상 기온, 올해 양파 가격은 얼마?

기상데이터와 농산물 생산성 예측(기상청, 농촌진흥청)

3-1

사각지대 속에 방치된 이웃들의 삶

잠재적 사회취약계층 일자리 창출 및 자립지원(남양주시, 국민연금공단)

남편을 여의고 두 명의 자녀와 단칸방에 살고 있는 Y씨의 삶은 넉넉지가 않다. Y씨는 환갑을 앞둔 데다 폐질환까지 있지만 자녀들 역시 장애를 갖고 있어 직업을 구하기 쉽지 않은 상황이다. 이에 Y씨의 경제적 자립과 지원을 위해 남양주시와 국민연금공단이 빅데이터를 기반으로 해결에 나섰다.

시야 밖의 이웃을 찾는 일, 데이터가 앞장서다

남양주시와 국민연금공단은 공공 데이터를 분석하여 사회취약계층의 취업을 선제적으로 지원하였다. 생활비를 지원하는 것보다 직업을 갖고 자립하는 편이 더 큰 사회적 효과를 내기 때문이다. 이를 위해 남양주시와 국민연금공단은 서로 다른 데이터를 “생활지역 단위의 집계 정보”로 합산·변환하여 도출된 수치를 이용해 우선 지원해야 할 취약계층을 선정했다. 도움이 필요한 우리 이웃들이, 눈앞에 나타난 것이다.



연계된 수치가 비추는 밝은 미래

남양주시가 제공한 데이터는 구직·구인신청목록, 취업자목록 등이며, 국민연금공단은 여기에 가입자와 사업장의 정보와 현황을 더했다. 이를 통해 취업현황을 분석할 수 있게 되었고 이는 정책수립의 밑바탕이 되었다. 실직현황과 실직기간의 교차분석, 실직자의 세대특성에 대한 상관분석 역시 진행되어 실직자의 현황을 파악함과 동시에 알맞은 직업훈련계획 및 취약계층 우선정책 수립에 큰 역할을 했다. 마지막으로 관내외의 사업장 분석을 통해 실직자들에게 알선할 사업장을 선정하는 작업이 진행되었다.

시너지를 일으키는 데이터, 사람을 일으키다

그동안 모수(母數)가 적은 중앙정부의 통계자료만 의지하던 남양주시에 이 프로젝트는 혁신적 성과를 가져다주었다. 6개월 동안 약 100명(남성 70명, 여성 30명)의 신규취업이 이루어졌는데, 이는 연간 총 18억 원의 가계경제 및 지역경제 활성화를 뜻한다. 게다가 각 공공기관 간 데이터 공유를 통해 국민의 생활을 더 나은 방향으로 이끌 수 있다는 사실을 알린 모델이 되기도 했다. 특히 경제적 어려움에 처한 계층을 선제적으로 지원하고 국민연금 가입과 연계해 국민연금 사각지대 해소에도 크게 기여할 것으로 기대된다.



3-2

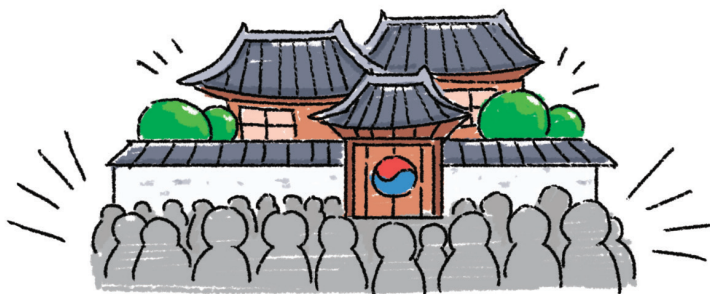
한옥마을, 관광객으로 인산인해

전주 한옥마을 관광분석을 통한 경제활성화(전주시)

여름방학을 맞아 전주 한옥마을로의 가족 여행을 준비하는 회사원 P 씨는 여행 계획을 세우면서 고민에 빠졌다. 2년 전 어린이날 연휴에 전주를 찾았다가 주차도 제대로 못하고 한옥마을에서 다른 관광지로 찾아가기가 어려웠기 때문이다. 어떻게 하면 쾌적하고 쉽게 관광명소들을 둘러볼 수 있을까?

