

예측 치안의 헌법적 논의 Constitutional Discussion of Predictive Policing

이 병 규(Lee, Byeong Gyu)*

ABSTRACT

This paper has recently dealt with the constitutional perspective of predictive policing in the course of the change of police work due to the remarkable development of information and communication technology(ICT). Especially, I wondered how big data is used in the police law enforcement and crime prevention and there is no constitutional problem in it.

In the first place, we examined the US case of how policing is used in police affairs. The local predicted data using data such as crime location and time, and the personal forecasting security by the interpersonal target system. Then, we examined whether the predictive security based on these big data infringed the individual fundamental rights. For example, applying the results obtained by data mining to individuals, that is, using data mining results to derive the nature or trend of a particular individual, infringes on the basic rights of the individual.

Finally, when predictive polices were introduced and carried out using these big data, we considered what constitutional problems are and how to control them. The application of data mining results to individuals needs to establish effective control methods such as individual and collective consent because they involve issues related to constitutional principles that respect individuals beyond simple personal information protection concerns.

Key words: predictive policing, ICT, big data, data mining, constitution, personal information protection

* 동의과학대학교 조교수, 법학박사

I. 서론

지난 수십 년간 우리 사회와 생활환경을 놀랍게 변화시킨 ‘정보혁명(Information revolution)’¹⁾은 이제 새로운 단계에 접어든 것으로 보인다.²⁾ 1970년대 인텔사(Intel)의 마이크로프로세서(microprocessor) 개발과 그것에 의한 계산능력(the power to compute)의 향상(제1단계), 네트워크화와 그것에 의한 연결능력(the power to connect)의 향상(제2단계)에 이은 새로운 단계라고 할 수 있다. 그것은 데이터(data)와 예측능력(the power to predict)의 향상에 의한 것으로 그 자체로 ‘빅 데이터 혁명(Big Data Revolution)’이라고 하는 사회적 전환기라고 할 수 있다.³⁾ 이에 의해 데이터, 쇼핑, 의료, 투표, 법집행, 테러 대응, 사이버 보안을 포함한 모든 인간 활동과 결정이 빅 데이터에 의한 예측(predictions)으로부터 영향을 받고 있다. 이러한 현상에 의할 때 이 시대를 주도하는 핵심 기술은 하드웨어나 소프트웨어가 아니라 데이터라는 것⁴⁾에 주목할 필요가 있다. 오늘날 어느 분야에서나 등장하는 4차 산업혁명은 이러한 정보통신기술의 최종적 결정체라고 할 수 있다.

정보통신기술의 비약적 발전은 경찰 업무 영역에도 적지 않은 변화를 가져오고 있다. 이미 일부 국가에서는 인공지능(Artificial Intelligence, AI) 기술을 활용한 빅 데이터 플랫폼을 통해 범죄나 사고를 예측하고 방지하는 프로젝트가 진행 중이다. 미국 시카고 경찰은 지난 2017년부터 범죄 예측 시스템을 도입해서 시간이나 계절과 같은 주기 정보, 날씨나 지역 경제, 과거 범죄 데이터의 분석을 통해 범죄가 발생하는 일정한 규칙을 도출한다. 이 규칙에 따라 범죄가 발생할 가능성이 높은 지역을 표시해서 경찰에게 미리 알려준다.⁵⁾ 일본 나가와현 경찰은 AI를 이용한 새로운 단속 시스템 도입을 검토하고 있다. 즉 범죄나 사고 발생을 예측하여 수사나 범죄 예방에 활용한다는 계획으로 2020년 도쿄올림픽과 장애인올

1) 정보기술의 발전에 의해 사회나 생활이 변혁하는 것으로 정보기술(Information Technology, IT)과의 관계성으로부터 ‘IT혁명’이라고 부르기도 한다. 이 용어는 영국의 과학자이자 마르크스주의자인 존 버널(John Desmond Bernal)이 마르크스주의의 범주에서 처음으로 사용했지만 오늘날은 그것과는 별개로 많은 영역에서 널리 사용되는 용어로 정착했다(<http://www.keiei.ne.jp/keyword/2764/>).

2) Neil M. Richards and Jonathan H. King, Big Data Ethics, 49 Wake Forest L. REV. 393, 397(2014).

3) Id. at 393. 최근 이러한 현상을 ‘4차 산업혁명(The Fourth Industrial Revolution)’이라고 명명하면서 이 용어의 범주에서 이해하는 경향이 있다. 증기기관을 1차, 전기기관을 2차, 제조업 자동화를 3차 산업혁명으로 보고, 인터넷에 의해 모든 기기가 결합하는 단계를 4차 산업혁명이라고 부른다. 4차 산업혁명은 주로 제조업을 중심으로 사물인터넷(Internet of Things, IoT)이나 인공지능을 도입하고 자율적·자동적·효율적으로 제조 공정이나 품질 관리를 하며 에너지 절약 등을 통해 새로운 사업의 고도화를 목표로 하는 것으로 원래는 독일의 산업체와 정부의 공동프로젝트로 추진한 산업고도화의 개념인 ‘Industry 4.0’을 지칭한다. 따라서 4차 산업혁명은 정보혁명의 고도화 내지는 새로운 단계로의 진입으로 볼 수 있을 것이다. 4차 산업혁명의 개념 정의에 대하여는 김덕현, “지금이라도 4차 산업혁명 대응정책 재정립해야”, STARTUP4, 2018. 10. 12 참조.

4) 김화중, 4차 산업혁명과 데이터 가치체계, 헬스케어 ICT정책 Vol. 06, 헬스케어미디어연구소, 2017. 3. 49면.

5) 디지털타임즈, 2019. 1. 29.

림픽 개막까지 시범 운용을 목표로 한다. 또한 일본은 경찰청 차원에서 AI 기술을 수사에 도입하기 위한 실증 사업에 나선다고 한다.⁶⁾ 싱가포르의 ‘디지털 트윈(Digital Twin)’ 기술을 통해 도시 전체를 가상의 3D 환경으로 구현하고 여러 가지 시나리오를 시뮬레이션 함으로써 범죄를 미리 예방한다는 계획⁷⁾을 실행 중에 있다.

예측 치안과 관련한 경찰 업무의 이러한 움직임은 우리나라에서도 확인할 수 있다. 최근 경찰은 ‘빅 데이터 통합 플랫폼 구축 전략’ 수립을 통해 전 치안 영역에 빅 데이터 기술을 접목할 계획이라고 한다. 그 첫 단계로 국내외 치안데이터 활용 사례를 수집하고 데이터 현황을 분석해서 경찰에 최적화된 ‘빅 데이터 통합 플랫폼 모델’을 구축할 방침이라고 한다. 특히 AI, 딥러닝(Deep Learning) 기술이 적용된 빅 데이터 플랫폼 모델 구축을 위해 치안 데이터의 표준화를 추진한다.

이에 본고는 경찰의 법집행이나 범죄 예방 영역에 초점을 맞추어 거기서 빅 데이터가 어떻게 이용되는지, 특히 이를 전제로 한 데이터 마이닝과 그 적용 — 이른바 예측 치안(predictive policing) — 의 헌법적 문제를 고찰한다.⁸⁾ 이미 다수의 논자는 빅 데이터가 다양한 형태로 경찰 수사에 극적인 변화(dramatic change)를 가져올 것⁹⁾이라고 한다.¹⁰⁾ 이하에서는 경찰의 빅 데이터 이용의 구체적 실례를 소개하고, 그것이 어떤 점에서 경찰활동에 극적인 변화를 가져올 수 있는지, 또한 이것이 헌법적으로 어떻게 문제될 수 있는지 — 단지 개인정보보호의 문제만 발생하는지 등 — 검토하고, 경찰에 의한 빅 데이터 이용 내지 데이터 마이닝 적용 등에 대한 적절한 통제 방법을 생각해본다.

II. 예측 치안의 형태

지금까지는 개별 경찰관이 과거의 통계를 분석하고 업무상 경험이나 감각으로 경찰활동을 해왔다고 할 수 있다. 그러나 앞으로는 경찰 영역에서도 정보통신기술(ICT: Information

6) 産経ニュース, 2018. 1. 29.

7) ICT 융합 동향, 정보통신산업진흥원, 2018. 2. 18면 이하 상세.

8) 여기서 보다 근본적 차원에서 ‘예방원칙’과 헌법이나 ‘예방경찰’과 헌법의 관계를 다루기는 어렵다. 예방원칙의 헌법적 차원에서의 접근은 예방국가, 안전기본권, 국가의 위험 방어, 경찰의 위험 통제 등 국가 전 영역에 미치는 쟁점들을 포함하기 때문에 별도의 논의를 필요로 한다.

9) Elizabeth E. Joh, Policing by Numbers: Big Data and the Fourth Amendment, 89 Wash. L. Rev. 35, 37 (2014); AI와 4차 산업혁명이 법조계 미칠 영향 좌담, 법률신문, 2017. 4. 21.

10) 현재 경찰청이 활용 중인 빅 데이터 시스템에는 형사사법종합정보시스템(KICS), 순찰차 배치시스템(IDS), 지리적 프로파일링시스템(GeoPros) 등이 있으며, 도입 예정인 시스템에는 빅 데이터 기반 범죄분석 시스템, CCTV 영상검색 고도화 및 신원확인 기술, 스마트 신호운영시스템 등이 있다. 특히 올해는 빅 데이터 기반의 첨단 범죄 분석 프로그램인 ‘클루(CLUE)’를 시범 운영할 예정이다(치안정책연구소, “스마트치안지능센터 기본구상 연구”, 2016. 12. 43-46면 참조).

and Communication Technology)을 활용한 과학적 근거에 기초한 예측 정보를 통하여 범죄나 사고 발생 가능성이 높은 시간대나 지역을 추출하여 경찰 자원을 효과적으로 활용하고, 불법 사안을 미연에 방지하기 위한 고도의 경찰 활동을 전개하는 것이 가능해질 것으로 생각된다. 이하에서는 이러한 ICT 기술을 활용한 미국의 예측 치안 사례를 살펴본다.

