

www.eai.or.kr

EAI 워킹페이퍼

지식 및 거버넌스 부문 : AI 시대 남북관계의 리스크와 기회

최인호(경기대학교)

지식 및 거버넌스 부문 : AI 시대 남북관계의 리스크와 기회

최인호

경기대학교 교수

I. 서론

지금까지 AI에 대한 논의는 주로 군사·경제 영역에 집중되는 경향이 있고, 그 영역의 변화가 눈에 띄게 빠른 것도 사실이다. 그러나 보다 근본적인 변화는 오히려 지식 및 거버넌스 영역에서 진행되고 있다. AI가 가져오는 가장 깊은 전환은 지식이 생산, 유통, 축적되는 방식 자체의 변환이며, 이 변화에 적응할 수 있는 능력이 군사력과 경제력의 토대를 결정한다. 산업혁명 이후 국력의 일차적 기준이 영토에서 자원으로, 자원에서 자본으로, 자본에서 인적자본으로 이동해 왔다면, Kissinger et al.(2024)이 논의한 AI 및 컴퓨팅 자본 및 지식의 변화는 새로운 국력 기준의 부상을 시사한다. 이를 ‘지식력’ — 데이터를 처리하고 새로운 지식을 생산하는 능력 — 이라고 할 수 있는데 이 글은 이러한 인식 위에서, 지식 및 거버넌스 부문이 AI 시대 남북관계의 리스크와 기회를 이해하는 데 부차적이라기보다 핵심적인 변수가 될 수 있음을 다룬다.

논의의 출발점에서 두 가지 구분이 중요하다. 첫째, ‘데이터 통제’와 ‘지식 생산 통제’는 동일한 문제가 아니다. 권위주의 정권은 국경 내 데이터를 통제할 수 있지만, AI 시대의 지식 생산은 글로벌 오픈소스 모델, 분산된 컴퓨팅 자원, 국경을 가로지르는 연구 네트워크라는 인프라 위에서 일어난다. 따라서 데이터를 봉쇄해도 그 외부에서 폭발적으로 진행되는 지식 생산을 멈출 수 없으며, 봉쇄가 강할수록 자국 체제는 그 흐름에서 완전히 배제되는 훨씬 더 복잡한 딜레마에 직면한다. 둘째, ‘AI를 활용한 통치 강화’와 ‘AI 시대에 적합한 거버넌스 능력’도 다른 차원의 문제다. 안면인식, 국경감시, 딥페이크 공작과 같은 AI의 도구적 활용은 기존 통제 체계의 효율화에 가까운 반면, 후자는 AI가 변화시키는 새로운 지식질서에 적응할 수 있는 사회 및 제도적 능력을 요구한다.

여기서는 후자에 1차적 초점을 두되, 전자(AI를 통한 통치 강화)가 ‘기술 종속의 딜레마’ 등의 형태로 거버넌스 능력을 어떻게 잠식하는지도 함께 검토한다.

글로벌한 AI 지식질서의 변환과 북한의 현황을 지식 및 거버넌스 측면에서 분석할 필요가 있는데, 우선 AI가 지식질서에 가져오는 철학적 전환의 성격을 분석하고, 그 안에서 북한이 구조적으로 뒤처질 수밖에 없는 이유, 격차의 자기강화적 확대, 그리고 정보통제 비용의 변화 전망을 다룬다. 다음으로 AI를 통치, 감시, 사회관리의 도구로 활용해 온 북한의 실태와, 그에 따르는 ‘기술 종속의 딜레마’라는 구조적 모순을 검토한다. 이후 교육, 연구, 의료 세 도메인에서 북한의 AI 지식-거버넌스가 단·중기에 어떻게 전개될 수 있는지를 도메인별로 살펴본다. 마지막으로 이러한 도메인별 협력의 누적 효과로서 AI 커먼스 형성과 그에 결부된 거버넌스 변화의 단·중기 전망을 종합적으로 검토한다.