한옥 마을 관광객 1,000만 시대

우리나라 전통문화를 느낄 수 있는 전주 한옥마을이 대한민국 대표 관광지로 급부상하면서 곧 1,000만 관광객 시대에 성큼 다가섰다. 하지만 관광객이 급증하면서 주차, 숙박, 쓰레기 등의 문제점이 불거지고 있다. 또한 관광객 동선 역시 일부 지역에만 집중되면서 관광객 불편이 계속 늘고 있어 관광환경 개선이 절실한 상황이다. 이에 전주시는 한옥마을 관광객 패턴을 분석하여 맞춤형 관광정책을 개발하고, 한옥마을을 타 지역의 관광 거점으로 연계·확대할 수 있는 방안과 해결책을 마련하고자 했다.



다양한 요인을 연계한 빅데이터 분석

행정자치부는 전주시와 함께 SKT 통신데이터와 현대카드 매출데이터를 사용하여 전주시와 전주한옥마을 관광객 특성 및 주요 유입지, 한옥마을 유입인구당 매출효과 등을 분석했다. 또한 인근 지역 연계관광 코스를 개발하여 지속 성장 가능한 관광 정책을 수립하기 위해 SNS·블로그 등 키워드 분석과, 날씨·유가 등의 분석을 통해 다양한 관점에서 지역관광 활성화에 기여하기 위한 빅데이터 분석 사업을 진행하였다.

한옥마을 관광 분석 통한 주변 관광지 연계

전주 한옥마을은 축제가 많은 5월·10월에 관광객이 가장 많았고, 매출 규모는 연간 1,150억 원으로 집계됐으며, 한옥마을 관광객 분산을 위해서는 전주시 남·북, 동·서를 연결한 교통수단이 필요한 것으로 분석되었다. 이를 통해 관광객 수요와 동선이 예측 가능하게 되었고, 관광 트렌드 변화 대비 방안, 관광 압력 분산 방안 등 다양한 정책 수립에 반영할 수 있게 되었다. 이와 더불어 코레일은 인근 지역 간 KTX 연계형 관광 상품을 만들어 내놓았다. 이제 관광객은 유용한 관광 정보를 제공받아 만족스러운 여행을 즐길 수 있게 되었고, 지자체 또한 좋은 평가를 받아 일석이조의 효과를 볼 수 있게 되었다.



3-3

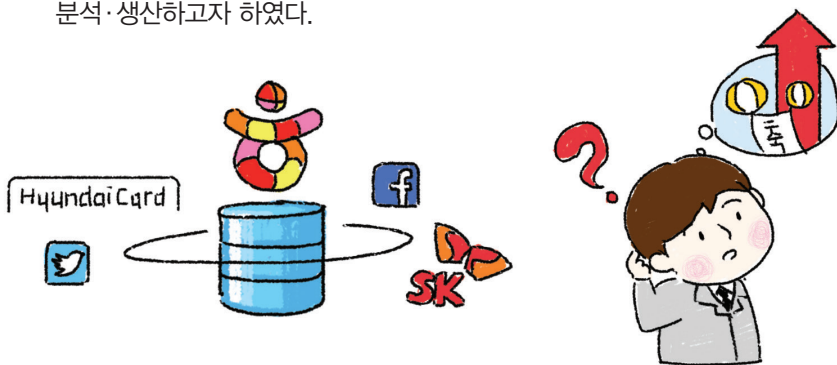
축제의 계절, 주말이 즐겁네!

