제3전선, 정보전쟁 AI 정보전

뇌파 해킹해 사람 생각까지 도청, AI 스파이 시대 온다

최성규
고려대 연구교수

17일 우크라이나 군인들이 날아오는 러시아 드론을 향해 총격을 가하고 있다. 미국 민간 정보회사가 개발한 '가디언AI'는 전쟁 발발 4개월 전에 러시아의 침공을 예측했다. 〈EPA=연합뉴스〉

인공지능(AI)의 등장으로 인류는 또 한 번 문명사적 전환기를 맞고 있다. 산업·금융·학문 등 전 분야에서 대전환이 이루어지고 있다. 정보도 예외가 아니다. 아직 본격적인 AI 정보전 시대는 열리지 않았지만, 분명히 문턱은 넘어서고 있다. 초기 단계인데도 충격파가 만만치 않다. 인간의 한계를 넘는 대량 정보처리 능력은 반박 불가이다. 사이버전과 여론전은 AI의 가세로 몇 배 더 강해졌다. 생성형 AI의 발전으로 AI 머신이 다른 AI 머신을 스파이하는 'AI 대 AI 정보전'도 꿈틀대고 있다. 바야흐로 인류는 그간 경험하지 못한 미증유의 AI 정보전 시대를 맞고 있다.

전문가들도 이구동성이다. 바이든 행정부 시절 백악관 안보부보좌관을 역임한 뉴버거(A Neuberger) 스탠퍼드대 교수는 올해 1월 포린어페어스에 '스파이와 AI' 기고를 통해 AI가 거대한 정보혁명을 몰고 오고 있다고 공언했다. 미 국가지리정보국(NGA)의 빈치(A Vinci) 전 기술국장도 2020년 포린어페어스 기고에서 AI 주도형 정보는 인간과 비교되지 않을 정도로 탁월하다며 정보 격변을 예고했다.

러 장교 현금 인출량 변화까지 찾아 반영

정보 현장에서도 AI 정보의 실험적 모습이 조금씩 드러나고 있다. 팔색조처럼 화려하기도 하고 몬스터처럼 괴물 같기도 하다. 그 실험장이 된 우크라이나 전장이 이를 잘 보여주었다.

미국 민간정보회사 롬버스파워(Rhombus Power)의 가디언AI가 그중 하나다. 분쟁 예측과 전략적 판단을 지원하기 위해 개발된 AI인데, 2021년 10월 러시아의 우크라이나 침공을 4개월 전 예측했다. 그 과정이 AI시대 정보전 모습을 미리 보여주었다.

가디언AI는 먼저 과거 러시아의 군사침공 사례부터 현재 러시아 내부 동향까지 모든 정보를 다 수집했다. 정보의 종류도 상업위성에서 언론·SNS까지 다양했다. 이렇게 이질적이고 방대한 정보들을 융합해 분석해 보니 곳곳에서 침공 징후가 나타났다.

러시아가 훈련이라고 주장한 병력·장비의 우크라이나 국경 이동은 2014년 러시아가 크림 반도를 침공할 때와 90% 이상 흡사해 침공 준비라고 해석했다. 국경 지역으로 이동한 러시아 장교들의 현금 인출량이 300% 이상 증가한 것도 찾아냈는데, 이것도 러시아의 병력 이동은 훈련이 아니라 전쟁용 장기체류 증거로 봤다.

언론과 SNS에서도 침공 징후를 찾아냈다. 러시아 언론과 친정부 블로거들이 "우크라이나에서 러시아계 국민에 대한 학살과 인종청소가 자행되고 있다"는 보도를 쏟아 내자, 러시아가 우크라이나 침공을 정당화하기 위한 명분 축적용 여론전이라고 분석했다.