II. 북한의 지식체계와 거버넌스에서의 AI 활용 현황과 변화 가능성

1. AI 지식질서의 철학적 특성: 신체 증강에서 인지 변환으로

인류의 기술사에서 변혁적 발명은 인간의 신체 기능을 증강하는 것이었다. 바퀴와 엔진은 근력을, X선과 현미경은 시야를, 전화기는 목소리를 확장했다. AI는 이 계보에서 이탈한다. 이전의 기술이 팔이나 눈과 같은 신체 기능을 확장했다면, AI는 인지 — 인간이 세계를 파악하고 지식을 생산하는 방식 — 자체에 개입한다(Kissinger et al. 2024). 이것이 AI가 이전의 어떤 기술혁명과도 본질적으로 다른 첫 번째 이유다.

다만, 이것이 인간 두뇌 전체의 기능을 기계가 복제했다는 뜻은 아니다. Yuk Hui(2026)가 칸트 철학의 렌즈를 통해 분석하듯, 현재 AI의 핵심 작동 원리는 경험적 데이터에서 패턴을 추출하고 확률적으로 가장 그럴듯한 다음 결과를 예측하는 것이다 — 18세기 흄(David Hume)의 경험주의를 극한까지 기계화한 것에 가깝다. 거대언어모델은 본질적으로 인류가 생산한 텍스트 전체를 하나의 압축 알고리즘으로 응축한 것이며, 그 위력은 이 압축의 규모와 속도에서 나온다. 현재 AI는 주관적 책임의 주체가 아니며, 그것이 산출하는 가설과 규범의 정당화에는 여전히 인간의 판단이 남는다. 다만 AI는 천문학적 양의 경험적 연관에서 인간이 발견하지 못한 패턴을 포착할 수 있다.

그러나 바로 이 점이 AI를 이전의 기술혁명과 구분 짓는다. 경험주의의 기계화가 충분한 규모에 도달하면, 기계는 인간이 이해할 수 없는 과정으로 참인 것을 산출하기 시작한다. Kissinger

등은 이를 “인류의 어떤 다른 시대에도 낯선 방식으로 지식과 이해를 분리하는 것”이라고 진단한다(Kissinger et al. 2024, 46). AlphaFold2가 높은 정확도의 단백질 구조예측 성능을 보였고(Jumper et al. 2021), 이후 AlphaFold DB가 2022년 2억 개가 넘는 예측 구조를 공개했을 때(EMBL-EBI 2022), 그 결과는 검증 가능하지만 그 과정은 인간의 주관적 경험 밖에 있다. 계몽주의 이래 500년간 ‘내가 재현할 수 없으면 참이라 알 수 없다’는 원칙 위에 세워진 지식 체계에 대해, AI는 ‘재현할 수는 없지만 참인 것’이라는 새로운 범주를 만들어내고 있다.

Hui(2026)는 이 변환을 ‘흙 기계(Humean machine)’라는 개념으로 정리한 바 있다. 현재의 AI, 특히 거대언어모델은 추상적 추론 기계보다는 경험주의적 패턴 추출 기계, 즉, 천문학적 규모의 경험 데이터에서 연관 패턴을 찾아내 재귀적으로 압축 및 갱신하는 시스템에 가깝다는 것이다. 따라서, 이 기계가 산출하는 지식의 질은 합리주의적 공리나 천재적 설계자의 통찰보다는 얼마나 광범위하고 다양한 데이터에 접근할 수 있는가에 의해 일차적으로 좌우된다. AI 시대의 지식 생산이 글로벌 오픈소스 모델, 다국적 연구자 피드백, 다 언어, 다 영역 데이터, 외부 컴퓨팅 인프라에 대한 접근 — 요컨대 글로벌 지식 네트워크와의 연결성 — 위에서 비로소 의미 있게 작동한다는 점은 이 글의 분석에서 중요한 출발점이 된다. 한 국가가 자국 영토 내의 데이터를 통제하는 것은 가능하지만, 이 연결성을 봉쇄하면서 동시에 글로벌 수준의 지식 생산을 따라가는 일은 현실적으로 쉽지 않다. 광명망 인트라넷, 외부 정보 차단, 자기검열적 데이터 환경, 국제 공저의 협소한 채널로 특징지어지는 북한의 폐쇄적 구조는 — 다음 절에서 좀 더 구체적으로 살펴보듯 — 흙 기계가 작동하기 위한 이 전제와 부합하기 어렵다. 글의 서두에서 제시한 ‘데이터 통제와 지식 생산 통제의 구분’이 정책적 의미를 얻는 것도 바로 이 지점이다.