빅데이터 활용 문화·관광 축제 분석(한국관광공사)

초등학생 아이들을 둔 직장인 씨는 아이들을 데리고 우리나라 고유의 모습을 보여주고 싶어 국내 여행을 떠나기로 마음먹었다. 씨는 마침 지역축제 분석 결과를 담은 뉴스를 접하게 되었고, 어린이 사이에서 인기인 함평 나비축제에서의 일정을 계획했다. 덕분에 씨네 가족은 잊지 못할 추억을 만들게 되었다.

세계화를 꿈꾸는 문화관광축제

축제는 무형의 관광자원으로서 지역 경제 활성화 및 지역 이미지 개선의 효과적인 수단으로 주목받고 있다. 우리 정부에서는 관광산업의 경쟁력 향상 및 지역경제 활성화를 위해 관광상품성이 큰 축제를 “문화관광축제”로 선정하여 지원 중이며, 경쟁력 있는 세계적인 축제로 육성하고자 한다. 이에 따라 축제의 실증적이고 과학적인 성과 측정 필요성이 대두되면서 한국관광공사에서는 빅데이터 기술을 통해 각 축제별 특성과 향후 행사 육성에 활용할 수 있는 유의미한 정보를 분석·생산하고자 하였다.

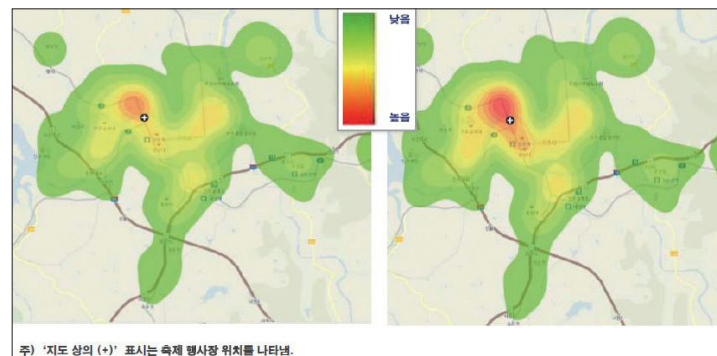


문화관광축제의 유동 및 유입인구

관광공사는 SKT의 통신 데이터, 현대카드의 결제 데이터, 소셜미디어 자료 등 2,200TB(테라바이트) 분량의 빅데이터를 활용해 선정된 문화관광축제 중 우수한 축제인 진주 남강 유등축제, 김제 지평선축제 등 총 16개의 문화관광축제를 선별하여 축제 성과를 분석하였다.

요즘 뜨는 축제는 무엇일까?

축제 기간 동안 16개 지방자치단체의 일평균 방문객은 약 13만 4,000명으로 평소보다 12%가 늘었다. 축제 방문객은 해당 지역 지자체에서 유입되는 경우가 가장 많았다. 축제 기간 해당 지자체의 일평균 소비업종 매출은 약 16억 원으로 평상시보다 7.4% 증가했음을 확인할 수 있었다. 이처럼 문화관광축제의 성과분석을 통해 축제별로 외부 유입 인구 규모 및 현황 등을 파악할 수 있어 향후 지방축제 관련 정책을 수립하는 데 시사점을 얻을 수 있을 것으로 기대된다. 또한 축제에 대한 만족도 향상뿐 아니라 국민들의 소중한 여가시간 활용에도 큰 도움이 될 전망이다.



평상시 유동인구 밀도와 축제기간 유동인구 밀도

3-4

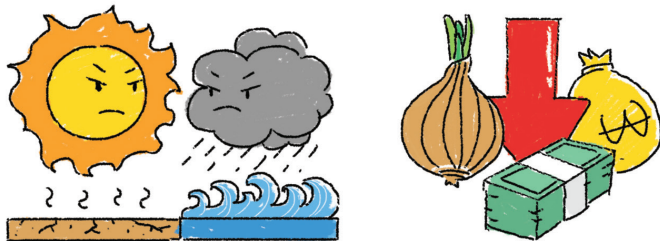
지난겨울 이상 기온, 올해 양파 가격은 얼마?

