

영향 받는 사람의 목소리

연속 워크숍



공권력 시의 현황과 대응- 경찰 시와 출입국 시

일시 2026년 7월 14일(화) 오후3시

장소 민주사회를위한변호사모임 대회의실(지하1층)
서울특별시 서초구 서초대로46길 74 ※온라인 생중계(줌) 병행 예정

사회 장여경(정보인권연구소 상임이사)

발제 박병욱(민주주의법학연구회, 제주대학교 행정학과 교수)/ 경찰분야
이일(공익법센터 어필, 변호사)/ 이주분야

토론 랑희(공권력감시대응팀 활동가)
최호웅(민주사회를위한변호사모임 디지털정보위원회 위원장, 변호사)
고기복(모두를 위한 이주인권문화센터 대표)
허오영숙(한국이주여성인권센터 대표)

주관 디지털정의네트워크
정보인권연구소

공동주최

공권력감시대응팀, 공익법센터 어필, 민주사회를위한변호사모임 디지털정보위원회,
민주주의법학연구회, 외국인이주·노동운동협의회,
인공지능 책임성과 공공성 강화를 위한 시민사회 공동행동, 한국이주여성인권센터

후원 아름다운재단

AI와 경찰활동

박 병 옥*

- I. 들어가며
- II. 경찰활동에 AI 활용이 내포하는 법적 문제점
- III. 경찰의 AI 활용 관련 독일 연방헌법재판소 2023년 결정
- IV. 경찰의 위협판단과 AI
- V. 경찰영역 고위험 인공지능에 대한 규제
- VI. 나가며

I. 들어가며

오늘날 ‘인공지능(Artificial Intelligence)’의 줄임말인 ‘AI’는 단순히 인간이 설정한 프로그램에 따라 인간이 하던 고된 노동을 대체하는 로봇의 단계를 넘어 ‘빅데이터(Big Data)’를 활용하여 일정한 알고리즘에 따라 자동적으로 필요한 정보를 취사선택하여 유의미한 관계, 규칙패턴을 발견하고 지식을 추출함으로써 인간의 인지적, 지적 판단 영역을 대체 내지 보완하는 기술로 이해되고 있다. 또한, AI는 무엇보다 스스로 학습할 수 있는 능력을 가지고 있고 인간의 행동을 모방한다는 점에서 기존의 ‘사물인터넷(Internet of Things)’ 논의를 넘어선다.

경험, 편견, 인지능력의 한계에 노출되어 있는 인간의 부족함을 보완할 수 있는 AI의 인지, 논리적 추론능력이 인간 상호관계, 특히, 인격권, 사생활, 정보자기결정권 등 기본권과 관련하여 논란을 일으키지 않는 가운데 인간을 위한 유익한 서비스를 제공하는데 이용된다면 이는 AI 기술발전의 긍정적 측면이라고 할 수 있을 것이고, 실생활에서 충분히 적극적으로 활용되어도 좋을 것이다¹⁾. 하지만, AI의 인지, 판단, 학습을 위한 과정에서 인간의 행동, 생활이 감시된다거나, 나아가 이러한 AI 인지, 판단, 학습의 결과가 인간의 사생활, 일반적 행동의 자유 등 기본권 행사에 제약으로 작용한다면 이와 같은 부분은 조심히 다루어질 필요가 있다. 심한 경우라면 이와 같은 AI

* 제주대학교 행정학과 부교수

1) 물론 이와 같은 인간을 위한 AI의 서비스 제공도 기존에는 인간이 담당하던 일자리, 즉, 생활의 경제적 기반이 되는 직업을 줄인다는 점에서 항상 좋은 측면만 있는 것은 아니다.

의 사용은 제한되어야 할 것이다.

일반적으로 경찰영역에 AI를 활용하는 것은 범죄피해자 또는 범죄·위험에 대한 잠재적 피해자로서의 일반 시민의 입장에서는 그들을 보호하기 위한 것으로서 유익한 것이 되는 경우가 많겠지만, 가해자 또는 위험악기자의 입장에서는 그 개인의 기본권을 제한하는 측면이 쉽게 연상될 수 있다. 왜냐하면, 후자의 경우 AI가 가해자 또는 위험악기자의 위험한 행동(위험방지의 측면) 또는 과거의 행적(수사의 측면)을 분석하기 위해서 가해자 또는 위험악기자 개인 또는 주변 관련자들을 관찰, 감시하거나, 필요한 정보를 수집, 처리하는 과정이 필요할 것이기 때문이다. 이런 활동에 따라 침해되는 기본권으로 인간의 인격권, 명예, 프라이버시, 개인정보자기결정권 등을 쉽게 떠올릴 수 있다.

위험방지 및 범죄예방 영역에서 AI를 활용한 ‘예측적 경찰활동(Predictive Policing)’ 뿐만이 아니라 범죄수사, 가정폭력·아동학대 영역에서 접근금지 등 재범우려자에 대한 경찰처분을 위한 AI의 활용이 오늘날 법적인 논란의 소재가 되고 있다. 이 글에서는 경찰활동에 있어서의 AI의 활용을 중심으로 다루겠지만, 법원에서도 재범위험성을 감안하여 각종 보안처분을 행하는데 있어서 전세계적으로 AI가 활용되고 있는바 이에 대해서도 필요한 경우 참고하도록 한다.

인간의 인지적 한계를 극복하여 주는 AI의 활용은 인지, 판단 과정에서의 효율성 강화를 중심으로 놓고 본다면 경찰 위험방지 활동에서의 생명·신체 보호 기능의 강화, 수사 영역에서의 실체적 진실 발견의 증진을 가져올 수도 있다. 하지만, AI는 필연적으로 개인정보가 포함될 수도 있는 자료를 분석, 평가 내지 학습을 하는 과정, 그리고 그러한 과정에 따른 결론을 적용하는 과정에서 각종 기본권 침해의 위험성을 내포하고 있는바, 이러한 위험성이 일정한 기준에 따라 적절하게 통제되는 경우에 비로소 AI의 활용이 허용될 수 있다고 생각된다.

2022년 경찰청이 발표한 「경찰미래비전 2050」은 AI의 장점에만 주목하여 그것이 내포하고 있는 인격권, 사생활, 명예, 개인정보자기결정권에 대한 침해 등에 대하여 경시하고 있다는 우려를 지울 수 없는바 이 글에서는 자동화된 데이터분석, 즉, AI의 자료 분석, 평가 내지 학습이 가지는 기본권 침해를 고려하여 그에 대한 법적인 기준을 고민해 보고자 한다.

II. 경찰활동에 AI 활용이 내포하는 법적 문제점

1. 경찰활동의 광범위성

경찰활동의 범주는 매우 광범위하다. 중요시설경비, 혼잡경비, 집회시위 안전관리, 교통안전관리, 주취폭력 대응 등 위험방지 영역 뿐만이 아니라 범죄수사, 그리고 위험방지와 범죄수사의 중간 영역이라고 할 수 있는 범죄행위에 대한 예방적 대처의 영역을 포괄한다.

이와 같이 경찰활동의 범주가 넓다 보니 그 활용이 논의되고 있는 AI의 투입으로 인한 개인의 기본권에 대한 침해의 정도는 매우 다양할 수 밖에 없다. 예컨대, 예를 들어, 도로의 노후화, 폭우 등으로 도로파손, 도로침수, 산사태가 발생할 수 있는 도로구간을 AI가 빅데이터를 종합하여 분석, 판단해서 해당 경로를 우회하도록 하는 경찰의 교통안전 예방활동은 주로 사물에 상태에 대한 자료의 분석, 평가이기 때문에 언급한 기본권에 대한 침해가 언뜻 떠오르지 않는다. 또한, 이태원 참사와 같은 경우 수많은 112 신고전화가 있었음에도 불구하고 인간경찰이 인간의 오감의 한계 아래에서 신고 전화에서 들려오는 숨이 막혀가는 사람의 신음, 헛호흡 등을 인지하지 못하였지만, AI가 학습을 통해 전화통화에서 이를 감별하고 급박한 위험 상황이라느 점을 알려주는 경우 이 경우도 AI의 학습 과정 중에 발생할 수 있는 기본권 침해보다는 요구자의 생명, 신체에 대한 보호를 효율적으로 잘 할 수 있겠다는 안도감이 먼저 연상된다.

반면 AI 순찰차, AI 드론, AI 로봇을 활용한 인공지능통합관제 차원에서 “수상한 행동이나 그 밖의 주위 사정을 합리적으로 판단하여 볼 때 어떠한 죄를 범하였거나 범하려 하고 있다고 의심할 만한 상당한 이유가 있는 사람”(경찰관직무집행법 제3조 제1항 제1호)에 대하여 몸짓, 얼굴표정, 동선 등을 바탕으로 불심점을 학습하여 관찰, 매칭(matching), 추적, - 여기까지는 주로 AI가 하게 될 것임 - 나아가 정지 및 질문 - 이 단계는 현재로서는 인간경찰이 전담할 것임 - 까지 행하게 되는 경우 이는 인격권, 명예, 프라이버시, 일반적 행동자유권, 개인정보자기결정권 등의 다양한 기본권 침해가 예상된다.

가택 내의 신음소리, 단말마의 비명, 물건이 물체에 부딪히거나 깨어지는 소음을 분석, 평가하여 경찰이 위험방지를 위한 긴급출입을 하게 되는 경우(경찰관직무집행법 제7조 ① 경찰관은 제5조제1항·제2항 및 제6조에 따른 위험한 사태가 발생하여 사람의 생명·신체 또는 재산에 대한 위해가 임박한 때에 그 위해를 방지하거나 피해자를 구조하기 위하여 부득이하다고 인정하면 합리적으로 판단하여 필요한 한도에서 다른 사람의 토지·건물·배 또는 차에 출입할 수 있다) 오원춘 사건에서 보여지는 것처럼 인간의 오감의 한계를 극복하고 인질을 구조할 수 있는 계기가 될 수 있다는 점이 먼저 연상된다. 다만, AI의 판단이 잘못된 것이라면 언급한 기본권에 더하여 주거의 자유에 대한 침해가 연상되기도 한다.

아동 실종의 경우에도 유사하다. AI가 아동이나 주변의 수상한 사람의 대화, 몸짓, 행동, 동선 등을 분석, 평가하는 경우 아동을 찾아서 부모에게 돌려보내주는 목적 자체는 국가의 가장 기본적이고도 필수적인 임무겠지만, 학습단계에서 누군가가 국가로부터 의심의 대상이 된다면 기본권 침해가 우려될 수 밖에 없다.

범죄수사의 영역에서는 과학기술이 발전되어 온 뒤 오랫동안 과학기술을 활용한 프로파일링에 있어서 언급한 기본권 침해가 문제가 되었는데 AI와 같은 고도로 발전된 툴의 사용은 기존의 프로파일링보다 더 증대된 기본권 침해의 위험이 예정되어 있다.

2. 「경찰미래비전 2050」 발표를 계기로 살펴본 AI의 경찰활동 활용이 가지는 문제점

〈표1〉 「뉴노멀 치안 선도」 분과 주요추진과제

첨단 통합 112시스템	▶ 신고자의 음성·위치·주변의 소리·과거 신고 등을 분석하여 위험도 예측 ▶ 치안정보·관제센터 영상 등 통합 분석, 현장출동 경찰관에게 실시간 전송
인공지능 통합 관제센터	▶ 순찰차·드론·로봇 등이 수집한 정보를 토대로 위험 징후 자동 분석 ▶ 범죄 위험지역 관제센터에 현출, 순찰 드론 우선 출동시켜 범죄 차단
지능형 로봇·드론· 웨어러블 장비	▶ (로봇·드론) 위험지역을 자율적으로 순찰하고 실시간 정보를 전송 ▶ (첨단 웨어러블 장비) 인공지능 치안보좌 시스템 활용, 경찰관 신체적 역량 증진
범죄분석 플랫폼	▶ 인공지능이 범죄·사고 데이터, 과거 사건처리 이력 등 데이터를 종합 분석하여 유용한 △기법정보 △수사단서 △가용자원 제공, 수사 정확성·효율성 제고
치안민원·행정 자동처리	▶ 경찰행정·수사민원 관련 음성인식 대화형 챗봇(Pol-Bot 개발), 민원 편의 증진 ▶ 단순 반복 민원·행정처리를 자동화·전산화(RPA)하여 인력 효율성 제고
메타 경찰청 인공지능 메타경찰	▶ 메타버스 플랫폼 내 실시간 신고·상담이 가능한 '메타경찰청' 구축 ▶ 가상공간 내 순찰 기능을 탑재한 인공지능 메타경찰 개발, 가상공간 치안유지

출처: 경찰청 보도자료. 2022.9.29.

2022년 9월 경찰에서는 4차 산업혁명에 수반하는 새로운 위협의 등장 등 미래치안 환경 변화에 대응할 새로운 경찰 패러다임으로 「경찰 미래비전 2050」을 제시하면서, 위와 같이 인공지능 기술을 활용한 각종 치안활동 시스템을 구상하고 있음을 밝혔다.

위의 내용들을 살펴보면 인공지능을 활용한 자료의 분석, 평가를 통하여 주로 인지, 판단, 정보상의 영역에서 - 최대치는 순찰 - 인간 경찰의 활동이 도움을 받도록 되어 있는 것이 대부분이다. 이미 그 자체로 많은 기본권 침해우려가 제기될 수 있다. 그런데, 여기에 더하여 첨단 웨어러블 장비는 인공지능 치안보좌를 넘어 인간 경찰관의 신체적 역량을 증진을 통하여 불심검문 대상자, 위협야기자에 대한 물리적 사용까지 상정하고 있다는 점에서 더 심각한 기본권 침해가 우려되고 있는 내용이 담겨있다고 보여진다.

인공지능 기술별로 제시된 위 계획은 경찰업무별로 매칭되어 제시된 것이 아니라 인공지능 기술 활용의 필요성과 그에 대한 제한 기준을 제시하기 위해서는 경찰업무별로 제시하는 것이 좋을 것이라 생각되어 다음에 기준을 제시하고자 한다. 아래 내용은 2023년에 발표된 다른 연구자의 논문의 내용을 채택하여 발제자 나름대로 상당부분 각색한 것이다²⁾.

① 위험방지 및 범죄예방을 위한 예측적 경찰활동으로서 시간, 장소 중심 순찰

AI 알고리즘으로 다양한 원 데이터를 분석하여 패턴을 식별하고 범죄가 발생할 가능성이 있는 장소, 시간을 예측할 수 있을 것이다. 이미 발생한 범죄의 종류(예컨대, 주취폭력범죄, 마약범죄, 성범죄 등), 시간, 장소 등을 분석해 범죄발생 확률을 분석하는 범죄예측분석 시스템(pre CAS, Predictive Crime Risk Analysis System)의 경우 이미 발생한 범죄, 장소 및 시간이 중심이 된다면 그 분석 및 평가 과정에서는 기본권 침해 가능성은 커보이지 않는다. 또한 이를 토대로 범죄 유형별 빈발 시간, 빈발 지역을 감안하여 순찰력을 집중하더라도 큰 기본권 문제는 없다고 생각된다.

한발 더 나아가 이미 AI를 통해 분석된 특정 범죄유형과 관련된 특정 지역, 특정 시간대의 범죄빈발정도를 감안하는 가운데 CCTV 영상을 활용하여 개인의 얼굴정보 등 신상정보가 아닌 특정 지역, 특정 시간대에 특정 범죄에 있어서 전형적으로 나타나는 이동경로(Bewegungsmuster), 폭력적 행동성향(Verhaltensmuster) 등을 AI를 활용하여 그 위험성에 대하여 분석, 평가하고 그러한 사람을 대상으로 경찰의 불심검문의 대상으로 설정하는 것은 개인의 일반적 행동자유권에 대한 침해는 될 수 있겠지만 개인정보자기결정권에 대한 침해까지는 아니라고 생각된다³⁾.

2) 송진순, 로봇경찰의 의사결정과 법집행에 있어 윤리적 요구: 블랙박스, 편향 그리고 거버넌스적 윤리 프레임 워크, 입법과 정책 제15권 제2호, 2023, 90쪽 이하.

3) 다만, 이 때의 행동성향이 걸음걸이 등 개인의 특정할 수 있는 행동적 특성과 관련되는 경우는 생체정보(민감정보)로서 별도의 엄격한 법적 근거가 필요할 것이다.

위에서 언급한 경우는 기본권 침해가능성이 커 보이지 않기 때문에 경찰관직무집행법 제8조의2(정보의 수집 등), 「경찰관의 정보수집 및 처리 등에 관한 규정」(대통령령), 개인정보보호법 제15조 제1항 제3호(‘공공기관이 법령 등에서 정하는 소관업무의 수행을 위하여 불가피한 경우’)에 따라 허용될 수 있을 것으로 생각된다.

다만, 인종, 국적 등 인적 구성을 중심으로 범죄발생 유형, 장소, 시간을 분석하는 것은 차별금지의 차원에서 엄격히 금지될 필요가 있다.

② 112 신고 시스템에의 활용

112 신고는 범죄 범죄현장에 대한 신고 또는 위험방지 상황에 대한 신고가 대부분이고, 형이 확정되거나 확정되지 않은 상태에서 수배된 자에 신고는 그렇게 많지 않다. 점에서 아래에서 소개될 독일 연방헌법재판소의 기준을 따른다면 위험방지 영역에서 AI를 활용하여 사람, 인적 그룹, 기관, 조직, 대상(Objekte) 및 사물(Sache)에 관한 관계 또는 맥락을 생산하거나, 의미없는 정보나 인지는 제외하고, 의미있는 것은 세부적인 인지가 알려진 사실관계에 귀속시키고, 저장된 자료를 통계적으로 평가하는 것은 가능하다고 생각된다.

다만, 위험방지 영역이라 할지라도 사람, 인적 그룹 등의 측면에서는 구체적 위험이 존재하지 않는한 정보자기결정권을 침해하는 AI활용은 제한되어야 할 것이다. 구체적 위험이 존재하는 경우에는 위험방지를 위하여 비책임자가 관여될 수도 있기 때문에 경찰법상 범죄행위에 대한 예방적 대처를 위한 AI 활용에 관한 연방헌법재판소 결정의 취지와 달리 구체적 위험이 존재하는 경우라면 비책임자, 나아가 형사사건의 증인과 관련된 정보도 위험방지를 위하여 기여할 수 있다면 AI를 통하여 활용될 수 있을 것이다.

만약 신고된 특정 사람의 범죄적 행동의 예측과 관련하여 AI를 활용하고자 하는데 그가 전과자라면 그의 범행 리스크를 AI로 분석/평가할 때 그의 전과자료를 기준으로 활용하여 그가 교화·개선에 따라 사회복귀를 하는 것을 방해해서는 안된다. 이 경우 전과자료 등을 근거로 한 AI 활용은 엄격한 내용적, 절차적 근거 하에서만 위험상황, 사건, 장소 중심으로 제한적으로만 허용되어야 할 것이다.

이런 점에서 독일의 경우 특정한 사람, 예컨대, 전과자 중심의 AI 활용 예방적 경찰활동은 아직까지 행해지고 있지 않다⁴⁾. 반면 미국의 주에서는 실시간 범죄센터

4) Julia Fricke, Big Data und künstliche Intelligenz – Chancen und Risiken für die Polizeiarbeit der Zukunft. 2020.5.27.27.

<https://ksv-polizeipraxis.de/big-data-und-kuenstliche-intelligenz-chancen-und-risiken-fuer-die-polizei-arbeit-der-zukunft/>

(RTCC: Real Time Crime Center)에서 잠재적인 범죄자의 위험도를 분석하여 등급화 하기도 하는바⁵⁾, 이는 기본권 보장 차원에서 문제가 있는 방식이라 생각된다⁶⁾. 독일의 이런 사정과 비교해보면 우리나라의 사람 중심의 경찰 우범자 관리는 많은 문제가 있고⁷⁾, 여기에 AI를 활용하는 것은 더 큰 문제가 될 것이다.

112 신고를 살펴보면 많은 경우 사람의 전화, 오늘날은 주로 스마트폰 신고에 대하여 사람인 경찰관이 위험상황을 판단하여 출동여부 및 출동요소를 결정하게 된다. 그런데, 신고자와 신고접수 경찰관 간에 대화만으로는 현장의 상황, 분위기를 제대로 파악할 수 없는 경우가 많다. 왜냐하면, 신고자는 범인으로 인하여 큰 소리로 말하거나, 논리적으로 말하는 것이 제약된 상태일 수가 있기 때문이다. 이런 경우에 깨지는 소리, 발이나 장비를 이용하여 강제로 문을 열려는 소리, 술집 등에서 다중 화자의 동시발화, 또는 늦은 밤 범인에게 칼로 위협받는 상황에서 신고자의 음성은 들리지 않지만 지하철역 안내방송과 같은 장소추정 음성은 야간근무 중 생리적 수면욕, 피곤함, 그리고 인지적 한계에 노출되어 있는 112 신고 접수자에게는 인식이 어렵겠지만 AI를 활용하여 이런 부분을 분석할 수 있다면 경찰활동에 획기적 도움을 받을 수 있는 영역이라고 생각된다. 이외에도 신고자가 나이가 어리거나 지능이 낮아 정확하게 의사표현을 하지 못하는 경우, 주취상태이거나 극도로 흥분한 경우, 경찰관과 신고자가 동시에 말하는 경우 등에 AI의 분석, 평가 기능은 112 신고와 관련된 위험 상황 여부를 판단하는데 큰 역할을 할 것으로 생각된다⁸⁾.

오원춘 사건의 경우 오원춘이 피해자가 문을 잠그고 도망해 있는 방실의 문고리를 강제로 열려는 소리가 나중에 112 녹취록 분석을 통해서 확인되었지만, 신고자와 대화에 집중하고 있는 112 신고접수자가 이런 정황까지 인지하는 것은 무리가 있었을 것이다. 또 하나 이 사건에서 문제가 된 것은 피해자인 발신자의 위치특정이었는데, 당시 오원춘 집 인근까지는 범위를 축소하였지만 정확한 위치를 특정할 수 없어 야간 시간대 강제출입에 실패하여 피해자의 목숨을 살릴 수 있는 기회를 놓쳐버리기도 하였다. 이 인질납치사건으로 피해자가 사망한 이후에야 비로소 경찰도 긴급구조기관

5) 황정용, AI의 112 상황실 적용 및 기대효과 - 신고접수와 경찰 언어모델 개발을 중심으로 한 국치안행정논집 제21권 제2호, 2024, 185쪽.

6) 112 신고를 넘어서 형사사법 영역에서 알고리즘을 토대로 유죄판결 받은 사람들의 재범위험성을 평가하는 방식에 대해서는 아래에서 후술한다.

7) 우리나라 우범자 관리의 형식적 근거는 예규에 불과한 「우범자 첩보수집 등에 관한 규칙」이었다. 2020. 12.9. 개정으로 「주요 강력범죄 출소자등에 대한 정보수집에 관한 규칙」으로 제명이 바뀌었으나 여전히 예규에 불과하다. 이미 출소하고 보호관찰이 끝난 대상자는 사회로 복귀하였다고 보아야 하는만큼 이들을 대상으로 경찰이 일정기간동안 해당 규칙에 근거하여 감시하는 것은 뚜렷한 근거없이 기본권 침해가 큰 상태에서 사람 중심으로 감시하는 방식이라 할 것이다.

8) 황정용, 위의 논문, 188쪽.

인 소방에 준하여 발신자의 기지국 위치정보, GPS 정보, WI-FI 위치정보를 확인할 수 있는 권한을 가지게 되었고, 이에 따라 경찰관서에도 위치정보시스템(LBS, Location Based Service)이 갖추어 지게 되었다⁹⁾. 기존의 「위치정보의 보호 및 이용 등에 관한 법률」은 경찰을 발생한 피해를 수사하는 수사기관의 측면에서만 고려하고 피해가 발생하기 이전 위험한 상태에서 기능하는 위험방지 기관이라는 점을 등한시 하였던 것이다.

112 신고와 관련하여 위치정보는 지금은 LBS 시스템을 통하여 어느 정도 상세한 신고자 위치를 알 수 있게 되었지만, 특정 지점의 위치정보 외에도 AI를 활용한다면 해당 지점까지 도로 상태는 어떤지, 해당 지점 주변은 유흥가, 상가, 주택 등 어떤 특성을 가지는지를 알 수 있게 되어 112 신고출동 대응의 효율성을 높을 수 있을 것이다.

또한, AI를 통해 피해자의 상태를 확인할 수 있다면 이를 토대로 필요한 조치, 예컨대, 소방, 의료기관 등과의 연계를 파악하여 단순히 위험방지, 범죄예방 및 범인검거를 넘어선 활동에도 기여할 수 있을 것이다. 또한, 대피가 필요한 화생방, 폭발물 테러 등이라는 점이 확인된다면 중독, 화재, 폭발에 대비하여 시민들에게 위험을 알려주고 이에 대비하도록 할 수 있을 것이다¹⁰⁾.

아직까지 112 신고는 영상신고가 아니라 음성신고이기 때문에 신고자 주변의 무관한 사람의 초상권, 개인정보자기결정권을 침해할 소지는 적다고 생각된다. AI 를 통

9) 「위치정보의 보호 및 이용 등에 관한 법률」 [시행 2012.11.15.] [법률 제11423호, 2012. 5.14., 일부개정]

제29조(긴급구조를 위한 개인위치정보의 이용)

① 「재난 및 안전관리 기본법」 제3조제7호에 따른 긴급구조기관(이하 “긴급구조기관”이라 한다)은 급박한 위험으로부터 생명·신체를 보호하기 위하여 개인위치정보주체의 긴급구조요청이 있는 경우 (중략) 위치정보사업자에게 개인위치정보의 제공을 요청할 수 있다

② 「경찰법」 제2조에 따른 경찰청·지방경찰청·경찰서(이하 “경찰관서”라 한다)는 위치정보사업자에게 다음 각 호의 어느 하나에 해당하는 개인위치정보의 제공을 요청할 수 있다. 다만, 제1호에 따라 경찰관서가 다른 사람의 생명·신체를 보호하기 위하여 구조를 요청한 자(이하 “목격자”라 한다)의 개인위치정보를 제공받으려면 목격자의 동의를 받아야 한다.

1. 생명·신체를 위협하는 급박한 위험으로부터 자신 또는 다른 사람 등 구조가 필요한 사람(이하 “구조받을 사람”이라 한다)을 보호하기 위하여 구조를 요청한 경우 구조를 요청한 자의 개인위치정보
2. 구조받을 사람이 다른 사람에게 구조를 요청한 경우 구조받을 사람의 개인위치정보
3. 「실종아동등의 보호 및 지원에 관한 법률」 제2조제2호에 따른 실종아동등(이하 “실종아동등”이라 한다)의 생명·신체를 보호하기 위하여 같은 법 제2조제3호에 따른 보호자(이하 “보호자”라 한다)가 실종아동등에 대한 긴급구조를 요청한 경우 실종아동등의 개인위치정보

10) 황정용, AI의 112 상황실 적용 및 기대효과 - 신고접수와 경찰 언어모델 개발을 중심으로 한국치안행정논집 제21권 제2호, 2024, 184쪽.

한 음성분석을 통해서도 신고자와 주변 인물의 내밀한 관계를 알 수 있게 되는 경우에는 개인정보보호와 관련된 내용을 준수하여야겠지만 일단 112 신고라는 긴급상황, 특히 코드 0, 코드 1 상황에서는 위험방지라는 공익적 필요성이 더 크다고 보여지므로 AI를 통한 분석, 평가는 적법하다고 보여진다. 다만, 이 경우에도 개인정보보호법에 따른 목적 외 이용·제공제한(제18조), 정보주체 이외로부터 수집한 개인정보의 수집 출처 등 통지(제20조), 개인정보의 열람(제35조), 개인정보의 정정·삭제(제36조) 등 개인정보보호주체의 권리를 보장하는 규정을 둘 필요가 있다.

③ 자율주행차 및 드론을 이동수단으로 하여 각종 범법행위, 위험야기 우려 행위 감시 및 관찰

AI 기술로 거리, 공원 및 건물과 같은 공공공간에서의 폭력 등 범죄행위, 교통사고 현장, 안전벨트 미착용, 끼어들기 등 교통위반 행위를 감시, 관찰할 수 있다면 범법행위 검거에 기여할 수 있을 것이다. 그러나, 일반적 행동의 자유, 개인정보자기결정권의 관점에서 본다면 인간의 모든 행동이 모든 공개된 장소에서 관찰, 감시될 우려가 있다는 점에서 이는 기본권 행사에 심각한 위축효과를 가져오기 때문에 공개된 장소라 할지라도 특별히 범죄가 빈발할 것으로 예상되는 지역, 위험야기가 우려되는 지역으로 한정할 필요가 있다고 생각된다.

AI 알고리즘으로 바로 자동차 번호판 인식기술이나 안면인식 기술을 활용하여 개인의 신원을 확인하는 것은 경찰관직무집행법상 금지된다. 경찰관직무집행법에 따르면 단순한 범죄혐의, 단순히 위험을 야기할 우려가 있다는 것만으로는 경찰관이 강제로 개인의 신원을 확인할 수 있도록 하고 있지 않기 때문이다. 만약 그것이 경찰활동으로 허용될 수 있으려면 특별히 고도의 위험상황, 신원확인 필요성에 대한 엄격한 요건이 경찰관직무집행법에 구체적으로 규정되어야 한다. 그렇지 않고서는 불가능하다.

일단 AI의 투입이 고려되지 않은 상황에서도 요구되는 각종 경찰상 감시, 관찰 조치에 대한 독일의 경찰법 규정 사례를 통해 그와 같은 조치의 기준 내지 요건을 먼저 알아보도록 한다. 그러한 연후에 이러한 조치들에 AI가 투입되는 경우의 허용 기준 및 요건에 대해서 고민해보도록 한다.

③ - 1 : AI 투입이 고려되지 않은 경우에도 경찰 감시, 관찰에 요구되는 조건들

a. 시각 기술적 수단의 공개된 투입을 통한 정보수집¹¹⁾

(1) 경찰은 범죄행위의 방지를 위하여 CCTV로 공개적으로 접근가능한 장소를 관찰할 수 있고, CCTV를 통해 전송된 사진 및 영상물을 녹화할 수 있다. 다만 다음과 같은 요건이 충족되는 경우에만 가능하다.

1. 해당 장소에서 반복적으로 범죄행위가 행하여졌거나 장소의 상태가 범죄행위의 발생을 용이하게 하는 경우로서 해당 장소에서 향후 계속적인 범죄행위가 행해질 것이라는 가정이 사실관계들을 통하여 합리화되는 경우

2. 관련된 사실관계를 통하여 장기 3년 이상의 법정형으로 위하되는 중대한 범죄가 약속, 준비되거나 행해질 것이라는 가정이 합리화되는 경우 (이하 생략)

(2) 위 제(1)항에 따라 수집된 데이터는 범죄행위의 추적을 위해 필요하거나 누군가가 미래에 범죄를 행할 것이라는 가정이 사실관계를 통해 합리화되는 경우가 아니고 그 보관이 범죄행위의 예방적 대처를 위해 필요한 경우가 아니라면 해당 자료는 최장 14일 동안만 저장될 수 있다.

(3) 시각기술적 수단의 투입을 통한 자료수집의 실시에 대해서는 경찰관서장이 결정한다.

(4) 제(1)항에 따른 조치들은 기록보관되어야 한다. 보관기간은 1년으로 제한된다. 보관기간이 만료되기 전 적절한 시점에 위 제(1)항에 따른 전제조건이 계속 충족되는지 심사되어야 한다. 이 경우 매 1년 단위로 보관기간의 연장이 허용된다.

b. 은밀한 단기간 관찰(kurzfristige Observation)

지속적으로 24시간 이내 또는 지속적이지 않고 시간간격을 두고 끊어서(예컨대, 일주일 간격을 두고) 그 관찰시간의 합이 2일 이내의 단기간 관찰을 법관의 명령이 없이도 경찰관서장의 명령으로도 행해질 수 있기는 하지만, 그 동인(이유, 근거)은 구체적 위협의 방지 또는 구체적 혐의에 대한 범죄의 석명을 위한 것에 한정된다.

위험방지를 위한 단기간 관찰을 통한 정보수집은 브란덴부르크, 브레멘, 함부르크, 노르트라인-베스트팔렌, 작센-안할트의 경찰법에 규정되어 있다.¹²⁾ 이 6개 주를 제외한 나머지 10개의 주에서는 정보수집과 관련된 일반적 수권조항으로 단기간 관찰이 가능하다. 이와 같은 단기간 관찰은 그러나, 위험방지 목적을 위하여 경찰책임자를 상대로 가능하며, 해당 조치가 없으면 경찰상의 임무가 곤란해질 경우에만 허용된다는 제한이 붙는다. 기본적으로 구체적 위협이라는 경찰법상의 요건이 충족된 경우에

¹¹⁾ 노르트라인 베스트팔렌 주 경찰법 제15a조 제1항에 기반하여 설명하였다.

¹²⁾ § 32 Abs. 4 PolG Bbg; § 32 Abs. 3 BremPolG; § 9 Abs. 4 DVPolG HH; § 16a Abs. 3 PolG NRW; § 17 Abs. 6a SOG LSA.

한하여 단기간 관찰이 허용될 수 있다.¹³⁾

c. 은밀한 장기간 관찰(längerfristige Obeservation)

또한, 단순히 수상한 거동, 위협야기 우려라는 패턴분석만으로 AI 기술을 활용하여 특정 개인을 상대로 한 장기간 관찰을 하는 것은 허용될 수 없다¹⁴⁾¹⁵⁾¹⁶⁾. 독일의 경

13) Petri, H. Informationsverarbeitung im Polizei- und Strafverfahrensrecht, in: Lisken/Denninger(Hrsg), Handbuch des Polizeirechts, 2007, Rn. 226; 同旨 Honnacker/Beinhofer, BayPAGM, § 33 Rdnr. 3; Brenneisen/Staack, Die Videoüberwachung nach allgemeinem Polizeirecht, DuD 1999, S. 449 f.; Klopfer/Breitkreutz, Videoaufnahmen und Videoaufzeichnung als Rechtsproblem, DVBl. 1998, S. 1152.

