

국정원감시네트워크

디지털정의네트워크, 민주사회를 위한 변호사모임, 민주주의법학연구회,
참여연대, 천주교인권위원회, 투명사회를 위한 정보공개센터, 한국진보연대

수 신 각 언론사 정치부·사회부
발 신 국정원감시네트워크 (담당 : 오병일 디지털정의네트워크 대표 02-774-4551
antiropy@digitaljustice.kr 이은미 참여연대 권력감시2팀장 02-723-5302 tsc@pspd.org)
제 목 [보도자료] AI 행동계획, 사이버AI 보안 국정원 권한 확대 우려
날 짜 2026. 1. 2. (별첨 포함 총 5 쪽)

보 도 자 료

AI 행동계획, 사이버AI 보안 국정원 권한 확대 우려

국감넷, AI 관련 국정원의 권한 제한 필요 의견 제출

- 국정원감시네트워크(이하 국감넷)는 오늘(1/2) 대통령 직속
국가인공지능전략위원회에 「대한민국 인공지능 행동계획」(안)에서 국정원에
사이버·AI 안보 권한을 부여한 데 따른 국가기관의 시민 감시 우려를 담은
의견서를 제출했다.
- 대통령직속 국가인공지능전략위원회는 지난 2025년 12월 16일 「대한민국
인공지능 행동계획」(안)(이하 AI 행동계획)을 공지하고, 2026년 1월 4일까지
의견을 수렴 중이다. AI 행동계획은 국정원에 ▲ AI를 활용한 사이버 위협
(딥페이크, 악성코드 생성 등)에 대해 기술적으로 선제적 탐지·차단하는
독자적인 ‘사이버 AI 안보 플랫폼’ 구축을 위한 정보화 전략 계획 수립 및
‘딥페이크 등 허위 정보 판별 기술’ 고도화 계획 마련 (행동계획 11) ▲ 공공분야
대상으로, 과기정통부, 개인정보위 등과 협력하여, 민간 AI 보안 전문기업과
화이트해커가 참여하는 AI 안전성·프라이버시 레드티밍 체계 및 상시 모의해킹
테스트베드 구축(행동계획 11) 등 국정원에 사이버 AI 보안 권한을 확대하는
정책권고와 타 부처와의 협력을 권고하고 있다.
- 시민사회의 지속적인 노력으로 국정원의 국내정보 수집 권한은 일부
제한되었지만, 국정원은 여전히 사이버안보정보와 공공 정보통신망에 대한
사이버보안 업무를 담당하며 이를 매개로 국내 및 민간 영역에 끊임없이

개입하려고 시도하고 있다. 그러나 **AI** 행동계획은 국정원이 '사이버 **AI** 안보 플랫폼' 구축 등 **AI** 사이버보안의 컨트롤타워 역할을 맡도록 하고 있을 뿐만 아니라, **K-사이버보안 LLM** 구축 협력, 허위정보 판별 기술, 민간 클라우드 정책 방안에 대한 검토, 개인정보 불법정보 방지 대책에 협력 등 민간의 기술 생태계나 정보통신망에 관여할 수 있는 여지를 마련해 놨다. 특히 "**AI**를 활용한 사이버 위협 (딥페이크, 악성코드 생성 등)에 대해 기술적으로 선제적 탐지·차단하는 독자적인 '사이버 **AI** 안보 플랫폼' 구축"이야말로 공공의 정보통신망에만 국한하지 않으며, 민간 정보통신망에 대한 감시와 개입 권한을 국정원에 부여하는 것이다.