기상데이터와 농산물 생산성 예측(기상청, 농촌진흥청)

농산물 수매가를 결정하는 것은 B팀장에게 매년 어려운 도전이다. 날씨에 따라 농산물의 물량과 가격이 좌우되기 때문이다. 따뜻했던 지난 겨울의 날씨로 인해 양파 가격이 폭락할 것으로 예상되는데, 이러한 고민은 기온, 강수량 등 기상 데이터를 분석해 농산물 가격을 예측할 수 있다면 해소될 수 있을 것이다.

빅데이터로 농산물 가격과 소비자 물가를 잡아라!

기상이변이 '이변'이 아닌 일상이 되어버린 요즘, 농산물의 생산량과 출하시기를 예상하는 일은 점점 더 어려워지고 있다. 특히 가뭄과 폭우·폭염 등이 잦아지면서 농산물을 생산하는 산지에서는 날씨에 따라 작황이 크게 좌우되는 일이 늘어나고 있다. 이렇게 기상상황이 급변함에 따라 요동치는 것은 바로 소비자 물가이다. 농사를 그저 '하늘의 일'이라는 생각으로 아무 일도 하지 않을 수만은 없다는 데에, 기상청과 농촌진흥청이 뜻을 함께했다. 하늘을 움직일 수는 없었지만, 그동안 누적된 데이터는 움직일 수 있었다.



기상기후 데이터와 농산물 가격 데이터를 분석

기상청과 농촌진흥청은 농작물에 영향을 미치는 기상요인을 분석하여 '농작물 생산성 예측 모형'을 개발하고, 서민경제의 안정화를 위해서 기상기후 데이터와 농산물 가격 데이터를 분석했다. 두 인자 간 상관관계 분석을 통해 상관성을 도출한 것이다. 수입 농산물의 경우는, 수입국 기상과 수입 농산물 정보의 상관성을 분석하여 농산물 수입 정책, 농산물 수급 분석 모델의 정확성을 제고하였다. 향후 분석 모델의 정확성을 높인다면 농민의 의사결정에도 큰 도움을 줄 수 있을 것으로 기대하고 있다.

기상을 예측하면 생활이 안정된다

기상청과 농촌진흥청이 개발한 분석 모델의 정확성이 높아지면, 앞으로 날씨가 소비자 물가에 미치는 영향은 크게 줄일 수 있다. 농작물의 단위면적당 생산량을 예측하게 되면, 가격이 안정되어 서민경제에도 긍정적인 영향을 미치고 농민의 정부에 대한 신뢰 역시 향상될 것으로 기대된다. 게다가 기상청은 데이터 활용 모델을 통해 향후 기상 정보를 다양한 분야에 적극적으로 활용할 수 있게 되었으며, 농촌진흥청에서는 기후 변화에 따른 병충해 발생에 대해 선제적 대응을 할 수 있는 기반을 마련하게 되었다.



4

더 정확하고 신속해지는, “스마트 행정”

4-1. 열심히 일한 당신! 당신의 임금은 안녕하십니까?

근로감독 사업장 선정 과학화(고용노동부)

4-2. 공동주택 관리비 지출내역 확인할 길 없어... 속 터지는 입주민들

공동주택관리비 빅데이터 분석(경기도, 국토교통부)

4-3. 도로 위에는 언제 터질지 모르는 ‘지뢰’가 있다

인공지능(딥러닝) 기반의 도로포장파손 실시간 탐지시스템 개발 및 적용
(한국도로공사 ICT센터)

4-4. 귀 기울여 듣는 시민의 목소리

민원분석을 통한 갈등문제 해결(포항시)

4-1

열심히 일한 당신! 당신의 임금은 안녕하십니까?

근로감독 사업장 선정 과학화(고용노동부)

곧 태어날 아기를 위해 야근과 주말근무를 마다하지 않던 L씨. 당장의 생계를 이어가기도 힘들었지만 회사에 사직서를 낼 수밖에 없던 이유는 지난 6개월간 월급을 전혀 받지 못했기 때문이다. L씨는 현재 밀린 월급은 물론 퇴직금까지 받지 못하는 상황이다. L씨의 답답한 마음을 풀기 위해서는 어떻게 해야 할까?