가디언AI는 이런 식으로 약 50개 지표에서 침공 징후를 찾아내 "전쟁 가능성 80%"라고 예측했다. 서방 정보 당국이 러시아의 정치·경제 리스크 등을 들어 침공 가능성을 낮게 평가한 것과 대비됐다. AI가 예측 게임에서 정보 당국을 이긴 셈이다. 물론 가디언AI의 80% 예측치는 실제 사용에는 제약이 따른다. 국가의 존망이 걸린 전쟁 예측은 1%의 불확실성도 용납되지 않기 때문이다. 그러나 가디언AI의 예측 모델은 AI시대 정보전이 어떤 모습일지를 보여주었다는데 의의가 있다.

공개정보·위성정보 등 이질적인 모든 정보를 실시간 교차 검증해 더 빠르고 더 정확하게 분석했다. 산더미 같은 대량 정보 속에서도 핵심정보는 놓치지 않는 예리함도 보여주었다. 오로지 상대가 무엇을 하고 있는지만 객관적으로 추적·분석해 관성과 편향성에 얽매이는 인간의 약점도 극복했다. 특히 정보를 수집해 분석하고 예측하는 전 과정을 분절 없이 처리해 국가위기 조기경보를 실시간 자동화하고 정량화할 수 있는 혁신의 길을 보여주었다.

이처럼 AI의 정보 잠재력이 확인되자 방첩정보, 국익정보 등 정보의 모든 영역으로 스며들기 시작했다. 중국이 AI와 안면인식 기술을 결합해 중국 내 미국 스파이들을 실시간 감시하는 등 방첩정보 활용은 이미 일상화되고 있다.

기존의 여론 정보전도 AI의 가세로 더욱 강력해지고 있다. AI가 상대의 정서와 분위기는 물론 개인의 신념까지 파악해 맞춤형 메시지로 공격하기 때문이다. 심지어 맞춤형 설득으로 상대의 생각까지도 변화시킬 수 있다는 것이 여러 실험을 통해 확인됐다. AI시대 여론 정보전은 국가를 분열시키는 소리 없는 무기가 될 수 있다는 경고다. 올해 초 미 국가정보장실(ODNI)이 AI를 통한 중국·러시아의 대미 여론전이 국가기능 약화로 이어질 수 있다고 경고한 것도 이 때문이다,

사이버 정보전도 더욱 고도화되고 있다. AI가 수백 개의 인터넷 보안 취약점을 신속하게 스캔해 최적의 공격 루트를 찾아 주기 때문이다. 더욱 주목되는 것은 사이버 공격과 방어를 AI가 스스로 수행하는 자율형 정보전을 선보인 점이다.

2024년 러시아가 AI를 활용해 우크라이나 대통령의 위조 영상을 유포하자, 우크라이나도 AI 시스템을 이용해 위조 영상을 자율적으로 차단했다. AI가 공격·방어를 자율적으로 수행한 초보 단계의 'AI 대 AI 정보전'이다. 이 단계가 더 진화하면 AI가 사이버 해킹을 통해 서로 상대방 국가시스템에 숨어 들어가 상대가 무엇을 하고 있는지 알아낼 수 있다. AI 머신이 다른 AI 머신을 스파이하는 본격적인 AI 대 AI 정보전이 온다는 의미다.

미국 기업이 개발한 '가디언AI'
러 우크라 침공 4개월 전 예측
상대 움직임만 객관적 추적·분석
관성·편향 보이는 인간 약점 극복
아직은 80% 정도 예측 정확성 보여
전통적 인간정보 중요성도 여전

미 국가지리정보국(National Geospatial-Intelligence Agency)
상공에서 촬영한 지리·지형 정보를 수집해 안보용 비밀지도 등을 만드는 미 5대 정보기관 중 하나다. '하늘 위 CIA'로 불린다.

민간정보회사(Private Intelligence Agency, PIA)
상업적으로 정보를 수집·분석해 주로 정부에 판매하는 기업으로 미국 사이언스 어플리케이션스 인터내셔널(SAIC), 스트랫포(Stratfor) 등이 있다.