4. 예산과 국회의 통제를 제대로 받지 않는 초법적 기관인 국정원에 공공 정보통신망에 대한 사이버보안 권한은 물론 국가 **AI** 사이버전략 수립 권한까지 부여하는 것은 사이버 공간에서의 국가 감시를 허용하는 것과 다름없고, 국정원에 이러한 역할을 부여하려면, 최소한 다른 정부 부처와 동일한 수준의 예산 통제와 국회의 실질적 감독을 받도록 해야 한다.
5. 또한 인공지능 기본법은 '국방 또는 국가안보 목적으로만 개발·이용되는 인공지능'을 적용 대상에서 제외하고 있고, 현재 입법예고된 시행령 역시 국방부나 국정원이 지정하는 업무를 적용 배제 대상으로 규정하고 있다. 그러나 **AI** 행동계획은 국방 분야의 경우 '(가칭) 국방 **AI** 기본법'을 '26년 2분기까지 제정하도록 하고 있으나 국정원이 담당하는 '국가안보 목적' **AI**는 여전히 제외되어 있다. 국정원이 국가안보를 명분으로 개발, 활용하는 **AI**에 대해서는 아무런 법적인 규율이 적용되지 않는다. **AI**가 과거의 다른 감시 수단에 비해 훨씬 은밀하고 광범위한 감시를 가능하게 한다는 점에서, 국정원이 아무런 규율 없이 **AI**를 개발·활용하도록 두는 것은 매우 위험하다. 아울러 국가안보 **AI**를 규율할 수 있는 법제가 조속히 마련돼야 하며, 이 역시 행동계획에 포함돼야 한다.
6. 이에 국감넷은 의견서를 통해 국정원의 권한에서 사이버보안 권한을 삭제하고, **AI** 행동계획에서도 국정원의 역할을 해외 정보기관으로서 타 부처에 협력할 역할로 제한해야 한다고 강조했다. 끝.

■ 붙임. 「대한민국 인공지능 행동계획」(안)에 대한 의견서

■ 붙임

「대한민국 인공지능 행동계획」(안)에 대한 의견서

국가인공지능전략위원회가 2025년 12월 16일, 공지한「대한민국 인공지능 행동계획」(안)(이하 AI 행동계획)에 대해 다음과 같이 의견을 제출합니다.

AI 행동계획은 국정원에 다음과 같이 행동계획의 정책 권고를 이행하거나 타 부처와 협력하도록 하고 있습니다.

- K-사이버 보안 LLM의 지속 학습과 ‘사후 대응’에서 ‘예측·예방’으로의 전환을 위해, 정부 및 공공기관에 AI 기반의 이상(anomaly) 탐지 체계를 구축하기 위해 과기정통부에 협력(행동계획 10)
- AI를 활용한 사이버 위협 (딥페이크, 악성코드 생성 등)에 대해 기술적으로 선제적 탐지·차단하는 독자적인 ‘사이버 AI 안보 플랫폼’ 구축을 위한 정보화 전략 계획 수립 및 ‘딥페이크 등 허위 정보 판별 기술’ 고도화 계획 마련 (행동계획 11)
- 공공분야 대상으로, 과기정통부, 개인정보위 등과 협력하여, 민간 AI 보안 전문기업과 화이트해커가 참여하는 AI 안전성·프라이버시 레드티밍 체계 및 상시 모의해킹 테스트베드 구축(행동계획 11)
- 주요국 정보기관과 글로벌 AI 안보 연대를 강화하여 AI 위협 관련 정보·기술 공유 지속 추진 및 AI 안보 현안에 대한 공동 대응 전략 수립 (행동계획 11)
- AI가 초래할 수 있는 국가 안보 위협인 CBRN (화학, 생물학, 방사능, 핵)에 대비하기 위해, 관련 기술개발과 국제협력 강화 (행동계획 11)
- 행안부·과기정통부와 함께 정부와 공공기관에서 민간 클라우드 이용을 제한하는 것으로 알려진 법과 제도를 모두 리스트업하고, 「AI 정부 인프라 거버넌스·혁신 TF」가 마련하는 정책방안을 토대로 이를 준수할 방안 마련 (행동계획 52)
- 타부처 사업과의 중복성 검토, 공공AI사업지원센터의 역할 및 운영 기반과 민간 제안 절차, 예산 확보 방안 마련을 위한 행안부 사업에 협력 (행동계획 61)

- 국방부의 "민군협력 기반 국방 AI 보안 혁신 로드맵" 수립에 협력 (행동계획 77)
- 개인정보위의 「AI 시대 개인정보 보호 및 불법 유통 방지 종합 대책」 수립에 협력 (행동계획 90)

그동안 시민사회의 지속적인 노력으로 국정원의 국내정보 수집 권한은 일부 제한되었지만, 국정원은 여전히 사이버안보정보와 공공 정보통신망에 대한 사이버보안 업무를 담당하고 있으며, 이를 매개로 국내 및 민간 영역에 끊임없이 개입하려고 시도하고 있습니다. 국정원감시네트워크는 국정원의 위상과 역할을 해외 정보기관으로 명확히 정립하고, 사이버보안 관련 국정원 업무를 다른 정부 부처로 이관할 것을 지속적으로 요구해 왔습니다.