위법 사업장 가리기 위해 빅데이터 활용

고용부는 사업장 자료와 근로자의 신고사건 자료를 바탕으로 근로 감독 우선 사업장을 분석했다. 여기에 근로복지공단의 고용·산재 사업장 정보, 건강보험공단의 보험료 체납 내역 등이 분석에 활용되었다. 이는 공공기관 자료를 융합시켜 새로운 데이터를 도출했다는 점에서 의의가 있다. 더욱이 본 사업은 5대 취약 유형(서면계약·임금체불· 최저임금·근로시간·약자보호)에 대한 종합 취약지수를 도출하고, 반복 적으로 위법 행위를 하는 사업장을 우선 감독할 수 있게 하였다.

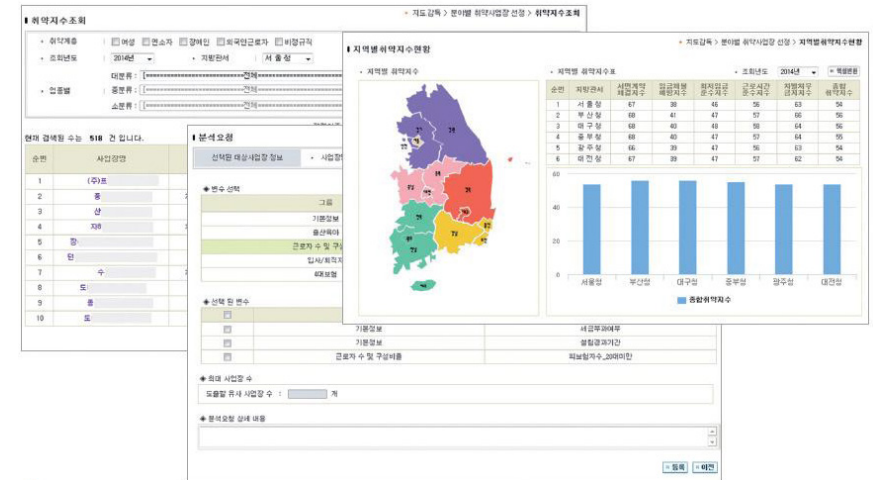


취약지수가 높은 사업장 우선으로 근로감독

근로감독 사업장 선정 과학화 사업을 통해 도출된 취약지수를 사용한 감독은 2016년 초부터 본격적으로 이뤄졌다. 성과는 상상 이상이었고, 근로감독관들도 이를 반기지 않을 수 없었다. 예전에는 사업장을 방문해서 허탕을 치는 경우가 많았는데 이러한 사례가 이전보다 훨씬 줄어들었기 때문이다. 실제로 한 근로감독관은 “빅데이터 분석을 통해 행정상 손실을 많이 줄일 수 있게 됐다.”고 밝혔다.

안전하고 올바른 근로환경 개선과 정착을 위해!

정부는 2018년까지 취약지수 사용률을 50%까지 단계적으로 확대 할 계획이다. 현장검사의 절반 이상을 취약지수에 바탕을 뒀 진행하겠다는 의미다. 취약지수를 이용한 근로기준법 위반 사업장 적발률 또한 80%까지 향상시킬 계획이다. 이를 통해 정부는 3년간 1,461억 원의 임금체불 감소 효과 등 근로자를 위한 안전하고 올바른 근로환경 개선과 정착을 기대하고 있다.



취약지수 현황 분석

4-2

공동주택 관리비 지출내역 확인할 길 없어... 속 터지는 입주민들

공동주택관리비 빅데이터 분석(경기도, 국토교통부)

안양시 A아파트에 사는 K씨는 옆 단지에 살고 있는 친구와 이야기하던 중 깜짝 놀랐다. 비슷한 평수에 가족 수도 똑같은데 관리비가 더 높았기 때문이다. K씨는 관리사무소에 항의했지만 관리사무소장은 오히려 큰소리를 냈다. K씨는 자신과 피해를 막기 위해 공동주택 관리비 지출내역 확인의 필요성을 절감했다.