이 전환은 지식 생산의 사회적 구조도 바꾼다. 산업화 시대의 지식 생산은 대학, 국책연구소, 대기업 R&D를 정점으로 한 톱 다운(top-down) 구조였다. AI 시대에는 분산된 행위자들이 글로벌 오픈소스 모델, 공개 데이터, 클라우드 컴퓨팅에 접근하여 자기 영역의 문제를 풀고, 그 결과가 다시 전체 모델에 환류되는 바텀업(bottom-up) 구조로 무게중심이 이동한다(백서인 2026). 이 구조의 결정적 특성은, 연결성과 데이터 품질이 혁신의 한계를 결정한다는 점이다. 망 및 보안 분리로 외부 모델과의 연결이 차단된 폐쇄적 구조에서는 데이터 품질과 피드백 루프 자체가 약화되어 혁신의 천장에 빠르게 봉착한다 — 이 패턴은 한국의 일부 기업·기관에서도 이미 관찰되는 현상이며, 폐쇄망을 체제 원리로 삼는 경우 그 한계는 훨씬 더 구조적이다.

지식 생산 양식의 전환은 일반인의 일상에 도달하는 분야에서 가장 직관적으로 확인된다. 농업 분야에서는 AI가 위성, 기상, 토양 데이터를 통합하여 개별 필지 단위의 영농 의사결정을 지원하고 있으며, Digital Green의 Farmer.Chat은 새로운 농법을 적용하는 비용을 농민 1인당 약 35달러에서 1달러 미만으로 낮췄다고 보고한다(Singh et al. 2024; Digital Green n.d.). 의료 분야에서는 AlphaFold가 단백질 구조예측 단계를 크게 단축해 초기 신약발굴을 가속하고 있다. MedGemma는 단일 GPU 운용과 모바일 적용 가능성이 있는 의료 AI 개발용 기반모델로, 이미지 판독 등 진단·치료에 있어 AI 도구로 활용되고 있다(Golden & Pilgrim 2025; Google n.d.). 교육 분야에서는 학습자별 인지 패턴에 적응하는 개인화 튜터링이 확산되며, 이는 ‘교사 한 명당 학생 30명’이라는 산업화 시대의 자원 제약을 부분적으로 우회할 가능성을 보여준다. 이 세 분야가 시사하는 것은, AGI 이전과 이후에 지식이 일상에 분배되는 양상 자체가 질적으로 달라진다는 점이다. 나아가 농업, 의료, 교육은 국가의 생산력 관리와 노동력 재생산에 직결되는 사회적 거버넌스의 핵심 영역이며, 따라서 이 변화는 곧 거버넌스 능력의 재정의의 의미를 지닌다.