1. 지역 예측 치안

경찰이 빅 데이터를 이용한 실례로 미국 캘리포니아 주 산타크루즈카운티의 예측 치안 프로젝트를 들 수 있다. 2011년 7월에 개시한 이 프로젝트는 과거의 방대한 범죄 데이터로부터 일정한 패턴을 추출·발견하고 차량·주거의 불법 목적 침입(burglaries)이나 차량 절도와 같은 재산범죄가 발생하기 쉬운 장소와 시간 — 핫스팟(hot spots)¹¹⁾ — 을 예측하는 컴퓨터 프로그램(PredPol: Predictive Policing)¹²⁾ 구축을 목적으로 한다. 이 프로그램에 의해 한정된 경찰 자원을 효과적으로 배분하는 것 — 범죄가 발생하기 쉬운 장소·시간에 경찰관을 집중 배치하는 등 — 이 가능해지기 때문에 예산 절감과 효과적인 치안의 양립이 가능하다고 한다.¹³⁾ 물론 그 이전에도 컴스탯(CompStat)¹⁴⁾과 같은 범죄 추적(crime

11) 핫스팟(hot spots, 범죄다발지대)을 분석하고 이를 범죄와 매핑(mapping)하는 기법이다. 과거의 범죄 발생 데이터를 근거로 범죄 위험이 높은 지점이나 지역을 예측하는 것으로 범죄는 모든 지역에 균일하게 분포하지 않고 일부 지점·지역에 편재한다는 사실에 기초한다. 범죄 매핑 중에서 가장 단순한 방법은 격자무늬(grid)를 범죄 발생 유무에 따라 색으로 구분하여 표시하는 것이다.

12) 2011년 퍼듀대학 조지 몰러(George Mohler) 교수가 제안한 EM(Expectation Maximization)법을 이용한 범죄예측알고리즘을 기본으로 한 범죄예측 어플리케이션 'PredPol'은 미국 전체 60개 이상의 경찰에서 도입했다. 이것은 '재발가능성이론', '근접반복이론', '환경적 요소'라는 세 가지 요소로 이루어지는 비공개 알고리즘에 의한 수학적 기법으로 처리하여 지도상에 범죄 발생 예측 장소를 150미터 사방에 프레임 표시하고 그 표시된 장소를 지정된 시간에 순찰함으로써 범죄를 억제한다.

13) 근접반복피해분석에 의한 범죄 예측 기법도 있다. 이것은 일부 범죄는 시간적·거리적으로 과거에 발생한 범죄와 근접하여 발생한다는 가설에 기초하여 범죄발생을 예측하는 방법이다. 환경범죄학에서는 반복 피해 현상, 즉 한차례 피해를 입은 물건·사람·장소는 다시 피해를 입을 가능성이 높다고 지적되어 왔다. 나아가 과거에 피해를 입지 않았더라도 근접하여 발생한 범죄의 표적과 유사한 속성이 나타나면 유사한 피해를 입기 쉽다는 근접반복피해 확인에 의해 범죄 예측에서 근접반복피해 분석이 유효해졌다. 실례로 캘리포니아주 산페르난도에서는 2004년부터 2년 간 최초 주택침입절도 발생 후 3시간 내에 피해 장소에서 200미터 내의 범위에서 100건 이상의 주택침입절도가 발생했다. 이러한 근접반복피해는 많은 국가와 지역에서 발생하고 있는 등 어느 정도 보편성이 확인되고 있다. 일본의 주식회사 Singular Perturbations 대표이자 도쿄대학 공간정보과학연구센터 객원연구원인 카지타 마미(梶田眞實) 박사는 한번 범죄가 발생하면 그 근방에서 범죄가 발생하기 쉽다는 범죄 캐스케이드 현상에 기초하여 '범죄발생밀도'를 계산하였다(범죄발생밀도 $(t)=g(t-t_1)+g(t-t_2)+0$), 참고로 t, x 에서 단위 시간·면적당 범죄발생 수(梶田眞實, 犯罪オープンデータをを用いた犯罪予測アルゴリズムとシステムの開発, 第2回官民ラウンドテーブル, 株式会社Singular Perturbations 2018. 2. 28 참고).

14) 뉴욕시경찰의 범죄 방지를 목적으로 하는 전략관리시스템으로 Comparison Statistics(비교 통계)의 약칭이다. 컴스탯은 정확하고 적시에 범죄 정보의 수집·분석, 효과적인 전술의 전개, 신속한 인원 배

tracking) 시스템은 존재하고 있었고, 또한 실제로 다수의 경찰에 의해 이용되어 왔지만 산 타크루즈 경찰의 프로그램(소프트웨어)은 범죄가 발생할 때마다 데이터가 축적되고 프로그램 내용이 일상적으로 갱신·검증되기 때문에 예측의 정밀도가 종전보다도 비약적으로 높아졌다고 한다.¹⁵⁾ 실제 이 프로그램을 도입한 2011년 7월과 전년 동월을 비교하면 불법침입 건수는 27%(70건에서 51건) 정도 감소했다고 한다.¹⁶⁾ 또한 컴퓨터 알고리즘 적용에 의해 인간의 경험이나 감으로는 알기 어려운 핫스팟이 나타나고 그것이 실제 범인 체포나 범죄 예방으로 이어지는 경우가 많다.

그 외에 범죄가 발생하는 사회적 맥락 등 과거의 범죄 시간이나 장소 이외의 요소를 고려하는 접근도 존재한다. 예컨대, 뉴저지 주 모리스카운티 경찰은 이른바 ‘위험면분석(RTM, Risk Terrain Modeling)’¹⁷⁾¹⁸⁾에서 ① 과거 불법목적침입, ② 최근 재산범죄로 체포된 자의 주거지, ③ 주요 간선도로와의 거리, ④ 젊은 층의 지역적 응집, ⑤ 집합주택 및 호텔 위치 정보 등을 이용하고 있다.¹⁹⁾ 이러한 위험면분석으로 인해 폭력범 및 재산범 건수는 크게 줄었다고 한다.²⁰⁾

또한 뉴욕시 경찰은 마이크로소프트사와 협력하여 ‘영역인식시스템(Domain Awareness

치, 엄밀한 후속 조치와 평가 등을 특징으로 들 수 있다. 컴스탯 도입에 따라 컴퓨터를 이용한 범죄 통계 분석이 이루어지고, 이러한 해석 결과는 주마다 행해지는 범죄전략회의에서 사용되어 전술의 전개, 인원 배치 및 경찰 업무를 평가하는데 중요한 역할을 한다(Paul E. O’Connell, Using Performance Data for Accountability: The New York City Police Department’s CompStat Model of Police Management, The Business of Government, 2001; 윤일홍·정진성, 한국경찰의 성과주의 도입과 그 시사점 - 뉴욕경찰의 컴스탯(Compstat) 시행결과 분석을 중심으로, 한국공안행정학회보 제40호, 한국공안행정학회, 2010 참조).

- 15) hot spot으로 제시되는 범위도 축소되어 현재는 약 150미터 사방의 정방형으로 지도상에 표시된다. 매일 15개의 spots이 경찰관에게 제시되고, 경찰관은 웹상에서 항상 그 위치를 확인할 수 있다. Zach Friend, Predictive Policing: Using Technology to Reduce Crime, FBI Law Enforcement Bulletin (Apr. 9, 2013),

<https://leb.fbi.gov/articles/featured-articles/predictive-policing-using-technology-to-reduce-crime>.

- 16) Jennifer Bachner, Predictive Policing: Preventing Crime with Data and Analytics, Washington, DC: IBM Center for The Business of Government, 2013. <http://www.businessofgovernment.org/sites/default/files/Management%20Predictive%20Policing.pdf>.

- 17) Leslie W. Kennedy et al., Risk Clusters, Hotspots, and Spatial Intelligence: Risk Terrain Modeling as an Algorithm for Police Resource Allocation Strategies, 27 J Quantitative Criminology 339, 342-343 (2011).

- 18) 위험면분석은 인간의 행동 패턴에 지리적 조건이 영향을 미치는 것을 고려하여 공간 역학의 관점에서 범인의 행동에 영향을 주는 지리적 특성(음식점, 주류 판매점, 주요 도로 등)이나 인구통계학적 특성(주민 구성, 연령·성별, 사회계층 등)에 착안한 기법으로, 음식점이나 도로와 같은 지리적 특성을 거의 고려하지 않고 범죄가 발생한 장소로부터의 거리 및 경과시간만을 대상으로 분석하는 근접 반복피해분석과 그 방법을 달리한다. Walter L. Perry et al, Predicting Policing: The Role of Crime Forecasting in Law Enforcement Operations, NIJ, 2013, at 19.

- 19) Joh, supra note 10, at 46.

- 20) Jeffrey S. Paul and Thomas M. Joiner, Integration of Centralized Intelligence with Geographic Information Systems, Geography and Public Safety, Oct. 2011, at 7.