그러나 AI 행동계획은 국정원이 '사이버 AI 안보 플랫폼' 구축 등 AI 사이버보안의 컨트롤타워 역할을 맡도록 하고 있을 뿐만 아니라, K-사이버보안 LLM 구축 협력, 허위정보 판별 기술, 민간 클라우드 정책 방안에 대한 검토, 개인정보 불법정보 방지 대책에 협력 등 민간의 기술 생태계나 정보통신망에 관여할 수 있는 여지를 마련해 놓고 있습니다. 이를 통해 기존에 국정원이 사이버보안 권한과 관련하여 갖고 있던 문제가 AI 시대에도 확대재생산될 위험이 있습니다. 국정원의 사이버보안 역할 만큼이나 AI 사이버보안을 왜 비밀 정보기관이 맡아야 하는지 의문이 아닐 수 없습니다.

국정원은 공공 정보통신망에 대한 사이버보안 권한을 민간의 정보통신망으로 확대하기 위해 호시탐탐 노리고 있던바, "AI를 활용한 사이버 위협 (딥페이크, 악성코드 생성 등)에 대해 기술적으로 선제적 탐지·차단하는 독자적인 '사이버 AI 안보 플랫폼' 구축"이야말로 공공의 정보통신망에만 국한하지 않으며, 민간 정보통신망에 대한 감시와 개입 권한을 국정원에 부여하겠다는 것이나 마찬가지입니다. 국정원이 위협 예측 기반의 능동적 방어라는 명분으로, 불법이 명확하지 않은 민간의 트래픽에 대해 선제적으로 탐지, 차단, 관여할 수 있다면, 민간 정보통신 공간에 대한 사이버 감시와 프라이버시 침해의 우려가 큼니다. 사이버보안을 위한 이러한 대응이 필요하다면, 권한 남용의 여지가 적고 정치 사회적 감독이 가능한 일반 행정기관에서 맡도록 해야 할 것입니다.

이는 비단 국정원이 시민감시와 정치 개입을 해왔던 암울한 과거를 갖고 있기 때문이 아닙니다. 문제는 현재도 국정원은 예산 및 국회 감시를 제대로 받지 않는 초법적인 기관이라는 것입니다. 해외 정보 수집이라는 극히 제한적인 역할이 아니라, 왜 민간 네트워크와 연결될 수밖에 없는 공공 정보통신망에 대한 사이버보안, 나아가 국가 AI

사이버전략에 대한 수립 권한을 통제받지 않는 정보기관에게 부여해야 하는지 의문입니다. 만일 국정원에 이러한 역할을 맡겨야 한다면, 다른 정부부처 수준의 예산 및 국회의 통제를 받도록 해야 할 것입니다.

한편, 인공지능 기본법은 '국방 또는 국가안보 목적으로만 개발·이용되는 인공지능'을 적용 대상에서 제외하고 있고, 현재 입법예고된 시행령은 주로 국방부나 국정원이 지정하는 업무를 적용 배제 대상으로 규정하고 있습니다. 즉, 국방부나 국정원의 업무에는 인공지능 기본법이 적용되지 않는다는 것이다. 그나마 AI 행동계획은 국방 분야의 경우 '(가칭) 국방 AI 기본법'을 '26년 2분기까지 제정하도록 하고 있습니다. 그러나 '국가안보 목적' AI 분야로서 국정원이 담당하는 분야는 여전히 제외되어 있습니다. 다시 말하면, 국정원이 국가안보를 명분으로 개발, 활용하는 AI에 대해서는 아무런 법적인 규율이 적용되지 않는다는 것이고, 앞으로 이를 규율할 법제를 만들 계획도 없다는 것입니다. AI는 과거 다른 감시 수단에 비해 훨씬 은밀하고 광범위한 감시가 가능하도록 하는 바, 국정원에 아무런 규율도 없이 AI를 마음대로 개발, 활용하도록 하는 것은 너무나 위험하다. 국가안보 AI를 규율할 수 있는 법제가 조속히 마련되어야 하며, 이 역시 행동계획에 포함되어야 합니다.

이에 국정원감시네트워크는 국정원의 권한에서 사이버보안 권한을 삭제하고, AI 행동계획에서도 국정원의 역할을 해외 정보기관으로서 타 부처에 협력할 역할로 제한할 것을 요구합니다. 끝.