14) 독일의 모든 주경찰법, 그리고 연방경찰법에서는 24시간 이상의 지속적인 관찰 또는 지속적이지 않지만 일정한 날짜 간격 (예컨대, 일주일 간격을 두고) 하루에 일정 시간씩 행해져서 도합 2일 이상 행해지는 지속적인 관찰은 반드시 절차적 통제가 가해져야만 하는 민감한 법익침해로 보고 있다. 많은 주의 독일 경찰법에서는

1. 중대한 의미를 가지는 범죄행위를 저지를 것이라는 가정이 사실관계를 통하여 합리화되고(und)

2. 그와 같은 범죄행위의 예방을 목적으로 사실관계의 설명이 필요하거나 또는(oder) 장기간 관찰 수단을 투입하지 않으면 불가능해지거나 매우 어려워지는 경우에 한하여

① 구체적 사실관계를 통하여 중대한 의미를 가지는 범죄행위를 저지를 것이라고 예상되는 사람, 또는 그러한 범죄에 공범으로 참가할 것이라고 생각되는 사람,

② 위험해진 시설 또는 설비에 대한 (경찰)책임자를 상대로 행해질 수 있다. (해당 조문은 메클렌부르크 포어포먼주 공공의 안녕 및 질서에 관한법 33조 제1항 제1호 및 제2항을 요약한 것임)

15) 독일경찰법 모범초안 8c조 제2항 제1호는 ‘장기간 관찰’을 통해 경찰이 정보를 수집할 수 있는 권한을 부여하고 있는데, 이 경우 그 대상은 첫 번째로 행위책임자, 상태책임자 및 비책임자, 두 번째로 ‘장래에 범죄를 범할 것이라는 근거가 있는 자’에 한정된다. 그러나, 이와 같은 대상자를 상대로 무턱대고 장기간 관찰할 수 있는 것이 아니라 다음과 같은 하위 요건에 해당되는 경우에 인정된다.

1호: ‘현저한 위협의 방지를 위하여’,

2호 a: ‘사실상의 근거에 기초하여 성폭력 관련 범죄, 특수절도, 무기 등을 사용한 특수절도 등의 범죄, 영업적 장물죄, 사기, 사회부조금 부당수령, 환경범죄에 예방적으로 대처하기 위하여’

2호 b: ‘조직범죄에 예방적으로 대처하기 위하여’

언급한 독일 경찰법 규정들은 기본적으로 구체적 위협의 존재를 요구하고 있고, 다만 2호 a 의 경우 구체적 위협의 요건에 조금은 못미칠 수는 있겠지만 최소한 구체적 사실관계를 통하여 그 사람이 언급한 범죄와 관련된 위험한 행위, 범죄행위를 할 것이라는 점이 근거지어지는 개별적인 경우에 한하여 장기간 관찰이 허용될 수 있다.

16) 독일 경찰법 외에 독일 형사소송법 제163f 조에도 장기간 관찰 규정이 있는데 이는 중대한 의미를 가지는 범죄행위에 대한 충분한 사실적 근거가 존재하는 경우에 또는 다른 방법으로는 그와 같은 중대한 범죄의 설명이나 주거지의 탐색이 본질적으로 어려워지는 경우에 행해질 수 있다. 다만, 형사소송법상 단기간 관찰의 경우 경찰의 일반적 수사권, 즉, 구체적 혐의에 따른 범죄행위를 수사하고자 하는 경우 허용될 수 없다. 이는 이 발제문에서 언급하고 있는 예방적 경찰활동과는 다른 구체적 범죄혐의의 수사와 관련된 내용이다.

우 장기간 관찰의 경우 그것이 경찰법에 따른 것이건 형사소송법에 따른 것이건 법관의 명령을 필요로 한다. 장기간 관찰용으로 AI를 사용한다면 구체적 사실관계를 통하여 범죄가 행해질 것이라는 것이 합리화되고, 그 예방을 위하여 사실관계의 석명이 필요한 경우에 한하여 행해질 수 있을 것이다.

단기간 관찰이건 장기간 관찰이건 위에서 언급한 요건을 갖추었다고 할지라도 AI의 자료 분석, 평가 내지 학습과정이 내포하고 있는 편견, 오류의 가능성이 통제되어야 하고 구체적 사실관계를 통하여 구체적 위험이 발생할 것이라는 것이 합리화되는지, 구체적 사실관계를 통하여 범죄가 행해질 것이라는 것이 합리화될 수 있는지에 대한 AI의 결론이 합리적으로 설명되고 이유제시될 수 있어야 할 것이다.

d. 경찰 수배(노르트라인 베스트팔렌주 경찰법 제21조)

개인 및 이제까지 그에 의하여 행해진 범죄행위를 전체적으로 관찰하였을 때 형법상 구성요건에 의해 보호되는 범의들을 직접적으로 침해하는 특별히 중대한 범죄행위가 행해질 것이고, 경찰 수배가 범죄행위의 예방적 대처를 위하여 필요하다는 가정이 특별한(즉, 개별적, 구체적) 사실관계를 통해 합리화되는 경우 경찰은 개인정보, 특히 개인의 신원이 그에 의하여 사용된 자동차의 번호판을 수집할 수 있다.

e. 관찰 결과의 주민에의 통보

단순히 이미 발생한 범죄의 종류, 장소, 시간에 따라 범죄빈발 시간, 장소를 특정하는 것이 아니라 거동이 수상해보이거나 위험해 보이는 사람들의 행동을 중심으로 모니터링, 프로파일링, 추적하는 것이라면 경우에 따라 - 아직 범죄가 기수에 이르지 않았음에도 불구하고 - 해당 사람을 지목하여 인근 주민들에게 통보하는 것은 별도의 기본권 침해를 나타내는 것이므로 명확한 법적 근거가 경찰관직무집행법에 규정될 필요가 있다고 생각된다. 이 경우는 법관에 의한 영장 통제가 필요하다고 생각된다.

③ - 2 : 위의 감시, 관찰 조치에 AI 투입이 고려되는 경우 요구되는 조건들

우리나라의 법제 사정에서는 위에서 언급한 모든 경찰 감시·관찰 조치들은 경찰관 직무집행법 제8조의2(정보의 수집 등)을 근거로 범죄·재난·공공갈등 등 공공안전에 대한 위협의 예방과 대응을 위한 것이라면 쉽게 허용될 수 있다고 해석될 정도로 경찰정보수집 활동을 위한 기준이 포괄적이고 불명확하다. 현재 많은 국내 판례에서는

개인정보보호법 제15조 제1항 제1호에 따른 정보주체의 동의가 없더라도 같은 법 같은 조 같은 항 제3호의 ‘공공기관이 법령 등에서 정하는 소관업무의 수행을 위하여 불가피한 경우’라는 매우 느슨한 조건만 충족된다면 개인정보를 수집, 이용할 수 있는 것으로 해석하는 경우가 많다.

사람의 신원을 확인가능하게 하는 얼굴이 포착될 수 있는 CCTV 동선추적은 초상권, 정보자기결정권 측면에서 매우 논란의 소지가 클 수 있겠지만, 112 신고 등과 관련하여 도로교통상의 위험방지를 위하여 난폭운전 또는 일반적인 주행경로와 다른 특이한 운전 등 특정한 행동이나 이동과 관련된 지표를 토대로 AI를 활용하여 CCTV를 분석하는 것은 위에서 언급한 규정으로 충분히 가능하지 않을까 생각된다. 112 신고를 통하여 차량을 이용한 납치신고가 되어 이를 검거 및 인질의 생명·신체를 구하고자 하는 경우 범죄의 중대성, 급박성을 감안했을 때 CCTV 분석을 통하여 차량번호 및 신원확인까지도 위 규정으로 가능할 것으로도 생각된다.

그런데, 사실 위에서 언급한 개인정보보호법 제15조 제1항 제3호는 개인정보보호와 관련하여 개별 기관이 소관업무 수행을 위하여 관련 개인정보침해 규정을 만들 수 있다는 것을 의미하는 것이지 그 자체로는 해당 기관의 목적 수행을 위한 개인정보 침해 수권규정이 아님에도 불구하고 이렇게 폭넓게 해석하고 있는 것이다.

예컨대, 위 독일 경찰법이 제시하는 경찰 감시, 관찰 규정들에서 공통적으로 요구되는 개별적, 구체적 사건에서 구체적 위험악기 상황의 방지로 한정하는 경우라 할지라도 우리나라 법제(특히, 개인정보보호법, 경찰관직무집행법) 하에서는 구체적 위험을 야기하는 행동을 하는 것으로 보이는 차량의 흐릿한 번호판을 인식, 판독하여 경찰 내부에서 차량번호를 공유하거나 위험악기자의 인상착의(특징)를 공유하는 것은 불심검문(경찰관직무집행법 제3조), 위험발생의 방지(제5조), 범죄의 예방과 제지(제6조)를 위한 요건이 갖추어진 경우에 한하여 위험방지를 위한 준비행위로서 경찰관직무집행법 제8조의2(정보의 수집 등)을 근거로 허용될 수 있을 것처럼도 보인다.

이렇게 폭넓게 해석하는 경우 관련 법률규정들이 인간에 의한 CCTV의 단순한 시각적 활용에 대한 제한 뿐만이 아니라 심지어 자율자동차 및 드론과 연계된 AI를 경찰활동에 투입하는 것이 내포하는 기본권 침해를 충분히 제한할 수 있을 것인지는 매우 의심스럽다.

이와 같이 아직 자동화되지 않은 기존의 경찰상 감시, 관찰 조치에 대해서도 느슨한 정보보호규정의 해석이 적용되는 우리나라에서는 그보다 더 심각한 기본권 침해 가능성을 가지고 있는 AI를 통한 자동화된 분석, 평가 내지 학습을 위해 수집, 제공될 수 있는 개인정보가 어떤 것인지, 그 개인정보가 제공되는 경우 어떤 형태로 (예컨대, 익명, 무기명) 제공되어야 하는지, 그리고 그렇게 AI가 분석, 평가, 즉 학습한

정보를 다른 사안에 매칭시켜 관련 위험성을 판단함에 있어 매칭시켜도 좋은 것은 무엇이고 매칭시켜서는 안되는 것은 무엇인지 그 기준이 전혀 제시되어 있지 않다고 생각된다.

AI의 투입을 통해 대폭 증대되는 위 각종 경찰 감시, 관찰 조치들의 기본권 침해성은 분석, 평가 알고리즘이 신뢰성이 있는 것인지, 투명성이 있는 것인지, 그리고, 학습의 결과 및 매칭을 통해 얻은 판정 결과가 합리적으로 설명될 수 있는 것인지를 관점은 현재의 우리 법제에서는 전혀 고려하지 못하고 있다. 합리적으로 이유제시될 수 없는 경찰의 모든 조치는 법의 세계에서는 허용될 수 없는바, AI가 자료의 분석, 평가를 통해 얻은 인지, 판단 및 그에 따른 대상자와의 매칭에서 위험성이 있다는 결론을 합리적으로 이유제시할 수 없다면 이는 법의 세계에서 설 자리가 없다. 이러한 AI의 학습결과 및 그에 따른 매칭 결과는 나중에 설명할 단순한 “위험의 의심”에 불과한 것으로서¹⁷⁾ 이를 근거로 한 은밀하거나 강제적인 정보수집은 허용될 수 없고 그 의심점에 대하여 사실확인의 대상으로 삼을 수 있을 뿐이다. 이 경우에도 대상자가 왜 국가측으로부터 그런 의심을 받아 경찰상 감시, 관찰의 조치의 대상이 되어야 하는가에 대한 이유제시는 AI 알고리즘 및 이를 활용하는 경찰관이 개인에게 설명할 수 있어야 한다. 그렇지 않다면 이는 법적으로 허용될 수 없다.

④ 범죄수사 과정에서의 조사 및 증거 분석에 AI의 활용

우리나라 경찰수사기관은 현재 지문, (신발)족적, 차량 바퀴흔적 자료를 시스템화 하여 활용하고 있다. 미래에 행해질지도 모르는 범죄행위의 예방적 대처가 아닌 과거에 있었던 범죄행위의 수사에 AI를 활용하는 것은 그것이 개인정보, 특히 지문 등 생체정보와 같은 민감정보를 활용하는 경우는 특별하고도 엄격한 개인정보 관련 보호조치가 요구되어야 한다고 생각된다. 이외의 영역에서는 과거의 범죄를 위한 수사 영역에서는 예측적 경찰활동의 경우보다 폭넓게 AI의 자동화된 분석, 평가 능력이 활용될 수 있을 것으로 생각된다.

⑤ 재범위험성 평가시 AI의 활용

재범위험성 평가는 과거의 불법에 대한 책임에 기초하는 제재라기보다는 장래의 위

¹⁷⁾ 최근에 발표된 AI와 관련된 공법논문은 예측적 경찰활동을 경찰법 이론에 대입하기 위해서는 (1) 경찰관의 경험, (2) 사안에 대한 진단, (3) 손해 발생에 대한 예측, (4) 경찰권 행사의 네 가지 단계에 대한 검토가 필요하다고 진단하고 있다. 박원규, 인공지능, 빅데이터, 그리고 경찰: 예측적 경찰활동에 대한 공법적 검토, 공법학연구 제24권 제1호, 2023, 334쪽 이하.

험성으로부터 행위자를 보호하고(즉, 교화·개선)하고 감시를 통해 재범 위험으로부터 사회를 방위하고자 하는 보안처분의 전제로서 행해진다¹⁸⁾. 위와 같은 보안처분은 원칙적으로 법원(판사)가 결정하도록 되어 있다(소위 보안처분 사법부 결정주의)¹⁹⁾. 이와 관련해서 국내의 한 연구에서는 AI의 예측능력과 재범예측 알고리즘의 헌법적 문제를 미국의 State v. Loomis 판결을 중심으로 행한 것이 있는데²⁰⁾ 이는 경찰에서 가정폭력, 아동학대 범죄 대응에 있어 100m 이내 접근금지, 전화연락금지 등 긴급입시 조치를 행함에 있어 재범위험성을 평가하는 과정에도 큰 차이없이 참고될 수 있을 것으로 생각된다.

참고로 현행법상 법원이 행하는 재범위험방지 및 사회방위를 위한 보안처분에는 「치료감호에 관한 법률」상의 보호처분인 치료감호, 「보호관찰 등에 관한 법률」상의 보호관찰, 「보안관찰법」상의 보안관찰, 「특정 범죄자에 대한 보호관찰 및 전자장치 부착 등에 관한 법률」상 성범죄자에 대한 전자장치 부착명령, 「성폭력범죄자의 성충동 약물치료에 관한 법률」에 의한 성충동 약물치료 등이 있다.

재범위험성 평가도 과거의 범죄에 대한 수사 내지 제재라기보다는 범죄자가 미래에 행할 수 있는 재범위험성을 예측하는 것이라는 점에서 예측 경찰활동(Predictive Policing)의 범주에 해당된다. 언급한 State v. Loomis 판결을 중심으로 AI의 예측능력 및 재범예측 알고리즘의 헌법적 한계를 지적하는 논문에서는 “남성 → 갈색 눈동자 → 경범죄 이상의 범죄경력 → 문신한 사람”은 “중범죄를 범할 가능성이 크다”는 식의 AI의 자료 분석, 평가는 권력적, 편향적 시각이 반영될 수 있다고 한다. 즉, 인종(흑인), 학력(저학력), 재산정도(가난한 사람) 등 특정한 공통된 속성을 가진 집단에 포함되는 자는 일반적으로 중범죄를 저지를 가능성이 크다는 확률을 나타내는 것임에도 불구하고, 그 확률이 아무리 높아도 동일 집단에 속하는 특정한 개인은 전혀 중범죄를 범할 가능성을 가질 품성이 없음에도 불구하고 국가측의 의심의 대상이 된다는 점에서 언급한 사회적 평가기준에 따른 권력적, 편향적 시각이 AI의 재범위험성 판단에 반영될 수 있다는 점을 신랄하게 지적한다²¹⁾. 이는 논리적으로 보아서도 범죄

18) 대법원 2010. 9. 30. 선고 2010도6403 판결 【근로기준법위반·도박개장】

【판시사항】

[1] 형법 제62조의2 제1항에서 규정한 ‘보호관찰’의 법적 성격 및 준수사항 부과와 허용 한계

[2] 근로기준법을 위반한 피고인에 대하여 형의 집행을 유예함과 동시에 집행유예기간 동안 보호관찰을 받을 것을 명하면서 “보호관찰기간 중 선거에 개입하지 말 것”이라는 내용의 특별준수사항을 부과한 원심판단을 정당하다고 한 사례

19) 보안처분에 대한 “사법부 관할의 원칙”은 프랑스, 독일, 미국에서 확립되어 있다. (박지현, 보안관찰법에 관한 연구 - 사상법에 대한 보안처분 부과와 법리상 문제점을 중심으로, 서울대학교 대학원 법학석사학위논문, 1999, 46쪽 이하)

20) 이병규, AI의 예측능력과 재범예측 알고리즘의 헌법 문제 - State v. Loomis 판결을 중심으로, 공법학연구 제21권 제2호, 2020, 169쪽 이하.

와 본질적으로 관련없는 사안으로 범죄위험성을 추론하는 것으로 오류에 해당될 뿐만 아니라 인간 그 자체를 범죄적 행동과 관련없는 다른 지표로서 범죄퇴치의 객체로 전락시킨다는 점에서 인간을 목적으로 대하지 않고 수단으로 대함으로써 법학의 전제하고 있는 칸트의 정언명령에 정면으로 위반한다고 생각된다. 편향적, 권력적 기준으로 집단적 속성을 특징으로 하여 재범위험성을 추론할 경우 범죄와 관련없는 무고한 선량한 개인을 잠재적 범죄자로 낙인찍는다는 점에서 무죄추정의 원칙에도 반한다.

위와 같이 편향적이고 권력적으로 증명되지 않는 집단적 속성에 바탕하여 AI가 개인에 대한 재범위험성을 추론할 경우 그것은 효율성을 말하기 이전에 그 자체로 오류이다. 또한, AI 알고리즘의 복잡성을 핑계로 (소위 블랙박스화 문제) 왜 그러한 결론에 도달하였는지 스스로도 논리적으로 설명하지 못하고 매칭의 대상자(AI 매칭에 따라 위험성이 있다고 판단될 경우 경찰조치의 대상이 될 우려 아래 놓여지게 되는 사람)에게도 논리적으로 이유제시하지 못한다는 이와 같은 정보수집에 기한 자료의 자동적 분석, 평가 및 매칭은 그 자체로 법적인 이유제시 의무 위반으로 법적 수단으로는 허용될 수 없다고 할 것이다. 블랙박스화 경향은 기술적 측면이고 알고리즘 보호는 기업의 경영적 측면에 불과하지²²⁾ 그것이 AI에 의한 재범위험 매칭 대상자에 대한 법적인 설명의무를 면하여 주는 것은 결코 아니다.

불심검문의 경우 그 요건이 여러 가지로 설명될 수 있겠지만 핵심은 국가측이 왜 대상자에 대하여 죄를 범하였거나 범하려 하고 있는지에 대한 의심을 가졌는지에 대하여 합리적으로 설명될 수 있을 경우에 비로소 그와 같은 국가적 활동은 정당화 될 수 있다. 국가가 개인에 대하여 가지는 의심에 대한 합리적 이유제시가 행해지지 않거나, 그 이유제시가 정당화될 수 없다면 이는 그 자체로 내용적으로도 적법절차의 측면에서도 위법한 경찰활동이 되고 그에 대하여 개인이 저항하고 심지어 폭력을 행사하더라도 공무집행방해죄에 해당되지 않는다는 것이 우리나라 법원의 일관된 태도이다. 이런 불심검문(불심검문도 결국은 의심점에 대한 질문을 통한 국가 측의 정보수집활동이다)과 비교해볼 때 AI의 자동적 자료 분석, 평가 내지 학습의 결론과 그에 따른 매칭의 결론이 합리적으로 설명될 수 없다면 이는 그 자체로 법적으로 아무런 의미가 없는 기술적 매칭에 불과할 뿐이다.

미국의 일부 주에서는 법관이 피고인의 보석여부 결정, 재판시 형벌에 부과되는 보

21) 이병규, 위의 논문, 173쪽.

22) COMPAS (Correctional Offender Management Profiling for Alternatives Sanctions의 약자. 대체적 제재를 위한 교정적 범죄자 관리 프로파일링) 시스템을 개발한 노스포인트사는 재범위험성 평가에 사용되는 알고리즘의 상세한 내용을 공개하지 않는다고 한다. 이병규, 위의 논문, 176쪽.

안처분여부 및 기간을 결정함에 있어 AI에 의한 자료의 자동적 분석, 평가 결과에 따른 재범위험성을 고려한다고 한다. AI 재범위험성 예측에서 고려되는 요소가 범죄 경력, 고용상황, 학력, 연령, 성별, 사상, 가정환경, 친구관계 등이라고 하는데 이러한 요소 또는 그 요소의 합이 특정인의 특정 범죄의 재범과 관련될 수 있다고 논리적으로 설명될 수 없는한, 그리고 특히 집단적 특성, 사회적 편견에 노출되어 있는 지표는 그 자체로 합리적 이유제시에 있어 문제가 될 수 있는 소지가 다분하다는 점에서 AI의 재범위험성 평가는 그것을 받아보는 사람이 법관이 되었건 경찰관이 되었건 그 자체로 의심의 대상이 되어야 하는 참고자료로 보아야지 법관이 그에 종속되어서는 안될 것이다.

이런 관점에서 위스콘신주 대법원에서 LOOMIS 사례에서는 COMPAS의 사용이 적법절차의 원칙을 위반하지는 않았다고 판단하기는 하였지만, COMPAS 알고리즘의 정확성이 검증되지 않는 현 상황에서는 COMPAS의 평가결과 활용을 위해서는 엄격한 사용목적, 사용방법상의 한계가 설정되는 경우에 비로소 적법할 수 있다고 판시한 것은 의미가 있다. 그리고, 구금형(실형) 선고 여부, 형의 경중 판단, 피고인 보호 관찰 판단에는 COMPAS의 사용을 금지하고 COMPAS 사용시 해당 시스템은 작동 원리가 비공개이고, 개별 평가가 아니라 집단 평가이므로 주의하라는 문구를 삽입할 것을 결정한 것도 위에서 제시한 관점이 상당 부분 반영된 것이라 생각된다²³⁾.

범죄수사 영역에서 AI를 사용하는 경우에도 위스콘신주 대법원의 취지가 고려되어야 할 것이다. 이는 위 ③ - 2에서 언급한 정보상의 경찰활동에 AI의 활용에서도 이미 설명한 바와 같이 크게 다르지 않는 판단기준이라 할 것이다.

III. 경찰의 AI 활용 관련 독일 연방헌법재판소 결정

- 헤센주 경찰법 및 함부르크 경찰정보처리법상 자동적 자료 분석/평가 조항 -

§ 25a HSOG 자료분석을 위한 자동적 적용	§ 49 HmbPolDVG 보유자료의 평가를 위한 자동적 적용
(1) 경찰관청은 형사소송법 제100a조 2항에서 언급되는 범죄행위의 예방적 대처를 위해서 (Alt. 1), 또는 연방 및 주의 존립과 안전, 개인의 생명, 신체, 자유의 존립과 안전, 또는 그	(1) 경찰은 형사소송법 제100a조 제2항에서 언급되는 범죄행위의 예방적 대처를 위해서 (Alt. 1) 또는 연방, 주의 존립과 안전, 개인의 생명, 신체 또는 자유의 존립과 안전, 그 유지

²³⁾ 정보인권연구소, 인공지능과 정보인권 - 해외 논쟁사례(2): 차별, 이슈리포트 <정보인권> 2023-12(통권 제16호), 2023, 57쪽, 60쪽.

유지가 공공의 이익의 관점에서 명령되는 의미있는 물건의 존립과 안전에 대한 위험의 방지를 위해서(Alt. 2), 또는 동일한 무게의 환경상의 손상이 예상되는 경우에는(Alt. 3) 이유가 제시되는 개별적 사례들에서 자료분석을 위한 자동화된 적용을 할 때 개인정보를 재처리할 수 있다. (2) 제1항에 따른 재처리의 틀 내에서 특히 사람, 인적 그룹, 기관, 조직, 대상(Objekte) 및 사물(Sache)에 관한 관계 또는 맥락이 생산될 수 있고, 의미없는 정보나 인지는 제외될 수 있고, 세부적인 인지가 알려진 사실관계에 귀속될 수 있고, 저장된 자료가 통계적으로 평가될 수 있다. (3) 자료분석을 위한 자동화된 적용의 설치 및 변경은 관서장의 명령을 통해서 또는 그로부터 위임을 받는 공무원의 명령을 통해서 행해질 수 있다. 헤센주 정보보호담당관은 제1문에 따른 설치 및 중대한 변경 전에 청문할 권한을 가진다. 지체의 위험(Gefahr im Verzug)이 있는 경우에는 청문은 나중에 행해질 수 있다.	가 공익적인 관점에서 명령되는 사물의 존립 및 안전에 대한 위험의 방지를 위해서(Alt. 2) 이유가 제시되는 개별적인 사례들에서 자료평가를 위한 자동화된 적용을 할 때 경찰 정보 시스템에 저장된 개인정보를 재처리할 수 있다. (2) 제1항에 따른 처리의 틀 내에서 사람, 인적 그룹, 기관, 조직, 대상(Objekte) 및 사물(Sache)에 관한 관계 또는 맥락이 생산될 수 있고, 의미없는 정보나 인지는 제외될 수 있으며, 세부적인 인지가 알려진 사실관계에 귀속될 수 있고, 저장된 자료가 통계적으로 평가될 수 있다. (3) 제1항에 따른 자동화된 적용의 설치 및 변경은 경찰청장 또는 그 대변인의 명령을 통해서 수행될 수 있다. 함부르크 정보보호 및 자유 담당관은 제1문에 따른 설치 또는 중대한 변경 전에 청문할 권한을 가진다. 지체의 위험(Gefahr im Verzug)이 있는 경우에 청문은 나중에 행해질 수 있다.
---	--

1. 개요

독일 연방헌법재판소는 헤센주 경찰법(Das Hessische Gesetz über die öffentliche Sicherheit und Ordnung, HSOG) 제25a조 제1항 제1호(Alt. 1) 및 함부르크 경찰정보처리법(Das Hamburgische Gesetz über die Datenverarbeitung der Polizei, HmbPolDVG) 제49조 제1항 제1호(Alt.1)를 위헌으로 결정하였다²⁴⁾. 해당 조항들은 범죄행위에 대한 예방적 대처(양 경찰법률 해당 조항 Alt.1) 또는 위험방지(양 경찰법률 해당 조항 Alt.2)를 위하여 경찰에게 자료분석(헤센) 또는 자료평가(함부르크)의 측면에서 한 번의 클릭으로 광범위한 개인정보가 저장된 다양한 데이터베이스를 AI의 도움을 통해 활용할 수 있는 분석 소프트웨어 시스템의 이용과 관련된 것이었다. 해당 조항은 경찰의 이와 같은 AI 시스템 이용과 관련된 특별한 법적 근거를 규정해 놓았는데, 분석플랫폼으로 이제까지 연결되어 있지 않은 자동화된 데이터 및 데이터

²⁴⁾ Bundesverfassungsgerichtsurteil vom 16. Februar 2023, 1 BvR 1547/19, 1 BvR 2634/20.

원천(data source)을 연결하고 이미 경찰이 보유하고 있는 데이터를 검색기능을 통해 체계적으로 이용하도록 할 수 있게 하였다.

자세히 살펴보면 해당 규정들은 그 이유가 정당화될 수 있는 개별 사례들에서 독일 형사소송법 제100a조 제2항에 따른 중범죄에 대한 예방적 대처를 위하여(Alt.1) 또는 특정한 법익에 대한 위협의 방지를 위하여(Alt.2) 경찰로 하여금 자료분석(헤센) 또는 자료평가(함부르크)의 측면에서 저장된 개인정보를 자동화된 기술로 처리할 수 있도록 하는 권한규정을 마련하였다. 이와 같은 방식으로 개인, 인적 그룹, 기관, 조직, 대상(Objekte) 및 사물(Sachen)에 대한 관계 또는 상관관계(맥락)가 생산되고, 의미없는 정보나 인식들은 제외되어서, 파악되는 인식에 알려진 사실관계가 정리되고 저장된 정보가 통계적으로 분석, 평가될 수 있도록 하였다.

실제로 헤센에서는 헤센주 경찰법 제25a조(§ 25a HSOG)를 활용하여 „hessenDATA“라는 분석 플랫폼이 매년 수천회 사용되었다. 반면, 함부르크 경찰정보 처리법 49조(§ 49 HmbPolIDVG)는 독일 연방헌법재판소 결정 때 까지 함부르크 경찰이 실제로 적용한 바는 없었다.

2. 독일 연방헌법재판소의 판단

가. 기본권 제한 여부

독일 연방헌법재판소는 범죄행위에 대한 예방적 대처를 위하여 저장된 원자료가 자동분석/평가 시스템에 의하여 처리된다면 이는 독일 기본법상의 일반적 인격권과 관련된 개인의 정보자기결정권을 헌법적으로 충분한 침해문턱(Eingriffsschwelle)을 제시하지 못한채 침해하는 것이라고 보았다. 특히, 이전에 별도로 저장된 자료를 사용할 때 뿐만이 아니라 나아가 이를 이용하여 자동화된 분석/평가 시스템을 통하여 창조될 수 있는 기본권과 관련되는 새로운 지식(Wissen)을 획득할 때도 기본권 침해가 있을 수 있다고 보았다.

따라서, 이런 상황에서 범죄행위를 위한 예방적 대처를 위한 자동화된 자료분석/평가는 헌법적인 정당화 사유를 필요로 한다고 보았다. 이와 같은 자동화된 자료분석/평가가 기본적으로 가능하지만, 해당 권한과 관련된 구체적인 범위(Reichweite)에 따라 그 요구조건이 설정될 때 비례의 원칙과의 부합성을 반드시 고려하여야 한다. 이때 관련 요구조건은 해당 권한과 관련된 구체적인 범위에 따라 다양할 수 있다²⁵⁾.

문제된 사안에서 헤세주 경찰법 제25a조 제1항 제1호(Alt. 1) 및 함부르크 경찰정

²⁵⁾ BVerfGE 165, 363, 388.

보처리법 제49조 제1항 제1호(Alt. 1)는 경찰의 범죄행위를 위한 예방적 대처시 자료 분석/평가(즉, AI의 사용)에 대한 합당한 침해문턱(Eingriffsschwelle)이 제시되지 못하였으므로 헌법에 위반된다. 다만, 위 두 주의 경찰 관련 법률들에서 위험방지와 관련된 권한들(Alt.2)은 헌법에 위반되지 않는다²⁶⁾.

나. 비례의 원칙 위반 여부

1) 목적의 정당성

해당 권한 규정은 정보기술의 발전상 하에서 중범죄에 대한 예방적 대처에 실효성을 높인다는 점에서 정당한 목적에 기여한다. 왜냐하면, 이를 활용하지 않았다면 경찰 데이터베이스(Datenbestand, ‘경찰 원자료’로 번역가능)에 경찰에 알려지지 않은 채 존재할 수 있었던 장래에 발생할 중범죄에 대한 근거가 획득될 수 있기 때문이다. 특히, 최근 경찰은 테러, 극단적 범죄 및 조직범죄, 중범죄 영역에 있어서 점점 디지털 미디어와 커뮤니케이션 수단이 사용되고, 그 질(Qualität)과 형태(Format)별로 점점 다양한 정보에 노출되는 상황에 직면하여 있다.

2) 수단의 적합성

경찰 데이터베이스에 존재하는 정보는 시간적인 압력으로 인하여 적시에 수동으로 획득될 수 없고, 따라서 자동화된 자료분석/평가는 성공적인 경찰활동을 위하여 큰 의미를 가진다. 예방적 범죄행위에 대한 예방적 대처의 효과를 증대시키기 위해서 자동화된 자료분석/평가는 적합한 수단이 된다.

3) 최소침해의 원칙

그리고, 범죄행위의 예방적 대처에 있어서 기본권을 덜 침해하는 다른 방식으로는 그 정도 수준으로 획득될 수 없는 관련된 인식/인지(Erkenntnisse)가 자동화된 자료분석/평가를 통해 추론될 수 있기 때문에(erschlossen werden können) 해당 수단은 최소침해의 원칙에도 부합한다²⁷⁾.

4) 협의의 비례의 원칙(상당성)

결국 중요한 것은 관련된 특별 권한규정과 관련된 상당성의 원칙의 문제이다. 개별

²⁶⁾ BVerfGE 165, 363, 388.

²⁷⁾ BVerfGE 165, 363, 389.

적으로 얼마나 엄격하게 요구조건이 설정되어야 할지는 해당 수단의 침해정도에 따라 결정될 수 밖에 없다. 법익균형성과 관련된 내용이 길기 때문에 이 부분은 항목을 바꾸어서 설명하도록 한다.

다. 헌법적 정당화를 위한 요건

1) 목적구속의 준수 및 목적 변경을 위한 요건

먼저 자동화된 자료분석/평가 시스템의 침해정도는 기존 연방수사청법 결정²⁸⁾에서 어느 정도 윤곽이 형성된 목적구속 및 목적변경의 원칙과 관련되어 설명될 수 있다.

① 목적준수 계속 이용

입법자들은 정보수집을 위한 기준이 되는 절차를 넘어서 자료이용과 할지라도 애초 자료의 수집목적의 측면에서 계속된 이용(*weitere Nutzung*)이라면 허용하고 있다(소위 목적을 유지한 정보의 계속 이용 *zweckwahrende Weiternutzung*).