System, DAS)²¹⁾을 개발하여 사용하고 있다. 이 소프트웨어는 “시내에 설치된 약 3,000개의 감시 카메라, 200개 이상의 번호판 자동인식장치, 2,000개 이상의 방사전센서, 경찰이 보유한 데이터베이스와 같은 다른 원천의 정보를 상시 수집·연결·분석하여 시내의 ‘잠재적 위협(potential threats)’을 발견하는데 이용한다.²²⁾ 이 시스템은 개별적인 범죄 분석으로는 알기 어려운 방법으로 사람·물건·장소의 관련성을 명확히 하고 이러한 정보에 경찰이 실시간 접근할 수 있도록 하는 것이다.²³⁾ 뉴욕시 경찰은 이 소프트웨어를 이용하여 2013년 11월에 개최된 뉴욕마라톤 - 그해 4월 보스턴마라톤 후 테러의 표적이 된 - 의 전체 구간을 감시할 수 있었다.²⁴⁾

2. 개인 예측 치안

비행 안전 확보를 이유로 공항에서는 이른바 ‘No Fly List’²⁵⁾ 해당자의 동정(identification)이 행해진다. 이 리스트 작성·준비 단계에서 데이터 마이닝(data mining)이 이루어지는지는 분명하지 않지만, 근년의 테러에 대한 미국 정부의 태도를 감안하면 자동예측과정(automates prediction processes)이 적용되었을 것으로 생각된다. ‘No Fly List’ 절차와는 다르지만 미국 국토안보부(Department of Homeland Security, DHS)는 실제 입국 관리에 관하여 데이터 마이닝에 기초하는 예측 모델링을 이용하고 있다.²⁶⁾ 국토안보부에 의한 데이터 마이닝 보고는 대인적 표적화 시스템(Automates Targeting System-Persons, ATS-P) 모듈에 대하여 언급하고 있으며, 국경 침입자의 위협을 예측하기 위하여 정부가

21) 영역인식시스템은 감시 카메라를 네트워크화하여 필요한 지역의 영상을 7,000여개에 달하는 경찰 소유 및 상업시설의 방범 카메라나 순찰차에 부착한 카메라 영상에서 호출할 수 있는 특징이 있다. 상세한 기능은 알 수 없지만, 예컨대 붉은 셔츠를 입은 사람만을 표시할 수 있도록 지시하면 모든 감시 카메라에서 해당하는 영상만을 내보낼 수 있다고 한다. 도로나 경찰 차량에 부착한 번호판인식장치에서 특정 차량의 위치 정보를 확인할 수 있으며, 차량위치정보는 과거 수개월 간에 걸쳐 검색이 가능하다고 한다. 또한 이것은 감시카메라나 차량위치정보에 더하여 소셜미디어에서 발표를 시사하는 투고나 과거 범죄데이터와 같은 다양한 정보까지 집약하여 그 정보를 범죄 종류 등으로 분류하여 범죄정보센터(Real Time Center, RTCC) 화면에 반영하여 관리할 수 있다. 또한 범죄정보센터에서는 영역인식시스템을 사용하여 복수의 사건 관련성과 같은 범죄 수사나 사건이 발생할 가능성이 있는 지역에는 사전에 순찰차를 순회시키는 등 다양한 방식으로 범죄를 단속하고 있다.

22) Joh, *supra* note 10, at 48-49.

23) Id. at 49.

24) Michael Schwirtz, After Boston Bombings, New York Police Plan Tight Security at Marathon, N. Y. Times (Nov. 1, 2013), <https://www.nytimes.com/2013/11/02/sports/video-surveillance-to-be-a-key-component-of-marathon-security.html>.

25) 테러리스트 등 중요 범죄 관련자의 항공기 이용을 제한하기 위하여 미국 정부가 작성하는 리스트로 국내 항공사도 이 제도를 2017년부터 도입하여 시행하고 있다. 우리나라는 항공보안법 제23조 제7항에서 승객의 탑승을 거부할 수 있는 근거를 마련해두고 있다.

26) Tal Z. Zarsky, Transparent Predictions, 2013 U. Ill. L. Rev. 1503, 1515 (2013).

보유한 여러 종류의 데이터베이스가 분석되고 있다는 것을 보여준다.²⁷⁾

경찰 실무와 관련해서 주목할 만한 것은 시카고 경찰(Chicago Police Department, CPD)에 의한 ‘편의고지프로그램(Custom Notification Program)’이다. 여기서는 다양한 경험적 데이터²⁸⁾ — 정보자유법(Freedom of Information Act, FOIA)에 의한 정보공개청구가 배척되며 사용되는 데이터의 자세한 내용도 알려져 있지 않다²⁹⁾ —로부터 폭력행위자와 피해자를 예측하는 알고리즘을 개발하고, 그 적용을 통해 이러한 자를 구체적으로 정리한 리스트 — ‘heat list’라고 부른다 — 가 작성되고 있다.³⁰⁾ 경찰은 heat list에 오른 한 사람 한 사람을 직접 방문(편지를 보내는 경우도 있다)하고, 장래 범죄에 손을 댈 때 어떤 결과가 따른다고 경고하며, 이와 함께 그 사람이 받을 수 있는 사회적 서비스(직업훈련, 주택 공급 등)도 함께 고지한다.³¹⁾ 시카고 경찰은 이러한 노력에 의해 범죄가 억제되고 있다고 하며, 실제로 60명에 대한 방문 내지 개입에 의해 그들이 새로운 중범죄에 손을 대지 않고 있다고 한다.³²⁾ 이러한 효과가 주장되는 반면에 “heat list는 범죄 전(pre-crime) 범집행을 현실화하는 차별적 도구”라는 비판도 있다.³³⁾ 실제 중범죄 등에 연루된 적이 없는 사람이 리스트에 게재되고, 그로 인해 돌연 경찰의 방문을 받는 경우도 있으며, 리스트에 게재된 다수의 사람은 게재되었다는 사실 자체에 대하여 당혹감을 느낀다고 지적한다.³⁴⁾ 이러한 예측 치안의 부정적 측면에 대하여는 나중에 검토하기로 한다.

아직 실험 단계이지만 이미 복수의 전문가 내지 기술자가 특정인이 장래 중죄를 범할지를 일정한 정밀도로 예측할 수 있는 소프트웨어를 개발했다고 한다. 예컨대, 백그라운드체크(background check)를 업무 내용으로 하는 인텔리우스(Intelius)사는 회사가 보유한 방대한 데이터 — 중범죄·경범죄 경력, 교통위반 경력, 젠더, 눈·피부색, 문신의 유무 등 — 에 기초하여 누가 중죄를 범할지를 ‘합리적 정확성(reasonable accuracy)’을 기초로 결정하는 알고리즘을 개발했다고 한다.³⁵⁾ 이 프로그램은 적극적인 설정 하에서는 1980년대 이후

27) Id.

28) 예컨대, 살인 희생자의 지인이 살인에 연루될 가능성이 9배라는 것과 같은 데이터도 포함된다. Kristal Hawkins, Heat list’ brings Minority Report-style police attention for likely offenders, Chicago Crime Library (Feb. 24, 2014), <http://www.crimelibrary.com/blog/2014/02/24/heat-list-brings-minority-report-style-police-attention-for-likely-offenders-in-chicago/index.html>.

29) Aaron Rieke, David Robinson, and Harlan Yu, Civil Rights, Big Data, and Our Algorithmic Future (Upturn 2014), https://centerformediajustice.org/wp-content/uploads/2014/10/Civil-Rights-Big-Data_Our-Future.pdf.

30) Jeremy Goner, Chicago police use ‘heat list’ as strategy to prevent violence, Chicago Tribune reporter(Aug. 21, 2013), <https://www.chicagotribune.com/news/ct-xpm-2013-08-21-ct-met-heat-list-20130821-story.html>. 2014년 보도 시점에는 400명 이상이 리스트에 게재된 것으로 나타난다.

31) Rieke, supra note 24, at 18. 시카고 경찰은 ‘경고’보다도 사회서비스 고지라는 점을 강조하여 ‘편의 고지프로그램’이라고 부르는 것으로 생각된다.

32) Id.

33) Hawkins, supra note 19.

34) Id.

켄터키 주의 재판기록상 중범죄자(5만1,246명)를 정확히 동정하는 한편, 2,220명의 비중범죄자를 잘못 동정했다고 한다.³⁶⁾ 또한 소극적인 설정 하에서는 37,842명의 중범죄자를 정확히 동정한 것에 대하여 잘못 동정한 비중범죄자는 152명에 그쳤다고 한다.³⁷⁾ 인텔리우스사는 이러한 위양성(false positive)이 줄어들면 경찰에 의한 실제 사용 가능성은 매우 높다고 한다. 물론 대인적 표적 예측 치안이 제대로 사용되는지의 문제는 전문가나 기술자의 일이 아니라 법학자의 일이라고 할 수 있다.

그 외에도 펜실베이니아대학에서 통계학·범죄학을 가르치는 리처드 버크(Richard A. Berk) 교수는 가석방 중인 자가 살인에 관여하는지를 예측하는 소프트웨어를 개발했다고 한다.³⁸⁾ 버크에 의하면 6만 명 이상의 데이터를 이용하여 작성한 이 소프트웨어는 가석방된 경우의 살인 관여를 75% 이상의 확률로 예측할 수 있다고 한다.³⁹⁾ 메릴랜드 주에서는 버크의 협력 하에 아동 학대를 하는 가족을 예측하는 프로그램을 개발 중이라고 한다.⁴⁰⁾

Ⅲ. 예측 치안의 헌법적 쟁점

이상에서 미국의 경찰에 의한 빅 데이터 이용과 데이터 마이닝의 실패를 살펴보았다. 이 하에서는 이것을 헌법적 관점에서 고찰한다. 그에 앞서 ‘빅 데이터’나 ‘데이터 마이닝’과 같은 용어가 어떤 의미를 가지는지 알아보고, 그 고찰 범위 — 경찰에 의한 정보처리 중에서 어떤 행위를 문제로 삼아야 하는가 — 를 명확히 한다.

