

AI 시대 디지털 독재 방지와 국민주권 실현을 위한 법무정책적 제언 : 헌법적 비상권력의 알고리즘화 및 국민 환원을 중심으로

한국형사·법무정책연구원
강석구 선임연구위원

AI 대전환에 따른 디지털 독재 위기에 대응하기 위해 권력자의 AI 악용(정보 차단, 여론 조작, 피지컬 AI 동원 등)을 원천 차단하는 법무정책적 패러다임 전환이 필요하다.

이를 위해 비상권력 발동의 근거인 모든 AI 데이터와 로직의 국회·사법부 공개를 의무화하고, 살상 무기 체계의 인간 통제 원칙과 국가의 알고리즘 직접 개입 금지를 명문화해야 한다. 또한, 권력을 감시하는 '역(逆) 알고리즘'과 레드팀을 구축하고, 국회 승인 없이는 계엄 발동 시스템 가동이 불가하도록 '코드화된 헌법'을 구현해야 한다.

#디지털 독재 #국가비상권력 #역(逆) 알고리즘 #코드화된 헌법 #데이터 주권

I. 서론: 아날로그 헌법의 한계와 디지털 독재의 도래

AI 대전환과 국민주권 급진전에 따른 헌법적 위기

- 최근 비상계엄 사태 이후 국민주권 의식이 급진전되고 AI 대전환이 현실화됨에 따라, 정부 국정과제에 반영된 '4년 연임제 및 비상계엄 규정 손질' 등 권력 구조 개편론을 넘어선 미래지향적 헌법 담론의 필요성이 대두되고 있음
- 기존의 아날로그식 권력구조 개편 논의만으로는 급격한 기술 변화를 포섭하기에 제도적 한계가 존재하므로 AI 시대의 주권 실현과 기술 독재 방지를 위한 실질적인 보완 콘텐츠 및 법무정책을, 국책연구기관으로서 선제적으로 제안함

○ 기존 비상계엄 규정의 아날로그적 한계와 AI 기반 ‘디지털 독재’ 가능성

- 현행 헌법의 비상권력 통제 장치는 군대 동원 등 물리적 무력 차단에만 국한되어 진화된 디지털 위협에 무력함
- 비상사태 선포 시 권력자가 AI 알고리즘을 장악할 경우 실시간 정보 차단, 여론 조작, 피지컬 AI(드론, 안면인식 등)를 동원한 지능형·보이지 않는 권력 집중이 현실화될 소지 다분
- 이에 AI 시대의 비상권력 통제는 ‘AI 자체의 통제’를 넘어, ‘권력자가 AI를 통해 견제 없는 독재를 실현하는 것’을 원천 차단하는 기술적·법적 방어벽 구축으로 패러다임을 전환해야 함

II. AI 악용형 국가비상권력의 3대 오용 범주

- AI 시대 국가비상권력 통제안의 핵심은 AI의 권력 행사 자체가 아니라, 권력자가 AI를 도구 삼아 견제 없는 독재를 실현하는 것을 방어하는 데 있음. 비상권력의 총위는 크게 세 가지 영역으로 구체화됨

○ 정보 및 내러티브 통제권

- 전망 : 비상계엄 선포 시 전통적 언론 제한을 넘어, AI 알고리즘을 통한 실시간 정보 차단 및 고도화된 여론 조작 가능
- AI의 역할 : 딥페이크 기반 가짜 뉴스 유포, 특정 키워드 자동 검열, SNS 활동 실시간 모니터링 및 계정 차단
- 통제안 : 국가의 AI 알고리즘 직접 개입권(마스터 키) 보유를 헌법적으로 금지하고, 비상시 알고리즘 중립성을 검증할 민간 독립 기구를 상설화함

○ 물리적 강제력 및 감시권

- 전망 : 드론, 자율주행 순찰차, 안면 인식 시스템 등 ‘피지컬 AI’를 동원한 계엄군 무인 운용 현실화
- AI의 역할 : 시위대 자동 식별·추적, 자율 무기 체계를 통한 무력 진압, 국가 통신망 강제 점유
- 통제안 : 비상사태 시에도 ‘살상 무기 체계의 인간 통제 원칙’을 법률로 명문화하고, AI 감시 데이터의 사후 즉시 소거를 의무화함