동일한 법익의 보호를 위한 동일한 임무 범위 내에서 동일한 법익의 소추나 예방을 위해 동일한 관청에서 행하는 자료의 모든 계속적 이용(*jede weitere Nutzung*)을 위해 매번 새로이 고려되어야 하는 목적구속의 원칙에 기본적으로 자료수집을 위한 기준이 되는 - 예를 들어 위험방지 영역에서의 충분히 구체화된 위험, 형사소추에서의 질적인 범죄혐의와 같은 - 침해문턱에 대한 요구조건들이 속하는 것은 아니다. 다시 말해서 구체화된 위험 또는 질적인 범죄혐의에 대한 요구(*Erfordernis*)는 애초 관련 정보가 수집되어도 좋다는 동인(계기)를 정하는 것이지 수집된 관청의 자료가 계속 이용되어도 좋다고 허용하는 목적을 정하는 것은 아니다.

요컨대, 애초 정보수집을 위한 요구조건 - 위험방지 영역의 정보수집을 위한 충분히 구체화된 위험, 형사소추에서의 질적인 범죄혐의 - 이 충족되지 않더라도 이와 같은 애초의 목적이 유지된 계속적 이용은 동일한 관청에서 동일한 임무수행의 측면에서 그리고, 동일한 법익보호를 위하여 정당한 고려의 대상이 될 수 있는바, 기존에 수집된 정보를 단순한 단서(*bloßer Spurenansatz*)로서 계속 이용하는 것은 허용된다²⁹⁾. 즉, 이 이런 경우에는 목적구속의 원칙에 위반되지 않는다.

²⁸⁾ BVerfGE 141, 220 ff.

²⁹⁾ BVerfGE 165, 363, 391 f.

② 목적변경을 통한 추가이용

입법자들은 애초 수집목적과 다른 목적으로의 자료의 계속적 이용을 허용하기도 한다(소위 애초 수집목적과는 변경된 정보의 계속적 이용. zweckändernde Weiternutzung). 비례의 원칙 심사의 기준으로서 가정적 자료수집(hypothetische Datenneuerhebung)이라는 기준이 적용된다. 만약 개별 사례에서 구체적인 수사단서가 유사한 정도의 무게를 가지는 범죄의 발견을 위하여 또는 적어도 유사한 정도의 무게를 가지는 범익에 대한 중대한 위협의 방지를 위하여 얻어질 수 있는 정보와 관련된 것이라면 입법자들은 경찰관청에 목적이 변경된 정보의 계속적 이용을 허락할 수 있다. 이와 같은 경우에 있어서 그 수단의 기본권 침해성이 매우 큰 주거감청(Wohnraumüberwachung) 및 온라인 수색(Online-Durchsuchung)으로부터 얻어진 자료에 대해서는 더욱 더 엄격한 전제조건 하에서 목적변경이 허용된다.

③ 소결

사안의 § 25a HSOG 과 § 49 HmbPolDVG 는 개인정보를 목적을 유지하는 가운데서도, 그리고, 목적을 변경하는 가운데서도 계속 처리할 수 있도록 하고 있다. 해당 규정들은 기본적으로 자료의 원천이나 애초의 수집목적은 구분하지 않고 광범위한 자료의 이용을 허락하고 있다. 독일 연방헌재판소는 목적구속의 원칙의 헌법적 요구조건을 충족시키기 위하여 현재와는 다른 방식으로 충분히 규범적으로 명확한 규정들이 정립되어만 한다고 보았다.

2) 자동화된 데이터 분석을 위한 추가적 요건

나아가서 한 번 적법하게 수집되어 저장된 자료라 할지라도 자동적인 분석/평가를 통해서 고유한 기본권 침해와 관련된 무게를 가질 수 있다. 왜냐하면, 한 번 획득되어 저장된 자료를 자동적 자료 분석/평가 시스템을 통해서 계속 처리하는 것은 애초 수집시 침해정도를 넘어서는 부담적 효과를 가질 수 있기 때문이다. 이러한 측면에서 상당성의 원칙과 관련하여 자동적 자료분석/평가 권한에는 현재보다 더 나아간 요구조건이 세워져야 한다.

① AI의 활용 데이터 분석의 독자적 기본권 침해성

자료 분석/평가 시스템은 기본적으로 새로운 인지(認知, Wissen)를 생산하는 것을 목적으로 한다. 해당 기관은 자신이 획득한 자료를 기반으로 정보기술적으로 가능한 모든 수단을 활용하여 더 나아간 인지를 획득할 수 있고, 분석을 통해 새로운 관련성 및 맥락을 도출해낼 수 있다. 물론 경찰이 한번 획득된 인지를 그 자체만으로 추적단서 내지 수사단서로 또는 경찰이 보유하는 다른 정보와 결합시켜 계속적인 수사의 출발점으로 이용하는 것은 드문 일은 아니다. 그러나, 자동적 분석/평가는 광범위하고 복잡한 원정보들(Informationsbestände)의 처리를 가능하게 하기 때문에 기존의 방식을 훨씬 넘어서는 것이다. 그 외에도 새로 투입된 분석수단에 따라 기존에 보유하는 정보의 연결된 평가를 통해서 - 그렇게 하지 않았다면 접근할 수 없었던 - 새로운 인격 관련 정보가 획득될 수도 있다. 결국 해당 수단은 기존 수단보다 자료에 내재되어 있는 정보를 보다 더 심도있게(intensiver) 도출해낸다. 해당 수단은 자료에 내재되어 있는, 그러나 맨처음에는 연결이 되지 않아 감추어져 있던 개인에 관한 인지(Erkenntnisse)를 이끌어 낼 뿐만이 아니라, 적당하게 투입하는 과정을 통해 개인에 관한 프로파일링(„Profiling“)에 근접할 수 있다. 왜냐하면, 관계와 맥락에 대한 자료와 알고리즘에 따라 산출된 가정이 주변환경으로부터 포함해서 고려될 수 있다면 소프트웨어에 기반한 새로운 기술들은 개인의 모습(Bild)에 관한 완전한 그림을 도출해낼 수 있기 때문이다. 목적 구속의 원칙은 자동적 자료분석/평가가 가지는 기본권 침해의 이러한 양상에 대하여 충분한 고려를 할 수 없다.

② 침해의 정도에 따른 차등적 헌법적 요구 설정

이러한 점에서 법적인 구성에 따라 고유한 침해정도가 완전히 차등적일 수 있기 때문에 자동적 자료분석/평가에 대한 헌법적인 요구는 다양할 수 있다.

aa) 일반적으로 정보의 자기결정권에 대한 침해의 무게는 유형, 범위, 그리고 자료의 생각할 수 있는 이용 및 오용의 위험에 따라 정해지게 된다. 그 외에도 자료분석/평가의 허용된 수단이 침해정도에 영향을 미친다. 특히 자료비교(Datenabgleich)의 복잡한 방식의 활용은 상당한 침해무게를 가질 수 있다. 경찰이 실질적으로 모든 가능한 정보기술 수단을 활용하여 이용가능한 자료(Daten)로부터 광범위한 인식을 추출해내고(abschöpfen), 그로부터 새로운 상관관계를 이끌어내고, 나아가 다단계 분석을 통하여 새로운 혐의정황(Verdachtsmomente)을 도출해낸 다음, 이에 기반하여(hieran) 경찰이 새로운 분석단계 또는 적극적 조치를 이어나갈 수 있게 된다면, 자동적인 자료분석/평가에 따른 당사자에 대한 불이익은 매우 심대할 수 있다. 이외에도 자료비교

(Datenabgleich)의 복잡한 형태에 있어서는 개인의 법익 보호, 감독적 통제, 그리고 이를 위해 필수적으로 요구되는 실수를 인지하거나 교정할 가능성의 관점에서 투입된 알고리즘의 추론가능성(Nachvollziehbarkeit)에 대한 어려움이 존재한다.

전체적으로 자동적 자료 분석/평가 수단은 개인에 관한 인지가 넓어지면 넓어질수록, 깊어지면 깊어질수록, 오류 취약성(Fehleranfälligkeit) 및 차별 취약성(Diskriminierungsanfälligkeit)이 더 높아지면 높아질수록, 그리고, 소프트웨어에 기반한 연관(연결고리, Verknüpfung)이 더 어렵게 추론가능하면 할수록 더 침해적이게 된다³⁰⁾. 앞에서 제기되는 여러 요인에 의한 침해 강도가 더 크면 클수록 그만큼 그와 같은 수단의 정당화를 위해서 더 엄격한 조건이 요구된다고 할 것이다.³¹⁾

bb) 허용되는 침해전제들에 대한 헌법적인 요구조건들은 자료(데이터)의 유형과 범위, 그리고 평가방법의 제한에 대한 규정들을 통해서 입법자에 의해 조정가능한 침해강도(Eingriffsintensität)와 부합한다. 침해정도(Eingriffsgewicht)에 부합하는 협의의 비례의 원칙(상당성 원칙)의 요구조건들은 해당 수단으로 보호하고자 하는 법익 뿐만이 아니라 침해의 문턱(Eingriffsschwelle), 그러니깐, 투입되는 수단의 동인(계기)을 감안하여 설정된다.

자동적 적용(automatisierte Anwendung, 즉, 자동적 자료분석/평가를 말함)이 당사자의 개인정보자기결정권에 중대한 침해를 가능하게 한다면, 이는 단지 좁은 전제조건들 하에서 합리화될 수 있다. 예를 들어 침해강도가 큰 당사자가 모르는 가운데 행해지는 은밀한 감시수단에 대해서는 일반적으로 이와 같은 좁은 전제조건들이 설정되어야 한다. 자동적 적용은 은밀한 감시수단의 허용요건에서처럼 일단은 단지 특별히 중대한 법익 – 예컨대, 생명, 신체, 개인의 자유 등 – 의 보호를 위한 충분히 구체화된 위험(hinreichend konkretisierte Gefahr)이라는 침해문턱이 제시되는 경우에만 허용된다³²⁾. 이 때 충분히 구체화된 위험이라는 요건은 언급한 중대한 보호법익에 대한 구체적 위험의 발생에 대한 최소한의 사실적 근거가 존재할 것을 전제로 하는바, 단순히 일반적 경험으로는 충분하지 않다. 충분히 구체화된 위험은 손해를 야기하는 인과관계가 아직 충분한 개연성으로 예상되지 않는 경우라 할지라도 특정한 사실관계들을 통해서 개별 사례에 매우 중대한 법익에 대한 침해가 암시되는 경우 존재한다고 할 수 있다³³⁾.

30) BVerfGE 165, 363, 404 f.

31) 정애령, 예측치안을 위한 데이터 활용의 정당성과 헌법적 한계, 공법연구 제54집 제1호, 2025, 388쪽.

32) BVerfGE 165, 363, 410 f.

33) BVerfGE 165, 363, 411.

구체적 위험이 있고, 적어도 중대한 무게를 가진 법익의 보호에 기여한다면, 예를 들어 적어도 중대한 의미를 가지는 범죄행위의 예방(Verhütung)에 기여한다면 적은 무게를 가지는 침해는 이미 정당화될 수 있다. 반대로 높은, 매우 중요하거나 또는 특별히 중대한 법익의 보호에 관련 수단이 기여하는 한 아직 구체화된 위험의 수준에 못미치는 침해의 문턱으로도 충분할 수 있다³⁴⁾.

AI 판단을 위해 관련지을 수 있는 자료들이 법적으로 유형과 범위에 있어서 하나의 방식으로 단순화될 수 있다면, 그리고 자동화된 분석 또는 평가의 가능한 방법들이 처음부터 제한될 수 있다면 그와 같은 권한에 터를 잡은 수단들은 관청이 자동화된 적용없이 더 노력을 들여 더 천천히 현실적으로 획득할 수 있는 것보다 당사자의 인격적 생활영위에 더 깊게 엿보는 것(Einsicht)까지 이르지 않을 것이거나, 또는 그 권한들은 처음부터 개인관련 정보를 생산하는 것 없이 위험하거나 위험해진 장소를 확인하는 것을 목적으로 할 수 있게 된다. 이와 같은 경우 자동화된 자료의 처리를 정당화하기 위하여 목적구속의 원칙은 이미 충분히 준수될 수 있다. 범죄행위의 예방적 대처를 위해서 완전히 동인(계기)없는 경찰관청에 의한 자동화된 개인정보의 평가는 헌법적으로 허용되지 않는바, 헌법적인 목적구속의 원칙의 준수는 침해동인(계기)이 있는 경우에 준수될 수 있다³⁵⁾.

cc) 만약 규범적으로 명확하게, 그리고 개별 사안에서 충분히 명확하게 허용되는 자동적 분석, 평가가능성이 좁게 제한된다면, 그래서 수단의 침해강도가 현저히 낮아진다면 특별히 중요한 법익에 대한 최소한의 구체화된 위험이라는 침해의 문턱은 포기될 수 있다³⁶⁾. 기본적으로 입법자들은 자신들과 행정 사이에서 이와 같은 규율임무(Regelungsaufgabe), 즉, 법정립 임무를 분배할 수 있다. 입법자들은 자료의 형태와 범위의 제한을 위해서, 그리고 자료처리방법의 제한을 위해서 법률유보를 유지하는 가운데 전체적으로 충분한 규율이 행해질 수 있도록 담보해야 한다(sicherstellen).

직접적으로 입법자들에 의해 규범화되지 않는 측면의 규정을 위해서 법규명령에 대한 수권이 고려될 수 있다. 나아가 여기서 입법자들은 행정으로 하여금 법에서 또는 법규명령에서 규정된 제약(Vorgabe)들을 추상적-일반적 형태로 계속 구체화할 것을 의무지을 수 있다. 어떤 경우든지 행정규정을 통한 구체화는 그러나 법적인 근거를 필요로 한다. 여기에서 입법자들은 개별 사례에서 규정들의 적용을 위해 기준이 되는

³⁴⁾ BVerfGE 165, 363, 411.

³⁵⁾ BVerfGE 165, 363, 412.

³⁶⁾ BVerfGE 165, 363, 413.

구체화, 표준화가 관청의 측면에서 추론가능하게 문서화되고 공포될 것을 담보해야 한다(sicherstellen).

③ 소결

제시된 일반적인 기준에 따르면 데이터나 수단에 개방적이게 형성된 헤센주 경찰법 제25a조 제1항 제1호, 함부르크 경찰정보처리법 제49조 제1항 제1호의 권한규정들의 특별한 침해강도는 잠재적으로 매우 높아서 이와 같은 규정들은 헌법적으로 엄격한 전제조건들을 충족해야만 한다. 해당 규정들은 경찰에 단 한번의 클릭으로 개인, 그룹, 그리고 주변환경(Milieus)에 대한 포괄적인 프로파일을 생산할 수 있도록 허용하고, 또한 자동화된 평가가 수많은 법적으로 관여되지 않은, 어떠한 맥락 속에서 자료를 남긴 사람들에 대하여 경찰을 잘못된 자취(흔적)로 데려가는 경우 당사자들로 하여금 계속된 경찰상 수단들을 감내하도록 한다. 따라서, 특별히 중요한 법익에 대한 구체화된 위협이 위 규정에 따른 자료의 자동적 분석/평가를 위한 전제로서 요구된다.

aa) 위 두 규정들은 자료분석 또는 평가에 있어서 투입되는 자료의 유형과 범위를 전혀 제한하지 않는다. 해당 조항들은 자동화된 자료분석/평가를 위해 자료의 어떠한 유형, 그리고 어떠한 원자료들(Datenbestände)이 이용가능한지 규정하고 있지 않다. 특히, 해당 조항들은 자료분석/평가의 대상이 되는 사람 중에서 범죄행위를 행하거나 그러한 사람과 특별한 관계에 있다는 가정을 위한 동인(계기)를 제공하는 사람과(vgl. §2 ATDG), 반대로 처리한 자료나 기지국 대역(Funkzelle) 조회 결과 많은 자료가 관계되어 있지만 중대한 범죄행위의 예방적 대처를 위한 수단을 위하여 어떠한 동인(계기)도 제공하는 않는 사람(예컨대, 피해자, 증인 또는 경찰상 비책임자 등)의 자료를 구분하지 않는다. 즉, 해당 규정들은 자료의 자동적 분석/평가의 대상이 되는 사람이 중대한 범죄행위를 할 수 있다거나, 이와 같은 사람과 특별한 연계 하에 있다거나 하는 조건을 제시하지 않고 있다. 해당 규정들은 경찰 수사수단에 노출될 수 있는 (unterziehen) 관여되지 않은 사람의 자료의 폭넓은 관련지움(Einbeziehung)을 허용한다³⁷⁾.

bb) 문언(Wortlaut)에 따르면 해당 규정들은 이외에도 자동화된 자료분석, 평가의 광범위한 방법들을 허용한다. 입법자들은 어떠한 분석과 평가의 방법이 허용되는지를

³⁷⁾ BVerfGE 165, 363, 419.

국한하지 못하였다. 사안의 규정들은 “데이터 마이닝”³⁸⁾ 시 스스로 학습하는 시스템, 즉, AI의 적용까지도 가능하게 하였다. 이와 관련하여 특히 개방적인 검색과정이 허용된다. 자료평가 또는 분석은 방대한 자료 속에서 - 그를 통해 계속된 자동화된 적용의 도움으로 결론이 도출될 수 있는 - 통계적인 유의성(Auffälligkeit)을 발견하는 것을 목적으로 한다. 해당 규정들은 또한 얻을 수 있는(erzielbaren) 검색결과와 관련하여 어떤 것도 제외하지 않는다. 문언에 따르면 검색결과 - “예측적 경찰활동(predictive policing)”의 의미에서 사람에 관한 위험성 언명(Gefährlichkeitsaussagen)에 이르기까지 - 기계적인(maschinelle) 사실관계 평가 하에 놓여져 있을 수 있다. 말하자면 자료분석, 평가를 통해서 해당 방법이 아니면 접근하지 못하는 새로운 인격 관련 정보가 생성될 수 있다. 얻을 수 있는 새로운 지식의 잠재적인 범위는 그 적용에 있어서 침해를 약화시키는 규정들을 통해서 그 범위가 국한되지 않는다³⁹⁾.

함부르크에서 입법자들은 “데이터 분석”이라는 용어 대신에 “데이터 평가”라는 단어를 사용함으로써 광범위한 적용을 배제하려는 시도를 행하였다. 단지 특정한 검색 기준들 (bestimmte Suchkriterien)에 의한 자동적인 적용을 통해서 부합여부가 입증된다는 헌법적으로 충분한 언명은 해당 함부르크 경찰정보처리법 규정에서 “데이터 분석” 대신에 “데이터 평가”라는 단어를 사용했다고 해서 성공하지 못하였다⁴⁰⁾.

cc) 헌법소원의 대상이 된 규정의 권한들은 무제한적인 자료평가 기술이 현재 존재하지 않는다는 사정을 통해서 헌법적으로 관련성 있게 제한되지 못하였다. 계속적인 기술발전의 결과 기능확장이 가능하다면 헌법적인 요구들은 기본적으로 이와 같이 기능확장된 기술에 따라 형성된 침해가능성에 따라 정해지게 된다. 해당 규범의 침해 무계가 권한의 범위가 제한된다는 입법자의 단순한 상상 또는 법적인 권한의 이용가능성이 포괄적으로 사용하지 않을 것이라는 행정의 의지에 따라 기준이 정해지는 것은 아니다⁴¹⁾.

함부르크 경찰법 제25a조 제1항 제1호 및 함부르크 경찰정보처리법 제49조 제1항 제1호의 규정들은 충분한 침해문턱을 제시하지 못하였기 때문에 협의의 비례의 원칙의 요구를 충족하지 못한다.

38) 데이터 마이닝(data mining)은 대량의 데이터에서 유용한 정보를 추출하는 것을 말한다. 데이터 마이닝은 다양한 통계적 기법, 수학적 기법과 인공지능을 활용한 패턴인식 기술 등을 이용하여 데이터 속에서 유의미한 관계, 규칙 패턴 등에 대한 규칙을 발견한다. (다음백과 2026.7.5. 검색).

39) BVerfGE 165, 363, 428 f.

40) BVerfGE 165, 363, 429.

41) BVerfGE 165, 363, 429.

1. 형사소송법 제100a조 제2항에 언급된 범죄행위에 대한 예방적 대처의 목적으로 자동적 자료분석 또는 평가를 할 수 있도록 하는 해당 규정들은 제시된 침해의 강도와 관련하여 침해동인(계기)이 비례의 원칙에 부합하지 않게 넓고 따라서 위헌적으로 규정되었다. 또한 이유제시되는 개별 사례가 있어야 한다는 양 규정의 넓은 기준은 여기에서 어떤 자세하게 내용적으로 확정된 것을 전혀 담고 있지 못하다. 헌법재판소의 구두 변론과정에서 헤센주 경찰실무에서의 “개별 사례들”이라는 전제가 이해되고 적용되는 좁은 개념이 설명되기는 하였다. 하지만 이미 행해진 범죄행위 또는 적어도 이미 행해진 범죄행위에 대한 사실관계를 통해 입증되는 혐의가 관련되고 그로부터 미래에 대한 진단이 도출된다. 한편으로는 과거에 이미 형사소송법 제100a조 제2항에 따른 범죄가 행해졌다는 것으로부터 시작해야 한다. 다른 한편으로는 이와 같은 상황을 근거로 미래에 계속되고, 누적될 범죄행위가 예상되어야만 한다.

헤센주 적용실무의 자세한 형태와 관계없이 헤센주 실무의 개념은 처음부터 최소한 구체화된 위험 및 그 위험방지에 적합한 자료의 확인을 목표로 하고 있지 않기 때문에 헤센주 경찰법에서 헌법적인 요구는 그 당시에 이미 충족되지 않는다. 헤센주 경찰법 제25a조 및 함부르크 경찰정보처리법 제49조에서의 권한들의 자료 및 방법은 열린 형식 때문에 앞서 언급한 헌법적 요구가 필요하다.

헌법소원의 대상이 된 규정들은 형사소송법 제100a조 제2항이 언급하는 범죄카탈로그만으로 그 위험구성요건이 포착된다고 해서 충분한 침해문턱을 제시하고 있는 것은 아니다. 비록 입법자에게 침해전제조건의 규정을 위하여 사전적 범죄구성요건을 행하는 위험과 연결시키는 것이 헌법적으로 허용되지 않는 것은 아니지만, 입법자는 특히 모든 개별 사례에서 범죄구성요건을 통해 보호되는 법익을 위한 구체적 또는 구체화된 위험이 존재한다는 것을 담보해야 한다. 그러나, 여기에서는 그런 부분이 결여되어 있다.

2. 헤센주 경찰법 제25a조 제1항 제1호 및 함부르트 경찰정보처리법 제49조 제1항 제1호의 범죄행위에 대한 예방적 대처는 법적 정의에 따르면 범죄행위의 예방 뿐만 아니라 미래의 범죄행위의 소추를 위한 배려까지도 포함한다. 경찰 원자료들(Datenbestände)은 자동화된 자료평가의 방법으로 미래의 석명임무와 수사절차에 대한 인지를 얻기 위해서 이용되어야 한다. 해당 규정들로부터 최소한 구체적 위험 또는 구체화된 위험이 존재해야 한다는 사정이 주어져 있어야 한다는 것이 도출될 수 없다. 해당 규정들에는 모든 침해동인(계기)을 제한시키는 구체화가 결여되어 있다⁴²⁾.

⁴²⁾ BVerfGE 165, 363, 431.

C. 헤센주 경찰법 제25a조 제1항 제1호는 새로운 입법규정이 2023.9.30. 한(限) 제정될 때까지 유효하다. 입법자가 국가적 임무수행을 위한 권한을 존중한다는 의미의 관점에서, 그리고 헤센주 경찰실무에 대한 의미 때문에 무효로 판단하기 보다는 제한된 유효기간 부여가 감내되어야 한다.

그러나, 제한된 기간동안 해당 규정에 대한 지속적 유효명령과 관련된 기본권의 관점에서 입법자를 통해 새로운 규정이 선결되지 않았기 때문에 제한된 기준을 요한다. 새로운 입법이 있기 전까지 헤센주 경찰실무에서 선택된 개념을 기반으로 헤센주 경찰법 제25a조 제1항 제1호의 권한은 다음과 같은 경우에 제한적으로 적법한 것으로 허용될 수 있다. 형사소송법 제100b조(온라인 수색) 제2항의 의미에서의 특별히 중요한 범죄행위가 행해졌고, - 생명, 신체, 또는 연방 및 주의 존속에 위해가 되는 - 미래에 계속되고 누적되는 범죄행위가 예상되는 그와 같은 개별 사례에 존재하는 범죄 혐의의 구체적인 정황을 근거로, 이와 같은 전제조건 존재, 그리고 독립적으로 표현되는 언명을 통해서 예상되는 범죄행위의 예방을 위해 사용되는 자료의 구체적인 적합성(Eignung)이 정당화될 수 있는 경우에, 그리고, 주거감청, 온라인수색, 전화감청, 교통자료조회, 장기간 관찰, 은밀한 신분위장 공무원 또는 신뢰관계 있는 사람의 투입, 또는 앞의 수단들과 비교가능한 정보자기결정권에 중대한 침해를 통해서 얻어지는 정보가 정보분석에 관련되지 않는다는 것이 보증되는 경우에 한하여.

함부르크 경찰정보처리법 제49조 제1항 제1호는 무효이다. 여기에는 해당 규정에 대한 제한적 유효명령을 요청하거나 정당화시킬 상황이 존재하지 않는다.

④ 평가

대상 판결은 경찰의 범죄의 예방적 대처를 위한 자동화된 데이터 분석은 독자적 침해성을 가지므로 목적 구속의 원칙 등 정보보호법상 요건 외에도 기본권 침해 정도에 따라 요건을 구체적이고 차등적으로 규율할 것을 요구하였다.

연방헌법재판소는 자동화된 데이터 분석은 침해정도(데이터의 유형, 범위, 분석방법에 따라 결정)가 높을수록 헌법적 정당화를 위한 요건(보호법익의 중대성, 침해기준에 따라 결정)이 엄격하여야 하며, 침해정도가 낮을수록 헌법적 정당화를 위한 요건이 완화될 수 있다는 취지의 결정을 하였다⁴³⁾.

43) 박원규, 범죄예방 목적 자동화된 데이터 분석에 관한 독일 연방헌법재판소 판례 분석 - BVerfG 2023.2.16. 선고 1 BvR 1547/19, 1 BvR 2634/20 판결, 형사정책연구 제36권 제1호, 2025, 13쪽.

IV. 경찰의 위협판단과 AI

1. 경찰활동의 법치국가적 통제기준으로서의 위협 개념

독일의 경찰법은 위협 개념을 기준으로 경찰임무의 개시여부, 나아가 경찰권한 사용여부가 결정될 수 있도록 법치국가적 통제기제를 발전시켜왔다. 즉, 위협한 상황에 이르러서야 비로소 경찰에게 임무가 주어지고, 경찰법에 위협상황 별로 개별 권한규정을 규정하여 경찰이 위협방지 임무수행을 위하여 개인의 기본권을 제한할 수 있는 경찰수단에 대한 근거를 마련해 놓았다. 그리고, 위협은 그 성격상 모든 위협상황을 유형별로 포섭하거나, 개별적 위협상황에 대처하기 위한 모든 권한을 모두 유형별로 포섭하여 규정할 수 없기 때문에, 즉, 이 경우는 개별 위협상황에 따른 개별적 권한을 규정해놓은 개별적 수권조항만으로는 적절하게 위협방지에 대처할 수 없기 때문에 독일 경찰법은 이에 대비하여 별도로 일반적 수권조항을 마련해 두고 있다⁴⁴⁾⁴⁵⁾.

어찌되었건 개별적 수권조항에 의하든, 일반적 수권조항에 의하든 기본적으로 구체적 위협이 존재하지 않는 경우, 좀 더 넓게 보면 위협과 관련되지 않은 경우에⁴⁶⁾ 경찰이 개입하는 것은 경찰의 임무를 벗어난 일이 되고, 이 경우의 권한행사는 위법한 것이 되도록 설계되어 있다⁴⁷⁾.

일반적으로 경찰작용, 특히 그 중에서도 개괄적 수권조항에 근거한 경찰작용의 정당화 근거가 되는 ‘구체적 위협’은 ‘어떤 상황의 정상적인 진행을 방지하게 되면, (정상적인 인과관계에 따를 경우) 가까운 장래에 어떤 상태 또는 행위가 공공의 안녕 또

44) 예를 들어 노르트라인 베스트팔렌주 경찰법 제8조 제1항에서는 “경찰은 - 이 법에서 정한 제9조부터 46조까지의 권한이 특별히 규정하지 않고 있는한 - 개별 사례에서 공공의 안녕 또는 질서에 대한 구체적 위협을 방지하기 위하여 필요한(notwendige) 조치를 취할 수 있다”고 규정하고 있다.

45) 원칙적으로 일반적 수권조항은 구체적 위협의 방지를 위해서만 그 근거규정이 될 수 있다. 그러나, 최근 위협의 의심상황과 관련된 확인조치(Gefahrerforschung)를 하기 위한 법적 근거가 없는 현재의 경찰법제 상황에서 일반적 수권조항을 근거를 유추해석의 근거규정으로 삼을 수 있다는 견해가 주장되고 있다(김성태, 위협에 대한 의심과 위협여부의 확인, 행정법연구 제51호, 2017, 168쪽 이하).

46) 대부분의 개별적 수권조항들은 구체적 위협을 권한발동의 전제요건으로 하지만, 불심검문 등과 같은 경우에는 아직 구체적 위협이 존재하지 않은 경우에도 경찰이 단계적으로 이를 확인하기 위한 조치, 권한을 행사할 수 있도록 설계되어 있기도 하다.

47) 이러한 사정을 독일의 한 경찰법 교과서는 “단지 위협이 있는 경우에만 경찰질서관청은 무엇인가를 행하는 것이 허용된다(Nur bei Gefahr dürfen die Polizei- und Ordnungsbehörde handeln)”, “단지 위협이 있는 경우에만 경찰질서관청은 (시민의) 자유를 제한하는 것이 허용된다(Nur bei Gefahr dürfen sie die Freiheit beschränken)”라고 표현하고 있다(Pieroth/Schlink/Kniesel, Polizei- und Ordnungsrecht, 7. Aufl. 2012, S. 57).

는 질서에 손해를 가져올 충분한 개연성이 있는 경우’를 말한다⁴⁸⁾.

개괄적 수권조항이 아닌 대다수의 개별적 수권조항도 구체적 위험을 경찰작용, 즉, 경찰개입의 정당화 근거로 규정하고 있는 경우가 대부분이지만, 불심검문에서 보여지는 것처럼 경찰개입을 위한 근거와 절차가 명확하게 제시된다면 반드시 구체적 위험을 요구하지 않는 경우도 있다. 하지만, 불심검문과 같은 경우도 구체적 위험과의 관련성은 분명하게 존재한다. 예를 들어 우리나라 「경찰관직무집행법」 제3조에서 불심검문의 요건으로서 ‘수상한 행동이나 그 밖의 주위 사정을 합리적으로 판단하여 볼 때 어떠한 죄를 범하려 하고 있다고 의심할 만한 상당한 이유가 있는 사람’이라는 문구에서 제시되는 상황은 구체적 위험 상황에 이르지 않을 수도 있지만, 위험 상황과의 관련성은 충분히 인정될 수 있다. 그리고, 독일의 경우도 노르트라인 베스트팔렌주 경찰법의 신원확인 규정(제12조 제1항)⁴⁹⁾을 보면 명확하게 ‘1. 위험의 방지를 위하여(zur Abwehr einer Gefahr)’ 신원확인을 할 수 있도록 하는 조건도 있지만, ‘2. 사실관계들을 고려했을 때 a) 그 장소에서 사람들이 중대한 의미를 가진 범죄를 약속, 준비 또는 행하려는 가정을 합리화시키는 장소에 어떤 사람이 머물고 있다면’이라는 조건도 있다. 즉, 후자의 조건이 충족되면 경찰이 정당하게 신원확인을 할 수 있다는 것이다. 후자의 조건을 살펴보면 구체적 위험의 개념요소인 ‘가까운 장래에’, ‘손해를 가져올 충분한 개연성’이 바로 인정된다고 보기 어렵기 때문에 구체적 위험에 해당된다고 볼 수는 없다. 그럼에도 불구하고 해당 문구를 보면 위험개념과의 관련성이 상당한 정도로 유지된다는 점은 누구나 쉽게 인정할 수 있다. 구체적 위험 개념이 되었건 아니면 그에는 못미치는 우리나라 경찰관직무집행법 불심검문의 언급한 요건 또는 독일 노르트라인 베스트팔렌주 경찰법 신원확인 중 제12조 제1항 제2호 a)의 요건이 되었건 경찰이 위험을 확인하기 위해서는 이를 판단할 수 있는 일정한 사실관계 내지 상황이 전제되고 주어져야 한다.

나아가 아래에서 설명하는 것처럼 위험상황에 대한 판단은 인간인 경찰관이 행하게 되고 따라서 어쩔 수 없이 일정한 한계에 노출되어 있다.

48) 서정범, 경찰행정법, 세창출판사 2020, 111쪽; Drews/Wacke/Vogel, Martens, Gefahrenabwehr, 9. Aufl. 1986, S. 202; Gusy, Polizei- und Ordnungsrecht, 7. Aufl, 2009, Rn. 108; BVerwGE 45, 51, 57.