속도, 정확성, 효과를 높인 '공동주택 관리비 감사'

행정안전부는 2015년 국토교통부, 경기도 등과 함께 '공동주택 관리비 투명성 제고' 사업을 진행했다. 빅데이터를 통해 공동주택의 난방비·전기료 등 관리비 47개 항목과 관련 데이터를 분석한 결과, 일부 공동주택의 경우 무려 2.2배의 관리비가 부과되었다. 경기도는 안양시의 160개 공동주택 단지를 대상으로 공동주택 관리비 분석 모델과 부당 지수를 만들고, 이를 바탕으로 광명시, 수원시, 안양시 등 경기 3개 지역의 5개 단지를 대상으로 현장실사를 실시하였다. 놀랍게도 결과가 비교적 정확히 들어맞았다.

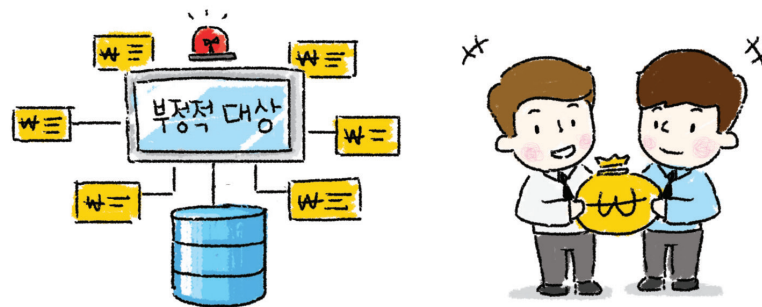


말 많은 아파트 관리비, 빅데이터로 해결!

관리비뿐만 아니라 공동주택 입찰에 참여한 사업자의 입찰 금액과 사업자가 하도급 업체에 계약한 공사 금액 데이터를 비교 분석해 금액 차이가 큰 '부적정 대상 리스트'도 만들어 냈다. 사업자의 입찰 금액과 하도급 업체와의 공사 계약 금액 간의 차이를 통해 입찰 비리가 추정되는 사업자를 도출해낸 것이다. 실제로 부적정 리스트에 있는 업체가 입찰에 참여한 광명시, 수원시 등의 2개 단지를 조사한 결과, 입찰 과정에서도 역시 문제가 있는 것으로 밝혀졌다.

빅데이터로 알아본 관리비 부조리 상태

지자체에서 전체 공동주택 관리비 실태를 점검하기란 매우 어렵다. 하지만 공공 빅데이터 분석을 통해 보다 효율적으로 관리비 과다 부과와 입찰 과정 비리 등을 포착할 수 있게 되었다. 이에 경기도는 관리비를 과다 부과하는 관리사무소에 대해 지속적인 감사활동을 벌일 수 있게 됐다. 행정안전부와 경기도는 공동주택 관리비 분석 모델을 향후 경기도 31개 시·군으로 확대하고 활용해 나갈 계획이다. 이로써 국민들이 부당하게 지출하던 관리비가 줄어들고 공동주택 입찰 부조리 사례도 점차 줄어나갈 수 있게 되었다.