이 전환은 국력의 측정 기준까지 바꿀 수 있다. 경제, 사회, 정치적 발전의 근간이 되는 과학기술 혁신 자체가 AI에 의존하게 되고, 나아가 AI 스스로가 혁신을 추동하는 환경이 형성되고 있다. Co-Scientist가 최소한의 배경정보만으로 세균의 유전자 전달 메커니즘에 관한 가설을 생성하고, 독립적으로 진행 중이던 유전체 및 실험 연구의 발견과 2일 만에 일치시킨 사례(Gottweis et al. 2026), AlphaEvolve가 4×4 복소행렬 곱셈의 기존 최선 알고리즘을 56년 만에 개선한 사례(Novikov et al. 2025)는 지식 생산의 속도와 방식이 변화하고 있음을 보여준다. 박사 및 전문가 수준의 다분야 추론을 측정하기 위해 2025년 초 출시된 ‘Humanity’s Last Exam’ 벤치마크의 최상위 모델 정답률이 1년 만에 10% 이하에서 38.3%로 상승한 점 또한 특정 추론 벤치마크에서 모델 성능이 빠르게 향상되고 있음을 보여준다. 이러한 성능 향상의 기저에는 인프라 측면의 가속이 자리한다. 글로벌 AI 컴퓨팅 용량은 2022년 이래 연 3.3배 확장되었으며(Stanford HAI 2026, 29, 95), 이는 흠 기계 작동의 물리적 전제가 한 세대 안에 빠르게 두터워지고 있음을 보여준다. AI 시대의 지식질서에서는 연구 인력이나 예산의 격차보다 AI 도구에 대한 접근성과 지식 생산의 인식론적 조건 자체가 핵심 변수가 된다. UNDP(2025)가 산업혁명의 ‘대분기(Great Divergence)’에 비유하며 ‘다음 대분기’를 경고하는 것은 이 맥락에서다.

물론 “북한은 현재 LLM도 자체 개발하지 못하는 수준인데 AGI나 지식질서를 논하는 것이 다소 성급한 것 아닌가”라는 반론도 가능하다. 그러나, 지식력이 군사력 및 경제력의 토대로 자리잡아 가는 환경에서는 AGI 단계에 도달하지 못하는 국가가 지식 생산 능력의 누적 격차로 인해 군사·경제 영역의 실질 통치 기반 자체가 잠식될 가능성을 배제하기 어렵다. 즉, AGI는 단순한 선택적 지향보다는 장기적으로 외면하기 어려운 압력으로 작용할 수 있다. 북한이 현재 그 단계에 미치지 못한다는 사실은 이 글의 전제일 뿐 논의의 무용성을 의미하지는 않으며, 오히려 그 격차가 어떤 구조에서 발생하고 어떻게 자기강화되는가가 이 글의 분석 대상이다.

2. 북한이 지식질서에서 구조적으로 뒤쳐질 수밖에 없는 이유

문제의 핵심은 기술 역량의 물리적 한계보다, AI가 지식 생산 인프라로 작동하기 위한 인식론적 조건이 북한의 통치 논리와 잘 양립하기 어렵다는 데 있다. 앞 절에서 살펴본 AI 지식질서의 네 가지 특성 — ① 글로벌 데이터, 모델 및 연구 네트워크에 대한 개방적 접근, ② 자유로운 시행착오와 피드백 루프, ③ 분산적 행위자들의 바텀업 혁신, ④ 재현 불가능한 결과를 수용하는 지식론적 유연성 — 은 모두 수령 체제와 폐쇄망을 핵심 통치 원리로 삼는 북한 체제의 작동 방식과 어긋난다. 따라서 북한이 지식질서에서 뒤쳐지는 것은 단순한 기술 부족의 문제보다 체제 원리 차원의 문제로 볼 수 있다. 이하에서는 그 구조적 요인을 세 가지로 정리한다.