49) § 12 NRW PolG. § 12 Identitätsfeststellung

(1) Die Polizei kann die Identität einer Person feststellen,

1. zur Abwehr einer Gefahr,

2. wenn sie sich an einem Ort aufhält, von dem Tatsachen die Annahme rechtfertigen, dass

a) dort Personen Straftaten von erheblicher Bedeutung verabreden, vorbereiten oder verüben,

b) sich dort Personen treffen, die gegen aufenthaltsrechtliche Strafvorschriften verstoßen,

c) sich dort gesuchte Straftäter verbergen,

2. 위험의 판단주체: 성실하고, 합리적이며 전문적 지식을 가진 경찰공무원

경찰의 임무개시 및 경찰권한의 기준으로 제시해놓은 위험은 주어진 일정한 사실관계, 상황 하에서 인간인 경찰공무원의 관점에서 판단되는 것이다. 물론, 이 때 위험판단을 하는 경찰공무원은 성실하고(*gewissenhafter*), 합리적인(*besonnener*), 그리고 전문적 지식을 가진(*sachverkundiger*)⁵⁰⁾ 경찰공무원⁵¹⁾을 말한다.

하지만, 성실성, 합리성, 전문적 지식을 가진 경찰공무원이라 할지라도 판단근거가 되는 사실관계, 상황이 충분하게 주어지지 않는다면 위험한 상황이라고 인정하는 어려운 경우가 많다. 특히, 구체적 위험개념의 구성요건인 ‘개연성’을 판단함에 있어 위험판단의 주체가 되는 평균적인 경찰관은 현재 자신이 가지고 있는 ‘지식상태(Wissenstand)’에 의존할 수 밖에 없을 뿐만이 아니라⁵²⁾, 한 인간으로서 인지적 한계에 노출되어 있다. 또한, 성실성, 합리성, 전문적 지식을 가진 경찰공무원의 개념에는 평균적인 경찰공무원의 정당한 내지 수궁할 수 있는 경험칙(*Erfahrungssätze*)도 포함되게 된다⁵³⁾. 요컨대, 경찰공무원은 지식, 경험칙 없이는 상황진행의 인과관계(*Kausalität des Geschehensablaufs*), 즉, 어떤 상황이 손해로 발전하게 될 개연성을 제대로 판단할 수 없다⁵⁴⁾.

많은 사람의 생명까지도 좌지우지 하였던 이태원 참사, 청주시 오송 궁평2 지하차도 사고⁵⁵⁾는 지식, 인지, 경험칙에 의존하는 인간으로서의 경찰에 의한 위험상황 판단의 특성을 잘 보여준다. 이태원 사고의 경우 사고 발생 3시간 40분 전인 18시 34

⁵⁰⁾ Schenke, Polizei- und Ordnungsrecht, 2009, S. 44.

⁵¹⁾ 실질적 경찰 개념에서는 경찰공무원 뿐만 아니라 일반공무원도 위험판단의 주체가 될 수 있을 것이다.

⁵²⁾ Pieroth/Schlink/Knise, Polizei- und Ordnungsrecht, 2010, S. 68; Poscher, Gefahrenabwehr – Eine dogmatische Rekonstruktion, 1999, S. 125 ff. 특히, Poscher 의 저서는 AI 관련 논의가 본격적으로 시작되기 전에 출판된 것임에도 불구하고 경찰상 위험진단의 근거가 ‘객관적인 지식지평(objektiver Wissenshorizont)’이라는 점을 선구적으로 지적하였다.

⁵³⁾ 예컨대, 이태원 참사와 관련하여 코로나가 종료된 이후라 해방감을 만끽하고자 청년들이 이태원 축제에 많이 몰릴수 있다는 것과 판단은 수궁할 수 있는 경험칙적 위험판단일 것이다.

⁵⁴⁾ Albers, Die Determination polizeilicher Tätigkeit, 2001, S. 34; Preuß OVG, PrVBl. 1894/1895, S. 125 f.

⁵⁵⁾ 2023년 7월 15일 충청북도 청주시 오송읍의 궁평2지하차도가 폭우로 인해 침수되어 차량 17대가 고립되어 14명이 사망한 사고이다. 사고 당일 오전 8시 30분경 청주시 오송읍 508번 지방도에 위치한 궁평2지하차도에서 550여m 떨어진 철골 가교 끝의 제방 둑이 터졌고 인근 미호강이 범람하면서 충청북도 추산 6만 톤에 달하는 엄청난 양의 물이 단 2~3분 만에 지하차도로 들어왔다. 이로 인해 오전 8시 40분경 해당 지하차도가 완전히 침수됐다. 사고 당일 아침 7시 2분 최초 112 신고, 7시 58분 2차 신고가 있었지만, 경찰 순찰차는 사고현장에는 완전 침수(아침 8시 40분) 후인 9시 2분에 현장에 도착하였다. (오송 참사 1주기 ① ‘오송 참사’ 공소장 입수...경찰은 그날 어디서 무엇을 하고 있었나, 뉴스파타 2024.7.11. 자 기사)

분에 압사 언급을 하며 최초 112 신고가 접수되었으며, 사고 발생 직전까지 112 신고가 11건이 들어왔다고 한다. 결론적으로 경찰은 - 물론, 실질적 위험방지는 경찰뿐만 아니라 구청 등 행정기관의 임무에도 포함된다 - 이태원 사건에서 현장의 위험상황을 제대로 평가, 판단하지 못하여 적절한 경찰상 대응을 제때 하지 못하였고, 159명의 생명의 침해라는 참사로까지 이어졌다. 청주 오송 궁평 2지하차도 사건의 경우 112 신고와 관련하여 경찰은 당일 아침 7시 2분 최초 접수를 비긴급상황인 코드 3으로, 당일 아침 7시 58분 2차 신고를 역시 비긴급상황인 코드2로 분류하여 대응하였다. 비긴급 코드로 지정되었음에도 112 지령실로부터 상황을 전달받고 침수가 진행되는 상황을 인지하고 있던 순찰차는 나름대로 분주하게 움직였지만, 정작 112 신고를 받고 순찰차가 출동한 곳은 침수사고 현장인 궁평2 지하차도가 아닌 사고현장으로부터 1.2 km 떨어진 궁평 지하차도였다고 한다⁵⁶⁾.

이런 사건들은 경찰의 과실여부를 판단하는 것을 떠나 경찰 위험상황 판단이 사실관계, 오감을 포함한 인지력, 경험칙, 지식 등에 얼마나 좌우되는지 극명하게 보여준다. 하지만, 경찰법상의 위험개념 자체가 이러한 내용을 모두 알고 있고 포섭하고 있다는 것은 다행스러운 일이다.

위험 개념은 구체적 위험 외에도 외관상 위험, 오상위험, 위험의 혐의와 같은 다양한 위험 개념들이 존재한다. 이와 같이 다양한 위험개념이 존재한다는 사정은 위의 사고들이 보여주는 것처럼 제한된 사실관계, 상황 속에서 인간으로서의 경찰관이 - 즉, 성실하고, 합리적이며, 전문적 지식을 가진 경찰관이라 할지라도 - 지식적, 인지적 한계에 노출되어 있는 상황을 방증하고 있다. 다양한 위험 개념의 종류에 따라 경찰의 개입여부, 그리고 경찰의 권한행사 방식도 달라지게 때문에 다양한 위험 개념은 매우 중요한 의미를 가진다. ‘구체적 위험’개념은 위에서 살펴보았으므로 외관상 위험, 오상위험, 위험의 의심에 대해서 간단히 일별해보도록 한다.

가. 외관상 위험

외관상 위험이란 성실하고, 합리적이며, 전문적 지식을 가진 전형적이고(*typische* n)⁵⁷⁾, 객관화된(*objektivierten*)⁵⁸⁾ 경찰관이 사전적 관점(*ex-ante*)⁵⁹⁾에서 사실관계를 합리적으로 평가해 보았을 때(*verständiger Würdigung des Sachverhalts*)⁶⁰⁾ 가까운 장래에 손해가 발생할 고도의 개연성이 있는 상태, 즉, 위험한 상태라고 판단하였지만, 사

56) ‘오송 지하차도 침수’ 직전 112 신고 2번…경찰 대응 미흡, 뉴스 1, 2023.7.18. 자 기사.

57) 해당 표현은 Pieroth/Schlink/Knise, Polizei- und Ordnungsrecht, 2010, S. 70.

58) 해당 표현은 Schenke, Polizei- und Ordnungsrecht, C.F.Müller, 2009, S. 44 f.

59) 모든 위험판단은 손해가 현실화되기 이전에 행하는 사전적 (*ex - ante*) 판단이다.

60) Poscher, Gefahrenabwehr - Eine dogmatische Rekonstruktion, 1999, S. 125 ff.

후적으로(ex-post) 경찰개입 시점인 위험판단 당시의 상황을 돌아보았을 때 실제로는 위험이 존재하지 않았던 경우를 말한다.

예를 들어 사실은 영화촬영 상황이었지만 경찰관이 누군가 골목길에서 칼을 빼어 들고 분노에 찬 표정으로 타입에게 접근하는 것을 보고 그 사람을 제압한 경우와 같은 상황이 외관상 위험에 해당된다⁶¹⁾. 사실감있게 현장을 촬영하기 위하여 영화감독도 카메라 등 영상장비도 모두 영화촬영 현장인 거리에서 감추어진 상태, 예컨대, 고층 빌딩의 한 방실에서 해당 장면을 촬영하고 있었다고 상상해볼 수 있다.

이 경우 성실하고, 합리적이며, 전문적 지식을 가진 전형적이고(typischen), 객관화된(objektivierten) 경찰관이 경찰개입 시점에서 해당 상황을 합리적으로 판단했을 때 위험의 존재를 긍정할 수 있었던 것이다. 즉, 이 경우 특정 경찰관뿐만이 아니라 위 요건에 해당되는 다른 경찰관들도 경찰권 발동 시점에 위험한 상황이라고 판단하였을 것이기 때문에 그와 같은 위험판단은 정당한 것이고, 그에 따른 경찰권 행사도 적법하다. 다만, 사후적으로 위험상황이 아니라는 것이 밝혀지는 경우에 손실보상 등의 문제가 고려될 수 있을 것이다.

위험은 사후적인 (ex-post) 판단이 아니라 손해가 발생하기 이전의 사전적 (ex-ante) 판단이므로 이와 같은 외관상 위험은 진정한 위험과 다를 바 없다거나, 심지어 진정한 위험(ecthe Gefahr)이라고 말해지기도 한다⁶²⁾.

나. 오상 위험

외관상 위험과 달리 오상위험에서는 경찰개입 시점에서 당해 경찰관이 어떤 상황을 주관적으로 위험하다고 판단한 경우이다. 예를 들어, 위 외관상 위험 사례에서 전날 술에 취하여 다소 판단력이 떨어진 특정한 경찰관이 평균적인 주의력을 발휘하지 않아 주변에 모자를 쓴 영화감독 등 스텝들과 영상장비, 방송사 차량 등이 있음에도 불구하고 이를 고려하지 않고 만연히 ‘누군가 골목길에서 칼을 빼어 들고 분노에 찬 표정으로 타입에게 접근하는’ 상황에만 매몰되어 위험한 상황이라고 판단을 한 경우이다. 이 경우는 해당 경찰관이 위험판단을 그르친 것으로 이와 같은 오상위험 상황에서의 경찰의 제압행위는 위법한 것이 되고, 손해배상의 문제가 고려될 수 있다.

외관상 위험과 달리 오상위험에서는 해당 경찰관이 사실관계를 합리적이며, 전문적으로, 그리고 객관화된 상태에서 행하지 못한 판단 결함이 있는 것이다. 그로 인하여 경찰관과 경찰관청은 오상위험에 따른 잘못된 경찰권 행사에 대하여 배상책임, 심한 경우에는 과실치상 등과 같은 형사상 책임을 질 수도 있겠지만, 언급한 사례를 살펴

61) 서정범, 경찰행정법, 세창출판사 2020, 116쪽.

62) Schenke, Polizei- und Ordnungsrecht, C.F.Müller, 2009, S. 44.

보면 제한적으로 주어진 사실관계 하에서 평균적인 경찰관이 경찰개입시점에 제대로 위험을 판단하는 것이 어려울 수 있다는 것 또한 부인될 수 없다.

이처럼 위험 개념은 주어진 사실관계를 토대로 한 인지적, 지식적 판단에 따라 결정된다. 그리고, 이와 같은 경찰의 위험판단 결과에 따른 경찰권 행사가 적법하게 될 수도 있고 위법하게 될 수 있을 정도로 중요한 의미를 가진다.

다. 위험의 의심

위험의 의심은 경찰관이 한편으로는 위험이 있다고 판단할 근거를 가지고 있지만, 또다른 한편으로는 그 근거만으로는 불충분해서 위험이 존재하지 않을 수 있다는 것을 경찰관 스스로 자각하고 있는 경우이다. 위험의 의심의 경우 경찰이 경찰권을 발동하려고 하는 시점에 경찰 스스로 위험이 존재하지 않을 수도 있다는 점을 자각하고 있다는 점에서 앞서 설명한 외관상 위험, 오상위험과는 구분된다⁶³⁾.

예컨대, 위험의 의심사례로 만우절에 누군가가 경찰관서에 전화하여 영업시간 중에 백화점을 폭파하겠다고 협박전화를 하는 경우를 생각해 볼 수 있다.

또 하나의 사례를 든다면 - 해당 사례는 우리나라에서는 제도적 의미의 경찰의 임무가 아닐 수도 있다 - 경찰관이 어떤 토지소유자가 자기집 뒷마당의 토지에 병에 담긴 검은 액체를 뿌리는 것을 본 경우를 들 수 있다. 만약 이 경우 이 검은 액체가 오래된 석유라면 토양오염과 관련하여 위험이 존재한다. 하지만, 병에 담긴 검은 액체가 인분, 가축분뇨를 발효시킨 비료였다면 위험이 존재한다고 볼 수 없다. 이 상황에서 경찰관은 위험이 존재하는 것이 맞는 것인지, 위험이 존재하지 않는 것이 맞는 것인지 알 수 없다. 일반적으로 독일 경찰법(예컨대, 노르트라인 베스트팔렌주 경찰법 제41조)에서는 현재의 위험이 존재하는 경우에만 타인의 주거에 출입할 수 있다. 경찰관은 위험의 존재여부를 알 수 없지만, 그것을 확인하려고 한다⁶⁴⁾.

위험의 의심은 경찰관이 주어진 사실관계 위에서(auf der Grundlage des vorliegenden Sachverhalts) 주의깊게(sorgfältig), 합리적으로(besonnerner), 그리고, 경험과 지식을 동원하여 사실관계를 평가하였음에도 불구하고 (verständiger Würdigung des Sachverhalts) 위험진단에 있어 불확실성이 존재하는 상황이다. 즉, 해당 경찰관이 위험이 있는지 없는지 확신을 가지지 못하는 상황이다⁶⁵⁾. 위험의 의심 상황에서 경찰관은 위험이 있다고 생각할 만한 불확실한 사실관계를 인식하고 있지만, 동시에 사실관계를 합리적으로 평가해보면(verständiger Würdigung des Sachverhalts) 위험의 존

63) 서정범, 경찰행정법, 세창출판사 2020, 118쪽 이하.

64) Gusy, Polizeirecht, 6 Aufl. 2006, 92쪽 이하의 사례이다.

65) Pieroth/Schlink/Kniesel, Polizei- und Ordnungsrecht, 2010, S. 71.

재를 긍정할 만한 충분한 근거는 가지고 있지 못하다⁶⁶⁾.

이 경우에 경찰은 구체적 위험을 확인하지 못하였기 때문에 구체적 위험을 근거로 성립된 경찰법상의 경찰권한을 발동할 수 없다. 따라서, 일단은 구체적 위험이 있는지 없는지의 판단을 목적으로 사실관계를 좀 더 밝히기 위하여 위험여부 확인조치(Gefahrerforschungseingriff)를 행하게 된다⁶⁷⁾. 위험여부 확인조치는 구체적 위험을 근거를 이를 방지하기 위한 종국적인 조치를 말하는 것이 아니라 불확실한 내지 석명되지 않는 사실관계(ungewissen bzw. noch nicht geklärten Sachverhalts)를 밝히기 위한 잠정적인 조치를 말한다. 위험의 의심상황에서 구체적 위험이 있는지 여부를 확인하기 위한 조치로는 일반적으로 질문, 조사, 정보조회 등을 들 수 있을 것이다. 하지만, 만약 침해되는 법익이 크다고 인정될 수 있는 경우라면 경우에 따라서는 예외적으로 위험의심을 구체적 위험과 동일하게 취급하여 위험 상황에서도 같은 조치를 받을 수 있게 법적으로 규정하는 것도 가능하다⁶⁸⁾. 예컨대, 독일 감염병예방법 제16조 제1항은 위험과 위험의 의심의 경우를 함께 그 요건으로 하면서 관할 행정청이 감염병 방지를 위하여 필요한 조치를 발할 수 있음을 규정하고 있다⁶⁹⁾. 현실적 상황과 법적 필요에 따라 생각된 다양한 위험 개념이 감염병예방법 영역에서처럼 구분없이 포괄적으로 사용되는 가운데 행정관청이 권한을 행사하는 것은 어디까지나 예외적이라는 것을 주지하기 위하여 사족을 단다면 실질적 위험방지법인 감염병예방법에서의 이와 같은 상황은 감염바이러스의 전파가능성과 그로 인한 막심한 생명, 신체의 피해 발생 우려 때문이라는 점을 짚고 넘어가도록 한다.

위험의 의심 상황과 관련하여 “사실관계가 (위험이 존재한다는) 가정을 합리화한다면(wenn Tatsachen die Annahme rechtfertigen)” 또는 “(위험이 존재한다는) 사실적인 근거가 있다면(wenn tatsächlicher Anhaltspunkt vorliegt)”과 같은 표현이 사용되는데 이는 불확실 내지 석명되지 않은 사실관계를 밝히기 위한 위험여부 확인조치를 위한 권한을 설정될 때 사용되는 표현이다. 독일의 경찰법에서는 구체적 위험의 단계에 이르지 못하는 위험의 의심상황을 전제한 규정으로 앞에서 예시한 노르트라인 베스트팔렌주 경찰법상의 신원확인 규정(제12조 제1항 제2호 a)목)을 들 수 있을 것이다.

⁶⁶⁾ Schenke, Polizei- und Ordnungsrecht, C.F.Müller, 2009, S. 46.

⁶⁷⁾ 김성태, 위험에 대한 의심과 위험여부의 확인, 행정법연구 제51호, 2017, 157쪽 이하.

⁶⁸⁾ 同旨 김성태, 위의 논문, 169쪽.

⁶⁹⁾ 감염병환자가 아닌 감염병의사환자, 감염병의심자는 경찰법에서 의미하는 ‘구체적 위험’, 즉, 국민의 생명·건강에 손해를 발생시킬 충분한 개연성이 있는 것은 아니지만 독일 감염병예방법은 감염병의 특수성을 고려하여 이들도 경찰상 책임자의 범위 내로 포섭하고 있다고 한 논평으로는 박원규, 감염병 영역에서의 경찰권 행사 - 코로나19 대응에 있어 경찰작용의 법적근거와 한계, 법학논총 제37권 제4호, 2020, 97쪽.

한 견해는 “경찰법적 의미에서 위험이라는 것도 그 손해가 현실화되기 이전까지는 불확실성에 사로잡혀 있는 진단에 해당되는 것이므로 위험은 의협의 의심과 구조적으로 다르지 않고 위험의 의심은 완전한 구체적 위험에서 단지 질적인 측면에서 마이너스가 되어 있는 상태이다”라고 표현하기도 한다⁷⁰⁾. 결국 이는 구체적 위험은 사실관계를 통해 위험진단이 일반적으로 수궁할 수 있을 정도로 확실하게 이루어졌지만, 위험의 의심은 아직 일반적으로 수궁할만큼 위험에 대한 진단이 이루어지지 않은 상태라는 것을 말하는 것이다. 요컨대, 이 견해에 따르면 위험이든 위험의 의심이든 판단 및 논리구조는 같지만 의협의 의심은 통상적으로 위험이라고 확정하기 위한 사실적 근거의 측면, 논리적, 인지적 판단의 측면이 결여되어 있는 상태인 것이다.

위에서 제시된 위험의 의심 사례인 만우절 사례를 고려한다면 위험판단의 주체로서의 평균적인 경찰관은 당사자의 진술의 진위여부를 확정적으로 판단할 수 있는 지식, 인지능력이 부족할 수 밖에 없다. AI를 동원하여 전화발신 위치가 어디인지 확인하고 전화를 건 사람이 누구인지, 주소가 어디인지, 그리고, 발신자가 최근 어떠한 행동을 하였는지, 예컨대, 무기를 제조하기 위하여 인터넷을 검색하였는지 무기를 구매하기 위하여 총포상을 들렀는지 등등을 확인, 조사할 수 있다면 사실관계, 관련 지식, 인지적 요소, 앞의 것들을 종합한 인과관계의 측면이 보강되어 위험의 의심은 어느 정도 해소될 수 있을 것이다. 위험의 의심 상황에서 AI의 도움으로 위험의 존재를 긍정할 수 있는 개연성을 발견할 수 있게 된다면 구체적 위험을 인정할 수 있게 되는 것이고, 경찰은 이에 따라 범죄의 예방과 제지, 위험발생의 방지 등과 같은 경찰상 조치로 나아갈 수 있게 될 것이다. 이렇게 된다면 궁극적으로 위험방지의 효과성, 효율성이 훨씬 강화될 수 있는 것이다.

3. 소결

위에서 살펴본 것처럼 위험 개념, 나아가 외관상 위험, 오상위험, 위험의 의심은 관련된 사실관계 뿐만이 아니라 위험판단의 주체인 평균적인 합리적인 경찰관의 인지적, 지식적 한계에 기반하여 성립되는 개념이다. 그런데, 위와 같은 한계에 노출되어 있는 위험개념을 근거로 다양한 위험개념에 따라 각각 경찰개입의 여부 및 개입시 권한행사 방법이 틀려지게 된다는 것은 정확한 위험판단이 얼마나 중요한지를 역설한다고 할 것이다.

위와 같은 다양한 위험을 판단하는 과정에서 AI는 인간의 오감(시각, 청각, 후각, 미각, 촉각) - AI의 오감은 위험판단의 첫 번째 관문인 사실관계 인식에 엄청난 도움

⁷⁰⁾ Gärditz, Strafprozess und Prävention, 2003, S. 35.

을 주게 된다 -, 나아가 인간으로서 한 개인의 사고능력, 지식의 한계를 간단히 뛰어넘으리라는 것은 충분히 인정될 수 있다. 물론, 오늘날 AI가 아직도 상당한 오류 가능성을 가지고 있고, 그리고, 아직도 여러 영역에서 인간의 능력에 못미치는 경우가 많은 것도 사실이다⁷¹⁾.

하지만, 이러한 오류를 줄여나간다면 AI는 인간의 인지적 한계를 보완해서 현장의 경찰관으로 하여금 더 정확한 위험판단을 하도록 도울 가능성은 열려 있다. 현장의 경찰관이 경찰임무의 출발점이자 경찰권 발동의 기준이 되는 위험과 관련된 판단을 함에 있어 이러한 AI의 도움을 받을 수 있다면 ‘외관상 위험’이라고 판단하게 될 범위는 훨씬 줄어들어 경찰력을 불필요하게 투입할 일이 줄어들 것이다. 그렇게 된다면 경찰은 다른 일에 경찰력을 활용할 수 있게 된다.

그리고, AI의 도움을 받을 수 있다면 경찰관의 주관적인 판단에 의존하여 오상위험이 존재한다고 판단할 일이 줄어들게 되어 이 경우 또한 불필요한 경찰력 투입이 줄어들 뿐만 아니라 애매한 상황에서 엉뚱한 사람을 상대로 경찰력을 투입하지 않게 되어 경찰의 공신력이 높아질 것이다. 위험의 의심과 관련해서도 인간인 경찰관 입장에서 부족한 판단근거를 AI가 그의 인지능력, 사고력, 지식으로 보완하여 그것이 구체적 위험이 있는 상황인지 아닌 상황인지를 인간인 경찰보다 더 정확하게 판별하여 경찰권이 효율적이고 적절하게 행사될 수 있도록 도움을 줄 수 있을 것이다.

AI의 긍정적인 측면을 종합해보면 AI는 인간으로서 평균적이고 합리적인 경찰관의 경험적, 인지적, 지식적 한계를 보완하여 경찰임무 개시 및 경찰권 발동의 근거가 되는 위험상황을 더 정확하게 판단하여 경찰력이 더 효율적으로 이용될 수 있도록 도와줄 수 있고, 또, 인간으로서의 인지, 지식, 사고의 한계에 노출된 경찰관의 위험판단으로 엉뚱한 사람을 상대로 경찰권을 행사하는 일이 줄어들게 되어 경찰의 신뢰도를 제고하게 할 수 있다.

그런데, 여기서 우리가 잊지 말아야 할 것은 경찰법 이론에 따르면 위험판단의 주체는 여전히 성실하고, 합리적이며, 전문적인 지식을 가진 경찰관이라는 것이다. AI가 방대한 지식을 가지고 있고 인지능력이 뛰어나다고 해서 AI의 분석에 따라 경찰상 위험을 전적으로 판단하기에는 아직도 AI는 여러 영역에서 한계가 있고 오류의 가능성이 존재하기 때문에 문제가 있다. 어느 정도의 오류는 감수할 수 있다고 항변할 수 있지만, 개인의 기본권 침해와 관련된 경찰권한 행사로 이어진다면 이는 치명적인 오류이고 이는 불가역적인 경우가 많다. AI의 경찰활동에의 투입목적은 인간의 지식적,

⁷¹⁾ AI에게 한국 수능을 풀게 했다, 결과는? 시사IN 2024.7.17. 자 기사. 해당 기사에서는 AI를 상대로 한국 수학능력시험 풀이 능력을 검증하는 실험을 해본 결과, AI는 언어 영역에서 평균보다 월등한 성적을 보였지만, 수리와 과학 추론에서는 수험생 평균보다 훨씬 남을 점수를 받았다고 한다.

인지적 한계, 나아가, 오상위험에서 보여지는 것과 같은 인간의 오류를 극복하여 보다 더 효율적일뿐만 아니라 보다 더 적절한, 다시 말해 기본권 친화적인 경찰권 행사를 하도록 하는데 두어져야 한다. 게다가 AI의 전적인 위험판단에 의존하여 경찰이 기계적으로 경찰권 행사 여부 및 권한행사 방식, 정도를 결정하게 된다면 AI가 경찰권 행사의 명령주체가 되고 인간으로서의 경찰은 이를 실행하는 보조적 존재에 머물게 될 것인바, 이것이 우리 인간이 바라는 AI의 활용모습은 아닐뿐더러 현재 경찰법 체계도 그렇게 구성되어 있지 않다.

AI가 경찰의 위험판단에 관여하더라도 경찰임무 개시, 경찰권한 발동을 위한 종국적인 위험판단은 성실하고, 합리적이며, - AI와는 다른 차원의 - 전문적 지식을 가진, 그리고, AI 보다는 인지능력이 떨어지는 인간으로서의 경찰관이 자신의 책임하에 수행하는 것이 타당하다. 만약 인간이 위와 같은 과정 속에서 경찰관이 잘못된 판단을 하여 경찰권한을 행사 또는 불행사 하였다면 그에 따른 책임은 경찰관의 몫이 되어야 할 것이다. AI는 어디까지나 기술적 보조수단이다.

경찰의 결정재량은 위험상황이 존재하는 경우라 할지라도 위험상황에서의 경찰력 투입이 오히려 상황을 고조, 긴장시켜 위험방지 목적달성, 효율성에 저해된다면 경찰력을 투입하지 않을 수 있도록 하고 있다. 물론, 관련된 상황 고조, 긴장 여부에 대한 판단마저도 AI가 할 수 있겠으나 결정재량 또한 인간경찰에게 주어진 것이지 AI에 주어질 수는 없다. AI는 인간경찰이 관련된 상황에 대해서 결정재량을 행사할 때 상황 고조, 긴장여부에 따른 결정재량을 행사할 것을 인간경찰에게 제안할 수는 있을 것이다. 이는 어디까지나 보조적, 참고적 제안이고 위험판단 및 그에 대한 설명의무는 경찰의 몫이 될 수 밖에 없다.

V. 경찰 영역 고위험 인공지능에 대한 규제

먼저 이 논의에 있어서는 앞에서 살펴본 독일 연방헌법재판소가 제시한 기준이 일정한 방향으로 제시될 수 있을 것이다. 즉, 자동화된 데이터 분석은 침해정도(데이터의 유형, 범위, 분석방법에 따라 결정)가 높을수록 헌법적 정당화를 위한 요건(보호법익의 중대성, 침해기준에 따라 결정)이 엄격하여야 하며, 침해정도가 낮을수록 헌법적 정당화를 위한 요건이 완화될 수 있다는 점을 감안하여 고위험 인공지능일수록 그 사용의 헌법적 정당화를 위하여 더 높은 기준이 설정되어야 할 것이다.

과학기술정보통신부와 한국지능정보사회진흥원(NIA)가 2026년 1월 공동발간한 「고영향 인공지능 판단가이드라인」에서는 범죄수사 및 체포에 관련해서만 고영향 인공지능의 적용범위 및 사례를 규정하고 있다. 범죄수사 및 체포는 구체적 범죄혐의

라는 전제가 충족되는 경우가 대부분일 것이므로 비교적 낮은 기준에서도 AI의 투입이 예상될 수 있을 것으로 생각될 수도 있지만, 다른 관점에서는 생명·신체에 대한 손해는 이미 진행된 상태라 적법절차, 무죄추정의 원칙에 따른 실체적 진실 발견이 중요한 의미를 가질 수 있다고 생각된다.

이런 관점에서 보면 「고영향 인공지능 판단가이드라인」 67쪽 i)에서 제시하는 범죄혐의는 - 어느 정도 구체화되어 있지만 - 특정 사건의 범인이 누구인지를 모르는 경우에 있어서 범인 검거를 목적으로 한 CCTV, 인터넷 사진·동영상 화면·화상캠을 활용한 무작위적 안면 이미지 정보의 수집 및 분석, 평가는 생체정보라는 점 외에도 무작위적이라는 점에서 높은 기본권 침해가능성을 가진다고 보아야 할 것이다. 이 경우 최소한 범죄의 중대성, 사건암장 방지의 필요성, 다른 방법으로는 범인의 신원을 특정하기 어려워 AI의 투입이 필요한 경우 등의 요건이 법률 규정에 제시되어야 할 것이다.

법률유보의 원칙을 개인정보보호법 제15조 제1항 제3호의 미약한 근거로 느슨하게 해석하는 우리나라의 판례 환경을 감안하더라도 경찰관직무집행법은 아니더라도 최소한 「경찰관의 정보수집 및 처리 등에 관한 규정」(대통령령)에 관련 요건이 제시되어야 할 것이다. 그런 요건이 없는 현재의 상황에서는 해당 요건에서의 안면이미지 정보 수집, 인식을 위한 AI의 활용은 위법하다고 생각된다.

책자 67쪽 ii)의 특정인의 범인여부 판단을 목적으로 한 인종, 신체적 특징 등을 추론하거나 추론하기 위하여 사용되는 생체정보 분류 인공지능 시스템 또한 인종, 생체정보는 매우 민감한 개인정보에 해당되는데 반하여 그 목적이 되는 특정인의 범인여부 판단과 관련하여 범죄의 중대성, 사건암장 방지의 필요성, 다른 방법으로는 특정인의 범인여부 판단 확인이 어려운 경우 등의 요건이 최소한 대통령령에 규정되어 있지 않는한 허용될 수 없다.

책자 67쪽 iii)의 경우는 민감정보인지 아닌지에 대한 지표가 전혀 없이 불특정 다수에 대한 감시라고만 표현하고 있어서 기본권 침해의 심각성이 전혀 제거되지 않고 있는바, 이동경로(Bewegungsmuster), 행동성향(Verhaltensmuster) 등 민간정보, 개인정보 관련성이 제한된 경우에 한하여 일정한 법정형 이상의 범죄수사를 위하여 필요하고 그렇게 하지 않을 경우 증거가 암장될 우려가 있는 경우와 같은 요건이 충족되는 경우에 한하여 AI 투입이 허용될 수 있을 것이다.

다만, 해당 요건은 고영향 인공지능영향평가의 대상으로 삼고 있다는 점에서 iii)에서 제시하는 무작위적 신원확인도 생체정보를 포함한다고 생각되는바, 이런 점을 요

건 설정에서부터 분명히 할 필요가 있다. 생체정보가 포함된다면 관련 요건은 위 i), ii)에 준하여 제시되어야 할 것이다.

특히, iii)의 경우 괄호에서 ‘단, 납치, 인신매매, 성적 착취의 특정 피해자에 대한 수색 및 실종자 수색인 경우는 제외’라고 하여 해당 범죄의 경우 범인 검거를 위하여 불특정 다수를 감시함으로써 피해자의 수색을 통해 피해자 구조(생명·신체의 보호)하는 경우 고영향 인공지능 평가를 받을 필요가 없는 것으로 상정하고 있는 경우인데 해당 죄명 및 피해자 수색이라는 설명을 통해 피해자 구조의 의도가 추론될 수 있다고 하더라도 경찰관직무집행법 제7조 제1항 긴급출입에서 제시되고 있는 “사람의 생명·신체 또는 재산에 대한 위해가 임박한 때에 그 위해를 방지하거나 피해자를 구조하기 위하여 부득이하다고 인정하면 합리적으로 판단하여 필요한 한도에서”, 그리고, “AI 투입없이 다른 방법으로 피해자 구조라는 목적을 달성하기 매우 어려운 경우”라는 요건을 최소한 대통령령에 규정할 필요가 있다고 생각된다.

위 i), ii), iii)의 경우 위에서 제시한 요건 외에도 본문에서 제시한 편향(편견), 오류, 블랙박스화 현상에 대한 통제는 고영향이든 아니든 모든 경우에 있어서 충족되어야 할 것이다.