35) Jordan Robertson, How Big Data Could Help Identify the Next Felon — Or Blame the Wrong Guy, Bloomberg (Aug. 15, 2013), <https://www.bloomberg.com/news/2013-08-14/how-big-data-could-help-identify-the-next-felon-or-blame-the-wrong-guy.html>.

36) Id.

37) Id.

38) Richard A. Berk et al., Forecasting murder within a population of probationers and parolees: a high stakes application of statistical learning, Journal of the Royal, Statistical Society (Series A), 172 (2009), <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.391.9120&rep=rep1&type=pdf>.

39) 빅토르 마이어 쾨버거·케네스 쿠키어 지음 / 이지연 옮김, 빅 데이터가 만드는 세상, 북이십일 21세기북스, 2013, 295면. 이것은 네 차례 중 한 차례는 판단 오류라는 의미이다.

40) 우리나라도 빅 데이터를 활용해서 학대를 당하거나 보호가 필요한 아동을 신속하게 찾을 수 있는 ‘e아동행복지원시스템’을 2018년부터 전국적으로 실시·운영하고 있다. ‘e아동행복지원시스템’은 아동의 장기결석 정보, 영유아 건강검진·예방접종 실시 정보, 병원 기록 등 빅 데이터를 모아 분석하고, 일정 수준 이상의 위험 인자가 발견되면 보호 필요 아동으로 추정해 각 읍면동으로 자동 통지하는 시스템이다. 통지를 받은 읍면동 공무원은 해당 아동의 집을 직접 방문해 양육환경을 확인하고, 복지서비스가 필요할 경우 드림스타트 등 서비스 제공기관에 아동학대가 의심되는 경우 경찰이나 아동보호전문기관에 연계한다. 매일경제, 2018. 3. 18; 이 시스템 구축 내용에 관하여는 최현수 외, 빅데이터를 활용한 e아동행복지원시스템 구축방안 기초연구, 보건복지부·한국보건사회연구원, 2016, 11, 201면 이하 참조.

1. 빅 데이터와 데이터 마이닝

먼저 빅 데이터(big data)⁴¹⁾는 주로 기술적 관점에서 “전통적인 데이터베이스시스템이 가지는 처리능력을 초월한 데이터”⁴²⁾나 “양, 속도, 종류가 모두 압도적인 정보 자원으로 고도의 통찰과 의사결정을 위한 효율적이고 혁신적인 정보처리 기술을 필요로 하는 데이터”⁴³⁾라고 정의한다. 다만 근년에는 기술적 관점에 사회적 관점을 가미한 정의가 유력하게 주장된다. 즉 “작은 규모에서는 해낼 수 없는 것을 큰 규모에서 실행하고, 새로운 지식의 추출이나 가치 창출에 의해 시장, 조직 및 시민과 정부의 관계 등을 변경하는 데이터”⁴⁴⁾라는 정의이다. 그것이 주는 광범위한 사회적 충격을 감안하면 데이터의 양 등에 더하여 거기서 얻을 수 있는 통찰 내지 추론·예측의 사회적 의미를 고찰하는 파악 방법이 필요할 것이다.

다음으로 데이터 마이닝(data mining)은 말 그대로 데이터(date)를 채굴(mining)하는 것을 의미하지는 않는다. 기본적으로는 데이터에서 지식을 채굴하는 것(knowledge mining from data)⁴⁵⁾을 의미하기 때문에 데이터 마이닝은 적절한 명칭(misnomer)이 아니라는 견해도 있다.⁴⁶⁾ 다만 여기서는 그 통속성에 따라 데이터 마이닝이라는 용어를 사용하고, 그 의미를 위와 같이 데이터에서 지식을 찾아내는 것으로 이해한다.

데이터 마이닝의 정의에는 여러 가지가 있다. “데이터를 고문해서 무언가 자백하게 하는 것”⁴⁷⁾이라는 회화화한 정의도 있지만, “데이터 속에 숨어 있는 패턴과 미묘한 관계성을 명확히 하고, 장래 결과 예측을 가능하게 하는 규칙성을 추론하기 위하여 통계학적 분석이나 모델링과 같은 데이터베이스 기술과 기교를 적용하는 것”⁴⁸⁾이라는 미국 회계감사원(Government Accountability Office, GAO)의 정의나 “방대한 데이터 세트(data set) 속에서 지금까지 알려지지 않은 유효한 패턴이나 관계성을 발견하기 위하여 통계학적 모델, 수학적 알고리즘, 기계적 학습법과 같은 세련된 데이터 분석법을 이용하는 것”⁴⁹⁾이라는 의회

41) 빅 데이터는 정보통신기술(ICT)의 진전에 의해 생성·수집·축적 등이 가능·용이해지는 다종다양한 데이터를 활용함으로써 가까운 미래의 예측 등을 통해 이용자의 개별 수요에 맞는 서비스 제공, 업무 운영 효율화나 신산업 창출 등이 가능하다.

42) 이 정의는 다음과 같이 이어진다. “데이터가 매우 크고 빨리 움직이기 때문에 당신의 한정적 데이터베이스 구조에는 적합하지 않다.” E. Dumbill, What is big data? : an introduction to the big data landscape, O'Reilly (January 11, 2012), <https://www.oreilly.com/ideas/what-is-big-data>.

43) IT Glossary: Big Data, Gartner, <https://www.gartner.com/it-glossary/big-data/>.

44) 빅토르 마이어 쾨버거·케네스 쿠키어, 위의 책, 14면.

45) Jiawei Han & Micheline Kamber, Data Mining: Concepts and Techniques 6(3Rd ed. 2011).

46) Id.

47) Jeff Jonas, What is Data Mining? Depends Who You Ask ..., Jeff Jonas (Sept. 8, 2006), https://jeffjonas.typepad.com/jeff_jonas/2006/09/what_is_data_mi.html.

48) U.S. Gen. Acct. Off., GAO-04-548, Data Mining: Federal Efforts Cover a Wide Range of Uses(2004).

49) Jeffrey W. Seifert, Crs Report for Congress: An Overview: June 7, 2005 - RI31798.

조사국(Congressional Research Service, CRS)의 정의가 주목된다.

2. 논의의 쟁점

여기서 주의할 것은 데이터 마이닝 자체가 개인의 권리와 직접 관련되지 않는다는 것이다. 앞에서 언급한 일반적 정의에서 짐작할 수 있는 바와 같이, 데이터 마이닝은 방대한 데이터 세트(이른바 ‘빅 데이터’) 중에서 통계적으로 유의미한 패턴이나 관계성을 추출·발견하는 것으로, 사용하는 개별 데이터가 누구의 것인가에 대하여는 관심이 없기 때문이다. 예컨대, 데이터 마이닝에서 분명한 것은 “무향료 로션 구입 경력 + 특정 보충제 구입 경력 + 큰 가방 구입 경력 → 당사자는 임신부일 가능성이 높다”⁵⁰⁾거나, “남성 + 녹갈색 눈동자 + 경범죄 경력 + 문신 → 당사자는 중죄를 범할 가능성이 높다”⁵¹⁾라는 패턴 내지 관계성이지만, “누가 임신부인지” 혹은 “누가 중죄를 범하기 쉬운지”가 아니다. 요컨대, 사용되는 빅 데이터가 익명화되어 있는 한 데이터 마이닝이 특정 개인과 결부되는 것은 거기서 추출·발견된 패턴 등이 개인식별정보(personally identifiable information)를 포함하는 데이터베이스에 ‘적용(apply)’되는 단계이다. 빅 데이터에서 찾아낸 위와 같은 패턴이 데이터베이스에 적용될 때 비로소 특정한 누군가의 경향이나 속성 등을 추단할 수 있다.

이에 의할 때 분명한 것은 개인의 권리와 직접 관련되지 않아도 되는 — 법적 분석 대상으로 하지 않아도 되는 — 데이터 마이닝도 존재한다는 것이다. 예컨대, 지역에 착안한 예측은 거기서 사용되는 빅 데이터가 익명화되어 있는 한 개인의 권리와 직접 관련되지 않는다. 거기서 표적이 되는 것은 어디까지나 ‘장소’나 ‘시간’이지 특정의 ‘개인’이 아니기 때문이다. 그러나 앞서 소개한 데이터 마이닝의 실례는 그렇다고 할 수 없다. 예컨대, 시카고 경찰의 ‘heat list’는 전통적인 지리적 범죄 매핑(mapping)을 넘어서 경찰이 조기에 개입해야 하는 개인을 특정하는 것이다(지역→사람). 데이터 마이닝 자체, 즉 데이터 세트에서 종전에는 알려지지 않은 패턴 내지 관계성을 추출·발견하는 것은 개인의 권리와 직접 관련되는 것은 아니지만, 여기서는 패턴 등이 개인에게 — 엄밀하게는 개인식별정보를 포함하는 데이터베이스에 — 적용되어 개인의 경향이나 속성이 판단된다.