첫째, 오픈 액세스 거부와 글로벌 지식 생산으로부터의 배제. 앞서 본 흙 기계적 작동 — 인접성·유사성·인과성에 따른 대규모 패턴 추출 — 은 Hui(2026, 10)가 설명한 머신러닝의 경험주의적 성격을 가리킨다. 이를 북한 사례에 적용하면, 대규모 패턴 추출이 의미 있는 결과로 이어지기 위해 글로벌 모델, 데이터 및 연구자 네트워크와의 개방적 연결이 중요하다는 것이 이 글의 분석적 확장이다. 그러나, 북한은 수령 체제 옹위라는 통치 원리상 이 연결성 자체를 받아들이기 어려운 위치에 있다. Baek(2024)은 북한의 국산 스마트폰이 약 700만 대 보급됐고 Mirae Wi-Fi 이용자가 인구의 약 20%라고 추정하지만, 이들 기기와 광명망은 글로벌 인터넷과 단절된 폐쇄 환경에서 운용된다. 글로벌 AI 모델이 인류 전체가 생산한 텍스트, 이미지, 코드를 훈련 데이터로 활용하는 것과 대비하면, 광명망 내부의 데이터는 양적·질적으로 비교가 어렵다. 이러한 폐쇄성은 단순한 기술적 결함이 아닌 체제 유지를 위한 의도적 설계의 산물이며, 이 때문에 인프라

보강만으로 해소되기 어렵다. 오픈 액세스를 거부하는 한 북한은 AI 시대의 지식 폭발에서 점진적으로 멀어질 수밖에 없다.

이러한 폐쇄성에 더해 북한의 낙후된 경제는 AI 개발을 선도하는 국가들과의 격차를 더욱 벌리고 있다. 인프라 측면에서 AI Index 2026에 따르면 미국에는 AI 데이터센터 5,427개소가 집계되어 2위국 대비 10배 이상의 격차를 보인다. 다만, 시설 수는 개별 센터의 크기, 용량 및 이용률을 반영하지 않는다. 북한의 GPU 자원은 가용 규모를 외부에서 추정하기 어려워, 흠 기계 작동의 물리적 전제가 매우 제한된 것으로 보인다. 투자 측면에서 2025년 AI 민간투자는 미국이 \$285.9B, 중국이 \$12.4B로 약 23배의 차이를 보였다(Stanford HAI 2026, 31, 182). 사용자 측면에서 Anthropic Economic Index(2026.1)는 Claude 사용과 1인당 GDP 사이의 강한 양의 상관관계를 보고했고, 별도의 2026년 3월 보고서는 인구보정 사용량 상위 20개국 비중이 45%에서 48%로 높아졌다고 분석했다(Anthropic 2026.3). 1인당 GDP가 1,000달러대에 머무는 것으로 추정되는 북한은 설령 외부 접근이 허용되더라도 산술적으로 의미 있는 사용자 기반을 형성하기 어려운 위치에 있다.

둘째, 톱 다운 통치 논리와 바텀업 지식 구조의 충돌. AI 시대의 지식 생산은 분산된 행위자들의 자율적 시행착오를 통해 작동하며, 이 행위자들은 실패를 자유롭게 공유하고, 비표준적 가설을 시도하며, 권위에 의해 사전 승인되지 않은 결론에 도달할 수 있어야 한다. 이 작동 양식은 북한이 지난 70여 년간 정착시켜 온 톱 다운 지식 통제 구조 — ‘주체사상에 의한 사상의 단일화’, ‘당의 유일사상체계 확립의 10대 원칙’, 그리고 김정은 시대의 ‘5대 교양’ — 와 정면으로 충돌한다. 기술적 인프라를 갖추는 것보다, 사상, 문화, 기술 등 ‘3대 혁명 역량’ 노선이 전제하는 위계적 지식 생산 구조 자체를 어떻게 재편할 것인가가 더 깊은 차원의 문제다. 이는 북한 체제의 정체성 위기와 직결되므로, 단기적·점진적 해소가 어렵다.