AI, 푸틴 침공 의지까지 알아내지는 못 해

윤리적 논란을 불러올 AI 정보전도 꿈틀대고 있다. AI와 뇌과학을 결합해 사람의 생각을 알아내려는 시도가 그중 하나다. 본래 사람이 생각할 때 발생하는 뇌파를 문자나 소리로 전환해 언어·시각 장애우들을 돕겠다는 뇌의학에서 출발했다. 그러나 윤리적 부담이 상대적으로 덜한 권위주의 국가들이 이를 정보와 군사에 응용하기 위해 꾸준히 연구하고 있는 것으로 알려져 있다. 이를 정보의 언어로 풀어쓰면 AI가 뇌파를 해킹해 사람의 생각을 도청하는 것이다.

이 모든 것들이 생성형 AI의 자기개선적(self-improving) 발전 때문에 가능하다. 하나의 혁신이 또 다른 혁신을 이끄는 자기발전적 속성을 가지고 있어 일단 궤도에 올라서면 가속적 진화가 가능해진다. 그렇게 되면 AI가 스스로 스파이 활동을 하는 AI 자율형 스파이 시대도 올 수 있다.

이처럼 AI 정보전은 이제 막 문을 연 단계로 파장이 곳곳에서 나타나고 있다. 우리 앞에 놓인 과제도 수두룩하다. 첫 출발은 AI 정보전에 대한 냉철한 인식을 토대로 백지 위에 국가정보를 새롭게 설계한다는 자세이다. 정보지형 자체가 변하고 있음을 직시해야 한다.

정보의 운영방식도 고민해야 한다. AI 시대는 구조화된 보고서 중심의 정보문화에서 벗어나 AI로 무장된 새로운 정보전에 익숙해져야 한다. 대북정보에 AI 활용도 검토할 필요가 있다. 북한의 기습도발, 핵 위협과 관련해 AI는 우리가 '생각할 수 없는 것'까지 생각해내는 통찰을 제공할 수 있다.

민간과 정보의 협력도 한 몸처럼 움직일 준비가 돼 있어야 한다. 미국처럼 AI 기술을 주도하는 민간의 국가정보 참여 확대로 정보생태계도 변하고 있기 때문이다. 오염되지 않은 품질 좋은 빅데이터도 대량 확보해야 한다. 데이터는 AI의 성능을 결정하는 '총알'로 비유될 정도로, 앞으로 AI 정보전의 승패를 가를 핵심요소이다.

그러나 AI 시대는 인간정보의 중요성도 비례적으로 증가한다는 것을 잊지 말아야 한다. 미 중앙정보국(CIA) 번스 국장이 2024년 1월 포린어페어스에 밝힌 것처럼 AI가 인간의 정보활동을 지원하는 뛰어난 수단은 될 수 있어도 비밀 정보활동 그 자체를 수행하는 데는 한계가 있기 때문이다. 특히 번스 국장은 우크라이나 전쟁 당시 실제 경험을 토대로 인간정보의 중요성을 다시 한번 강조했다.

인간정보를 통해 러시아 푸틴 대통령의 우크라이나 침공 의지가 어느 정도인지 내밀하게 파악했어야 했는데 이를 소홀히 해 초기 전쟁 가능성을 낮게 평가하는 실책을 범했다고 고백했다. 앞서 가디언AI의 '80% 전쟁 예측' 사례처럼 AI가 풀 수 없는 나머지 20%의 핵심 퍼즐은 결국 사람이 풀어야 한다는 이야기다. 번스 국장이 "세계 모든 정보기관이 새로운 시대에 성공하기 위해서는 전통적 인간정보와 새로운 첨단정보의 창조적 결합에 달렸다"고 설파한 것도 이 연장 선상이다.

이처럼 AI시대 정보전은 인간과 AI의 협업, 민간과 정보기관의 협업 등 전략적인 정보 협업이 중요해진다. 그 첫발은 관성적인 정보 칸막이부터 걷어 내는 것이다. 지금 서방처럼.

〈광주일보와 중앙 SUNDAY 제휴 기사 입니다〉

최성규 국가정보원에서 장기간 근무하며 국제안보 분야에 종사했다. 퇴직 후 국내 최초로 비밀 정보활동의 법적 규범을 규명한 논문으로 고려대에서 박사학위를 받았다.