이렇게 보면 경찰에 의한 빅 데이터 이용이나 데이터 마이닝은 법적으로 문제라는 식의 추상적이고 포괄적인 주장은 이에 관한 논의를 발전시키는데 별 도움이 되지 않는다. 오히려 법적인 검토가 필요한 것은 데이터 마이닝 자체라기보다는 그 결과 — 패턴이나 관계성 — 을 개인에게 적용하는 것, 즉 발견된 패턴 등에 들어맞는 누군가를 탐색하는 것이다. 물

50) 예컨대, 수건을 한꺼번에 많이 사거나 민감성 피부용 세제나 DHA를 포함한 비타민제, 대량의 보습제를 구입하는 사람은 임신부일 가능성이 매우 높다고 한다(찰스 두히그 지음 / 강주현 옮김, *겔리온*, 2012).

51) Robertson, *supra* note 39.

론 이러한 결론에도 몇 가지 문제점은 있다. 예컨대, ① 패턴 등의 분석·발견의 단계(데이터 마이닝 자체) — 익명 단계 — 와 ② 이러한 패턴 등의 적용 단계 — 현명 단계 — 로 분류했지만, 프로그램 혹은 시스템상 양자를 분리하기는 어렵다(추출하면서 적용하는 연속적 작업 공정)는 지적도 가능하다. 이와 같이 ①과 ②를 분리할 수 없거나 하지 않는다면 양자를 하나로 파악하고 이를 있는 그대로 — 즉 데이터 마이닝도 포함하여 — 법적 검토 대상으로 할 수밖에 없을 것이다. 그래서 법적인 관점에서는 양 단계를 의도적으로 구분하여 논의할 필요가 있다.

또한 본래 ‘적용’도 법적 검토 대상일 수 없다는 지적도 가능할 것이다. 실제로 개인에 대한 적용 단계에서도 경찰관과 대상자의 물리적 접촉은 없다. 적용 결과에 기초하여 경찰관이 실제 행동으로 옮기는 단계 — ③ 행동·개입의 단계 — 에 이를 때 비로소 대상자는 자신이 경찰의 관심사로 되어 있다는 것을 구체적으로 알게 되고, 또한 무언가 현실적인 불이익이 생기게 되기 때문이다. 그렇다면 법적으로 통제해야 하는 것은 적용 결과에 기초한 경찰관의 구체적 개입의 단계이고, 그 이전의 ①·②단계에서는 없다고 할 수도 있다. 이 점도 결코 간과해서는 안 된다. 이하에서는 ‘적용’이 어떠한 헌법적 권리와 가치에 저촉하는지를 검토함으로써 이러한 지적에 답하고자 한다.

3. ‘적용’의 침해적 성격

이하에서는 데이터 마이닝에 의해 얻은 결과(특정 패턴 내지 관계성)를 개인(개인식별정보를 포함한 데이터베이스)에게 적용하는 것, 즉 데이터 마이닝 결과를 이용하여 특정 개인의 성질이나 경향을 도출하는 것이 어떠한 헌법상의 권리나 자유, 이익을 침해하는지에 대하여 논의한다.

(1) 사생활의 보호 및 사회적 인격상에 관한 자기결정권

우선 들 수 있는 것은 헌법 제10조에 의한 일반적 인격권이다. 일반적 인격권의 보장 내용은 ‘사생활의 보호’와 사회적 인격상에 관한 자기결정권으로 나눌 수 있다. 인격의 자유로운 발현을 위해서는 한편으로는 인간이 독자적인 개성을 자율적으로 형성할 수 있는 개인적 생활 영역인 사생활의 보호를 필요로 한다.

사생활의 비밀이란 개인의 사생활 영역이 당사자의 의사에 반하여 공개되지 아니할 권리, 즉 사생활 영역으로부터 당사자의 의사에 반하여 정보를 수집하는 것에 대한 보호를 제공하는 기본권이다. 이러한 사생활 비밀은 사생활 정보에 관한 것이고, 궁극적으로 사생활 정보 자기결정권의 문제이다.

사생활이나 사적인 일을 몰래 엿보는 행위가 사생활의 자유를 침해한다는데 이론이 없다

고 하면, 데이터 마이닝 결과의 적용은 이미 보유한 단편적 개인 정보에서는 분명하지 않은 정보 주체의 사적인 일을 새롭게 알고자 하는 행위 — 데이터 개입을 통한 엿보기 — 로서 사생활의 비밀과 자유를 침해하는 것이라고 할 수 있다. 예컨대, 임신부인지 여부(특히 외견상 특징이 나타나지 않는 임신 초기)는 사적인 일에 속하는 사항으로 통상 본인의 동의 없이 이것을 알기 위해서는 높은 장애물을 넘어야 한다. 그러나 데이터 마이닝 결과의 적용은 언뜻 보면 상관성을 인식할 수 없는 개별 상품의 구입 경력 등에서 본인과 어떠한 접촉도 없이 그 사실을 알 수 있다. 물론 여기서의 결과는 그 정보 주체가 임신부일지도 모른다는 가능성에 불과하다는 반론도 가능할 것이다. 패턴의 적용에 의해 표출된 결과는 어떤 사람이 임신부라는 사실 내지 진실 그 자체가 아니라 분석자의 창조적인 지적 작업에 기초하는 평가라는 반론이다.

그러나 공개된 내용이 사생활상의 사실 혹은 사실인 것처럼 받아들여질 가능성이 있는 사항이거나 일반인의 감수성을 기준으로 할 때 당해 사인의 입장에서 공개를 원하지 않을 것으로 인정되는 것이라면 사생활의 자유의 침해로 볼 수 있다.⁵²⁾ 그렇다면 어느 정도 담보된 알고리즘에 의해 도출된 결과는 진실인 것으로 해석되는 정보로 이해할 수 있기 때문에 사생활 자유의 침해를 구성한다고 할 수 있다.

두 번째로 사회적 인격상을 형성할 수 있는 개인정보에 관한 자기결정권 침해가 문제될 수 있다. 개인의 인격은 사회 내에서, 즉 외부 세계와의 접촉과 상호작용을 통하여 형성되고 발현된다. 이 경우 사회는 개인에 관한 정보로부터 그에 대한 일정한 사회적 인격상을 형성하게 된다. 따라서 어떤 개인 정보가 어떤 관계에서 사회적으로 공개되는지, 즉 어떤 방법으로 처리되고 사용되는지의 문제는 매우 중요하다. 자유로운 인격 발현을 위해서 개인은 자신이 외부 세계에 어떻게 묘사되고 있는지를 스스로 결정할 수 있어야 한다.⁵³⁾ 그것은 본인이 자유의사에 기초하여 선택하고 행동하기 전에 그 개인을 통계적으로 판단하고 있기 때문이다.

앞서 검토한 임신부인지 여부를 추측하는 알고리즘의 적용은 대상자의 과거 내지 현재의 사실 혹은 대상자 자신이 알고 있거나 자각하고 있는 사실을 산출하고자 하는 것이었다. 다른 한편 앞서 소개한 'heat list'와 같이 어떤 자가 중죄를 범할지를 예측하는 알고리즘의 적용은 대상자의 미래에 속하는 사항 혹은 대상자조차 알 수 없거나 자각하지 못하는 사항을 예측하고자 하는 것이다. 이와 같이 본인조차 알 수 없는 그 자의 경향, 무의식적 본능(심적 대상)을 엿보고 장래의 행동을 예측하는 것은 과거나 현재의 사실을 산출하는 것과는 차원이 다른 문제를 제기하는 것이다. 이러한 개인에 대한 예측은 사람다움, 즉 정체성(identity)을 위협하는 것이라고 할 수 있다. 빅 데이터에 기초하는 분석은 자신이 그렇게 결정하기

52) 프라이버시 침해는 다수의 경우 허위의 사실이 뒤섞여 그것이 진실인 것으로 받아들여짐으로써 발생한다.

53) 한수웅, 헌법학, 법문사, 2011, 532면.

전에 내가 누구인지를 조정하고, 결정(determinate)하도록 하는 제도적 감시를 가능하게 하는 것⁵⁴⁾이다.

또한 스몰 데이터(Small Data)⁵⁵⁾에 기초하는 고전적 프로파일링(profiling)이 집단에 기초하는 분석을 중심으로 특정 집단에 대한 차별과 연결 짓는 것에 대하여, heat list에 기초하는 예측은 개인에 기초하는 치밀한 분석을 중심으로 하기 때문에 차별적 요소나 편견을 배제할 수 있다고 한다.⁵⁶⁾ 그러나 개인에 대한 보다 정확한 파악이 가능해지기 때문에 실제 <나=X>와 예측알고리즘에 의해 동정된 <나=X'>의 간극이 축소되고 - <나=X>가 <나=X'>로 된다 - 결국 <나=X>는 확률이라는 감옥에 갇히게 된다고 지적한다.⁵⁷⁾

이러한 견해에 의하면 데이터 마이닝 결과를 이용한 개인의 성향 파악이나 행동 예측은 헌법의 근본원리인 개인의 존중이나 인간 존엄의 원리에 반할 수 있다.⁵⁸⁾ 이는 자율적인 개인으로서의 주체성을 부정하는 것이라고 할 수 있다. 개인의 자율적·주체적 판단·선택이나 실제 노력을 배제하고 오로지 통계학적 예측에 따라 그 개인이 어떤 사람인지가 결정되기 때문이다.⁵⁹⁾ 빅 데이터에 의한 개인의 예측은 내가 어떤 사람인가를 결정하는 기본적 권리를 위태롭게 하는 것으로, 예측을 위한 적용은 헌법의 근본원리인 인간의 존엄성에 반하며, 또한 각자는 자신의 생각과 사고방식에 따라 삶을 결정하고 살아갈 수 있는 헌법의 본질적 원리에도 반한다.⁶⁰⁾

54) Richard and Kings, supra note 1, at 422.