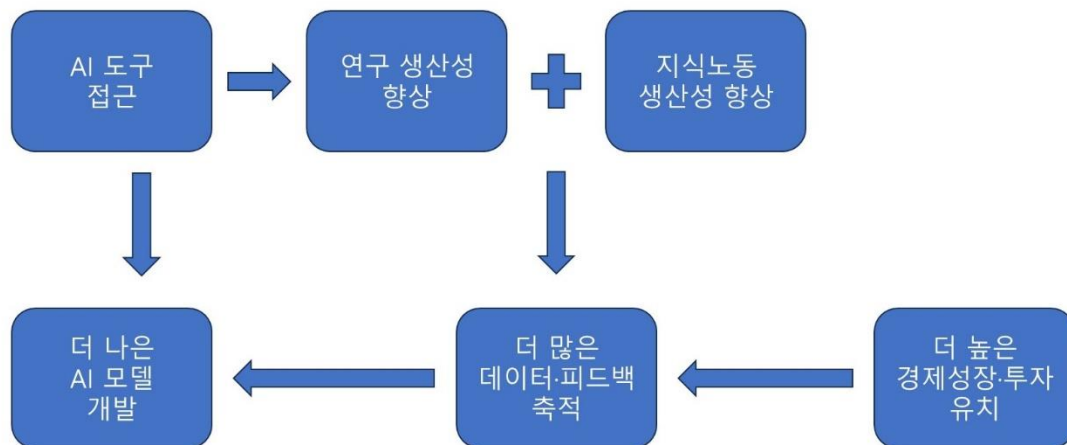
셋째, 인지-구현 괴리. 북한은 AI의 중요성을 분명히 인식하고 있다. 2024년 이후 로동신문의 AI 관련 보도가 급증했고, 제9차 당대회(2026.2)는 새 에너지, 우주, AI 같은 새로운 첨단산업 분야의 핵심기술 연구개발을 과업으로 제시했다(최현규 2026). 2026년 1월 로동신문은 김일성종합대학 인공지능학부장 한석민의 이름으로 심층학습과 생성형 AI의 기술적 특성을 해설했고(강진규 2026a), 3월에는 “인공지능을 사용할 줄 아는 연구사들의 생산성은 거의 2배로 올랐다”는 글로벌 수치를 인용하며 AI 활용 능력을 “과학기술인재가 갖춰야 할 필수적 자질”로 규정했다(강진규 2026b). 그러나 같은 시기 확인 가능한 실제 산출물은 이 담론의 수준에 미치지 못한다. Harold et al.(2022)은 공개 증거에 비추어 북한의 AI 연구가 비교적 초기 단계라고

평가하며, Kim(2024b)의 2017~2023년 분석에서도 북한의 잠재적 AI 관련 논문은 160개국 중 145위였다. 김일성종합대학 인공지능연구소가 “GPT 기술을 이용해 사람들의 정신노동까지 대신하고자 하는 높은 목표를 세웠다”고 밝혔으나 실제 구현된 기술은 Caffe 기반 새 퇴치와 CNN-LSTM 미세먼지 예측 같은 좁은 범위의 민간 응용 기술이다(북한ICT연구회 외 2025).

반면, 글로벌 AI 연구에서는 에이전트 시스템을 활용한 반자율적 과학 발견 시스템까지 출현하고 있다(Xin et al. 2025). 결과적으로 북한은 글로벌 AI의 최전선을 정확히 인지하면서도 구현 역량과의 괴리가 오히려 심화되는 구조에 갇혀 있다. 로동신문(2025.8.20)이 반도체 수출통제를 “경쟁적수들의 고도기술발전을 억제하려는” 것으로 비판한 것은 이 구조적 차단에 대한 위기의식의 표현으로 읽힌다.

3. 자기강화적 격차와 ‘지식종속’

이 괴리가 특히 우려되는 것은 AI 격차의 자기강화적 성격 때문이다.



이 순환에 참여하는 국가에서는 생산성 향상이 AI 역량 강화로 다시 이어지는 반면, 단절된 국가는 이 순환 자체에 진입하기 어렵다. AlphaEvolve는 Gemini 학습에 사용되는 특정 행렬곱 커널을 23% 가속했고, 이는 전체 학습시간을 약 1% 줄였다(Novikov et al. 2025). 이 사례는 AI가 AI 개발 과정의 일부를 개선하는 재귀적 루프가 현실화되고 있음을 보여준다. 제한된 소프트웨어 및 연구 과제에서 AI 에이전트의 50% 성공 시간지평은 약 7개월마다 두 배로 늘어났다(Kwa et al. 2025).