책자 67쪽 iv)의 경우 i), ii), iii)이 제시하고 있는 생체정보와 같은 민간정보가 활용되는 고영향 인공지능인지 여부가 기준이 되는 것이 아니라 “특정 사건의 범인 검거과정에서 최종 결정을 내리기 전에 인공지능시스템의 추천 결과를 인간이 검토하거나 평가할 수 없는 인공지능시스템을 사용하는 경우”라는 조건이 제시되어 있는 바, 해당 항목과 병렬적으로 제시될 것은 아니라고 생각된다.

또한, 해당 가이드라인은 범인 수사 및 체포에 대해서만 고영향 인공지능평가를 예정하고 있고 위험방지, 범죄행위의 예방적 대처(테러, 연쇄살인 방지 등)의 경우에는 생명, 신체를 보호하는 일이므로 AI 투입을 위한 요건이 제시되어 있지 않는데 마찬가지로 경찰관직무집행법 제7조 제1항 긴급출입이 제시하는 요건과 “AI 투입없이 다른 방법으로 피해자 구조라는 목적을 달성하기 매우 어려운 경우”라는 요건이 최소한 대통령령에는 제시되어야 할 것이다.

VI. 나가며

AI는 인간으로서 평균적 경찰관의 인지적, 지식적 한계를 극복하여 경찰관보다 더 정확하게 위험판단을 할 수 있는바, 이를 활용하게 되면 위험방지의 효율성에 기여할 수 있다. 범죄수사의 영역도 AI 활용의 요건은 다르게 설정될 수 있겠지만 언급한 AI의 장점은 마찬가지로 적용된다.

하지만, 아직까지 특정 영역에서는 AI의 인지, 사고능력이 인간의 그것에 뒤지는 경우도 많다는 것이 확인되고 있으며, 무엇보다 학습과정에서 인간의 역사적, 지역적, 생물학적 편견에 노출되어 있다는 약점이 있다. 예컨대 인공지능 학습에 사용되는 데이터가 그 활용맥락에 적합한 다양성과 대표성을 가지지 못하거나, 역사적 편향을 가지는 경우가 발생하고 있고, 측정과정에서의 편향(흑인 등 소수집단의 전과자들이 더 높은 체포율을 가지고 있기 때문에 백인 다수집단보다 재범율이 높은 것으로 측정되는 편향)도 나타나는데 그에 기반하여 AI가 도출하는 결론을 대표성을 가지는 합리적인 결론으로 보기 어려울 것이다⁷²⁾.

이에 더하여 AI가 자동적 분석, 평가 내지 학습을 통하여 최종적인 결정에 이르렀을 때 인간의 인지, 지식으로는 왜 AI가 그와 같은 결론에 이르렀는지 그 과정을 추론해내기 어렵고, 많은 경우 AI의 분석, 평가 결과물이 산출되는 과정이 투명하지 못하다는 점이 지적되고 있다(소위 ‘Black Box’로서의 AI)⁷³⁾.

인공지능 알고리즘이 불투명한 상황에서 이에 기반해 내려진 결정에 대하여 이의를 제기할 수 없다면 인간은 인공지능 시스템에 장착된 불투명한 알고리즘에 따라 영문도 모른채 평가받고 감시받는 상황에 처할 수 있게 되는데, 이는 인간의 인격권, 일반적 행동자유권 등 기본권, 자율성 보장이 심각하게 침해되는 상황이라 하지 않을 수 없다.

법적으로 합리적으로 이유제시되지 못하는 AI의 판단(자료의 자동화된 분석, 평가에 따른 결론) 및 그에 따른 경찰의 각종 정보상 조치, 나아가 물리적인 조치는 법의 세계에서는 결코 정당화될 수 없다.

현재 AI의 언급한 한계를 감안했을 때 단순히 바둑 프로그램 등과 같이 사람의 기본권에 직접적인 영향을 미치지 않는 영역에 AI가 적용되는 것은 큰 문제가 없을 수 있으나 사람의 기본권과 밀접한 영향을 미치는 사회보장자격 심사, 유·무죄 판단, 보안처분 등에의 활용은 그 자체로 이미 심각한 기본권 침해 우려 문제를 내포하고 있다.

AI를 활용하지 않은 단계의 경찰의 위험방지 및 범죄행위를 위한 예방적 대처를

72) 정보인권연구소, 인공지능과 정보인권 - 해외 논쟁사례(2): 차별, 이슈리포트 <정보인권> 2023-12(통권 제16호), 2023, 45쪽 이하.

73) 정보인권연구소, 인공지능과 정보인권 - 해외 논쟁사례(1): 불투명성, 이슈리포트 <정보인권> 2023-12(통권 제16호), 2023, 23쪽 이하.

각종 정보상 조치에 대해서도 독일의 경찰법은 세심한 기준을 제시하고 있다. 이에 더하여 이와 같은 정보상 조치를 위하여 AI 수단을 도입하는 경우 편견 및 오류가능성, AI의 판단은 어디까지나 경찰법상 위협의 의심에 불과하다는 점, 합리적인 이유 제시가 가능한 경우에 한해 AI의 판단이 ‘위험의 의심’으로나마 인정될 수 있고, 경찰의 확인과정을 거쳐 다음 단계의 경찰조치로 나아갈 수 있다는 점이 유의되어야 한다. AI의 활용으로 기본권 침해가 심각하게 우려되는 고영향 인공지능의 활용의 영역에서는 더욱 세심하게 이러한 AI의 한계가 고려되고 관련된 조건이 법적으로 – 최소한 「경찰관의 정보수집 및 처리 등에 관한 규정」(대통령령)에라도 설정되어야 한다.

무기명으로 된 자료를 바탕으로 AI를 활용한 범죄현상의 진행에 관한 추상적인 상황도(Lagebilder) 제시, 또는 장소 및 시간대와 관련된 범죄예측 활동 정도까지는 정보의 자기결정권을 비롯한 기본권 보호의 관점에서 큰 문제가 없을 수 있다고 생각된다⁷⁴⁾. 다만, 향후 기술이 발전할수록 무기명성을 제거하고 신원을 쉽게 재확인하게 되는 상황이 도래될 수도 있는바, AI 알고리즘 설계시 이러한 현상이 방지될 수 있도록 통제되어야 할 것이다.

또한, 불특정한 목적을 위하여 예비로 자료를 수집하는 것은 정보보호법의 원칙인 목적구속의 원칙에 부합하지 않는다는 점도 유의하여야 한다. 정보수집 당시의 목적은 정보처리시에도 준수되어야 하는바 AI의 활용에 있어서도 이 점이 고려되어야 한다⁷⁵⁾. AI를 활용한 최종적인 결과를 경찰의 정보처리 활동의 근거로 제시한다고 해서 그것만으로 정보보호의 원칙에 부합되는 것은 아니다⁷⁶⁾. 빅데이터라는 것이 속성상 어떠한 자료간의 연결도 가능하지만 최초 수집된 목적을 고려하여 목적 구속의 원칙이 존중되는 가운데 자료간 조합이 이루어져 최종적인 학습, 분석, 평가 결과물이 도출될 수 있도록 하여야 할 것이다⁷⁷⁾.

74) 이와 관련하여 개인정보보호위원회에서 인공지능 모델 학습과정에서 주민등록번호와 같은 중요 개인 식별정보를 사전에 제거하도록 권고하면서, 한국인터넷진흥원(KISA)에서 제공하는 개인정보 노출페이지(URL) 삭제·차단 정보를 반영하도록 한바 있다(개인정보보호위원회, 주요 인공지능(AI) 서비스 사전 실태점검 결과 발표, 2024.3.28.).

75) Hornung/Herfurth, Datenschutz bei Big Data – Rechtliche und politische Implikationen. In: König, C., Schröder, J., Wiegand, E. (Hrsg.). Big Data. Chancen, Risiken, Entwicklungstendenzen. Wiesbaden: Springer VS, 2018, S. 169.

76) Roßnagel/Geminn/Jandt/Richter, Datenschutzrecht 2016. „Smart“ genug für die Zukunft? Ubiquitous Computing und Big Data als Herausforderungen des Datenschutzes. Kassel: ITeG Wissenschaftliches Zentrum für Informationstechnik-Gestaltung, 2016, S. 122 f.

77) Kring, Big Data und der Grundsatz der Zweckbindung. In: Plödereder/ E., Grunske/L., Schneider/E., Ull/D. (Hrsg.). Informatik 2014. Bonn: Gesellschaft für Informatik e.V., 2014, S. 557; Roßnagel, Big Data – Small Privacy? Konzeptionelle Herausforderungen für das Datenschutzrecht. Zeitschrift für Datenschutz – ZD, 2013, S. 565; Schaar, Zwischen Big Data

수집된 정보의 목적변경 이용은 원칙적으로 처음부터 완전히 배제되는 것은 아니지만 목적 변경은 개별 전문법에서 특별한 수권규정으로 통해 규정되는 경우에만 정당화될 수 있다는 점이 망각되어서는 안된다⁷⁸⁾. AI에 의한 자료의 자동적 분석, 평가시에도 당연히 유의해야 할 항목이다.

경찰법상 위험판단의 주체는 인간으로서의 경찰공무원이다. 즉, 인지적, 지식적 한계에 노출되어 있는 성실성, 합리성, 전문적 지식을 가진 경찰공무원이 경찰법상 위험판단의 주체이고, 그러한 판단은 통계적·산술적 판단이 아니라 규범적 판단이다⁷⁹⁾. 이러한 경찰법상 위험판단시 창조적이고 효율적인 수단으로서의 AI의 도움을 받는 경우에도 AI가 산출하는 결과물의 바탕이 되는 데이터의 각종 편견 가능성 및 그에 따른 차별 및 오류가능성, 판단결과에 이르는 불투명성 등과 같은 한계를 고려하는 가운데 인간 스스로 주체적인 책임 하에 경찰상 위험판단을 행하는 것이 타당하다.

국가배상법의 관점에서 보면 국가의 내부기관인 공무원의 잘못된 행위에 대해서는 국가가 국가배상책임을 지도록 되어 있다. 이 때 AI의 경우에도 잘못 설계된 AI 알고리즘에 대해서 만약 국가가 그 알고리즘을 설계했다면 국가배상책임으로, 그리고, 만약 이를 민간이 설계했다면 제조물 책임의 입장에서 민간이 책임을 지도록 할 수 있을 것이다. 그러나, 국가의 존립과 관련된 안전관련 임무는 최소한 국가가 최종적인 책임을 지는 것이 타당하다. 민간 AI 알고리즘 개발회사에 제조물 책임을 물리는 경우에도 이를 활용하여 경찰활동 등 국가기능에 활용하였다면 국가의 책임은 면책될 수 없다.

AI의 법인격을 인정하여 잘못된 AI의 활용에 대하여 AI에게 법적 책임을 물리려는 시도가 알고리즘 제작자 또는 제조사의 책임을 회피하려는 수단으로 활용될 수 있다는 점도 경계되어야 한다. 기업이 - 법적 픽션으로서 - 법적인 사람, 즉, 법인이 될 수 있었던 가장 큰 동인이 자연인인 자본가의 책임회피라는 것을 잊어서는 안된다⁸⁰⁾. 이런 점에서 유럽평의회(Council of Europe) 인권위원장(Commissioner for Human Rights)이 “인공지능의 인권 침해에 대한 책임성과 책문성은 항상 자연인 또는 법인에게 있어야 한다”고 지적한 것은 경청할 필요가 있다.

und Big Brother – zehn Jahre als Bundesbeauftragter für den Datenschutz und die Informationsfreiheit. Recht der Datenverarbeitung – RDV2013, S. 225

78) Johannes/Weinhold, Das neue Datenschutzrecht bei Polizei und Justiz. Europäisches Datenschutzrecht und deutsche Datenschutzgesetze. Baden-Baden: Nomos, 2018, § 1, Rn. 150 ff.

79) 박원규, 인공지능, 빅데이터, 그리고 경찰: 예측적 경찰활동에 대한 공법적 검토, 공법학연구 제 24권 제1호, 2023, 337쪽 이하.

80) 톰하트만 지음/이시은 옮김, 기업은 어떻게 인간이 되었는가 - 인간의 권리를 탐하는 거대기업의 음모, 어마마마 2014.

다만, 개인정보보호법에서는 개인정보를 “해당 정보만으로는 특정 개인을 알아볼 수 없더라도 다른 정보와 쉽게 결합하여 알아볼 수 있는 정보”(법 제2조 제1호 나.목)로 규정하고 있는바, 해당 조문에서 “이 경우 쉽게 결합할 수 있는지 여부는 다른 정보의 입수 가능성 등 개인을 알아보는 데 소요되는 시간, 비용, 기술 등을 합리적으로 고려하여야 한다”라는 제한이 설정되어 있기는 하지만 인공지능환경에서는 이 기준에 따라 개인정보의 해당성이 광범위하게 충족될 수 있다는 점에 대한 비판적인 관점이 제시되고 있다는 점도 감안할 필요는 있어 보인다. AI가 본문에서 언급한 한계점을 가지고 있다 하더라도 다른 한편에서는 인간의 인지적, 지식적 한계를 극복하여 국가 임무의 효율적이고 적절한 수행에 기여할 수도 있다는 장점을 고려하자는 것이다. 그런데, 다른 입장에서는 AI 등을 활용한 기술고도화로 재식별 가능성이 높아지고 있다는 점에서 개인정보보호법에서 AI의 학습 과정에 대하여 사용되는 자료, 정보에 대하여 법적 보호를 강화하여야 한다는 입장도 있다⁸¹⁾. 위에서 제시된 경찰상의 활동에서 다양한 수준과 범위의 개인정보 침해성이 있을 수 있고 그에 따라 AI 활용에 대하여 적절한 침해기준이 설정될 수 있다는 점을 고려한다면 AI의 장점 활용과 개인정보보호 사이에서 적절한 균형점을 찾아나가야 될 것으로 생각된다.

81) 이상의 논의에 대해서는 박미사, 인공지능 학습데이터 활용을 위한 입법 과제, 공법연구 제53집 제4호, 2025, 111쪽 이하.

붙임 3

경찰 미래비전 2050 추진과제 목록

[뉴노멀 치안 분야 : 8개 도전과제, 17개 시행과제, 47개 실행과제]

도전과제	시행과제	실행과제	
		번호	과제명
1.경찰관의 신체적·지적 능력 보완·확장	①스마트 디바이스 도입 등 현장 대응력 강화	1-①-①	웨어러블 장비를 통한 현장 경찰관 지원
		1-①-②	친환경·스마트리류 등 경찰복제·장비의 첨단화 추진
		1-①-③	지능형 로봇을 활용한 효과적인 치안서비스 제공
		1-①-④	드론을 활용한 효율적인 치안현장 지원
		1-①-⑤	스마트 순찰차·디바이스를 이용한 첨단 치안활동 지원
		1-①-⑥	저위험 대체무기 등 도입을 통한 현장 대응력 강화
	②인공지능 활용 활성화	1-②-①	미래치안의 구심점 역할을 할 「치안 AI 기술」 개발
		1-②-②	현장력 강화를 위한 치안보좌시스템 마련
		1-②-③	AI 기반 온라인 이슈 분석 시스템 개발
2.데이터 기반 경찰활동 실현	③데이터 생태계 구축	2-③-①	미래치안의 핵심 자원인 치안분야 데이터 주도권 확보
		2-③-②	미래치안 선도를 위한 지능형 통합 데이터댐 구축
		2-③-③	중장기 데이터전략 수립을 통한 데이터 행정기반 구축
		2-③-④	데이터를 기반으로 한 총포·화약 디지털 행정 구현
		2-③-⑤	디지털 약자 편의성 제고를 위한 민원서비스 개선 추진
	④네트워크 선진화	2-④-①	초연결시대 원활한 업무지원을 위한 차세대 통신망 구축
		2-④-②	엣지 컴퓨팅 도입을 통한 현장 대응력 제고
		2-④-③	양자암호통신 방식을 통한 경찰시스템 보안성 강화
		2-④-④	자율화·지능화 기반 내부 사이버 보안관제 인프라 강화
3.안전하고 편리한 스마트 이동	⑤교통행체계 혁신	3-⑤-①	PAV 상용화에 선제적 대응을 위한 경찰 UAM 개발·도입
	⑥자율주행 전향적 도입	3-⑥-①	스마트 치안활동 구현을 위한 자율주행의 전향적 도입
		3-⑥-②	자율운행장비를 통한 치안현장 지원
4.국민의 평온한 삶을 위협하는 요인 대처	⑦가상공간에서 안정된 치안 확보	4-⑦-①	메타버스 공간 내 범죄 및 안전 문제 선제적 대응
		4-⑦-②	가상 자산과 모바일금융 범죄 대응
	⑧생활 주변 위협요인 입체적 감지 및 범죄 예방	4-⑧-①	범죄예방 환경 조성을 통한 미래 대응
		4-⑧-②	사이버테러(해킹·랜섬웨어) 등 新 안보위협 대응력 강화

도 전 과 제	시 행 과 제	실 행 과 제	
		번 호	과 제 명
5. 거버넌스적 안전활동 전개	9 연결·협업에 기반한 C&D	5-9-1	범정부 사회적 약자 종합 지원센터 신설
		5-9-2	치안 공공재의 저변 확대를 위한 청원경찰제 활성화
		5-9-3	공인탐정업법 제정을 통한 안전한 사실조사 서비스 제공
	10 지역 치안 균형 혁신	5-10-1	인구구조 변화에 대비한 지역 안전 협업체계 고도화 추진
		5-10-2	지역 치안 균형을 위한 네트워크 구축 및 재정 확충
6. 과학기술 기반 수사역량 강화로 공정사회 견인	11 미래기술 확보 및 숙련된 과학수사 인재 양성	6-11-1	미래 환경 변화에 대비한 첨단수사기법·장비 개발
		6-11-2	과학수사 전문인력 확보·양성
	12 과학기술을 활용, 수사 전문성과 효율성 제고	6-12-1	증거기반 수사 확립을 위한 과학수사 플랫폼 구축 및 고도화
		6-12-2	디지털 증거의 신뢰성 제고를 위한 디지털포렌식 역량 강화
		6-12-3	경찰 수사역량의 상향 평준화를 위한 지능형 사건처리 지원 시스템 도입
		6-12-4	과학기술 중심 수사로 전환하기 위한 '범죄분석 프로그램' 활용
	13 첨단 과학 수사로 도약하기 위한 토대 마련	6-13-1	인공지능·빅데이터 분석 활성화를 위한 「형사절차 전자화법」 개정
		6-13-2	과학수사 인프라 융합과 협업을 위한 거버넌스 추진
7. 소통 과 갈 등 조 정 으 로 신뢰사회 구현	14 가상공간 소통 및 인포데믹·정보편식 해소	7-14-1	메타버스 플랫폼을 활용한 선제적 경찰 홍보
		7-14-2	정보의 왜곡·편향 등 가짜뉴스 대응 체계 구축
		7-14-3	알고리즘 중립성 관련 사회적 속의 및 법·제도 정비
	15 핀포인트 경찰 활동 으로 사회 갈등 요인 조기 해소	7-15-1	소통·협업활동 강화로 집회시위 등 공공갈등 대응 역량 제고
		7-15-2	집단 폭력·불법 행위 현장의 완충장비 도입
		7-15-3	포용 사회 구현을 위한 선제적 외사정보 활동 전개
8. 인권 중심 의 사회적 가치 수호	16 디지털 윤리 확립	8-16-1	지능정보화사회 인격권 보장을 위한 디지털윤리현장 마련
		8-16-2	개인정보 보호를 위한 전문인력의 역량 강화
	17 인권보장 관리 역량 강화	8-17-1	인권 행정체계 구축으로 인권보장 관리 역량 강화

한국 이민행정의 AI의 전면적용에 관한 우려와 과제

이일(공익법센터 어필 변호사)

I. 배경

- 인공지능(Artificial Intelligence, 이하 'AI')은 앞선 발제의 관련분야와 마찬가지로 전 세계에서 국경관리와 이민행정의 거의 모든 분야에 영향을 미치고 있음. 예를 들어, 얼굴인식에 기반한 자동출입국심사, 비자·전자여행허가신청에 대한 알고리즘 위험평가, 자율작동 드론을 통한 국경 감시, 난민 심사 과정에서의 신원과 출신지 검증까지.
- 이민행정은 그 대상이 되는 개인 – 특히 이주민, 난민, 무국적자등 구조적으로 취약한 지위에 있는 비국민(non-citizens) –의 인권에 중대한 영향을 미치며, AI 도입은 차별의 자동화, 대규모 감시, 적법절차의 형해화, 강제송환금지의무 위반 등 심각한 인권 문제를 약할 수 있다는 우려가 제기되어 옴.¹
- “행정 이민 절차와 국가 안보 사이의 경계가 모호해지면서, 이동하는 사람들의 권리를 보호하고 국가 정책에 대한 대중의 감시를 촉진해야 할 법적 체계에 광범위한 예외가 생겨나고 있고, 대부분의 데이터 보호 체제는 국가 안보 목적으로 처리되는 데이터를 예외로 인정하고 있음. 소위 "대규모 이민자 유입"에 대응하기 위해 국가들이 비상 권한을 사용하는 사례가 증가하는 것도 같은 효과를 가져올 수 있고”,² 최근 몇 년 동안 미주, 남유럽 및 동유럽의 최소 11개국 이 이러한 목적으로 국경 당국 지원을 위한 군 병력 배치, 비호 권리 제한, 신속한 구금 및 추방, 신원 확인 및 수색 압수 권한 확대, 감시 드론 및 생체 인식 시스템 사용 확대, 국경 지역에 대한 일반인 및 언론 접근 제한 등이 포함됨.
- 한국에서도 법무부와 과학기술정보통신부가 추진한 인공지능 식별추적시스템 개발사업³에서 출입국심사 과정에서 수집된 내·외국인의 개인정보와 안면데이터 약 1억 7천만 건이 정보주체의 동의 없이 민간기업의 AI 학습용으로 제공된 사실이 확인되어 헌법소원이 제기되었으나, 헌법재판소는 2026년 2월 26일 이를 각하함. 2026년 1월부터 시행된 「인공지능 발전과 신뢰 기반 조성 등에 관한 기본법」(이하 'AI기본법')은 세계에서 두 번째로 제정된 AI에 관한 포괄적 법률임에도, 이주·출입국 영역을 고영향(고위험) AI의 범위에 포함하지 않는 등 규율의 공백이 크다는 비판이 많은데, 출입국당국은 이민행정에 AI를 향후 전면도입하겠다고 함.

¹ OHCHR and University of Essex, *Digital Border Governance: A Human Rights Based Approach* (September 2023), <https://www.ohchr.org/sites/default/files/2023-09/Digital-Border-Governance-A-Human-Rights-Based-Approach.pdf>

² <https://verfassungsblog.de/human-rights-and-digital-border-governance/>

II. 이민행정 AI의 개념과 범주

- 개념정의의 시도 : ‘이민행정 AI’란 국경관리, 출입국심사, 사증(비자) 및 여행허가 심사, 난민인정 심사, 외국인 체류관리 및 단속 등 이민행정의 전 과정에서 데이터를 분석·분류·예측하여 인간의 결정을 보조하거나 대체하는 자동화 시스템이라고 일응 정의할 수 있음.
- 이는 EU 인공지능법(AI Act)과 OECD가 채택한 AI 시스템의 정의, 즉 명시적·묵시적 목적을 위하여 입력값으로부터 예측·권고·결정 등의 출력을 생성하는 기계 기반 시스템이라는 정의에 상응.³ 여기에는 좁은 의미의 머신러닝 기반 시스템뿐 아니라, 생체인식정보를 처리하는 자동화 시스템과 규칙 기반(rule-based) 자동 심사 시스템도 포함.
- 이러한 기술군을 OHCHR은 디지털 국경거버넌스(digital border governance)라고 총칭하기도 함.

유형	내용	대표 사례
① 신원확인·생체인식	얼굴·지문·홍채 등 생체정보에 기반한 신원확인, 자동출입국심사, 생체정보 데이터베이스 구축	자동출입국심사대, Eurodac, 한국 AI 식별추적시스템
② 위험평가·프로파일링	국적·여행이력 등 데이터를 이용하여 개인의 '위험도'를 산정하고 심사를 자동화	영국 비자 스트리밍, EU ETIAS, 한국 K-ETA
③ 국경 감시	자율드론, 감시탑, 센서, 위성을 통한 국경지대 이동 탐지·추적	미국 남부국경 감시탑·무인기, Frontex 감시
④ 심사·결정 보조	비자·비호 신청의 자동 분류·처리, 거짓말탐지, 출신지·방언 검증	캐나다 Chinook, EU iBorderCtrl, 독일 BAMF 언어인식
⑤ 체류관리·집행	체류 외국인에 대한 위험분류, 구금·단속 결정 보조, 전자감독	미국 ICE 위험분류평가, GPS 위치추적

³ Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act), OJ L, 2024/1689, Art. 3(1); OECD, *Recommendation of the Council on Artificial Intelligence*, OECD/LEGAL/0449

- 유형별 문제점 : 유형 ①·③은 프라이버시와 개인정보자기결정권에 대한 대규모 침해 가능성이, 유형 ②는 차별과 인종 프로파일링의 위험이, 유형 ④·⑤는 적법절차와 강제송환금지원칙 등 국제인권규범 위반이 특히 문제될 수 있음.

III. 해외의 이민행정 AI 활용사례와 현황

1. 미국: 국경감시의 산업화와 CBP One

- 미국 국토안보부(DHS)와 관세국경보호청(CBP)은 세계에서 가장 광범위한 이민행정 AI 체계를 운영함. 남부 국경지대에는 AI 기반 자율감시타워(autonomous surveillance towers)와 소형무인기(sUAS)가 대규모로 배치되어 국경을 넘는 사람들의 이동을 자동으로 탐지·추적함⁴. 학계에서는 이러한 감시 인프라가 이주 경로를 더 위험한 사막지대로 우회시켜 사망자 증가와 연관된다는 실증 연구도 축적되고 있다고 함.
- 난민에 관한 심사·결정 영역에서 대표적인 것이 2025년 3월부터 운용되는 CBP One 모바일 앱. 미국 정부는 남부 국경의 출입국항에서 비호신청 예약을 받는 사실상 유일한 통로로 이 앱의 사용을 의무화하였는데, 신청자는 얼굴 실물인증(facial liveness) 테스트를 위한 사진 촬영과 GPS 위치확인을 요구하여 비호신청 접근권에 제약을 부과함.
- 이민세관집행국(ICE)은 체포된 외국인의 구금 여부 판단을 보조하는 위험분류평가(Risk Classification Assessment) 알고리즘을 운영해 왔는데, ICE가 알고리즘에 영향을 미쳐 대부분의 사건에서 구금이 권고되도록 운영되었다는 점이 언론과 시민사회의 조사로 드러나 책임성 논란이 제기되었음.⁵

2. 유럽연합: iBorderCtrl, ETIAS, Eurodac

- EU의 iBorderCtrl 프로젝트는 국경을 통과하는 외국인의 얼굴 표정을 분석하여 거짓말 여부를 판별하겠다는 '자동 기만 탐지 시스템(ADDS)'을 헝가리·그리스·라트비아 국경에서 시범 운용하는 것임. 이에 대하여 인권단체와 학계는 감정인식에 기반한 거짓말탐지가 과학적 근거가 결여된 유사과학이며 필연적으로 차별적이라고 강하게 비

⁴ 감시망의 급속한 확장이 인종적 편견에 기반하고 있다는 지적으로는, Tsion Gurmu, Hinako Sugiyama and Sobechukwu Uwajeh, "Decolonizing AI at the U.S. Border", *Global Voices Advox*, 25 April 2026, <https://advox.globalvoices.org/2026/04/25/decolonizing-ai-at-the-u-s-border/>

⁵ S. Priya Morley, "AI at the Border: Racialized Impacts and Implications", *Just Security*, 17 September 2025, <https://www.justsecurity.org/97172/ai-at-the-border/>

판하였고, 시스템의 오류 위험과 프라이버시 침해 문제도 제기⁶.

- 2025년부터 단계적으로 시행되는 유럽여행정보허가시스템(ETIAS)은 무사증 입국자에 대하여 사전에 신상·여행 정보를 제출받아 보안·불법이주·보건 위험을 자동 심사하는 제도로, 신청 정보를 위험지표(risk indicators)에 기반한 스크리닝 규칙과 대조하는 알고리즘 프로파일링을 핵심 요소로 삼음.
- 셋째, 비호신청자 지문 데이터베이스인 Eurodac은 2024년 개정을 통해 수집 대상 생체정보에 얼굴이미지를 추가하고, 수집 연령을 6세까지 낮추는 등 점차 이를 대폭 확대해왔는데, 아동을 포함한 이주민 전반에 대한 생체감시의 정상화라는 비판이 있음.⁷

3. 영국: 비자 스트리밍 알고리즘의 도입과 폐기

- 영국 내무부(Home Office)는 2015년경부터 비자 신청에 대해 자동으로 적색·황색·녹색 위험등급으로 분류하는 ‘스트리밍 도구(streaming tool)’를 운용해왔는데, 여기서 국적이 핵심 변수로 사용되어 특정 국적 신청자가 체계적으로 고위험으로 분류되고 심사가 자동화됨. 이주민 권익단체 JCWI와 법률기술단체 Foxglove가 이 알고리즘이 인종차별적이며 위법하다는 사법심사(judicial review)를 제기하자, 내무부는 2020년 8월 소송 계속 중 도구의 사용을 전면 중단하였음⁸.
- 과거의 차별적 집행 데이터가 알고리즘에 학습되어 차별이 미래에도 계속 재생산·증폭되는 소위 ‘피드백 루프’ 문제를 보여주는 대표적 사례로 평가할 수 있음.

4. 캐나다: Chinook의 문제와, 유의미한 자동화 의사결정 지침

- 캐나다 이민난민시민권부(IRCC)는 임시비자 심사의 적체 해소를 위하여 신청 정보를 스프레드시트 형태로 일괄 표시·처리하는 엑셀 기반 Chinook 시스템과 신청 분류(triage) 자동화 도구를 도입(정부는 자체적인 의사결정 알고리즘은 없다고 함). 도입 이후 특정 국가 출신 신청에 대한 거부율 급증과 정형화된 거부 사유의 남발이 관찰되면서, 심사의 실질이 형해화되었다는 비판을 직면하였음.

⁶ Migration Policy Institute, “The Increasing Use of Artificial Intelligence in Border Zones Prompts Privacy Questions”, 2 February 2022, <https://www.migrationpolicy.org/article/artificial-intelligence-border-zones-privacy>

⁷ Regulation (EU) 2024/1358 of the European Parliament and of the Council of 14 May 2024 (Eurodac Recast), OJ L, 2024/1358.

⁸ Foxglove, “Home Office says it will abandon its racist visa algorithm”, 4 August 2020, <https://www.foxglove.org.uk>

- 한편 캐나다 연방정부 차원의 「자동화 의사결정에 관한 지침(Directive on Automated Decision-Making)」(2019년 발효)은 행정기관이 자동화 시스템을 도입하기 전에 알고리즘 영향평가(Algorithmic Impact Assessment)를 의무적으로 실시하고, 영향 수준에 따라 인적 개입·통지·설명·감사 등의 안전장치를 차등 적용하도록 하여 주목할만 함.⁹

5. 독일 및 기타 사례

- 독일 연방이민난민청(BAMF)은 2017년부터 신원서류가 없는 비호신청자의 진술 출신지를 검증하기 위하여 음성 샘플을 자동 분석하는 소프트웨어를 사용했음. 그러나 음성과 국적은 일치하지 않으며 시스템의 오류율이 상당하다는 점에서, 오판이 곧 난민지위 거부와 강제송환으로 이어질 수 있는 고위험 용도에 검증되지 않은 기술을 사용한다는 비판 다수¹⁰.
- 네덜란드에서는 복지 부정수급 탐지를 위한 위험프로파일링 시스템 SyRI에 대하여 헤이그 지방법원이 2020년 2월 5일, 시스템의 불투명성과 저소득·이주민 밀집 지역에 대한 차별적 적용 가능성을 이유로 유럽인권협약 제8조(사생활 존중) 위반을 인정하고 그 사용을 금지하는 판결을 선고¹¹. 이민행정에 국한된 판결은 아니지만, 국가의 알고리즘 프로파일링에 대하여 법원이 인권규범을 근거로 제동을 건 선례로서 이민행정 AI 논의에서 반복적으로 인용될 수 있음.
- 이 밖에 호주·싱가포르 등도 생체인식 자동심사와 비자심사 자동화를 확대하고 있어, 이민행정의 AI 전환은 사실상 전 세계적 추세.

IV. 한국의 활용 사례와 현황 - 출입국관리의 국경과 사법심사의 국경의 겹을 AI가 더욱 벌리고 있음

1. 법무부 ‘AI 식별추적시스템’ 사건

- 한국에서 이민행정 AI를 둘러싼 가장 중요한 사건은 법무부·과학기술정보통신부의 ‘인공지능 식별추적시스템 개발사업’(2019~2021). 지문 기반의 공항 출입국심사 시스템을 데이터·AI 기반으로 고도화하는 한편, 민간기업에 법무부 보유 실증 데이터를 학습용으로 제공하는 것을 내용으로 했던 것임.

⁹ Government of Canada, *Directive on Automated Decision-Making* (effective 1 April 2019, as amended), <https://www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=32592>

¹⁰ Petra Molnar, *The Walls Have Eyes: Surviving Migration in the Age of Artificial Intelligence* (The New Press, 2024) 참조.

¹¹ Rechtbank Den Haag Judgment of 5 February 2020, ECLI:NL:RBDHA:2020:865 (*NJCM et al. v. the Netherlands*, 이른바 SyRI 판결).