55) 빅토르 마이어 쾰버거·케네스 쿠키어, 앞의 책, 293면.

56) See e. g., Joh, supra note 4, at 57-59.

57) 빅토르 마이어 쾰버거·케네스 쿠키어, 위의 책, 296면.

58) 빅토르 마이어 쾰버거·케네스 쿠키어, 위의 책, 36면.

59) 차별이 집단에 대한 편견이나 고정 관념에 의해 비과학적으로 개인을 결정짓는 것이라고 하면 알고리즘에 의한 예측은 빅 데이터의 힘에 의해 고도로 과학적이고 통계학적으로 개인을 결정짓는 것이라고 할 수 있다. 그것들은 개인의 존엄과도 관련한다.

60) NTT東日本은 2018년 6월 하순부터 벤처기업과 공동 개발한 'AI 경비원' 서비스를 출시하였다. AI 기능을 탑재한 카메라가 실시간 CCTV 영상을 분석해 점포 안팎 사람들의 수상한 행동을 분석한다. 범죄를 저지르기 전의 행동을 감지해 범죄를 사전에 예측하고 방지하는 것이다. 만약 특정인이 같은 장소 주변을 계속해 서성이거나 주위를 여러 번 둘러보는 행동을 하면 AI 카메라는 이상을 감지하고 장소나 사진 등을 점원의 스마트폰으로 전송한다. 통지를 받은 점원은 그 즉시 스피커를 통해 수상한 사람에게 "무엇을 도와드릴까요?" 등의 말로 대응할 수 있다. 카메라가 인파의 수상한 행동을 감지하는데 필요한 패턴 파일은 보유한 과거의 소비자 행동 데이터를 기반으로 만들어졌다. AI카메라는 머신러닝을 통해 범죄 행동의 특징과 제품 특성들이 다른 다양한 점포의 행동 데이터를 학습한다. 또 클라우드를 이용해 이 패턴 파일을 정기적으로 업데이트하면 고객층의 변화와 신종 수법의 출현이 동반되더라도 대응할 수 있다(<https://dcross.impress.co.jp/docs/news/000545.html>). 또한 2019년 4월 AGC, DeNA, NTT도코모는 버추얼 캐릭터를 활용한 경비시스템 시작기(試作機)를 개발했다고 발표했다. 각종 센서를 탑재하는 미러 디스플레이에 표시된 사람 크기의 3D캐릭터가 건물 내 접수 업무나 경계 감시를 하며, 2020년 봄 상용화를 목표로 한다(<https://www.itmedia.co.jp/news/articles/1904/25/news123.html>).

(2) 양심의 자유

이러한 예측은 헌법상 양심의 자유에도 저촉할 수 있다. 어떤 사람의 과거나 현재가 아니라 미래의 행동을 예측하는 행위는 행동의 전 단계인 내심 영역의 동향을 엿보는 행위라고 할 수 있기 때문이다. 헌법의 기본권 체계 내에서 양심의 자유의 기능은 개인적 인격의 정체성을 유지하는데 있다. 따라서 양심의 자유를 제한하는 행위는 개인의 윤리적 정체성을 침해하는 행위이다. 예컨대, 어떤 자가 중죄를 범할지도 모른다는 예측은 본인조차도 신경 쓰지 않는 범죄에 대한 무의식적인 욕구를 가시화하는 행위이다. 물론 양심의 자유의 보장 범위에 대하여는 여러 종류의 논의가 있지만, 만약 이를 넓게 파악하면 특정의 행위로 연결되는 내심 영역의 동향이나 경향을 — 빅 데이터를 활용한 높은 정밀도로 — 산출하고자 하는 적용 행위는 이 자유를 침해한다고 해석할 여지는 있을 것이다.

물론 데이터 마이닝 결과의 적용에 의해 도출되는 개인의 예측은 어디까지나 통계적인 추측이고, 그것에 의해 개인을 결정짓더라도 개인의 내심의 동향을 완전히 알 수는 없다 — 따라서 인간 존엄의 원리에 반하더라도 양심의 자유를 침해하는 것은 아니다 — 고 주장할 수도 있다. 그러나 인간이 합리적 정확성이 검증된 예측 알고리즘의 적용 결과에 대하여 얼마나 회의적·비판적일 수 있는지 의문이다. 비과학적 고정 관념으로부터 도피하지 못하는 인간이 과학적인 확률로부터 자유롭다고 할 수 있을지도 의문이다. 사실 인간은 의외로 데이터의 독재에 지배되기 쉽다는 지적은 경시할 수 없을 것이다.⁶¹⁾ 이러한 관점에 의할 때 현실의 $\langle 나=X \rangle$ 가 데이터에 기초하는 $\langle 나=X' \rangle$ 로 대신하거나 혹은 내심 영역에 장래 특정의 행동으로 결부하는 본능이 존재하는 것으로 받아들이는 것은 가능할 것이다.⁶²⁾

(3) 소결

이상의 논의에 의할 때 데이터 마이닝과 관련한 “분석→적용→개입” 과정 중에서 ‘적용’은 헌법상 자유와 권리를 침해할 수 있다. 그러나 문제는 적용 결과에 따라 경찰관이 실제로 어떻게 행동하느냐(개입 단계)이고, 거기서 비로소 권리이익과 구체적으로 충돌한다고 할 수 있다. 물론 이러한 견해에 있어서 적용 결과를 그대로 수용하더라도 제3자에 의한 통제가 미치기 때문에 경찰에 의한 빅 데이터 이용 내지 데이터 마이닝은 새로운 법적 문제를 유발하지 않는다고 생각할 수도 있다. 예컨대, 경찰관이 X가 중죄를 범할 가능성은 75%라는 적용 결과에만 기초하여 X에 대한 강제처분을 하는 경우 법원에 의한 심사가 개

61) 빅토르 마이어 쉰버거·케네스 쿠키어, 앞의 책, 301-302면.

62) 개인의 예측 결과를 다루는 것이 일반인이 아니라 전문가라면 이러한 결과를 냉정하게 받아들이고 $\langle 나=X \rangle$ 와 $\langle 나=X' \rangle$ 사이의 간극을 인식하는 것은 가능할지도 모른다. 그렇다면 데이터 마이닝 결과의 적용은 개인의 권리이익이나 인간 존중 원리에 저촉된다고 하기보다 단순한 정보활동으로 파악할 수 있을지도 모른다.

시되고 영장이 발부되지 않을 수 있다. 이와 같이 강제처분 단계에서 사법적 통제가 충분히 미친다면 적용이 X의 권리이익을 직접 침해하는 것은 아니라고 할 수 있다.⁶³⁾

그러나 법원의 통제가 미치는 것은 적용에 기초하는 경찰관의 행동이 가시화되는 단계라는 것에 주의해야 한다. 현상의 법제도를 전제로 하면 구체적으로 개입에 이르기 전의 표적화(예컨대, 본인이 의식하지 못하는 감시의 단계)에 대하여는 법원에 의한 실질적인 통제가 미치지 않는 것이 되기 때문이다. 이러한 귀결은 ‘불가시적 표적화’도 개인의 행동을 위축시킬 수 있다는 것을 근거로 하면 문제일 수 있다. 예컨대, 이탈(deviant)로 볼 수 있는 특징을 여럿 가지는 자가 적용에 의해 범죄 위험성이 높다는 예측을 받게 되면 경찰에 의한 중점 감시 대상이 된다는 것을 의식하고, 그로 인해 행동을 조심하거나 표준화할 가능성은 충분하다. 이렇게 볼 때 구체적인 개입 이전의 ‘적용’ 행위에 법적 문제점을 찾고, 이를 적절히 통제할 필요성이 있을 것이다.⁶⁴⁾

IV. 예측 치안의 통제

이상의 논의에 의할 때 데이터 마이닝에 의해 추출·발견된 패턴이나 관계성을 ① 정보 주체의 과거·현재의 사적 일을 엿보기 위하여 혹은 ② 정보 주체의 미래 행동을 예측하기 위하여 적용하는 것은 헌법적 권리와 가치를 훼손할 수 있다. 이하에서는 이에 대한 통제 방안으로서 집합적 동의에 대하여 알아본다.

1. 개별적 동의

개인정보보호법에 따라 경찰의 적용 행위를 정보 주체에게 사전에 고지하고 동의를 얻으면 애초에 기본권 침해는 인정되지 않는다. 그러나 경찰에 의한 데이터 마이닝 결과의 적용을 생각했을 때 개별적 동의라는 접근은 다음 두 가지 한계가 있다. 첫째는 동의를 얻기 위해서는 경찰이 정보 주체에게 적용에 관한 정보를 사전에 고지할 필요가 있지만 그것이 수

63) 적용에 의해 본인이 직접 피해를 입는 것은 아니다. 그러나 프라이버시 침해 행위와 실질적 손해가 어긋나는 경우는 적지 않은 것으로 생각된다. 예컨대, A가 B의 침실에 도청기를 설치했다고 하자. B가 이에 의해 충격을 받고 깊은 상처를 받으면 A에 의해 도청기를 설치했다는 것을 알고 난 뒤 프라이버시 침해는 도청기를 설치한다는 A의 행위 그 자체에 의해 이미 생긴다고 할 수 있다.