4-3

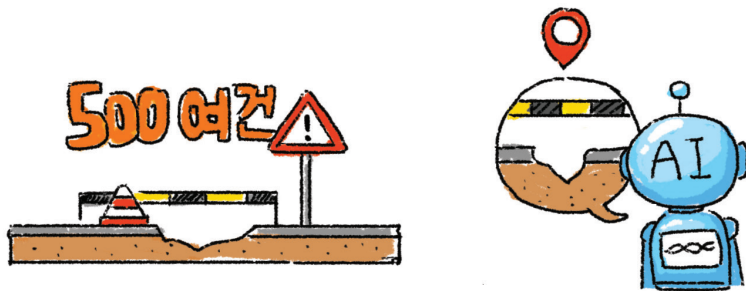
도로 위에는 언제 터질지 모르는 '지뢰'가 있다

인공지능(딥러닝) 기반의 도로포장파손 실시간 탐지시스템 개발 및 적용(한국도로공사 ICT센터)

전국을 영업처로 둔 R씨는 비가 온 후 고속도로를 달릴 때 신경이 곤두선다. 비가 온 다음 날이면 포트홀이 부쩍 늘어나기 때문이다. R씨는 포트홀 때문에 큰 사고가 날 뻔한 경험이 있기 때문에 도로가 위험천만한 지뢰밭처럼 느껴진다. 자칫 잘못하면 큰 사고로 이어지는 포트홀, 미리 예방할 수는 없을까?

포트홀, 선제적으로 발견해 대응하라!

최근 3년간 발생한 포트홀은 5만여 개로, 이로 인한 교통사고가 500여 건에 달했다. 이에 한국도로공사는 '인공지능(딥러닝) 기반의 도로포장파손 실시간 탐지시스템 개발 및 적용' 프로젝트를 진행했다. 지금까지의 대응은 고객의 제보, 안전순찰과 도로 유지보수 점검 과정에서의 발견이 전부였다. 따라서 포트홀 보수 대응팀이 현장에 도착할 때까지 많은 시간이 소요되어, 이는 포트홀로 인한 사고 가능성을 높이는 요소가 되었다. 한국도로공사는 이러한 맹점을 개선하기 위해 프로젝트를 시행했다.

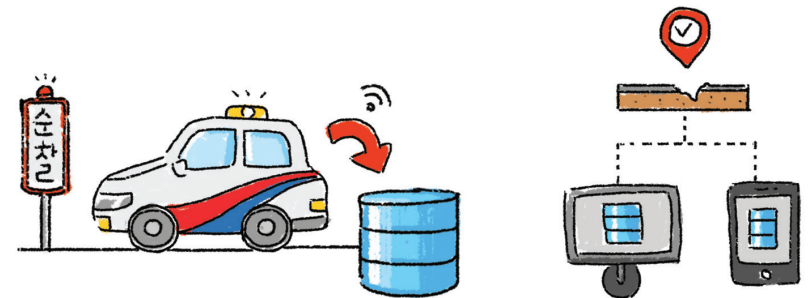


고속도로 데이터와 영상 빅데이터, 실시간 탐지를 가능케 하다

한국도로공사의 주된 업무 중 하나는 도로 유지보수 점검이다. 고속도로를 달리다 보면 작업 중인 순찰차를 어렵지 않게 발견할 수 있는데, 이 순찰차들로 하여금 노면을 촬영하여 데이터화 하는 작업을 시작했다. 이렇게 축적된 데이터를 바탕으로, 촬영된 포트홀은 그 특징과 파손 정도에 따라 분류된다. 이와 함께 포트홀의 정확한 위치와 깊이, 크기, 부피를 GPS 위치 정보와 함께 전송하게 되는데, '심각파손'으로 분류되는 경우 즉시 대응팀이 출동하게 된다. '빠른 대응'을 앞선 '선제적 대응'이 가능하게 된 것이다.

안전하고 빠른 도로, 빅데이터로 만들어 간다!

한국도로공사는 실시간 탐지 시스템을 기반으로 향상된 주행안전 서비스를 제공할 수 있게 되었다. 더불어 체계적인 도로포장 유지관리가 가능해졌으며, 그동안 2인 1조로 수기 입력하던 업무 방식을 단독 점검으로 전환해 효율이 향상되었다. 앞으로는 양방향 통신이 가능한 내비게이션과 휴대전화 등을 통해 포트홀 발생에 대한 정보를 받아볼 수 있는 시스템도 구축할 계획이다. 또한 민자고속도로와 지방도로로 확대한 후 지속적으로 자료를 축적해 신뢰도를 높여 해외도로 정비 사업에도 진출할 계획이다.