- 개인정보보호위원회의 2022년 4월 27일 조사 결과에 따르면, 법무부가 출입국심사 과정에서 수집·보유하던 내국인 5,760만 건과 외국인 1억 2천만 건, 합계 약 1억 7천만 건의 여권번호·국적·생년·성별 등 개인정보와 민감정보인 안면식별정보가 정보주체의 동의 없이 24개 민간기업에 AI 알고리즘 학습·검증용으로 제공됨.¹²
- 개인정보보호위원회는 얼굴정보가 민감정보에 해당한다고 보면서도, 출입국관리법상 '안전한 국경관리'라는 목적 범위 내의 이용이자 정상적인 처리위탁에 해당한다고 판단하고, 위탁사실 미공개 등을 이유로 법무부에 과태료 100만 원을 부과하는데 그침.
- 시민사회단체들은 이를 두고 국가가 공적 목적으로 수집한 민감정보를 사기업에 대 규모로 제공한 행위에 면죄부를 준 것이라고 강하게 반발함. 이후 내·외국인 청구인을 대리하여 2022년 7월 이 사건 개인정보 처리행위와 그 근거조항인 출입국관리법 관련 규정, 그리고 생체정보를 AI 개발 목적 활용으로부터 보호하지 않은 입법·행정부작위의 위헌확인을 구하는 헌법소원을 청구하였으나, 헌법재판소는 2026년 2월 26일 본안 판단 없이 권리보호이의 소멸 등을 이유로 청구를 각하함.

2. K-ETA와 자동화된 입국 전 심사

- 2021년 9월 시행된 전자여행허가제(K-ETA)는 무사증 입국 대상 국가 국민이 입국 전에 온라인으로 개인정보와 여행정보를 제출하여 여행허가를 받도록 하는 제도임.
- 법무부의 공개된 보도 자료에 따르면¹³ K-ETA 심사는 시스템이 생체정보(얼굴), 규제자 정보, 승객위험도의 세 항목을 자동으로 분석·심사하여 문제가 없으면 통상 30분 이내에 자동허가하고, 하나라도 문제가 있으면 분석관의 정밀분석을 거쳐 조건부 허가 또는 불허를 결정하는 구조. 즉 K-ETA는 위 유형분류상 '위험평가·프로파일링'(유형 ②)에 해당하는 자동화 심사를 제도의 중핵으로 삼음. 그러나 '승객위험도'를 산정하는 기준과 변수는 공개되어 있지 않고, 불허 결정에 대하여 구체적 사유가 고지되지 않으며 독립적인 불복 절차도 마련되어 있지 않아, 특정 국적 신청자에 대한 높은 불허율 논란과 함께 심사의 불투명성이 문제됨.
- 이와 별도로 1)승객정보사전분석시스템(APIS)을 통한 탑승 전 위험분석, 2)탑승전 사전 확인제도, 3)출입국항에서의 자동출입국심사대 및 안면인식 기반 스마트패스 등 생체인식 기반 자동화(유형 ①)도 전면적으로 확대되는 중임.

3. 'AI 기반 이민행정 혁신'의 전면화

¹² 참여연대 공익법센터, "출입국 얼굴인식 인공지능 식별추적 헌법소원 청구", 2022. 7. 21., <https://www.peoplepower21.org/PublicLaw/1898014>

¹³ 법무부 보도자료, "전자여행허가제(K-ETA) 시행 안내", 2021. 8., <https://www.immigration.go.kr/bbs/immigration/47/551356/artclView.do>

- 법무부는 최근 이민행정 전반의 AI 전환을 공식 정책 기조로 채택. 2025년 11월 28 일에는 기존의 정책 아이디어 발표회를 확대·개편한 제1회 'AI 기반 이민정책·행정 혁신 발표회'를 개최하여 입국심사에서 체류관리에 이르는 행정 전반에 AI 기술을 적용하는 방안을 논의하였음.¹⁴
- 더 나아가, 『2030 이민정책 미래전략』을 통하여 데이터·AI 기반 이민행정 고도화를 중장기 과제로 제시. 이는 심사관의 경험 중심 심사에서 데이터 기반 심사로의 전환, 체류 외국인에 대한 데이터 기반 관리 등 유형 ②·④·⑤ 전반으로 AI 활용이 확대될 것임을 예고하는 것임.
- 이러한 확대가, 후술하는 바와 같이 이민행정 AI를 규율하는 법제도적 안전장치가 사실상 부재한 상태에서 진행되고 있는데, 도입이 필요한 AI와 규율이 필요한 고영향 AI¹⁵에 관한 규범이 존재하지 않아 위험함.

V. 국제법 및 국제규범상의 기준

이민행정에서의 AI 오용을 통제하기 위한 규범은 크게 (1) 기존 국제인권조약상 의무, (2) 유엔인권기구의 해석·권고 문서, (3) 지역적·국내적 입법례로 분류해볼 수 있음. 아직 이민행정 AI만을 규율하는 단일한 구속력 있는 조약은 존재하지 않으나, 아래에서 보듯 각 층위의 규범을 종합하면 일응의 규범적 기준을 도출할 수 있음.

1. 국제인권조약상 의무

- 첫째, 「시민적 및 정치적 권리에 관한 국제규약」(ICCPR) 제17조는 사생활과 통신에 대한 자의적·불법적 간섭을 금지하며, 이는 국가의 생체정보 수집·처리와 감시 활동에 합법성·필요성·비례성 요건을 부과하는 근거임. 같은 규약 제2조 및 제26조의 차별금지와 법 앞의 평등은 국민·외국인을 불문하고 적용됨.
- 둘째, 「모든 형태의 인종차별 철폐에 관한 국제협약」(ICERD)은 인종·피부색·민족적 출신에 따른 구별·배제를 금지하는데, 인종차별철폐위원회(CERD)는 일반권고 제36호(2020)에서 법집행에서의 인종 프로파일링을 정면으로 다루면서, 알고리즘 프로파일링에 사용되는 데이터의 편향이 차별적 결과를 낳을 수 있음을 지적하고, 알고리즘 의사결정의 투명성 확보, 독립적 감독, 영향평가 등을 국가의 의무로 권고하였던 바, 이는 이민행정에서 특히 음미할 가치가 있음.¹⁶

¹⁴ 한국다문화인터넷신문, “법무부, AI 기반 이민행정 혁신 본격화... 입국심사·체류관리까지 스마트 전환 속도”, 2025. 12. 1.

¹⁵ 제2조 제4호. 사람의 생명, 신체의 안전 및 기본권에 중대한 영향을 미치거나 위험을 초래할 우려가 있는 인공지능시스템으로서 다음 각 목의 어느 하나의 영역에서 활용되는 것을 말한다.

¹⁶ Committee on the Elimination of Racial Discrimination, *General Recommendation No. 36 (2020) on Preventing and Combating Racial Profiling by Law Enforcement Officials*, UN Doc. CERD/C/GC/36 (17

- 셋째, 「난민의 지위에 관한 협약」 제33조의 강제송환금지원칙은 비호절차에 대한 접근을 기술적으로 차단하거나 오류 있는 자동화 결정으로 보호신청을 배척하는 경우 침해될 수 있음.¹⁷
- 넷째, 「아동의 권리에 관한 협약」 상 아동 최선의 이익 원칙은 이주 아동에 대한 생체정보 수집·프로파일링에 특별한 제약을 가함.

2. 유엔인권기구의 견해

- 첫째, 인종차별 특별보고관(E. Tendayi Achiume)은 2020년 유엔총회에 제출한 보고서에서 디지털 국경기술이 이주민·무국적자·난민 등 비국민에 대하여 외국인혐오적이고 인종차별적인 영향을 미치고 있음을 분석하고, 국가들에 대하여 인권영향평가 없이 도입된 국경기술의 사용 중단, 차별적 영향이 입증된 기술의 금지, 그리고 위험성이 규명될 때까지의 모라토리엄(moratorium)을 권고하였음¹⁸.
- 둘째, 유엔인권최고대표사무소(OHCHR)는 에식스대학교와 공동으로 70명 이상의 전문가 인터뷰를 통해 2023년 9월 『디지털 국경 거버넌스: 인권에 기반한 접근』 보고서를 발간함. 이 보고서는 생체인식, 대규모 상호운용 데이터베이스, 위험 스크리닝·프로파일링, 자동화 의사결정, 국가 간 데이터 공유, 민간부문 참여 등 국경기술 전반의 인권 영향을 분석하고,
 - 국가에 대하여 (i) 기술 사용 전에 그 필요성과 합법성을 입증할 것,
 - (ii) 인권과 근본적으로 양립할 수 없는 기술의 사용을 자제할 것,
 - (iii) 포괄적인 사전 인권영향평가를 실시할 것 등을 권고함.¹⁹
- 나아가 OHCHR은 이 보고서를 토대로 소위 『인권에 기반한 디지털 국경 거버넌스에 관한 원칙과 지침의 초안』 (guidance on human rights-based digital border governance)을 유엔사무총장의 지시에 따라 만들 것을 요구받았음(아직 미발표되었고, 과거 지침에는 디지털 국경 거버넌스가 없음²⁰), 이 지침에는 생체정보와 대규모

December 2020), paras. 31-39

¹⁷ 고문방지협약 제3조 및 ICCPR 제6조·제7조상의 비호 관련 의무와도 연결됨.

¹⁸ E. Tendayi Achiume, *Report of the Special Rapporteur on Contemporary Forms of Racism, Racial Discrimination, Xenophobia and Related Intolerance: Racial and Xenophobic*

¹⁹ OHCHR, "Digital Technologies at Borders: A Threat to People on the Move", 9 October 2023, <https://www.ohchr.org/en/stories/2023/10/digital-technologies-borders-threat-people-move>

²⁰ RECOMMENDED PRINCIPLES and GUIDELINES on HUMAN RIGHTS at INTERNATIONAL BORDERS https://www.ohchr.org/sites/default/files/Documents/Issues/Migration/OHCHR_Recommended_Principles_Guidelines.pdf

데이터베이스, 위협 스크리닝과 프로파일링, 자동화·AI 보조 의사결정, 국가간 데이터 공유와 상호운용성의 문제, 민간부문 참여를 핵심적인 주제로 다루면서, 권리 제한에는 명확하고 접근 가능하며 예견 가능한 법적 근거와 적절한 안전장치가 요구된다는 점, 국가안보가 인권에 대한 모호하거나 자의적인 제한의 구실이 되어서는 안 된다는 점을 확인할 것으로 보임.²¹

- 셋째, 유엔인권이사회는 「디지털 시대의 프라이버시권」 결의(2021)에서 원격 생체 인식 기술이 야기하는 중대한 프라이버시 침해 위험을 지적하면서, 각국이 국제인권법 준수를 보장할 수 없는 경우 공공장소에서의 원격 실시간 생체인식 사용에 대한 모라토리엄을 검토할 것을 촉구하였음.²²
- 넷째, 유네스코의 「인공지능 윤리 권고」(2021)는 194개 회원국이 채택한 최초의 글로벌 AI 윤리 기준으로서 비례성, 안전, 공정성·차별금지, 투명성·설명가능성, 인간의 감독과 결정, 책무성 등을 원칙으로 제시하며, 대규모 감시 목적의 AI 사용을 배제하여야 한다고 명시함.²³

3. 지역적·국내적 입법례: EU AI법과 유럽평의회 AI 기본협약

- 구속력 있는 입법 가운데 가장 중요한 것은 2024년 제정된 EU 인공지능에 관한 조화규칙의 제정에 관한 규정(Regulation (EU) 2024/1689)임. 위 법은 위험기반 접근(risk-based approach)을 채택하여, 제5조에서 '허용될 수 없는 위험'을 가진 AI 활용 — 공공장소에서의 법집행 목적 실시간 원격 생체인식(엄격한 예외 있음), 사회적 점수화, 인터넷·CCTV 영상의 무차별 수집(scraping)에 의한 얼굴인식 데이터베이스 구축, 직장·교육기관에서의 감정인식 등 — 을 금지하고(2025년 2월부터 적용), 부속서 III 제7항에서 이주·비호·국경관리 분야의 AI, 즉 (a) 거짓말탐지기 및 유사 도구, (b) 입국하려는 자의 안보·불법이주·보건 위험 평가, (c) 비호·비자·체류허가 신청 심사 보조, (d) 이주·국경관리 맥락에서의 개인 탐지·인식·식별을 고위험(high-risk) AI로 분류함.²⁴ 고위험 AI에는 위험관리체계, 데이터 거버넌스, 기술문서화, 로그 기록, 투명성, 인적 감독(human oversight), 정확성·견고성 확보 등의 의무가 부과되며

²¹ "Human Rights and Digital Border Governance: Opportunities and Challenges for the New OHCHR Guidance", *Verfassungsblog*, December 2025, <https://verfassungsblog.de/human-rights-and-digital-border-governance/>

²² Human Rights Council, *The Right to Privacy in the Digital Age*, UN Doc. A/HRC/RES/48/4 (7 October 2021), <https://digitallibrary.un.org/record/3896430?ln=en>

²³ UNESCO, *Recommendation on the Ethics of Artificial Intelligence* (adopted 23 November 2021), <https://www.unesco.org/en/articles/recommendation-ethics-artificial-intelligence>

²⁴ Regulation (EU) 2024/1689, Art. 5 and Annex III, para. 7.

(2026년 8월부터 전면 적용), 공공기관 등이 고위험 AI를 배치할 때에는 기본권 영향평가(fundamental rights impact assessment)를 실시하여야 한다고 규정함(제27조).

- 다만 유럽 시민사회의 #ProtectNotSurveil 연합 등은 위 규정이 이주 영역에서 법집행·국가안보 관련 예외와 허점을 두고 있어, 정작 이주민 보호에는 미흡하다고 비판하고 있기도 함.²⁵
- 한편 유럽평의회가 2024년 채택한 「인공지능과 인권, 민주주의 및 법의 지배에 관한 기본협약(The Framework Convention on Artificial Intelligence)」은 AI 시스템의 전 수명주기에 걸쳐 인권·민주주의·법치주의에 합치될 것을 요구하는 최초의 구속력 있는 국제조약임. 유럽 외 국가에도 가입이 개방되어 있고 일본도 가입함.²⁶

4. 일응의 기준

- 첫째(합법성), AI에 의한 개인정보 처리와 권리 제한에는 명확하고 예견 가능한 법률상 근거가 있어야 하며, 수집 목적 외 이용은 엄격히 통제되어야 함.
- 둘째(필요성·비례성 등 비례의 원칙), 덜 침해적인 수단으로 목적을 달성할 수 있는지 사전에 심사되어야 함.
- 셋째(차별금지), 국적·인종 등을 직접 변수로 하거나 대리변수를 통하여 특정 집단을 표적화하는 프로파일링은 금지되며, 시스템의 차별적 영향에 대한 지속적 모니터링이 가능해야 함.
- 넷째(특정영역에서는 전면적 금지), 공공장소 실시간 원격 생체인식, 감정인식 기반 거짓말탐지 등 인권과 근본적으로 양립하기 어려운 응용은 금지하거나 안전성이 입증될 때까지 유예하여야 함.
- 다섯째(투명성·설명가능성), 이용되는 시스템의 존재·목적·논리와 운영 통계가 공개되어야 하고, 당사자에게 결정 이유가 고지되어야 함.
- 여섯째(인적 개입), 개인에게 중대한 영향을 미치는 결정이 자동화 시스템에 의해 단독으로 내려져서는 안 되며, 유의미한 인간의 검토와 최종결정권이 보장되어야 함.
- 일곱째(사전 영향평가), 도입 전 인권영향평가·알고리즘 영향평가가 의무적으로 실시되고 그 결과가 공개되어야 함.
- 여덟째(감독과 구제), 독립적 감독기구의 상시 감독과 당사자의 효과적인 불복·구제 절차가 보장되어야 함.

VI. 한국의 이민행정 AI 도입에 관한 평가

²⁵ <https://protectnotsurveil.eu>

²⁶ <https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence>

1. 사업 차원의 문제

- 첫째, AI 식별추적시스템 사건은 위 규범적 기준의 거의 전 항목에 저촉됨. 출입국심사 목적으로 수집된 생체정보가 민간기업의 AI 개발이라는 목적 외 용도로 이용되었다는 점에서 합법성·목적구속 원칙이(기준 ①), 1억 7천만 건에 이르는 처리 규모와 대안적 수단에 대한 검토 부재에서 필요성·비례성이(기준 ②), 사업의 존재와 데이터 제공 사실이 정보주체에게 전혀 알려지지 않았다는 점에서 투명성이(기준 ⑤), 도입 전 인권영향평가가 실시되지 않았다는 점에서 사전 영향평가 의무가(기준 ⑦), 그리고 과태료 100만 원과 헌법소원 각하로 귀결된 사후 처리에서 감독과 구제가(기준 ⑧) 각각 문제됨. 특히 이 사건에서 정부가 법적 근거로 원용한 출입국관리법 제12조의²⁷는 외국인이 입국 시 본인확인 절차에서 생체정보를 제공하도록 한 조항일 뿐인데, “응하여야 한다”라는 규정에서 이것이 AI 학습용 대량 제공까지 허용하는 근거가 될 수는 없음.
- 둘째, K-ETA와 승객위험도 분석은 국제적으로 가장 논쟁적인 유형인 알고리즘 위험 프로파일링을 제도의 중핵으로 삼고 있음에도, 위험도 산정에 사용되는 변수와 기준이 공개되지 않고, 불허 결정의 구체적 사유가 고지되지 않으며, 국적별 불허율 등 운영 통계의 공개나 차별 영향에 대한 독립적 점검도 이루어지지 않았음. 이는 투명성(기준 ⑤), 차별금지 모니터링(기준 ③), 구제(기준 ⑧)의 관점에서 영국 스트리밍 알고리즘이나 ETIAS에 제기된 비판이 한국에도 그대로 적용될 수 있음을 의미함. 자동허가가 다수를 차지하는 심사 구조에서 인적 개입(기준 ⑥)이 ‘불허 방향’으로만 작동하고 ‘허가의 자동화’는 통제되지 않는 점도 검토가 필요함.
- 셋째, 법무부가 추진하는 이민행정 전반의 AI 전환은 심사·체류관리·단속 영역(유형 ④·⑤)으로의 확대를 예고하고 있으나, 개별 사업에 대한 인권영향평가, 알고리즘 공개, 당사자 권리보장 등의 안전장치가 제도화되어 있지 않은 상태에서의 확대는 위 문제들을 구조적으로 재생산할 위험이 큼.

2. 법제도 차원의 문제

- 첫째, 출입국관리법의 규율 공백이 있음. 현행 출입국관리법은 입국단계에서 생체정보의 수집(제12조의2 등)을 허용하면서도, 수집된 생체정보의 이용 범위, 보유기간, 제3자 제공, AI 개발 등 2차적 이용의 요건과 한계에 관하여 전혀 규율하지 않음.
- 둘째, 개인정보보호법 적용의 한계가 있음. 위 식별추적시스템 사건에서 개인정보보호위원회는 ‘안전한 국경관리’라는 넓은 목적 해석을 통하여 AI 학습용 이용을 목적 범위 내로 인정하였는데, 이러한 해석이 유지되는 한 공공기관이 보유한 대규모 민감

²⁷ ① 입국하려는 외국인은 제12조에 따라 입국심사를 받을 때 법무부령으로 정하는 방법으로 지문 및 얼굴에 관한 정보를 제공하고 본인임을 확인하는 절차에 응하여야 한다.

정보의 AI 전용(轉用)을 개인정보보호법으로 통제하기는 어려움. 2023년 개정으로 신설된 자동화된 결정에 대한 정보주체의 권리(제37조의2)²⁸는 의미 있는 진전이지만, 법률에 특별한 규정이 있는 경우 등에는 거부권이 제한되는 구조여서 출입국 영역에서의 실효성은 불분명함.

- 셋째, AI기본법의 구조적 한계도 있음. 2026년 1월 시행된 AI기본법은 고영향 AI와 생성형 AI에 대한 사업자 의무를 규정한 기본법제이지만, 이민행정의 관점에서 문제가 분명함.
 - EU AI법과 달리 완전히 금지되는 AI 유형에 관한 규정 자체가 없어, 공공장소 실시간 원격 생체인식이나 감정인식 기반 심사와 같은 최고위험 응용도 법률상 금지되지 않음.
 - 고영향 AI의 범위가 협소하여, EU가 명시적으로 고위험으로 분류한 이주·비호·국경관리 영역이 포함되어 있지 않음. 생체인식 관련하여는 범죄 수사·체포 목적의 분석·활용 정도가 포착될 뿐이어서, 출입국심사·전자여행허가·체류관리에서의 AI는 고영향 AI 규율의 사각지대에 놓여 있음.
 - 국방·국가안보 목적으로만 개발·이용되는 AI를 법 적용에서 전면 배제하는 조항(제4조 제2항)²⁹은, 국가안보를 이유로 한 광범위한 규율 회피의 통로가 될 수 있다는 점에서 국가안보가 인권 제한의 구실이 되어서는 안 된다는 국제기준과 충돌할 소지가 있음.

²⁸ 제37조의2(자동화된 결정에 대한 정보주체의 권리 등) ① 정보주체는 완전히 자동화된 시스템(인공지능 기술을 적용한 시스템을 포함한다)으로 개인정보를 처리하여 이루어지는 결정(「행정기본법」 제20조에 따른 행정청의 자동적 처분은 제외하며, 이하 이 조에서 “자동화된 결정”이라 한다)이 자신의 권리 또는 의무에 중대한 영향을 미치는 경우에는 해당 개인정보처리자에 대하여 해당 결정을 거부할 수 있는 권리를 가진다. 다만, 자동화된 결정이 제15조제1항제1호·제2호 및 제4호에 따라 이루어지는 경우에는 그러하지 아니하다.

② 정보주체는 개인정보처리자가 자동화된 결정을 한 경우에는 그 결정에 대하여 설명 등을 요구할 수 있다.

③ 개인정보처리자는 제1항 또는 제2항에 따라 정보주체가 자동화된 결정을 거부하거나 이에 대한 설명 등을 요구한 경우에는 정당한 사유가 없는 한 자동화된 결정을 적용하지 아니하거나 인적 개입에 의한 재처리·설명 등 필요한 조치를 하여야 한다.

④ 개인정보처리자는 자동화된 결정의 기준과 절차, 개인정보가 처리되는 방식 등을 정보주체가 쉽게 확인할 수 있도록 공개하여야 한다.

⑤ 제1항부터 제4항까지에서 규정한 사항 외에 자동화된 결정의 거부·설명 등을 요구하는 절차 및 방법, 거부·설명 등의 요구에 따른 필요한 조치, 자동화된 결정의 기준·절차 및 개인정보가 처리되는 방식의 공개 등에 필요한 사항은 대통령령으로 정한다.

²⁹ 인공지능 발전과 신뢰 기반 조성 등에 관한 기본법 제4조 ② 이 법은 국방 또는 국가안보 목적으로만 개발·이용되는 인공지능으로서 대통령령으로 정하는 인공지능에 대하여는 적용하지 아니한다.

- 법률이 시행령으로 고영향 AI의 추가 지정을 위임하였음에도 시행령(안)이 이를 전혀 구체화하지 않았다는 점³⁰, 나아가 정부가 규제 조항의 시행 유예까지 추진하였다는 점에 대하여 국가인권위원회와 22개 노동·시민사회단체가 잇달아 우려와 반대를 표명³¹.
- 넷째, 절차적 통제장치가 부재함. 캐나다의 알고리즘 영향평가나 EU의 기본권 영향평가와 같이 공공부문 AI 도입 전에 의무적으로 실시되는 영향평가 제도가 없음. 개인정보 영향평가(개인정보보호법 제33조)는 단지 개인정보파일 운용에 초점을 둔 것으로 - 정보주체의 개인정보 침해가 우려되는 경우에는 그 위험요인의 분석과 개선 사항 도출을 위한 평가 -에 불과하여, 알고리즘의 차별적인 영향이나 적법절차 문제를 포괄하지 못함. 일반적인 경우도 그러한데, 이민행정 AI에 특화된 독립적 감독 메커니즘, 당사자에 대한 고지·설명·불복 절차도 마련되어 있지 않음.

VII. 현 단계에서의 과제

- 규범적으로 이민행정에서의 AI 사용 일반의 금지가 규범적으로 요구된다고 볼 수는 없음. 개별 위험에 비례하는 차등적 통제를 할 필요가 있음. EU조차 이주 분야 AI를 금지한 것이 아니라 고위험으로 분류하여 조건부로 허용함.

1. 입법적 개선³²

- 첫째, 출입국관리법을 개정하여 생체정보의 처리에 관한 명확한 규율을 도입하여야 함. 구체적으로 수집 목적의 특정, 이용 범위의 한정(본인확인·심사 목적), 보유기간과 파기, 제3자 제공 및 AI 개발 등 2차적 이용의 원칙적 금지와 예외적 허용 요건(법률의 명시적 근거, 독립기구의 사전 심사), 그리고 처리 현황의 정기적 공개를 법률에 명시해야함. 이미 이는 식별추적시스템 사건에서 드러난 규율 공백을 메우는 최소한의 조치라고 할 수 있음.
- 둘째, AI기본법을 개정하거나, 시행령에 따라 (i) 고영향 AI의 범위에 이주·출입국·난

³⁰ 인공지능 발전과 신뢰 기반 조성 등에 관한 기본법 제2조 제4항 카. 그 밖에 사람의 생명·신체의 안전 및 기본권 보호에 중대한 영향을 미치는 영역으로서 대통령령으로 정하는 영역. 그러나 대통령령이 제정되지 않았음.

³¹ 참여연대 외, "[공동논평] 인공지능기본법 규제 시행유예안 반대한다", 2025. 9. 3., <https://www.peoplepower21.org/publiclaw/1999461>

³² "[의견서] 고영향 AI 해당 분야 구체적, 폭넓게 포함하고, 시민의 인권, 안전 보호할 안전조치 규정해야" <https://www.peoplepower21.org/publiclaw/1989089>

민 심사와 국경관리 분야를 명시적으로 포함하고, (ii) 공공장소에서의 실시간 원격 생체인식(엄격한 법관 통제 하의 예외 유보), 감정인식에 기반한 거짓말탐지·심사, 사회적 점수화 등 전면 금지 AI 유형을 신설하며, (iii) 국가안보 예외조항을 별도 입법을 통해 마련하거나 축소·정비하고, (iv) 제재의 실효성을 확보하여야 함.

2. 절차적 통제: 영향평가·투명성·인적 개입·구제

- 공공부문이 이민행정 AI를 도입하기 전에 인권영향평가와 알고리즘 영향평가를 의무적으로 실시하고 결과를 공개하도록 하는 제도를 도입하여야 함. K-ETA를 포함한 자동화 심사에 대하여 시스템의 존재와 심사 논리의 개요, 국적별 허가·불허율 등 운영 통계를 정기적으로 공개하고, 불허 결정 시 구체적 사유를 고지하며, 인간 심사관에 의한 재심사를 청구할 수 있는 불복 절차를 신설하여야 함.
- 아울러 개인의 권리·지위에 중대한 영향을 미치는 결정(입국불허, 비자거부, 난민심사, 강제퇴거, 구금)은 자동화 시스템의 출력만으로 내려질 수 없고 유의미한 인적 검토를 거치도록 법정화하고(Human in the loop), 시스템 로그의 보존과 감사가능성을 확보하여 사후 검증과 구제의 기반을 마련하여야 함.

3. 거버넌스: 독립 감독과 민간참여 통제

- 이민행정 AI에 대한 상시적 감독 권한을 명확히 하여야 함. 개인정보보호위원회의 공공기관 조사·제재 권한을 실질화하고, 차별에 관한 조사와 권고 기능을 이미 갖고 있는 국가인권위원회에 공공 AI의 차별 영향에 대한 조사·권고 기능을 제도적으로 부여하며, 국회 차원의 정기 보고 의무를 도입하는 방안을 검토할 수 있음.
- 문제가 되었던 AI에 관한 민간기업의 참여에 대하여는, 조달·위탁 계약 단계에서 인권실사와 보안·차별성 검증을 요구하고, 국가 보유 개인정보의 민간 AI 학습용 제공은 원칙적으로 금지하되 예외적으로 허용할 경우에도 가명·익명처리, 독립기구 사전 승인, 정보주체 통지를 요건으로 하여야 함. 이는 「기업과 인권에 관한 유엔 이행원칙」³³이 요구하는 국가의 보호의무와 기업의 인권 존중책임을 이민행정 AI의 맥락에서도 구체화하는 것으로 볼 수 있음.

4. 혁신과의 병행: 위험기반 차등 규제

- 다국어 안내 챗봇, 서류의 OCR과 자동 분류, 통·번역 지원, 대기시간 예측과 인력 배치 최적화 등 개인의 권리에 중대한 영향을 미치지 않는 저위험 응용은 오히려 이주민, 난민들의 권리 옹호를 위해서는 적극적인 활용을 촉진할 수 있음.

³³ 2011. 기업과 인권 이행 원칙: 유엔 "보호, 존중, 구제" 프레임워크의 실행 Guiding Principles on Business and Human Rights: Implementing the United Nations "Protect, Respect and Remedy" Framework <https://uhr.humanrights.go.kr/pub/uhrstd/1159>

- 그러나, 동시에 위험평가·프로파일링과 생체인식 등 고위험 응용은 위에서 제시한 안전장치를 조건으로 예외적으로 허용하고, 인권과 근본적으로 양립하기 어려운 응용은 전면적으로 금지하는 3단계 구조를 만들어야 함.

경찰의 AI 기술을 활용한 치안 현황

랑희_공권력감시대응팀

예측적 치안(predictive policing)은 발생 가능성이 있는 잠재적인 범죄행위를 예방하기 위한 목적으로 인공지능 및 빅데이터 등 지능정보기술에 기반한 예측 분석 결과를 치안행정에서 활용하는 것을 의미. 예측적 치안은 범죄 발생률을 낮추고 경찰 활동의 효율성을 제고할 수 있으리라는 기대에서 그 도입의 정당성 및 당위성이 적극적으로 논의되었음.¹⁾ 기술의 발달은 위험 예방을 위한 경찰 활동의 가능성을 확장했으나 AI를 활용한 범죄 예측 및 예방 시스템은 대량의 개인정보와 감시 기술을 필연적으로 동반함.

현재 경찰의 AI 치안사업에는 인공지능 CCTV와 같은 AI 영상분석 기술이 많은 부분에서 활용되는데 영상분석 기술에는 얼굴인식 등 생체인식 기술이 수반되기도 하고, 일상 공간의 광범위한 모니터링이라는 특성상 시민 프라이버시 침해 우려가 큼.

경찰은 2026년을 '치안 AI'의 원년으로 선언하고 치안업무 전반에 인공지능(AI)을 접목하기 위한 기반 조성하고 있음.²⁾ 이와 관련해 진행되고 있는 치안정책과 기술들을 살펴보고자 함.

1. 조직개편

미래치안의 핵심을 AI 분야로 보고 관련 업무 영역 전반에 이를 적용·확산하기 위해 경찰청 조직개편 진행. AI 기술을 활용해 치안 역량을 대대적으로 강화하겠다는 취지.

TF에서만 담당했었던 AI 업무를 미래치안국 전체로 확산, 미래치안정책국 산하 과 명칭에 '인공지능'을 넣음³⁾. AI 혁신 TF(태스크포스)도 정식 직제화함.

치안 AI 혁신 태스크포스(TF)는 2025년 경찰 내부 인력과 외부 전문가 등 6명으로 구성되었음. AI 기반 치안 체계 구축 필요성을 정부 정책에 반영하는 작업을 추진해 9개 단위과제와 31개 세부과제가 국정과제로 채택됨. 경찰은 이를 토대로 3대 전략과 25개 세부과제로 구성된 '치안 AI 혁신 종합계획'을 수립해 추진 중. 현장·수사·행정·국민 서비스 등 4개 분야에서 12개 선도 프로젝트 설정.

1) 정재연, 예측적 치안과 형사정책: 쟁점과 과제, 영남법학 제57호, 영남대학교 법학연구소, 2023

2) 발제문의 경찰 미래비전 2050 추진과제 목록에서 도전과제 1,2와 연결되어 있음을 추정할 수 있음.