○ 비상사태 선포 및 유지의 판단권

- 전망 : 위기 상황(사이버 테러, 대규모 재난 등) 정의 및 비상권력 발동 여부 결정 시 AI 시스템 의존도 심화

- AI의 역할 : ‘데이터상 현재는 국가 비상상태임’을 입증하는 정교한 시뮬레이션 결과를 권력자의 정치적 명분용으로 제공
- 통제안 : 비상권력 발동의 근거가 된 AI 데이터와 로직을 국회·사법부에 즉각 공개하도록 강제하며, 불투명한 ‘블랙박스’ 결정에 의한 계엄은 원천 무효로 규정함

Ⅲ. 통제 패러다임의 전환: ‘역(逆) 알고리즘’을 통한 권력의 국민 환원

- AI 시대에는 기술이 권력의 도구가 되는 것을 방어하는 것을 넘어, ‘권력을 감시하는 역(逆) 알고리즘’으로 기능해야 함. 즉, “**권력을 휘두르는 AI를 막는 것은, 그 권력을 감시하는 더 투명한 AI**”임.
- **상시적 레드팀의 헌법기관화**
 - 국가 비상권력 운영 시스템의 취약점을 상시 분석하고 오남용 시나리오를 가동하는 독립적 AI 감시 체계를 구축함 (예: 감사원 산하 혹은 독립 기구화)
- **코드화된 헌법(Encoded Constitution) 시스템 구현**
 - 비상권력 발동 프로세스를 조작 불가능한 블록체인이나 검증된 알고리즘에 탑재하여, 국회 승인 없이는 기술적으로 계엄령 발동(통신 차단, 군 출동 등) 시스템 가동이 불가하도록 ‘기술적 잠금 장치’를 구현함
- **AI 설계 알고리즘 헌법 적합성 평가**
 - 주권자인 ‘국민의 통제권’ 구현을 위해, 국가 운용 모든 AI 시스템은 비상상황에서도 헌법적 가치를 위반하지 않도록 설계되어야 하며, 이를 상시 감시하는 ‘시민 배심원단 AI’를 가동함
- **데이터 주권 기반의 권력 감시**
 - 비상권력 발동 시 수집되는 모든 데이터의 소유권은 국민에게 있음을 명시하고, AI를 통해 권력자의 데이터 접근 및 활용 경로를 실시간 트래킹함
- **시민에 의한 ‘실시간 거부권’ 보장**
 - 보편적 법률복지의 일환으로 전 국민 단말기에 AI 법률 비서를 탑재, 국가 비상권력 남용 시 즉각 법적 위반 사항을 전파하고 국민적 자기방어권(지식)을 보급함

IV. 국가 거버넌스 재설계를 위한 법무정책적 핵심 과제

○ AI 알고리즘 투명성 및 데이터 주권의 헌법적 명문화

- 비상권력 발동의 근거인 모든 AI 데이터·로직을 국회와 사법부에 즉각 공개 의무화
- 비상 데이터의 국민 소유권을 천명하고 권력자의 접근 경로를 시로 트래킹하는 헌법적 기반 마련

○ 국가 비상권력 발동 시 AI 시스템 운영의 법적 한계 설정

- 국가의 알고리즘 개입(마스터 키) 금지를 통한 언론·출판 자유 보장 및 계엄군의 '살상 무기 인간 통제 원칙' 명문화
- 불투명한 '블랙박스' 알고리즘 결정에 기반한 모든 비상조치의 원천 무효화

○ 디지털 개헌 및 (가칭) '디지털 비상권력 통제법' 제정

- 국회 승인 전 통신 차단 및 군 출동을 원천 통제하는 '코드화된 헌법' 가동 및 하위 법제 정비
- 외산 빅테크 AI의 비상권력 도구화 및 인프라 마비 위험에 대응하여, '한국형 법치주의'가 이식된 검증용 오픈소스 모델 개발을 국가 차원에서 지원함

○ 한국형사·법무정책연구원의 역할 전환: 'AI 사전 통제 알고리즘' 중점 연구

- 글로벌 빅테크에 비해 AI 개발 기술 자체는 뒤처졌을지라도, AI를 민주적으로 다루고 부리는(Governing) 법무정책적 철학만큼은 대한민국이 세계 최고가 될 수 있음
- 이에 연구원(KICJ)은 형사정책적 사후 처벌 위주의 관행에서 탈피, 주권자인 국민의 통제권을 기술적·제도적으로 보장하는 '법무정책적 사전 통제 시스템' 연구를 전면 주도해야 함