64) 이것은 데이터에 이끌린 구체적 개입에 대한 사법적 통제의 중요성을 부정하지 않는다. 법원은 강제 처분을 인정하거나 그 필요성, 합리성을 신중하게 고려하는 것은 물론, 원칙으로서 적용 결과, 즉 통계적 추측 이외의 근거를 찾아야 하고, 적용 결과에 중점을 두는 경우에도 (알고리즘) 전문가의 조언 등을 들으면서 데이터의 신뢰성이나 그 처리 과정의 완전성을 파악해야 한다. 다만 거기에는 처리 과정이나 알고리즘 등이 법원에 제시될 필요가 있다.

사상 비밀의 개시 등에 관련되고 범인의 도주나 증거 인멸 등 공공의 안전과 질서유지에 중대한 지장을 초래할 수 있다는 것이다. 또한 사전 고지를 하지 않더라도 실제 범인은 동의를 거부할지도 모르고 범인이 아닌 자(동의를 거부한다고 의심되기 때문에)는 동의를 거부할 수 없을지도 모른다(동의를 사실상 강제).

두 번째는 현실적인 어려움이 있다는 것이다. 여기서는 편의상 데이터 마이닝에 의해 추출·발견된 패턴 등을 개인에게 적용한다고 표현했지만, 엄밀히 말해서 그것은 개인식별정보를 포함한 데이터베이스에 대하여 적용하는 것을 의미한다. 데이터베이스에 패턴 등을 적용하고 거기서 위험성이 높은 사람을 추려 내는 것이 일반적이라고 할 수 있다. 그렇다면 적용은 최종적으로는 개인을 특정하는 것이라고 하더라도 현실적으로는 데이터베이스에 포함되는 모든 자를 대상으로 하게 된다. 데이터 마이닝의 결과는 빅 데이터에서 추출되고 빅 데이터에 적용되는 것이다. 이와 같이 생각하면 적용의 동의는 대상이 되는 데이터베이스에 포함되는 모든 자로부터 얻어야 하지만, 이것은 사실상 불가능하다.

결국 방대한 양의 데이터를 다루는 빅 데이터 시대에 “프라이버시는 죽었다(Privacy is dead)”라는 표현이 지나칠지 모르지만, 인간이 개별 정보의 동향을 파악하고 각각 본인으로부터 동의를 얻는 것(본인의 실질적·사실적 통제를 인정하는 것)은 사실상 불가능할 것이다. 그러나 이것이 이념으로서의 ‘자기정보통제’, 이념으로서의 ‘동의’를 부정하는 것이 되어서는 안 된다. 본인의 통제나 동의가 사실상 불가능하다고 해서 적용 행위의 권리 침해성이 부정되는 것은 아니다. 그렇다면 적용 행위의 침해성을 인정하고 그 헌법적 정당화의 필요성을 인정한 상황에서 이념으로서의 통제나 동의를 다른 형태로 실현하는 것이 요구된다고 할 수 있다. 즉 집합적 동의로서의 법률 제정과 본인의 통제 대신에 그 적절한 운용을 담보하기 위한 방법이나 구조를 마련하는 것이다.

2. 집합적 동의

앞서 언급한 이유 때문에 ‘적용’에 대한 개별적 동의가 어렵다면 적용의 침해적 성격을 인정하고 그 헌법적 정당화를 도모할 필요가 있다. 이 점에서 우선 생각할 수 있는 것은 ‘집합적 동의’라고 할 수 있는 ‘법률’에 의해 이를 민주적 절차에 따라 승인하는 것이다. 다만 이 법률의 제정은 개별적 동의의 대체라는 소극적 의미를 뛰어넘는 의미를 가질 수 있다. 예컨대, 개인의 행동 예측을 목적으로 하는 적용은 경찰에 의한 사전적·예방적 개입을 널리 허용하게 되고, 국가의 존재 방식, 사회의 존재 방식을 근본적으로 변경하는 것 — 이른바 예방국가로 전환하는 것 — 이 된다. 또한 데이터 내지 통계에 의해 개인의 행동을 예측하고, 예컨대 잠재적 범죄자로서 낙인찍는 것은 개인을 이성적 혹은 자율적 행동 주체로 생각하는 근대입헌주의 국가의 존재방식을 변경하는 것이기도 하다. ‘적용’이 인간의 존중이라는 헌법의 기본적 가치에 반한다는 것은 바로 이 점과 관련한다.

이와 같이 경찰에 의한 데이터 마이닝 결과의 적용은 개인의 자유와 권리를 침해할 뿐만 아니라 국가와 개인의 관계나 사회의 존재 방식을 변경하는 힘을 갖게 된다. 그렇다면 그 정당화에는 우리 자신의 동의, 달리 표현하면 안전을 위하여 경찰에 의한 사전적·예방적 개입을 수용한다는 우리 자신의 각오 표명, 즉 법률이 필요하다고 할 수 있다. 이리하여 집합적 동의로서의 법률은 현실적으로 취득이 곤란한 개별적 동의의 대체 수단으로서의 소극적 의미를 가질 뿐만 아니라 경찰도 법원도 아닌 우리 자신이 국가 내지 사회의 기본적인 존재 방식을 결정한다는 적극적 의미를 가질 수 있다.

경찰에 의한 데이터 마이닝 결과의 적용을 헌법상 정당화하는 데는 법률이라는 형식 외에 다음과 같은 장치들이 필요하다. 첫째, 적용을 실질적으로 정당화하는 이유가 요구된다. 예컨대, 절도를 신속하게 수사하거나 예방한다는 이유로 적용이 실질적으로 정당화될 수 있는가? 적어도 적용의 침해적 성격이나 인간 존엄의 원리와의 관계를 근거로 어떤 경우에 적용하는 것이 가능한지를 신중하게 검토하고 이를 법률에 명시할 필요가 있다.

둘째, 정보 주체인 본인에 의한 통제 대신에 적절한 운용(이유 내지 목적에 따른 적용)을 담보하기 위한 방법이나 구조를 마련할 필요가 있다. 예컨대 ① 보안시스템의 견고성, ② 징계처분 혹은 형벌에 의한 목적 외 이용·누설 등의 금지, ③ 감시기관 등 적절한 취급을 담보하기 위한 제도적 장치 등을 마련하고, 이들 조건이 충족되고 취급하는 본인 확인 정보가 남용·누설되는 구체적 위험이 없는지가 합헌성 판단의 조건이 될 수 있다. 경찰에 의한 데이터 마이닝이나 그 적용에 대하여도 이러한 구조적 조건이 필요하며, 취급하는 정보의 성질에 따라서는 이보다 엄격한 조건이 요구될 수도 있다.

예컨대, 데이터 마이닝 결과의 개인에 대한 적용은 그 자가 중죄 등을 범할 위험성이나 기능성에 관한 정보를 생산한다. 이 정보가 외부에 누설되었을 때 그 영향은 헤아릴 수 없으며, 그것만으로도 이러한 행위를 엄격히 금지할 필요가 있다. 또한 적용의 목적 외 실행이나 적용 결과에 대한 과잉 반응을 억지하기 위하여 적용하는 자와 그 결과에 접근할 수 있는 자를 축소함과 동시에 적용 결과는 어디까지나 통계적 추측에 불과하며 당연히 오차를 포함한다는 리터러시(literacy) 교육을 관계자에게 철저히 할 필요가 있다. 또한 감시기관의 전문성을 높이는 것도 중요하다. 데이터 마이닝이 공정하고 적절하게 행해지는지, 즉 데이터 마이닝에 의해 추출·발견된 패턴 등이 신뢰할 만한 것인지 등을 판단하는 데는 고도의 전문적 지식이 필요하기 때문이다. 사실 기술자·전문가에 의한 자의적 행위는 기술자·전문가가 아니고는 발견하기 어렵다.⁶⁵⁾

오늘날 컴퓨터는 무언가를 판단할 때 프로그램에 기술된 규칙에 따라 처리한다. 그래서 만약 컴퓨터에 이상한 움직임이 포착되면 프로그램에 어떤 규칙이 쓰여 있는지 점검하면

65) 기술자 혹은 코드 작성자(code writers)의 자의적 행위를 어떻게 예방하고, 그들의 권력화를 어떻게 통제할 것인가에 대하여는 Danielle Keats Citron, *Technological Due Process*, 85 Wash. L. Rev. 1249, 1254-1255 (2008).

된다. 그런데 빅 데이터 분석의 경우는 이와 같이 소급해서 조사하기는 대단히 어렵다. 알고리즘에 의한 예측은 매우 복잡하고 대부분의 사람들에게는 이해할 수 없는 것이 많기 때문이다. 따라서 빅 데이터 예측과 그 배후에 있는 알고리즘이나 데이터세트는 블랙박스화될 위험성이 있다. 책임 소재도 불분명하고 소급하여 조사하는 것도 불가능하기 때문에 신뢰하기 어렵다. 그렇게 되지 않도록 하기 위해서는 빅 데이터 감시와 투명화가 필요하며 그것을 위한 새로운 전문지식이나 제도를 마련해야 한다.⁶⁶⁾

그래서 알고리즘 전문가를 감시기관에 배치하는 것도 생각해볼 필요가 있다. 감시기관에게서 컴퓨터 사이언스나 수학, 통계학 분야의 전문가로 빅 데이터에 의한 분석, 예측 평가를 담당하는 알고리즘 전문가는 공평과 기밀 유지를 사명으로 하며, 정보원의 선택, 분석·예측 도구(알고리즘이나 모듈을 포함) 선정, 분석 결과 해석에 대하여 평가하고 문제가 발생한 경우 사용된 알고리즘이나 통계 기법, 데이터 세트를 조사하는 것이다.⁶⁷⁾ 경찰에 의한 데이터 마이닝 결과의 적용을 인정하는 경우에도 이러한 전문가를 감시기관 내에 배치하고 감시의 실질화를 도모해야 한다.