3) 미래치안정책과를 '인공지능정책과', 정보화기반과를 '인공지능기반과' 등으로 변경

미래치안정책국 조직 구성

과	부서명	전화번호	담당업무
치안인공지능정책과	미래치안기획계	(02) 3150 - 2241	중장기 미래치안전략 수립·조정 및 추진사항·성과 관리
	연구개발기획예산계	(02) 3150 - 2891	치안분야 과학기술 연구개발, 예산·결산 총괄
	인공지능정책계	(02) 3150 - 0503	인공지능 치안 정책 기획 및 법제 업무 총괄
	데이터정책계	(02) 3150 - 3205	빅데이터 플랫폼 운영, 데이터 분석·품질관리
치안인공지능기반과	정보자원관리계	(02) 3150 - 2244	정보화 기획·예산, 정보화사업관리, 통합지원
	디지털협력계	(02) 3150 - 2406	정보시스템(포털,온나라 등) 운영, 통합유지보수
	디지털인프라계	(02) 3150 - 2444	유·무선 통신, 재난안전통신망 및 업무용 컴퓨터 등 운영
	디지털보안계	(02) 3150 - 2741	정보통신보안, 개인정보보호, 보안관계
장비운영과	장비정책계	(02) 3150 - 2236	경찰장비 예산 및 규격 총괄, 피복 및 장구 구매보급
	장비운영계	(02) 3150 - 2237	차량·무기·드론 등 미래치안 장비 전담운영
치안인공지능서비스팀	인공지능서비스계	(02) 3150 - 1421	치안AI 서비스 개발 및 운영
치안연구개발산업과	안전연구개발계	(02) 3150 - 2526	안전기술 연구개발(R&D) 성과 및 연구관리
	수사연구개발계	(02) 3150 - 2108	수사기술 연구개발(R&D) 성과 및 연구관리
	치안산업계	(02) 3150 - 0675	치안산업 정책기획 및 진흥 업무 총괄

2. 치안 AI 혁신 종합계획

<추진 배경> ◦ 경찰청 국정과제 「치안 AI 혁신」을 체계적이고 안정감 있게 이행하기 위해 모든 관·부서가 참여하여 실천하는 치안 AI 종합계획(3개년) 수립·시행 ◦ 조직 내 숙의과정을 거쳐 대내외 환경분석 및 추진 전략을 도출하고, 25개 세부과제들로 구성된 종합계획을 수립하여 3년간 이행하고자 함 <치안 AI 혁신 종합계획 주요 내용> ◦ (비전) AI와 함께 더 스마트한 경찰, 더 촘촘한 국민 안전 ◦ (목표) △ 국민 안전 △ 조직 혁신 △ 국익 창출 ◦ 전략 및 추진과제 ① 12대 선도프로젝트 추진 - (민생 분야) AI 기반 차세대 112시스템 구축, AI 기반 지능형 사회적약자 보호 종합플랫폼 개발, 차세대 실종자정보시스템 구축 - (수사 분야) 수사 분야 3개 시스템 고도화 ※ 경찰 수사지원AI(KICS-AI), 전기통신 금융사기 통합분석 대응시스템, AI 기반 과학수사 - (현장 분야) AI 순찰로봇, AI 순찰드론 등 도입 - (행정 분야) 모두의 경찰관(경찰민원24 연계) 및 AI 기반 업무지원시스템 도입 등 ② 치안 AI 정책 선순환 - (AI 기술) AI 기술을 총괄하여 신규 사업기획, 서비스 자체개발 등 지원 - (데이터/인프라) 치안데이터를 AI 활용에 적합하게 관리하고, 공동활용 인프라 구축
--

- (인력/교육) 치안 AI를 '전문가만의 기술'이 아니라 초 구성원이 활용하는 업무도구로 정착
 - (산업/생태계) 치안 AI의 '도입기관'에서 나아가 산업과 시장을 만드는 마중물 역할 수행
 - (법/제도) 예상되는 AI의 기술적·윤리적 위험에 대해 '사전관리체계' 구축
- ③ 지속가능한 추진체계 마련
- (조직역량 AI 내재화) 본청 內 △ 'AI 총괄기능'은 미래치안국 △ 'AI 사업부서'는 개별 부서 △ 'AI 정책심의·조정'은 데이터기반행정실에서 담당
 - (외부 거버넌스 구축) △ 핵심과제를 국가AI전략위에 의제화하여 정책 추동력 확보 △ 공공AX사업 참여 활성화 △ 개방형 혁신 방식의 AI 협업프로젝트 발굴 및 추진
 - (추진체계 이행력 확보) AI 추진체계를 확립하고, 윤리·신뢰성 확보 등 기본법 상 책무를 구체화하여 이행력 확보를 위한 경찰청 AI 훈령 제정

3. 12대 선도프로젝트 추진 현황

○ 현장 순찰, 실종자 수색 등에 AI 도입⁴⁾

6월부터 AI 카메라와 드론을 탑재한 순찰차가 서울 강서·양천·영등포·구로·금천 등 서울 서남권 지역에 시범 도입. 차량 상부 전면과 좌·우측에 총 3대의 AI 카메라가 촬영 영상을 실시간으로 분석해 인파 속에서 흥기를 소지한 사람이나 화재로 인한 연기를 자동으로 식별. '가방을 멘 사람'처럼 특정 조건을 입력해 객체를 찾아내는 것도 가능

차량 상부에서 출격하는 드론은 90배 줌과 열화상 기능을 갖췄으며, 사람과 차량 등 77종의 객체를 식별할 수 있는 AI 기능이 장착

올해 순찰리포트 자동 작성 기능을 갖추고, 내년에는 비명·고성 등 음성인식 AI 기능을 도입한 뒤 2028년부터 112상황실과 차량 AI 간 원격 연동 기능까지 단계적으로 고도화할 계획. 2028년 이후에는 사람 형태의 휴머노이드 로봇을 AI 차량에 구비해 AI가 위험하다고 판단 경우 현장에 로봇을 투입.



4) 차세대 미래치안 순찰 모빌리티의 등장 <https://www.youtube.com/watch?v=ub-F0ESgBzQ>
 매일경제, 드론이 찾은 테러범, 로봇이 제압 ...'AI 폴리스' 뜬다, 2026.04.08.,
<https://www.mk.co.kr/news/society/12011680>

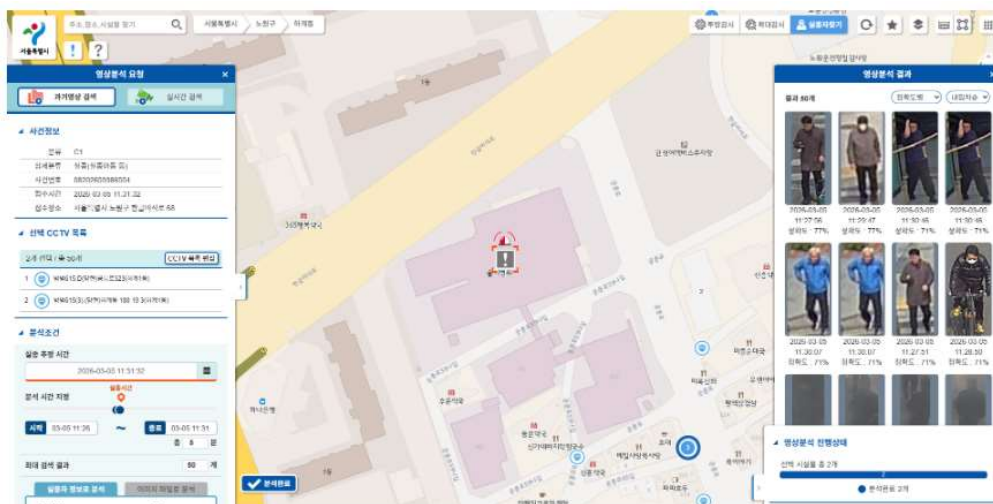
○ 서울 전역 약 12만 3천 대의 CCTV를 하나의 시스템처럼 검색할 수 있는 AI 기반 CCTV 고속검색시스템 올해 구축 완료 예정⁵⁾

AI 기반 CCTV 고속 검색 시스템은 실종 신고 시 인상착의, 이동 방향 등 객체 정보로 기반으로 CCTV 영상을 자동 분석해 대상자 이동 경로를 추적하는 영상 분석 장치

서울시는 서울경찰청 등 유관 기관과 협력을 통해 실종자 수색뿐 아니라 범죄나 사건 대응에도 활용 범위를 확대할 계획



고속검색시스템 개념도 및 서비스 흐름도



고속검색시스템 개념도 및 서비스 흐름도

○ 자율주행 AI 순찰 로봇

경기남부경찰청 분당경찰서는 성남시와 협력해 AI 자율순찰 로봇 4대를 서현역과 율동공원 등 치안 수요가 높은 지역 4곳에 투입해 운용 중.

수원팔달경찰서의 '팔달봇'은 비명소리나 몸싸움 등 이상행동이 포착되면 경고 방송을 송출하는 동시에 관제 센터와 인근 파출소에 실시간으로 신고.

고양경찰서는 '고양 폴리봇'을 운용. 보행 중 스마트폰 사용 자제, 보이스피싱 예방 안내 등의 홍보 방송을 송출하고 탑재된 카메라를 통해 112상황실에서 실시간 모니터링.

○ AI기반 차세대 112시스템 구축

AI 112 시스템 구축 과정에서 가장 중점을 두고 있는 목표는 신고 접수 직후 타 사건과의 연

5) 서울시, AI 기반 CCTV 고속검색시스템 구축 완료,

<https://scpm.seoul.go.kr/seoul-policy/evt0255>

계성 확인. 이를 위해 신고가 걸려온 전화번호 기반으로 구축되었던 데이터베이스(DB)를 신상 정보 중심으로 재편. 예를 들어 가정폭력 신고가 접수됐을 경우, AI가 거주자의 범죄 경력을 조회해 강력 범죄 전과자로 확인되면 현장에 출동하는 경찰관에게 방검복 등을 준비하도록 조언한다는 것.

112 신고가 접수되면 AI가 신고 내용을 자동으로 문자화한 후, 주변 소음과 목소리에서 감지되는 위기감, 신고자의 인적 사항 등을 분석해 현장의 위험도를 판단할 수 있도록 함. 외국어로 접수된 신고를 실시간으로 한국어로 번역할 수 있는 AI 기반 외국어 통역시스템 도입, 신고 접근성을 높이기 위해 기존 전화·문자 외 카카오톡, 라인 등 메신저를 활용해 112 신고를 할 수 있도록 신고 창구를 확대하는 방안도 검토.

도착 전까지 AI CCTV를 활용해 신고자 주변 환경을 실시간 영상으로 확인해 현장 상황을 사전에 최대한 파악할 수 있도록 하고 112 시스템에 챗GPT 등 생성형 AI를 도입해 경찰관이 현장에 도착하기 전까지 매뉴얼을 숙지할 수 있도록 함.

○경찰 수사지원 AI(KICS-AI)

지난해 11월부터 형사사법정보시스템(KICS)에 AI를 접목해 △문서 요약·번역·자료 검색용 대화형 AI △유사사건·판례 제공 △압수·수색·검증영상 신청서 초안 작성 등을 지원

2026년은 '수사 지원 AI(KICS-AI) 고도화 사업'으로 문서 작성 범위와 조사 기능을 확대할 예정. 이미지 문자 추출(OCR) 및 음성 인식 등 비정형 수사자료 분석 기능, 피의자·피해자·죄종·결정별 서식 초안을 작성, 수사 중인 사건의 유사 사례를 분석해 조사 질문 추천, 수사자료를 분석한 범죄일람표 작성 기능, 전국에 산재한 사건들의 계좌·전화번호·SNS 등을 비교 분석해 신종 범죄를 탐지하는 기능을 추가.

3단계(2027년) 사업은 3개년 사업의 최종 단계로 멀티모달 기능(영상분석 및 특정 장면 추출, 다수 이미지 내용 분석)과 영장신청서 완결성 검증, 통계원표 작성 등 수사의 완결성 확보와 종합적 자료 활용 체계 구축을 목표

○ 모두의 경찰관⁶⁾

모두의 경찰관은 경찰 민원 응대·접수·작성 지원과 경찰관의 민원 처리 업무를 인공지능(AI)으로 보조하기 위한 치안 특화 인공지능 서비스 개발 사업으로 경찰 민원 접수부터 문서 작성, 응대까지 전 과정을 AI로 지능화하는 게 목표. 올해 시범 운영을 거쳐 내년 전국 경찰관서로 서비스를 확대할 계획.

1년 차에는 민원답변 생성, 법령·매뉴얼 추천 인공지능(AI) 도우미 등 경찰관 업무지원 기능을 구축하고, 2년 차에는 국민이 직접 활용할 수 있는 경찰 민원 처리 인공지능(AI) 챗봇을 구현할 예정.

○ AI 통합 플랫폼 정보화전략계획(ISP)

AI의 효율적 운영을 위한 인프라 정비 사업으로 경찰 내부망 '폴넷' 기반 AI 통합 플랫폼 구축을 추진 중. 폴넷에 AI 검색과 문서 분석 기능을 도입하는 방안을 설계 중.

부서별로 구축돼 운영되던 그래픽처리장치(GPU)·거대언어모델(LLM)·광학문자인식(OCR) 등 인

6) 과학기술정보통신부와 정보통신산업진흥원(NIPA)이 주관하는 공공 인공지능전환(AX) 프로젝트 사업의 하나로 추진되며, 경찰청은 국민 체감 효과를 극대화하기 위해 경찰 민원 분야에 인공지능(AI) 기반 서비스를 우선 도입할 계획이다.

프라를 통합 관리할 수 있도록 체계를 만드는 사업. 플랫폼이 구축되면 판례 검색과 수사 자료 분석, 행정 문서 작성 등 내부 업무 전반을 AI로 지원하는 체계가 마련될 것으로 예상된다.

○ 경찰청 AI 훈령 제정⁷⁾

경찰청 인공지능의 개발 및 활용에 관한 규칙 [시행 2026. 6. 30.] [경찰청훈령 제1209호, 2026. 6. 30., 제정] 제1장 총칙 제1조(목적) 이 규칙은 「인공지능 발전과 신뢰 기반 조성 등에 관한 기본법」 및 같은 법 시행령 등에 따라 경찰청 및 「경찰청과 그 소속기관 직제」 제2조에 따른 소속기관이 치안 분야의 인공지능을 체계적이고 안전하게 개발·활용하기 위한 기준을 제시함을 목적으로 한다. 제3조(경찰기관의 책무) ① 경찰기관은 인공지능의 개발·활용 과정에서 사람의 생명·신체의 안전 및 기본권을 보호하고, 공공의 안녕과 질서 유지에 기여하도록 노력해야 한다. ② 경찰기관은 인공지능의 안전성과 신뢰성을 확보하고, 오작동이나 오류 등으로 인한 피해를 예방하기 위하여 인공지능을 지속적으로 관리해야 한다. ③ 경찰기관은 공정성과 책임성의 원칙에 따라 인공지능을 개발·활용해야 하며, 특정 개인이나 집단에 대한 차별이나 편향이 발생하지 않도록 해야 한다. ④ 경찰기관은 인공지능의 개발·활용 과정에서 치안 목적 수행에 필요한 최소한의 범위에서만 개인정보를 수집·이용해야 한다. 제4조(적용 범위) 이 규칙은 경찰기관이 수행하는 모든 인공지능사업에 적용한다. 제5조(기본법의 적용 제외 인공지능 지정) ① 기본법 제4조제2항 및 같은 법 시행령 제2조제3호에 따라 기본법의 적용 제외 인공지능 업무를 직접 기획·수행하는 부서의 장(이하 "소관부서의 장"이라 한다)은 경찰청장에게 기본법의 적용 제외 인공지능 지정을 요청할 수 있다. ② 경찰청장은 적용 범위의 적절성 등을 검토하여 기본법의 적용 제외 인공지능을 지정할 수 있다. ③ 소관부서의 장은 기본법의 적용 제외 인공지능 지정 사유가 해소된 경우 경찰청장에게 지정 해제를 건의해야 한다. 제2장 인공지능 정책 추진체계 제3장 인공지능사업 관리 제4장 인공지능윤리 및 신뢰성 확보 제18조(인공지능 투명성 확보) ① 인공지능사업부서의 장은 고영향 인공지능이나 생성형 인공지능을 이용한 시스템을 운영하려는 경우 기본법 제31조에 따른 투명성 확보 조치를 이행해야 한
--

7) 수사관들이 실무에서 지켜야 할 일종의 행동 강령인 'AI 윤리준칙' 제정도 추진

다.

② 제1항에 따른 투명성 확보 조치는 다음 각 호의 어느 하나에 해당하는 경우에는 적용하지 않는다.

1. 인공지능시스템의 명칭, 이용자 화면에 표시된 문구 등을 고려할 때 고영향 인공지능 또는 생성형 인공지능 기반 운용 사실이 명백한 경우
2. 경찰기관 내부 업무 용도로만 사용되는 경우
3. 그 밖에 인공지능시스템의 유형·특성이나 결과물의 내용, 이용형태 및 기술 수준 등을 고려하여 기본법 제31조제1항부터 제3항까지 중 전부 또는 일부에 대한 적용 예외가 필요한 사항으로 과학기술정보통신부 장관이 정하여 고시하는 경우

제19조(고영향 인공지능의 확인) ① 인공지능사업부서의 장은 인공지능시스템을 운영하려는 경우 그 인공지능이 고영향 인공지능에 해당하는지에 대하여 사전에 인공지능총괄부서의 장에게 서면 또는 전자적 방식으로 확인을 요청해야 한다.

② 인공지능총괄부서의 장은 다음 각 호의 사항을 고려하여 고영향 인공지능 해당 여부를 판단해야 한다.

1. 인공지능이 기본법 제2조제4호 각 목의 어느 하나의 영역에서 활용되는지 여부
2. 사람의 생명·신체의 안전 및 기본권에 초래할 수 있는 위험의 영향, 중대성, 빈도 및 활용 영역별 특수성

③ 인공지능총괄부서의 장은 고영향 인공지능 해당 여부 판단을 위하여 필요한 경우 외부 전문가의 자문을 받을 수 있다.

④ 인공지능총괄부서의 장은 제2항에 따라 고영향 인공지능 해당 여부를 판단하기 위하여 필요한 경우 기본법 제33조제1항에 따라 과학기술정보통신부장관에게 고영향 인공지능 해당 여부의 확인을 요청할 수 있다. 이 경우 인공지능사업부서의 장은 기본법 시행령 제25조제1항 각 호의 서류를 인공지능총괄부서의 장에게 제출해야 한다.

⑤ 인공지능총괄부서의 장은 경찰기관 내 고영향 인공지능의 목록을 정리하여 별도로 관리해야 한다.

제20조(고영향 인공지능 안전성·신뢰성 확보조치 등) ① 인공지능총괄부서의 장은 기본법 제34조제1항제1호에 따른 경찰청 위험관리방안의 수립·운영 업무를 총괄·지원한다.

② 인공지능사업부서의 장은 고영향 인공지능의 안전성·신뢰성을 확보하기 위해 다음 각 호의 업무를 수행한다.

1. 기본법 제34조제1항제1호에 따른 위험관리정책 수립(최신 기술 및 관리방법론이 적용될 수 있도록 주기적으로 갱신해야 한다)
2. 기본법 제34조제1항제2호에 따른 설명 방안의 수립·시행
3. 기본법 제34조제1항제3호에 따른 이용자 보호 방안의 수립·운영
4. 기본법 제34조제1항제4호에 따른 고영향 인공지능에 대한 사람의 관리·감독

③ 인공지능사업부서의 장은 제2항의 조치를 이행하고 그 근거를 문서로 작성하여 5년간 보관해야 한다.

④ 인공지능사업부서의 장은 제3항에서 작성한 문서를 인공지능총괄부서의 장에게 서면 또는 전자적 방식으로 제출해야 한다.

제21조(공공분야 인공지능 영향평가의 실시 등) ① 경찰청장은 인공지능사업을 수행하기 전에 「인공지능 및 데이터 기반 행정 활성화에 관한 법률」 제32조제1항에 따라 공공분야 인공지능 영향평가를 실시하고 그 결과를 공표해야 한다.

② 제1항에도 불구하고 경찰청장은 도입하려는 인공지능이 「인공지능 및 데이터 기반 행정 활성화에 관한 법률」 제32조제2항제1호부터 제3호까지의 어느 하나에 해당하는 경우로서 행정안전부장관과 협의한 경우와 같은 항 제4호에 해당하는 경우에는 공공분야 인공지능 영향평가를 실시하지 않을 수 있다.

③ 인공지능사업부서의 장은 제1항에 따른 영향평가를 수행하여 그 결과를 서면 또는 전자적 방식으로 인공지능총괄부서의 장에게 제출해야 하며, 인공지능총괄부서의 장은 이를 검토하여 주요 내용을 경찰청장에게 보고해야 한다.

④ 경찰청장은 필요하다고 판단하는 경우 「인공지능 및 데이터 기반 행정 활성화에 관한 법률」 제21조에 따른 전문기관 또는 같은 법 시행령으로 정하는 기관에 공공분야 인공지능 영향평가의 실시를 의뢰할 수 있다.

4. 자치경찰의 범죄예방 사업⁸⁾

○ 동국대와 서울시 자치경찰위원회의 '인공지능(AI) 활용 실종 아동 수색 및 환경설계를 통한 범죄예방(CPTED) 사업 지원 플랫폼'

인공지능 분석 기술을 활용해 학교 주변에서 발생할 수 있는 아동의 이상행동을 미리 감지해 범죄를 예방하고 아동 실종상황 발생 때 CCTV 등에 촬영된 흔적을 신속하게 탐지해 아동을 조기에 발견할 수 있는 기술을 개발

서울 시내의 공간 정보, 범죄 발생 통계 등을 통합 분석해 범죄를 예측하는 모델을 개발해 서울 경찰의 범죄예방 사업을 지원할 예정

○ 전주대와 전북 자치경찰위원회의 '순찰 로봇 및 영상 센서 활용한 안전관리체계 구축'

전주천 산책로 주변의 범죄를 예방을 위한 야간 적외선 감지, 인공지능 기술을 활용한 이상행동 탐지 등 기능을 개발해 자율순찰 로봇에 적용·운영. 사람이 갑자기 쓰러지거나 특정인을 강제로 끌고 가는 행동, 비명 등 이상 상황이 감지되면 즉시 관제센터로 영상과 경고 알림이 전송.

공중화장실 불법촬영 예방 시스템은 화장실 칸막이 위로 손이나 휴대전화가 올라오는 움직임 등 이상 행동이 포착되면 경고음이 울리고, 필요시 112 신고 체계와 연동.

5. 그 밖의 AI 활용 치안 기술

○ 성착취물 탐지 및 증거서류 작성 자동화 프로그램

경찰청 국가수사본부는 인공지능 기술을 활용한 '성착취물 탐지 및 증거서류 작성 자동화 프

8) 경찰청은 과학치안진흥센터와 협력해 자치경찰 연구개발 사업을 진행. 2022년부터 2027년까지 국비 107억 원을 들여 지역별 연구기관, 자치경찰위원회, 시·도 경찰청 등과 협력해 지역 특성에 따른 치안문제를 발굴하고 이를 해결할 수 있는 과학기술을 개발

로그램' 자체 개발을 완료하고, 전국 시도경찰청 및 경찰서 수사관들을 대상으로 배포. 인공지능 객체 탐지 기술(이미지나 영상파일에서 무엇이 어디에 있는지를 탐지하는 기술)과 백신 프로그램의 초고속 파일 검색 기술(사용자가 원하는 파일을 밀리초 단위로 찾아내는 기술)을 결합해 제작.

총재생 시간 403시간 분량의 동영상 4,215개(890기가바이트)를 탐색하여 성착취물 여부를 판단 하고 증거문서 형식으로 현출하는 작업이 기존에는 약 320시간이 소요되었으나, 본 자동화 프로그램을 사용하면 3시간 만에 완료

○ 관계성범죄 대응

현재 분산 관리 중인 가해자·피해자 데이터를 하나의 시스템으로 통합하고, 축적된 정보에 대한 분석을 기반으로 재범 위험성을 평가·감지하는 인공지능(AI) 기반 '사회적약자보호 종합플랫폼' 개발 중.

고위험 스토킹 가해자로부터 피해자를 선제적으로 보호하고 현장 대응력을 강화하기 위해 “경찰청-법무부 간 스토킹 잠정조치로 위치추적 전자장치를 부착한 가해자 위치를 등 실시간으로 위치추적·대응하는 시스템 구축 사업을 추진. 법무부 위치추적 관제센터에서 가해자 접근 등 위험경보 이관 시 경찰에 자동 접수·지령, 경찰은 출동 현장에서 가해자의 실시간 이동 경로를 확인하며 대응. 관제요원은 인근 CCTV로 영상을 확인.

○ 한국전자통신연구원(ETRI) '데자뷰(Dejaview)' 기술 개발

CCTV 영상, 범죄통계정보, 측위정보 등을 분석해 과거 범죄가 발생한 상황과 현재 진행 상황을 비교·분석해 범죄 위험도를 측정하고 예측하는 인공지능(AI) 기술.

적용 대상 및 방법에 따라 두 가지 기술로 구분할 수 있는데, 시·공간 중심 범죄예측 기술은 특정 장소, 특정 시간대에 어떤 유형의 범죄 발생 위험도가 높은지 분석. 과거 범죄 등의 통계정보를 바탕으로 학습된 AI가 실시간 CCTV 영상을 자동 분석해 범죄 상황과 유사도를 비교·측정. ETRI는 서울 서초구와 공동으로 지난 2018년부터 3년간 지역 내 32,656건의 CCTV 사건·사고 빅데이터를 분석해 범죄예측지도(PCM)를 개발함. PCM이 실시간 범죄 위험도를 화면에 제시해 관제사가 범죄예측지도를 통해 현재 지역별 범죄 위험도를 한눈에 파악할 수 있도록 함. 향후 경찰과 전국 지방자치단체 CCTV 통합관제센터 등에서 주로 활용될 전망.

두 번째 데자뷰 기술은 개인 중심의 재범 예측. 재범 우려가 큰 고위험군 전자감독대상자의 이동 패턴에 따른 위험도를 분석함. 전자감독대상자가 생업 등 사유로 인해 주기적으로 이동 제한 규정을 위반하면 인공지능 기술을 통해 이를 분석, 재범 위험도를 파악해 대응할 수 있도록 함.⁹⁾

6. 검토되어야 할 문제

○ 고영향 AI 판단 기준의 협소함

9)

https://www.etri.re.kr/kor/bbs/view.etri?keyField=&keyWord=&nowPage=6&b_board_id=ETRI06&year_gubun=&b_idx=19312

ETRI가 개발한 범죄 및 위험 예측 시스템 -2024년 제작 영상https://youtu.be/bx_v6lv8CLc

발제자의 「고영향 인공지능 판단가이드라인」에서의 요건에 대한 문제의식에 동의하며, 범인 수사 및 체포와 더불어 예측적 경찰활동도 고영향 AI에 포함되어야 하지 않는가라는 문제의식이 있음. 예측 치안을 위한 정보수집과 결합, 알고리즘 예측에 따른 경찰활동, 실시간 감시와 추적 등의 문제가 기본권과 밀접하게 연결되기 때문임.

경찰은 내부 점검을 통해 고영향 AI 해당 여부를 점검했는데 관련 가능성이 있는 과제가 21개로 분류됨. 이후 과학기술정보통신부의 고영향 AI 판단 가이드라인을 적용하는 과정에서 범 죄수사 목적의 생체인식 정보를 활용하는 AI와 관련된 2개 사업이 검토 대상으로 좁혀졌음.¹⁰⁾ 이는 고영향 인공지능 판단 요건이 협소하다고 해석할 수 있음.

○ 경찰청 AI 훈령에 대한 비판적 검토

경찰청 AI 훈령의 목적이 '치안 분야의 인공지능을 체계적이고 안전하게 개발·활용하기 위한 기준을 제시'하기 위한 것임을 밝히듯이 시민의 기본권 침해를 방지하기 위한 제도적 장치는 없음.

경찰기관의 책무는 명시되어 있지만 구체적으로 어떻게 실현할지와 책무를 위반했을 때의 조치, 업무 수행 과정에서 AI 활용 빈도가 높아질수록 의존 현상이 발생할 가능성이 있는데 과의존을 예방하기 위한 장치, 경찰이 최종판단을 해야하는 구조적 설계, 오류가능성에 대한 대책과 문제 점이 지적됐을 때 누가 책임을 부담해야 하는지 등 구체적인 문제와 관련해 훈령으로는 해결 되지 않음.

이는 인공지능기본법을 전제로 했기 때문에 한계가 있을 수도 있지만 민감한 정보를 다루고 경찰작용 자체가 기본권 침해에 영향이 큰만큼 치안영역에서의 인공지능은 제도적 보완이 필요할 수 있음.

○ 정보의 결합 추구하는 규제 제거 경향성에 대한 대응

초기의 치안 시스템은 경찰 빅데이터 기반의 패턴 분석이 대부분이었지만, 기술 발전과 함께 범죄·사건 정보에 실시간으로 센서나 웹 등에서 수집되는 데이터를 활용하여 특정 범죄나 위험 상황에 대해 예측할 수 있게 됨. 분석의 정확성을 높이기 위해 범죄 발생과 연결될 수 있다고 판단하는 가능한 많은 데이터 소스를 통합하려는 경향이 커짐. 소득 분포, 가구 규모, 건물 수, 교통망, 유흥 장소 등 특정 지역의 기반 시설과 사회경제적 특성뿐만 아니라 교통 데이터, 대중교통 이용 패턴, 날씨 변화, 소셜 미디어 활동 등 도시 내 인구 이동과 관련된 범죄 위험을 동적으로 모델링함.

경찰은 AI 기능에서 중요한 역할을 하는 데이터가 정부 부처별로, 부처 내에서는 각 부서별로 제각각 나뉘어 처리되고 있는 '데이터 사일로 현상' 해소를 향후 과제로 봄. 데이터 공유를 막는 규제 개정을 강조함.¹¹⁾ '공공의 안녕'이나 보다 정확한 예측을 이유로 데이터의 공유와 결합을 요구하는 것을 제어하기 위해 시민의 권리로서 어떤 통제가 가능할지 논의가 되어야 함.

10) 뉴시스, 현장 AI 과의존 우려에 내부통제 속도… 'AI 훈령' 제정, 2026.03.16., https://www.newsis.com/view/NISX20260313_0003547968

11) 아시아투데이, "명실상부 경찰 AI 컨트롤타워 될 것", 2026.04.19., <https://news.nate.com/view/20260419n13640>

토 론 문

민변 디지털정보위원회 위원장 최호웅 변호사(법무법인 덕수)

1. 지자체 지능형 CCTV 관제시스템에 대한 감사원 감사결과('26년 1월)

가. 감사결과 요지

- ▶ 감사원이 공개한 '인공지능 대비 실태' 감사 보고서에 따르면, 2024년 공공기관이 쓰는 AI 625개 가운데 228개(36.5%)가 CCTV카메라 영상을 분석해 이상 상황을 감지하는 '지능형 CCTV' 관련 AI였음
- ▶ 서울시 11개 자치구를 대상으로 한 지능형 CCTV의 탐지 정확도는 12~52% 수준
 - 소화전을 배회하는 사람으로 오인
 - 난간에 기대어 있는 사람을 침입 시도로 오인
 - 엎드린 개를 쓰러진 사람으로 오인
 - 팔짱을 끼거나 어깨동무한 사람들을 싸우는 것으로 인식

나. 지자체 CCTV 관제시스템에 대한 경찰 접근

- ▶ 지자체는 자신이 운영하고 있는 CCTV 통합관제센터에 경찰의 접속을 개인정보 보호법 제18조 제2항 및 경찰관 직무집행법 제8조의2(정보의 수집 등)에 따라 허용하고 있음
 - 접속사유는 순찰차영상제공, 교통사고, 소란, 절도 등 범죄, 위험방지, 화재 등임
- ▶ 감사원 결과 확인된 지능형 CCTV의 오류는 경찰 출동이나 불심검문, 추적 등 공권력 행사로 연결될 경우 정상적인 시민을 위험인물로 잘못 분류하여 개인정보자기결정권과 일반적 행동자유권을 침해로 연결될 수 있는 현실을 보여 주고 있음. 발제자의 지적처럼 수사기관이 개입할 수 있는 '구체적 위험성' 판단의 근거로 인공지능 분석을 활용함에 있어, 분석 오류의 위험성을 명심해야 할 것임
- ▶ 또한, 지능형 CCTV 문제를 단순한 학습데이터 부족이나 성능개선의 문제로 한정해서는 안 된다고 할 것임
 - 경찰 활동에 적용에 있어 인권영향평가, 탐지항목과 경찰연계의 범위의 법률상 제한, 잘못된 탐지에 대한 기록·감사 및 피해를 받은 영향받은 자에 대한 권리구제절차가 마련되어야 함

2. 현행 법률 규정만으로는 경찰 AI를 정당화 할 수 없음

- ▶ 발제자도 밝힌 바와 같이 개인정보 보호법 제15조 제1항 제3호 “소관 업무의 수행”이 인공지능시스템을 통한 경찰 직무 집행의 독립적인 수권조향으로 볼 수는 없다고 봄
- ▶ 경찰의 감시·추적·프로파일링은 개인정보를 단순히 처리하는 문제에 그치지 않음, 국가가 국민을 관찰하고 분류하며, 이를 근거로 불심검문·수사추적 등 권력적 조치를 할 수 있는가의 문제임. 따라서 일반 개인정보 처리 근거와 경찰권 행사 근거를 구분해야 할 것임
- ▶ 경찰 AI 활용에 대한 수권조향은 그 대상, 목적, 데이터 종류, 분석 방식, 사용 가능한 결과, 보유기간, 외부 제공, 사후통제까지 개별 법률(경찰관직무집행법 등)에 구체적으로 규정되어야 할 것임
 - AI 시스템을 사용해서는 안 되는 영역인지
 - 덜 침해적인 수단으로 목적을 달성할 수는 없는지
 - 오차와 차별이 통제되는지
 - 결과를 설명하고 대상자가 이를 다룰 수 있는지

3. 경찰 AI 활용에 있어 규제 제도 방향

가. 인공지능기본법 상 ‘고영향 인공지능’에 경찰 AI 활용 영역 확대

- ▶ 현행 인공지능기본법 상 경찰 직무 중 ‘범죄 수사나 체포 업무를 위한 생체인식정보의 분석·활용’에 한하여 ‘고영향 인공지능’ 영역으로 정의하고 있음
- ▶ 범죄 예측, 재범위험성 평가 등 역시 시민의 기본권에 중대한 영향을 미치거나 위험을 초래가 우려가 있는 영역임. 따라서 이러한 경찰 직무 수행을 위하여 인공지능 시스템을 활용하는 경우에도 인공지능기본법 상의 ‘고영향 인공지능’ 운용자의 책무를 부과해야 할 것임

나. 경찰 AI 활용에 의하여 영향받는 자에 대한 권리 보장

- ▶ 투명성 확보를 위한 경찰 AI 시스템 등록 및 공개
- ▶ 도입 전 인권영향평가와 개인정보 영향평가
- ▶ 경찰 AI 시스템에 대한 독립기관의 사전 심사와 통제 구조 마련
- ▶ 영향받은 자에게 AI가 경찰조치에 사용된 사실의 통지
- ▶ 영향받은 자에게 주요 판단요소와 데이터 출처에 대한 설명요구권 보장
- ▶ 영향받은 자에게 인간에 의한 재검토 요구권 보장
- ▶ 영향받은 자에게 오류 정정과 삭제청구권 보장
- ▶ 위법한 AI 시스템 사용으로 취득한 증거능력 배제

나. 집회·시위 영역은 별도의 금지 또는 고강도 규제

- ▶ 집회 참가자의 얼굴·동선·관계망·발언을 AI 활용을 통하여 분석하는 것은 개인정보자기결정권 뿐만 아니라 집회·결사의 자유와 표현의 자유에 직접적인 위축효과를 발생시킴. 치안을 위한

것이라는 일반적인 이유만으로 집회 참가자 전체를 분석하는 것에 대하여는 엄격한 제도적 제한이 필요할 것임

▶ 특히 집회·시위 영역에서 다음 행위는 원칙적 금지 대상으로 해야 한다는 의견임

- 집회 참가자에 대한 실시간 원격 생체인식
- 정치적 견해나 성향의 추론
- 집회 참가자 간 관계망 분석
- 장래의 불법행위 가능성으로 분석결과 반영
- 집회 참가 이력을 다른 경찰 데이터와 결합하는 행위

▶ EU AI Act 제5조 (1)항(h)는 경찰 등 법집행기관이 공개적으로 접근 가능한 장소에서 실시간 원격 생체식별시스템을 사용하는 것을 원칙적으로 금지하고 있음.