149개 모델의 복합 벤치마크 회귀에서는 최근 2년의 ECI 추세기울기가 이전 기간보다 약 1.85배 빨라졌다(Edelman & Lee 2025). SWE-bench Verified 등 코드 수정 벤치마크의 성능도 빠르게 향상되고 있다(Stanford HAI 2026, 75). 이는 소프트웨어 개발 과정의 일부를 자동화하는 역량이 강화되고 있음을 보여주는 보조 지표다. 이 순환에 참여하지 못하는 국가와의 격차는 선형이 아닌 지수적으로 벌어질 수 있다.

자기강화의 또 다른 한 축은 진입장벽의 비대칭적 상승이다. 프런티어 모델의 훈련 비용은 2016년 이후 연 2~3배 증가해 왔으며, 이 추세가 지속되면 2027년경 단일 훈련 비용이 10억 달러를 넘어설 가능성이 있다(Cottier et al. 2024). 인프라 측면에서도 Grok 4의 훈련 단계 탄소배출이 약 72,800톤, 글로벌 AI 데이터센터 전력 용량이 29.6GW(뉴욕주 피크 전력급 수요)에 이르는 등 차세대 모델 개발 단계의 자원 임계가 빠르게 상승하고 있다(Stanford HAI 2026, 14). 주목할 점은 추론(inference) 비용은 모델 및 하드웨어 최적화로 빠르게 하락하는 반면, 훈련(training) 단계의 자원 요구는 오히려 더 집중되는 비대칭이 누적되고 있다는 사실이다. 이는 오픈소스 모델의 활용은 점차 보편화되더라도 차세대 모델을 개발해 자기개선 루프에 진입하는 일은 더욱 소수 행위자의 영역으로 축소됨을 의미하며, 후술할 오픈소스 우회 변수의 구조적 한계와도 직결된다.

북한은 이 격차를 중국과 러시아 의존으로 보완하려 하고 있다. 그러나 이 경로는 단순한 ‘기술 격차’를 넘어 ‘지식종속(Knowledge Dependency)’ — 한 국가의 지식생산역량이 특정 외부 행위자에 구조적으로 종속되어 자율적 기술발전 경로가 제약되는 상태 — 으로 전환될 위험을 안고 있다.

지식질서의 관점에서 이 종속이 문제가 되는 것은 무엇보다도 중국산 AI 생태계에 편입될 경우 데이터 주권과 기술발전 방향이 중국의 영향 하에 놓일 수 있다는 것이다. Kim(2024b)의 광범위 키워드 검색은 북한 국내지를 제외한 2017~2023년 잠재적 AI 관련 논문 161편을 식별했으며, 이 가운데 중국 기관 참여 논문은 약 70편이었다고 보고하여, 이미 중국편향의 지식질서에 북한이 포섭되고 있는 모습을 보여준다. 더욱이 중국 AI 생태계 자체가 가지고 있는 한계도 있는데, 중국의 중앙집권과 검열이 실험과 혁신의 순환을 늦출 수 있기 때문이다(Dutt 2025). 중국편향에서 벗어날 수 있는 러시아 경로의 한계도 명확한데, 중국의 AI 산업 발전 속도에 비해 러시아는 아직 큰 성과를 보여주지 못하고 있으며, 중러 사이의 협력도 미미한 수준이다. Konaev et al.(2021)이 2010~2019년 영문 AI 논문을 분석한 결과, 중러 공동 AI 논문은 중국 전체의 0.1% 미만이었고 미중 AI 연구 협력은 중러의 129배에 달했다. 또한, Hughes & Tony A(2026)는 CRINK

국가들의 직접적인 다자 AI 정렬 증거는 확인되지 않으며 초기 단계의 양자 협력 징후만 관찰된다고 평가하여, 중국 경로를 우회할 다자채널 또한 한계가 있음을 보여준다.