한편 데이터 마이닝이나 그 적용을 공정하고 적절하게 행하기 위해서는 그 과정에 관한 정보를 일정한 범위에서 공개하고 투명성을 확보하는 것이 중요하지만, 그렇다고 해서 경찰 내부의 데이터 처리 과정 전체를 공개해야 하는 것은 아니다. 예컨대, 테러를 예측하는 패턴이나 그 요소($A+B+C+D \rightarrow$ 테러할 가능성이 높다)를 공개하면 잠재적 테러리스트는 요소가 되는 행동(A, B)을 피하거나 요소가 되는 속성(C, D)을 변경함으로써 표적화를 피할 수 있다.⁶⁸⁾ 또한 중죄를 범할지 예측하는 패턴의 요소로 인종 등이 포함될 때 그것을 공개하면 인종 차별을 조장할 수도 있다.⁶⁹⁾ 이렇게 볼 때 경찰에 의한 데이터 마이닝 전체 과정의 투명화가 반드시 필요한 것은 아니며 그 중 몇 가지는 의도적으로 애매하게 할 필요가 있다. 즉 적용하는 목적이나 이유, 결과 정보에 접근할 수 있는 자의 범위, 남용·누설한 경우의 제재 등은 법률에 명기하여 공적으로 개시할 필요가 있지만 예측 알고리즘의 상세한 내용 등에 대하여는 오히려 공공에 대하여 비닉해야 할 것이다. 그렇다면 예측 알고리즘의 적절성을 정보 주체나 공중을 대신하여 심사하는 조직으로 알고리즘 등을 망라하는 감시기관의 역할이 대단히 중요하다. 즉 투명성에 한계가 있기 때문에 감시기관의 역할이 대단히 중요한 것이다.⁷⁰⁾

66) 빅토르 마이어 원버거·케네스 쿠키어, 앞의 책, 326-327면.

67) 빅토르 마이어 원버거·케네스 쿠키어, 위의 책, 328-329면.

68) See Zarsky, *supra* note 16, at 1553-1560.

69) *Id.* at 1560-1563.

70) *Id.* at 1558-1562, 1563-1564, 1566.

V. 결론

본고는 빅 데이터나 데이터 마이닝을 이용한 예측 치안 방법이 도입·실시될 때 그 헌법적 과제가 무엇인지, 왜 그것이 문제가 되는지 고찰하고, 이에 대한 통제 방법을 제시하고자 했다. 데이터 마이닝 결과 개인에 대한 적용, 특히 개인의 행동 예측을 목적으로 하는 것은 단순한 개인정보보호의 문제를 넘어서 사생활의 자유나 양심의 자유, 자율적 존재로서의 개인을 존중하는 헌법의 근본원리에도 관계되는 문제를 포함한다.

따라서 빅 데이터시대에 자기 정보 통제의 사실상의 불가능성이나 곤란성을 이유로 헌법적 규율 필요성을 약화시켜서는 안 될 것이며, 이러한 시대적 상황에 걸맞은 실효적 통제 방법을 구축해야 한다. 아울러 우리나라 경찰의 예측 치안의 동향을 주의 깊게 살펴보면서 이에 관한 논의를 심화시켜나갈 필요가 있다.

참고문헌

- 김덕현, 지금이라도 4차 산업혁명 대응정책 재정립해야, STARTUP4, 2018. 12.
- 김화중, 4차 산업혁명과 데이터 가치체계, 헬스케어 ICT정책 Vol. 06, 헬스케어미디어연구소, 2017. 3.
- 빅토르 마이어 쾨버거 · 케네스 쿠키어 지음 / 이지연 옮김, 빅 데이터가 만드는 세상: 삶과 일, 그리고 생각하는 방식을 바꿔놓을 대혁명, 21세기북스, 2013.
- 성낙인, 헌법학, 법문사, 2019.
- 양종모, “인공지능 이용 범죄예측 기법과 불심검문 등에의 적용에 관한 고찰”, 형사법의 신동향 통권 51호, 대검찰청 검찰미래기획단, 2016. 6.
- 양종모, “인공지능 알고리즘의 편향성, 불투명성이 법적 의사결정에 미치는 영향 및 규율 방안”, 법조 66권 3호, 법조협회, 2017. 6.
- 윤영미, “양심의 자유의 내용과 제한”, 인권과 정의 제345호, 대한변호사협회, 2005. 7.
- 윤해성 · 전현욱 · 양천수 · 김봉수 · 김기범 외, “범죄 빅데이터를 활용한 범죄예방시스템 구축을 위한 예비 연구(Ⅰ)”, 한국형사정책연구원, 2014. 12.
- 이용걸, “인공지능과 수사”, 치안정책리뷰 제61호, 치안정책연구소, 2018. 9.
- 최현수 · 오미애 · 전진아 · 김용대 · 김정희 · 김솔휘 · 천미경, 빅데이터를 활용한 e아동행복지원시스템 구축방안 기초연구, 보건복지부 · 한국보건사회연구원, 2016. 11.
- 치안정책연구소, 스마트치안지능센터 설립 기본구상 연구, 2016. 12.
- 탁희성 · 박준희 · 정진성 · 윤지원, “범죄 빅데이터를 활용한 범죄예방시스템 구축을 위한 예비 연구(Ⅱ)”, 한국형사정책연구원, 2015. 12.
- 한수웅, 헌법학, 법문사, 2011.
- 행정자치부, 개인정보보호 법령 및 지침 · 고시 해설, 2016. 12.
- 大山智也 · 雨宮 護, 地理的犯罪予測の手法間比較—日本型犯罪予測手法の構築にむけた検討—, GIS-理論と応用 25(1), 地理情報システム學會, 2017. 6.
- 守山 正, 犯罪予測技法の展開: 近接反復被害分析を中心として, 政治・経済・法律研究 20 (1), 拓殖大學政治経済研究所, 2017. 9.
- 犯罪・交通事象・警備事象の予測における ICT活用の在り方に關する有識者研究會, 犯罪・交通事象・警備事象の予測における ICT活用の在り方に關する提言書, 警視廳, 2018.
- 梶田眞實, 犯罪オープンデータを用いた犯罪予測アルゴリズムとシステムの開発, 第2回官民ラウンドテーブル, 株式會社Singular Perturbations 2018. 2. 28.
- Elizabeth E. Joh, Policing by Numbers: Big Data and the Fourth Amendment, 89 Wash. L. Rev. 35, 37 (2014).
- Jeffrey S. Paul and Thomas M. Joiner, Integration of Centralized Intelligence with Geographic Information Systems, Geography and Public Safety, Oct. 2011, at 7.
- Jennifer Bachner, Predictive Policing: Preventing Crime with Data and Analytics, Washington, DC: IBM Center for The Business of Government, 2013.
- Leslie W. Kennedy et al., Risk Clusters, Hotspots, and Spatial Intelligence: Risk Terrain Modeling as an Algorithm for Police Resource Allocation Strategies, 27 J Quantitative

Criminology 339, 342-343 (2011).

Neil M. Richards and Jonathan H. King, Big Data Ethics, 49 Wake Forest L. REV. 393, 397(2014).

Walter L. Perry et al, Predicting Policing: The Role of Crime Forecasting in Law Enforcement Operations, NIJ, 2013, at 19.

Zach Friend, Predictive Policing: Using Technology to Reduce Crime, FBI Law Enforcement Bulletin(Apr. 9, 2013).

투고일자 : 2019. 09. 05

수정일자 : 2019. 09. 25

게재일자 : 2019. 09. 30

<국문초록>

예측치안의 헌법적 논의

이 병 규

본고는 근년 정보통신기술(ICT)의 비약적 발전에 의한 경찰 업무의 변화 속에서 행해지고 있는 예측치안(predictive policing)을 헌법적 관점에서 다루었다. 특히 경찰의 법집행이나 범죄 예방 영역에서 빅 데이터가 어떻게 이용되고, 또한 그것에 헌법적 문제는 없는지 고찰하였다.

이를 위해 먼저 경찰 업무 영역에서 예측 치안이 어떤 형태로 이용되는지 미국의 실례를 통해 알아보았다. 범죄 장소와 시간 등의 데이터를 이용한 지역예측치안과 대인적 표적 시스템에 의한 개인예측치안이 그러한 예이다. 그리고 이러한 빅 데이터에 의한 예측치안이 헌법상 자유와 권리를 침해하는지 살펴보았다. 즉 데이터 마이닝에 의해 추출·발견된 패턴이나 관계성을 정보 주체의 과거·현재의 사적 일을 엿보기 위하여 혹은 정보 주체의 미래 행동을 예측하기 위하여 적용하는 것은 헌법적 권리와 가치를 침해할 수 있다.

마지막으로 이러한 빅 데이터를 이용한 예측치안이 도입·실시될 때 헌법적 문제가 무엇이며, 그 통제 방법은 무엇인지 고찰하였다. 데이터 마이닝 결과의 개인에 대한 적용은 단순한 개인정보보호의 문제를 넘어서 개인을 존중하는 헌법원리에도 관계되는 문제이기 때문에 단순한 개별적 동의를 넘어서 집합적 동의, 즉 법률에 의한 민주적 절차에 따라 이를 승인하는 실효적 통제 방법을 통해 헌법적 정당화를 도모할 필요가 있다.

주제어: 예측치안, 정보통신기술, 빅 데이터, 데이터 마이닝, 헌법, 개인정보보호