한국 이민행정의 AI의 전면적용에 관한 우려와 과제

-GCM 23가지 목표에 비춰 본 이민행정과 AI(인공지능) 적용의 의미

-고기복(모두를 위한 이주인권문화센터 대표)

이주인권 영역에서 가장 포괄적인 국제협력 체계라고 할 수 있는 ‘이주에 관한 글로벌 콤팩트(GCM)’이 채택된 시기(2018년)에는 생성형 AI나 고도화된 인공지능 기술이 지금처럼 전면에 대두되지는 않았다. 하지만 현재 각국 정부는 출입국 관리의 효율성, 안보와 신속한 절차를 강조하며, AI 기술을 적극적으로 도입하는 명분(법적·정치적 근거)으로 GCM 23가지 목표를 들고 있다. 이들 목표는 안전하고, 질서 있으며, 정규적인 이주를 위해 국경 관리가 단순히 통제의 개념을 넘어, 합법적인 여행객과 이주민에게는 신속하고 원활한 절차를 제공하면서도 보안 위협에는 효과적으로 대응해야 한다는 원칙을 담고 있다. 이러한 원칙 아래에서 인공지능(AI)은 국경관리와 이민행정에 있어서 효율성과 정확성을 크게 향상시킬 수 있는 강력한 도구다.

문제는 AI를 이용함에 있어서 아직까지 법적·윤리적 한계와 기술 남용이나 불공정한 결과가 발생할 위험이 여전한데도, 효율성이라는 미명하에 ‘이주민 감시를 당연하게 여기는 분위기’가 각국 이민행정부에 만연하며, AI를 활용한 각종 통제 시스템이 광범위하게 운영되고 있다는 점이다. 각국 정부는 데이터 과학, 머신러닝, 자동화 등 AI 기술을 실제 이주 정책과 국경 관리에 깊숙이 활용하며 ‘AI 첨단 기술을 활용한 행정 효율화’라고 포장하고 있다. 하지만, 이것이 기술을 이용한 감시와 배제의 자동화가 아닌지 살펴 볼 필요가 있다.

GCM 23가지 목표 달성을 위해 AI는 ‘데이터 분석(목표 1)’, ‘정보 제공(목표 3, 6)’, ‘효율적 국경 감시(목표 11)’, ‘신속한 행정 심사(목표 12)’라는 강력한 도구로써 각국 이민행정에 이미 깊게 관여하고 있다. 관련 목표들의 지향점과 정부 논리를 비판적으로 살펴보면 다음과 같다.

■ GCM 목표 1: 증거 기반 정책의 기초로서 정확하고 세분화된 데이터를 수집하고 활용한다.

▶ 정부 논리: AI 알고리즘이 정교하게 작동하기 위해서는 대규모의 고품질 데이터(빅데이터) 축적이 필수적이다. 정부는 이 조항을 근거로 이주민의 체류 자격, 동선, 생체 정보 등을 광범위하게 수집·통합하여 ‘데이터 기반의 과학적 이민 행정’을 펴겠다고 강조한다. (법무부 『2030 이민정책 미래전략』의 핵심 기조이기도 하다.

▶ 비판: 데이터 수집의 목적구속 원칙이 무너지고 있다. 출입국 심사를 위해 수집한 생체정보를 AI 알고리즘 개발이라는 전혀 다른 목적으로 유용하고 있으며, 축적된 데이터가 이주민을 상시 감시하고 ‘잠재적 범죄자’로 인종 프로파일링하는 차별의 도구로 재생산되고 있다.

■ GCM 목표 11: 통합적이고 안전하며 조율된 방식으로 국경을 관리한다.

▶ 정부 논리: 출입국 관리 및 국경 안보의 효율성을 강화하기 위해 안면인식 기술, 생체정보 식별추적 시스템, 자율주행 드론 국경 감시 등을 도입하여 국경 검문 과정을 신속화하고 ‘불법 체류자 및 위험 인물’을 효율적으로 차단할 수 있다.

▶ 비판: 이 목표를 이행한다는 명목으로 한국 정부는 ‘법무부 AI 식별추적시스템 사건’을 일으켰다. 1억 7천만 건의 내·외국인 안면 데이터를 정보주체 동의 없이 민간 AI 학습용으로 넘긴 초법적 행위를 ‘안전한 국경 관리’라는 명분으로 정당화한 것이다.

■ GCM 목표 12: 적절한 스크리닝, 평가 및 연계를 위해 이주 절차의 확실성과 예측 가능성을 강화한다.

▶ 정부 논리: 이주민이나 입국 승객이 국가에 미칠 '위험도'를 사전에 스크리닝(국경 진입 전 필터링)하여 행정의 예측 가능성을 높이는 부분이다. K-ETA(전자여행허가제)는 데이터를 기반으로 승객의 위험도를 자동 분석·평가하여 입국 허가 여부를 사전에 가려내는 알고리즘 심사다.

▶ 비판: ‘확실성과 예측 가능성’이라는 표현과 달리, AI를 통한 자동 스크리닝은 철저히 베일에 싸인 ‘블랙박스 통제’가 된다. 위험도 산정 기준, 불허 결정의 구체적 사유 고지나 인간 심사관에 의한 재심사(Human in the loop) 기회가 박탈되어 적법절차가 형해화된다.

GCM의 정신은 본래 목표 7(이주 과정의 취약성 감소)이나 목표 17(차별 철폐)처럼 ‘인권 보장’에 방점이 찍혀 있다. 그러나 한국 정부는 GCM 이행 보고서나 2030 미래전략을 통해 목표 1, 11, 12(데이터·국경관리·스크리닝)만을 철저히 ‘국가 안보와 행정 편의주의적 효율성’의 관점으로만 편향되게 해석하여, 국내 이민 행정에 통제형 고위험 AI 기술을 무제한 도입하는 면죄부로 악용하고 있다.

한국 이민행정(2030 미래전략)과 AI

한국의 경우, 생체인식에 기반한 자동출입국심사, 사전 여행자 정보 공유 시스템, 전자여행허가제(K-ETA) 시행 등 국경관리에서 AI 시스템을 적극적으로 이용하고 있다. 법무부 출입국은 AI 기술 활용으로 입국 심사장에서의 대기 시간을 줄이고, 사전에 입국 부적격자를 걸러내는 국경 관리 시스템을 자랑하고 있다. 하지만, 높은 정확도를 자랑한다 할지라도 생체인식과 안면인식 기술은 개인의 프라이버시를 침해할 가능성이 있다는 점에서 개인정보 보호와 인권 침해의 위험을 동반한다. 특별히 AI 시스템의 효율성을 높이기 위해서는 방대한 데이터가 필요하다는 점에서 개인의 동의 없는 무분별한 정보 수집 가능성이 높다는 점에서 더욱 그러하다. 무엇보다 입국을 목적으로 하는 외국인이 목적국의 생체정보 요구를 무시할 수 없다는 점에서 이는 강제성을 갖고 있다고 봐야 한다.

또한, 이민행정에서 AI 자체의 편향과 그에 따른 알고리즘이 내리는 결정은 출입국의 누적된 자료¹⁾에 근거한다는 점에서 인종과 국적 등에 대해서도 편향적일 수밖에 없다. 그런데 이주민은 AI의 의사결정 과정에서 편향이나 알고리즘 오류(예: K-ETA 불허)에 직면해도 구체적 거부 사유조차 알지 못한 채 불공정한 결과를 수용할 수밖에 없다.

1) 과거부터 출입국이 수집해 온 출입국 심사, 체류 관리, 법 위반 단속 등과 관련된 모든 역사적 데이터(Historical Data)를 의미한다. 강제퇴거, 미등록 체류(불법체류) 기록 등이 특정 국적이나 특정 체류자격(비자)에 편향되어 쌓여 있는 현상과 직결된다. AI는 누적된 데이터에 특정 국적이나 인종에 대한 부정적 기록이 많다면, 이를 바탕으로 ‘해당 국적/인종의 입국은 위험도가 높다’는 규칙(알고리즘)을 스스로 도출하여, 입국 신청자의 국적, 나이, 직업 등이 과거 ‘누적된 자료’ 속 미등록체류 고위험군 데이터와 유사하면 자동으로 ‘불허’ 결정을 내린다.

CCTV나 드론 등을 활용한 공공장소에서의 상시 감시는 일반 시민에게 사생활 침해이다. 하지만, 이것이 이민 행정에 도입되면 특정 국적이나 인종을 잠재적 범죄자로 분류하여 특정 집단을 체계적으로 배제하는 ‘위험평가·프로파일링’으로 변질될 수 있다. 그렇다면, 효율성을 강조하는 출입국관리와 사법심사 등의 이민행정에서 AI로 인한 피해를 어떻게 예방할 것인가? 다국어 챗봇이나 통·번역 지원 등 이주민 권리 옹호를 돕는 ‘저위험 AI’는 적극 촉진하되, 인간을 감시하고 프로파일링하는 ‘고위험 AI’는 철저한 법적 통제 하에 두는 방향으로 정책을 지향하는 것이 바람직하다.

유감스럽게도, 지금까지 법무부가 보이고 있는 이민행정 방향을 보면, 고위험 AI의 위험성을 지나치게 간과하고, 낙관적으로만 보는 것 같다. 법무부는 2026 국제이주검토포럼(IMRF)에 보낸 보고서에서 ‘2026년 GCM 이행 공약2’에 “데이터 기반의 과학적 이민 행정 및 대민서비스 통합 플랫폼 구축”을 명시하고, 주요 추진 과제로 1) 외국인 행정정보 종합플랫폼 구축(2030) 2) 지자체 맞춤형 외국인 행정 데이터 분석 및 제공 3) AI 기반 출입국 심사 및 정보 분석 시스템 도입(2030) 4) 지자체 수요를 반영한 외국인 행정데이터 분석 및 제공 5) 빅데이터에 기반한 외국인 적정유입 규모 산정 및 경제효과 측정(2030) 6) 출입국 이민행정 대민포털 개편사업 추진을 제시했다

한국 정부는 최근 법무부의 『2030 이민정책 미래전략』과 ‘AI 기반 이민정책·행정 혁신 발표회’, ‘2026 GCM 이행 공약’ 등을 통해 AI 전면 도입을 예고하고 있으나, 한국 정책의 현실은 이주민을 ‘고위험군’으로 낙인찍고, 디지털 차별 자동화하고 있어 ‘규율 공백 속의 폭주’와 같다.

‘AI·빅데이터·임금자문위’로 포장된 선별·통제 정책

“고위험 외국인은 차단하고, 외국인 고용 성실기업은 우대합니다”.

법무부는 AI·빅데이터·생체정보를 활용해 “고위험 외국인”을 신속히 분류·입국 차단하고, 성실기업에는 인센티브를 제공하는 ‘K-Trust 기업 인증제’를 도입하겠다고 밝혔다. 이는 국가와 기업이 결합한 감시·선별 체제를 정당화하는 장치로, 기업 편의에 맞는 ‘순응형 노동자’만을 선별하는 구조를 강화할 위험이 있다.

더 나아가, AI와 생체정보를 이용해 “고위험 외국인”을 신속하게 분류·입국 차단하겠다는 계획과, 성실 고용 기업에게 인센티브를 부여하는 ‘K-Trust 기업 체류·고용 인증제’는, 국가와 기업이 결합한 감시·선별 체제를 정당화하는 장치가 될 수 있다. 성실 고용기업에 인센티브를 준다는 ‘K-Trust 기업 체류·고용 인증제’는 국가와 자본이 손잡고 이주민을 철저히 관리·선별하는 감시망이 될 수밖에 없다. 기업이 원하는 ‘말 잘 듣는 노동자’만 통과시키는 구조는 노동권 침해와 부당해고를 오히려 정당화하며, 이주민의 존엄을 짓밟게 된다. 국가가 감독 책임을 기업의 자율과 평판에 떠넘기는 이 제도는 이주노동자를 ‘인력 자원’이 아닌 ‘관리 대상’으로 고착화할 뿐이다. 이처럼 K-Trust 기업 체류·고용 인증제는 인종·국적·언어·종교 등에 따른 편견이 데이터·알고리즘에 내재화될 위험, 기업 편의에 맞춘 “순응형 노동자”만 선별되는 위험은 전혀 성찰하지 않고 있다.

법무부는 이 항목을 GCM 이행공약에도 ‘데이터 기반의 과학적 이민 행정 및 대민서비스 통합 플랫폼 구축’이라는 제목으로 제시했다. 이민정책에 AI와 생체정보로 “고위험 외국인”을 신속 분류·입국 차단하겠다고 명시한 것은 이주민 전체를 잠재적 위협으로 간주하는 적대적 시각의 결정판이다. 인종·국적·언어·종교에 뿌리박힌 편견이 데이터와 알고리즘에 고착화되면 차별은 더 이상 인간의 판단이 아니라

기계의 ‘객관적’ 결과로 위장된다. 이는 GCM의 인권·비차별 원칙을 정면으로 위배하며, 유럽이 생체정보 남용에 엄격한 금기를 세운 이유를 전혀 반성하지 않는 처사이다.

법무부가 내세우는 디지털 비자 신청 및 AI 기반 이민행정 서비스의 확산은 언뜻 현대적이고 효율적인 시스템처럼 보이지만, 그 속내를 들여다보면 행정의 본질적 대상인 ‘사람’은 사라지고 ‘기업 편의’만 남은 주객전도의 형국이다. 출입국 관리와 이민행정에서 마땅히 우선이어야 할 이주민은 정작 디지털 고도화의 주체가 아닌, 관리와 통제의 효율성을 높이기 위한 데이터의 파편으로 전락했다.

2030 이민행정 미래전략은 인력이 필요한 기업이 얼마나 더 빠르고 간편하게 외국인력을 수급할 수 있느냐에 초점이 맞춰져 있다. 행정 서비스의 디지털화가 이주민 개인의 권리 구제나 체류 안정성을 높이는 방향이 아니라, 기업의 인력난 해소를 위한 ‘속도전’에 동원되고 있는 것이다. 이는 국가가 이주민을 동등한 행정 서비스의 수혜자로 보지 않고, 오로지 산업 현장의 수요에 따라 유통되는 ‘공급물자’처럼 취급하고 있음을 적나라하게 보여준다.

더욱 우려스러운 점은 이러한 기술적 효율성이 이주민의 권리 보장을 위한 제도적 장치를 대체하고 있다는 사실이다. 시민단체가 꾸준히 문제 제기하고 있는 미등록 이주민 합법화 방안이나 고용허가제 독소조항 철폐와 같은 인권의 영역은 철저히 외면한 채, 인공지능을 활용한 감시와 선별 시스템에만 몰두하는 것은 기술을 통한 ‘교묘한 배제’에 불과하다. 모든 이주민이 법적 체계 안에서 보호받을 수 있는 디지털 권리가 보장되어야 함에도 행정의 시선은 오직 기업의 수익성에만 고정되어 있다.

진정한 의미의 디지털 이민행정 혁신은 기업의 서류 제출을 간소화하는 것에 그쳐서는 안 된다. 언어의 장벽을 넘어 이주민 스스로 자신의 체류 자격을 확인하고, 부당한 노동 현장에서 즉각적으로 권리를 주장할 수 있는 디지털 소통 창구를 마련하는 것이 우선이다. 이민행정의 서비스 대상은 기업만이 아니라, 이 땅에서 삶을 일구는 이주민에게도 동등하게 제공되어야 한다.

생체정보 오남용, 알고리즘 차별, 기업 중심 선별의 위험을 철저히 성찰하지 않은 채 추진되는 이 계획은 사회 갈등만 키울 것이다. AI·생체정보 기반 감시체제와 K-Trust 인증제는 즉각 폐기가 타당하며, GCM 취지에 맞는 인권 중심 이민정책 전환이 요구된다. 사람이 빠진 기술은 결국 더 정교한 통제의 도구가 될 뿐이라는 사실을 법무부는 직시해야 할 것이다

GCM 23가지 목표에 비춰 본 이민행정과 AI(인공지능) 적용의 의미

GCM 23가지 이행 목표와 국제인권규범이 지향하는 ‘인권 기반 디지털 국경 거버넌스’를 실현하기 위해 한국 정책은 혁신 방향을 전면 재조정해야 한다.

▶ **출입국관리법 개정을 통한 목적구속 엄격화:** 수집된 생체정보는 본인확인·심사 목적으로만 이용 범위를 한정하고, AI 학습용 등 2차적 이용은 원칙적으로 금지해야 한다.

▶ **AI기본법 내 이민·출입국 분야 ‘고영향 AI’ 명시 및 금지 유형 신설:** 감정인식 거짓말 탐지 등 인권 침해 소지가 다분한 응용은 전면 금지하고, 이민 분야 AI는 고영향 AI로 지정하여 위험관리체계를 의무화해야 한다.

▶ **독립적 인권영향평가와 인적 통제 보장 (Human in the Loop):** AI 도입 전 알고리즘이 초래할 차별성을 평가하는 제도를 도입하고, 입국불허, 비자거부, 난민심사, 강제퇴거 등 개인의 운명을 가르는 중대한 결정은 자동화 시스템 단독으로 내릴 수 없도록 하고, 반드시 인간 심사관의 유의미한 검토를 거치게 법정화해야 한다.

▶ **투명성 확보와 독립적 구제 절차:** K-ETA를 포함한 자동화 심사의 알고리즘 개요와 국적별 불허율 등 운영 통계를 투명하게 공개하고, 국가인권위원회에 공공 AI 차별 조사 권한을 실질적으로 부여해야 한다.

▶ **저위험 영역에서의 이주민 권리 옹호 활용:** 다국어 안내 챗봇, 통·번역 지원 등 이주민의 행정 접근성을 높이고 권리를 돕는 ‘착한 AI’ 영역의 활용은 적극 촉진하되, 프로파일링과 생체인식 등 고위험 기술은 3단계 위험 기반 규제로 철저히 통제해야 한다.

맺는 말

이민 결정은 본질적으로 재량적이고 미묘한 윤리적 판단을 요구하는 영역이다. AI는 예측을 할 수 있지만, 이해하는 것은 아니다. 따라서 이민 절차에서 요구되는 뉘앙스와 윤리적 판단은 여전히 인간의 능력 영역에 속한다. AI 기술이 정치, 경제, 사회, 행정에 많은 영향을 주고 있지만, AI가 초래하는 다양한 정책 의제를 다룰 수 있는 법률 기반은 제대로 마련되어 있지 않다. 국가 간 법적·기술적 차이로 인해 일관된 시스템 운영이 어렵다는 문제도 존재한다.

기술의 효율성이 인간의 존엄성보다 앞설 수는 없다. 한국 정부는 국경 통제의 편리함을 위해 취약한 처지의 이주민·난민을 AI의 시험대로 삼는 초법적 행태를 중단하고, 국제사회가 요구하는 투명성과 적법절차의 기준을 시급히 마련해야 한다. AI 기술이 인권과 윤리를 존중하는 방향으로 발전할 수 있도록 시민단체는 데이터 보호를 위한 강력한 규제를 촉구하며, 투명한 알고리즘 설계, 공정성과 신뢰성을 담보하기 위한 지속적인 모니터링을 꾸준히 해 나갈 필요가 있다.

허오영숙 (한국이주여성인권센터)

AI는 여러 면에서 기대와 우려를 동시에 주고 있다. AI와 관련하여 제기되는 문제는 성별, 장애, 인종, 연령 등 역사적·구조적 차별이 인공지능 시스템을 통해 재생산되고, 고착화될 수 있다는 점이다. AI가 학습한 과거 데이터로 인한 알고리즘의 편향성과 차별은 이미 주요한 논의가 되고 있다. 한국에서 이주관련 정책은 젠더화된 방식으로 나타난다. 이주민 구성에 따른 성별 차이에서 기본적으로 확인할 수 있다. 남성은 주로 전문/비전문 영역을 포함하여 노동영역에, 여성은 가족이나 서비스에 주로 편중되어 있다. 이주여성을 재생산, 돌봄 노동의 대상으로 상상하는 정책들은 지속되어 왔다. AI는 이런 편중을 강화시킬 가능성이 있다. (또한 오늘의 주요 논의는 아니지만) AI 산업과 관련하여 가장 위험하고 어려운 하층의 노동을 누가 할 것인가 하는 질문도 있다. 1차 전지 리튬 배터리 폭발 사고로 이주여성 노동자 15명을 포함하여, 23명의 노동자가 사망했던 아리셀 참사는 AI 산업과 간접적으로 연결되어 있다.

AI 시대 공권력, 이주여성을 위해 작동하게 될까?

AI가 인간의 어려움을 해결해 줄 것이라는 기대와 동시에 AI 발전이 국가의 통제에 사용될 것이고, 무엇보다 일상에서 젠더기반폭력에 사용될 것이라는 우려는 지속되고 있다. 어떤 제도이든 국가마다 서로 다른 사회경제적 배경과 역사적 맥락에 따른 경로의존성이 나타난다고 본다면, AI를 ‘활용’하는 한국 정부의 방식이 기존의 국가 통제적 특성을 바탕으로 할 것이라고 생각한다. 국가의 AI 활용 방향은 대체로 국경감시, 국가안보에 초점이 있고, 이는 필연적으로 기본권의 통제와 관련이 있다. 기술 발전이 가져올 젠더기반폭력의 지능화에 대응하는 범죄 수사에 적극적으로 활용될까 하는 의구심이 있다. 여성들은 안전하지 않다고 느끼지만, 젠더기반폭력에 대한 예방, 수사와 처벌은 여전히 기대에 못 미치고 있다. AI는 젠더기반폭력에 대한 역사와 사회적 흐름을 데이터화 할 것이고, AI도 젠더기반폭력에 허용적이지 않을까 하는 우려다.

이런 우려는 AI 시대 공권력이 여전히 젠더기반폭력 피해 이주여성을 위해 작동하지 않을 것 같은 예감으로 나아간다. AI 기술의 발전은 온라인에서 젠더기반폭력 피해자를 쉽게 양산할 가능성이 높다. AI를 통해 쉽게 젠더기반폭력물이 되는 딥페이크 영상물 같은 것을 만들 수 있다. 인종화된 판타지가 결합되어서 여성의 몸을 조작하고 유통하는 문제가 확산될 우려도 있다. 특정 인종, 특정 출신국과 같은 특정 집단의 여성을 성적으로 소비 가능한 이미지로 대량생산하고 소비하는 방식이 훨씬 강화될 가능성이 있다. 누군가의 얼굴을 합성하지 않는 방식으로든 성적 이미지물이나 성착취물은 가능하다. 현실에서 실질적인 피해자가 없을 수 있지만, 이주여성에 대한 모욕과 폭력임은 분명한 가상의 형태일 수도 있다. 사람의 얼굴, 신체를 성적 욕망 또는 수치심을 유발할 수 있는 형태로 편집, 합성, 가공하는 행위에 대해 제재하고 있으나¹⁾, 하지만 현행 법으로 다루지 못하는 영역은 늘 있다.

1) 성폭력처벌등에관한특례법 제14조의 2 (허위영상물 등의 반포 등)

온라인에서 이른바 ‘수배’가 내려진 이주여성들이 있다. 돈을 갚지 않아서, 말을 듣지 않아서 등등 온갖 이유로 이주여성들의 신분관련 서류가 얼굴 사진과 함께 공개되곤 한다. 이주여성들을 통제하고, 벌 주기 위한 방법 중 하나다.

베트남에 사는 한국남성들이 베트남 현지 여성들과의 성적 행위나 성 착취물을 모아 놓은 텔레그램 방에 대해 제보를 받은 적이 있다.

어느 것이든 수사는 잘 진척되지 않는다. 수사의 어려움이 경찰의 기술적 한계나 인력 혹은 역량의 한계일 수도 있고, 젠더기반폭력에 대한 관대했던 유구한 역사 때문일 수도 있고, 혹은 (외국인인) 이주여성 피해자에 대한 무관심일 수도 있다. AI 경찰 행정은 이주여성들의 피해를 잘 살피고, 적극적으로 수사하게 할까?

국제결혼 중개 산업의 성·인종차별과 AI 공권력?

2000년대 이후 급증한 한국 남성과 아시아 여성의 국제결혼은 상업적 결혼중개 산업과 긴밀히 연관되어 있다. 중개업은 한국남성의 욕망을 충족하는 순종적이고 어리고 예쁜 이주여성의 이미지를 노골적으로 창출해 왔다. 2006년 성차별적이고 인종차별적인 중개업 광고물들을 모아 국가인권위에 차별 진정한 것을 시발로, 이주여성운동은 지속적으로 국제결혼 중개업에 대한 비판적인 활동을 해 왔다. 현장의 문제제기에 따라 중개업에 대한 제도적 규제가 일정정도 이루어졌다. 하지만 여전히 중개업은 한국 남성들의 수요 창출을 위해 다양한 방식으로 작동한다. 유튜브는 국제결혼중개업 광고의 문제를 더 빠르고, 더 넓고, 더 자극적인 방식으로 확산시킨다. 유튜브 콘텐츠는 썸네일, 제목, 자막, 해시태그, 댓글, 좋아요, 알고리즘, 광고 수익이 함께 움직인다. 이 구조 안에서 이주여성은 “외국인 미녀” “순수한 외국인 아내” “한국 남자에게 반한 여성” “시골을 살릴 신부”같은 방식으로 소비된다. 이런 콘텐츠는 이용자 반응을 통해 더 확산된다. 클릭은 조회수로, 조회수는 추천으로, 추천은 다시 비슷한 콘텐츠의 생산으로 이어진다. 댓글은 단순한 반응이 아니라 2차 대상화의 공간이 되기도 한다. “어느 나라 여성이 더 좋다” “순하다” “예쁘다” “한국 남자에게 잘 맞는다”는 식의 반응은 영상 속 여성을 다시 평가와 비교의 대상으로 만든다. 한편 유튜브에서 성·인종차별은 노골적인 혐오 표현으로만 등장하지 않는다. 오히려 친근한 말투, 유머, 감동 서사, 가족·여행·국제커플 등 콘텐츠 속에서 자연스럽게 스며든다. 차별이 재미와 호기심으로 포장되고 있다²⁾.

만남에서 결혼까지 10일이 채 걸리지 않는 상업적 중개업을 통한 국제결혼이 전체 결혼의 50% 가까이 차지하는 결혼 상대국들이 있다. 국제결혼 중개업 규제 요구를 ‘관리’로 받았던 정부의 정책 기조는 중개업을 공식화함으로써 이주여성을 ‘사 올 수 있는 대상’으로 상상하게 만들었다. AI를 활용하는 산업들은 기존 체계를 답습하면서 이주여성에 대한 차별적인 이미지를 강화하게 될 것이다. 그리고 이러한 문제가 AI 공권력으로 해결될 것이라는 생각은 들지 않는다.

공권력과 AI 통번역³⁾

AI 통번역이 발전할수록 한국어 단일 사용의 진입장벽을 낮출 수 있을 것이고, 이주여성들에게 편의를 제공할 것이다. 수사기관(경찰, 검찰)에서는 통역사를 위촉하고 있으며, 법원은 법

2) 한국이주여성인권센터. 2022. 미디어 속 성·인종차별 모니터링 보고서

3) 허오영숙. 2026. 『AI와 성평등: AI가 가져올 사회변화, 젠더로 묻다 라운드테이블』.

정 통·번역인 인증 평가(사법통역사) 제도를 두고 있다. 통번역의 질이 균질하지 않거나 통역인에 의해 2차 가해 등의 문제가 되곤 한다. 체류를 위해 필수적인 방문처인 일선의 출입국외국인청에는 공식적인 통역 체계가 없다. 규모가 큰 일선 출입국외국인청에 유급 자원봉사 형태로 통역인이 있는 경우가 있지만, 요일마다 언어가 들쭉날쭉이다. 이주민의 삶에 영향을 줄 수 있는 필수적인 곳에 공적 통역시스템을 하지 않았던 정부의 모습이 학습된 AI는 이주민 편의를 고려할까?

출입국, 병원, 학교, 행정기관 등 일상 전반에서 AI를 통한 즉각적인 통번역은 생존의 도구가 될 수 있다. 한국어, 한국 제도와 행정에 능숙하지 않다는 이유로 독립적인 주체로 인정받지 못함이 나아질까 하는 기대를 하게 한다. 하지만 AI 통번역의 가능성이 바로 권리의 보장으로 이어지는 것은 아니다. AI 통번역과 사용이 확대될수록 국가는 전문 통번역을 통한 언어권 보장의 책임을 AI를 통하는 방식으로 회피할 우려도 있다.

젠더폭력 피해지원, 수사나 사법절차, 출입국 행정, 의료 현장 등에서 절차상 필요한 말을 정확하게 번역하는 것은 중요하다. 하지만 이주여성이 말로 표현하게 되는 감정과 맥락, 침묵, 두려움, 권력 관계를 읽고 설명하는 것은 반영되지 못하고 생략될 수 있다.

외국인 등록증, 일상의 신분 등록 시스템

한국사회 통제적 특성이 대표적으로 드러나는 신분등록제도를 보자. 한국은 1968년 주민등록 제도가 시행되었고, 지문날인이 강제되기 시작했다. 이와 관련하여 경찰청장이 주민등록증 발급신청서에 날인되어 있는 지문 정보를 보관·전산화하고 이를 범죄수사의 목적에 이용하는 행위가 헌법상 개인정보 자기결정권을 침해한다는 주장의 헌법소원은 2024년 기각된 바 있다⁴⁾.

출입국관리법 제5조 1절은 ‘외국인의 등록’으로 한국에서 90일 초과하여 체류하는 외국인은 등록하여 외국인등록증을 발급받도록 규정하고 있다. 등록 사항은 1) 성명, 성별, 생년월일 및 국적, 2) 여권의 번호·발급일자 및 유효기간, 3) 근무처와 직위 또는 담당업무, 4) 본국의 주소와 국내 체류지, 5) 체류자격과 체류기간, 6) 제1호부터 제5호까지에서 규정한 사항 외에 법무부령으로 정하는 사항이다⁵⁾.

지문날인을 포함하는 자국민에 대한 강제 등록 시스템이나 사실상 외국인의 거의 모든 개인정보를 등록하는 제도는 다른 나라에서 찾아보기 어려운 국가 통제 시스템이다. 식민지 통제, 이데올로기 쟁투와 독재 등 권위적이고 통제적인 시스템에 익숙했던 한국 사회는 지문날인 신분증의 강제에 대한 문제의식이 크지 않았다. 그것이 지금의 외국인등록 시스템을 가능하게 하는 바탕이 되었다고 생각한다.

AI와 경찰 활동, AI와 이민행정을 살펴 본 두 발제문 모두에서 우려하는 바가 기본권의 침해인 점은 한국의 통제적 국가 특성과 무관하지 않다고 생각한다.

AI 활용 이민행정이 아니어도..

4) 이희훈. (2024). 우리나라 주민등록법령상 지문날인제도의 쟁점 및 개선 방향 -헌법재판소 2024. 4. 25. 2020헌마542 결정의 평석을 겸하여-. 일감법학, 58, 185-212.

5) 출입국관리법 시행규칙 제47조(외국인등록사항) 법 제32조제6호의 규정에 의한 외국인등록사항은 다음과 같다. 1. 입국일자 및 입국항, 2. 사증에 관한 사항, 3. 동반자(「민법」 제779조에 따른 가족의 관계에 있는 사람으로서 동거하는 사람을 말한다)에 관한 사항, 4. 「초·중등교육법」 제2조 각 호의 어느 하나에 해당하는 학교에의 재학 여부, 5. 사업자 등록번호, 6. 직업 및 연간소득금액

법무부가 올해 발표한 <2030 이민정책 미래전략>에는 AI 활용 이민행정 서비스 개선이 등장한다. 디지털 사전 스크리닝과 AI를 분류, 심사에 활용하여 신속, 정확하게 비자, 체류허가 심사를 하겠다는 계획이다. 강화된 국경 관리는 이주할 수 밖에 없는 여성들의 이주를 막거나, 더 위험한 이주를 감행하게 만든다. 여성들의 이주 사유는 난민 인정 사유가 되기 어렵다. 대부분의 난민 인정 사유는 정치, 사회적 이슈이며, 주로 남성들이 참여가 가시화된다. 젠더기반 폭력으로 난민 인정을 받는 경우는 매우 드물다.

AI 활용은 단속과 구금에서도 고도화 될 것이다. 고도화된 단속과 구금에도 구금 시설은 여전히 열악하다. 시민모임 마중이 공개한 2025년 외국인보호소 구금 인원의 성별 현황을 보면, 여성 비율이 많은 달이 30% 정도, 낮은 달이 19% 수준이다⁶⁾. 생리대마저 요청시 지급하는 형태의 구금시설에서 여성은 더 생존하기 어렵다.

6) 화성외국인보호소 방문시민모임 마중. 2025년 마중 활동보고서