4-4

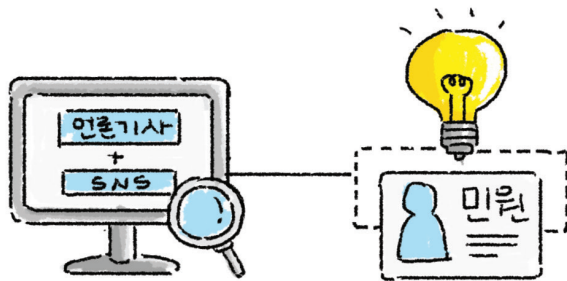
귀 기울여 듣는 시민의 목소리

민원분석을 통한 갈등문제 해결(포항시)

민원 콜센터 직원 U씨는 여름엔 악취 민원, 휴가철 주차 민원 등 특정 시기에 따라 민원이 비슷하다는 점을 알아챘다. 따라서 U씨는 지금까지 접수된 민원을 분석해 주민들이 필요한 민원을 미리 해결할 수 있는 아이디어를 냈다. U씨는 이러한 시도가 시민들의 만족도와 삶의 질을 높일 수 있을 것이라고 확신했다.

포항시, 민원 분석을 통한 갈등 문제 해결

행정안전부는 포항시 최근 10여 년간의 민원을 분석하기 위해 인터넷 기사, SNS 등을 수집하였다. 이를 통해 시민의 주된 민원을 파악하고, 민원의 경중 등을 분석하여 대응방안을 미리 마련할 수 있었다. 포항시의 최대 민원 이슈는 주차로 나타났으며, 민원이 집중된 지역을 집중 관리 지역으로 선정하여 주차 관련 해결방안을 우선순위로 행정에 반영하였다. 이처럼 빅데이터 활용의 가장 큰 장점 중 하나는 민원이 사회적으로 확산되기 전에 인지하여 적은 노력으로 문제 해결을 가능하게 할 수 있다는 점이다.



가장 작은 것에서 시작되는 분석

포항시에서 가장 먼저 실시한 것은 형태소 분석이다. 민원게시판에 시민들이 작성한 민원을 분석해, 문장을 이루는 가장 작은 단위인 형태소를 분류하기 시작한 것이다. 또한 민원 게시판의 지역별 이용현황을 연·월별로 분석해 가장 많이 사용된 키워드를 도출해냈다. 게시판 이용 빈도 역시 분석대상 중 하나로 이런 과정을 통해 최종적인 '화제어'를 뽑아내는 한편, 그 화제어의 '연관어'를 파악해 이를 시각화할 수 있었다. 즉, 하나의 화제어가 어떤 연관어가 있으며 언제, 어디서 많이 사용되었는지 시각화함으로써 반복되는 민원을 파악하고 이에 대해 선제적 대응을 할 수 있게 된 것이다.

데이터, 무엇이 먼저인지 정확하게 알려주다

데이터를 분석한 결과, 포항시는 지역 내 민원들 중 시급한 것들을 과학적으로 판단하고 이에 대한 우선순위를 합리적으로 정하여 한정된 예산과 인력을 효율적으로 사용하고 있다. 불법주정차 단속을 위한 CCTV 확대설치 및 단속문자 알람 서비스 구현, 관광지 인근의 학교 및 병원과 협의하여 여름철 임시 주차장 설치, 단속 예고 및 단속 강화 지역 설정 등 모두 포항시가 데이터 분석을 통해 마련한 해결책들이다.



워드클라우드

내 삶을 바꾸는
공공 빅데이터!



행정안전부

서울특별시 종로구 세종대로 209
공공데이터정책과 / 02-2100-4336
www.bigdata.go.kr

NIA
한국정보화진흥원

대구광역시 동구 첨단로 53
공공데이터활용지원센터 / 053-230-1542
www.nia.or.kr