여기에는 또 다른 차원의 종속 위험이 더해진다. 중국산 AI 모델 및 하드웨어에 대한 의존은 북한 주민의 데이터, 언어, 생체 정보가 중국 측에 유출·접근 가능해질 위험, 중국의 전략적 판단에 따른 공급 중단 리스크, 그리고 핵심 기술을 공유하지 않는 중국 측의 협력 관행으로 인한 자주적 기술발전 경로의 제약(백서인 2026)을 동시에 수반한다. 주체사상을 국가 정체성의 핵심으로 내세우는 북한에게 AI를 단일 외부 행위자에 대한 의존으로 도입하는 경로는 역설적으로 통치의 자율성을 약화시킬 수 있다. 다만, 다자 매개의 커먼스 형식으로 받아들이는 경로는 단일 의존을 분산시키면서 자주성을 새로운 형태로 재정의할 여지를 함께 연다(마지막 절의 커먼스 논의 참조).

Barasa et al.(2026)는 ‘AI 주권’이 완전한 자급자족이 아닌 ‘전략적 자율성’을 의미하며, 이마저도 글로벌 생태계 참여가 전제라고 분석한다. 국제제재와 자발적 네트워크 단절이 중첩된 북한에게는 이 전제 자체가 충족되지 않는다. 여기에 한 가지 환경적 변수가 더해지는데, AI Index 2026이 집계한 Foundation Model Transparency Index의 평균 점수가 1년 만에 58점에서 40점으로 하락한 사실이다(Stanford HAI 2026, 164). 이는 학습 데이터, 컴퓨팅, 역량, 위험성에 대한 정보공개 투명성이 낮아졌음을 뜻한다. 그 결과 후발국이 검증 가능한 정보에 접근하기 어려워지는 환경은 진입장벽을 높일 수 있으며, 북한의 고유한 폐쇄성과 중첩될 경우 지식종속 위험을 강화할 가능성이 있다. 이 점은 대북 AI 협력의 정책 환경을 평가할 때에도 함께 고려할 필요가 있다.

다만, 이 격차를 부분적으로 상쇄할 수 있는 변수가 하나 존재한다. DeepSeek R1을 비롯한 고성능 오픈소스 AI 모델의 확산이다. 오픈소스 모델은 자유롭게 다운로드, 수정 및 배포가 가능하며, 조달 흔적이 남지 않고 제재 체제를 구조적으로 우회한다. 최근에는 도메인 전문지식 없이 오픈소스 모델과 공개 데이터만으로 특정 산업 영역에서 빠르게 성과를 내는 사례가 글로벌하게 축적되고 있다. 북한이 이 경로를 통해 농업, 의료 및 교육 등 민간 영역에서 AI 활용 역량을 일정 수준까지 끌어올릴 가능성을 배제할 수는 없다. 그러나 이 경로 역시 근본적 한계를 안고 있다. 오픈소스 모델의 활용은 가능하더라도, 자체 데이터로 모델을 개선하고 차세대 모델을 개발하는 재귀적 순환에 진입하기 위해서는 컴퓨팅 인프라, 양질의 데이터, 글로벌 연구 네트워크에 대한 접근이 여전히 필요하다. 오픈소스는 격차를 줄이는 도구가 될 수 있으나, 격차를 해소하는 구조적 해법은 아닌 셈이다.

4. 정보통제 비용의 변화 전망

단기적으로(2025~2030), AI 기반 감시의 정교화와 드론 및 무인체계의 군사적 활용은 체제 통제력을 일정 수준 강화할 수 있다. 그러나 중기 이후에는 세 가지 변화가 이 균형을 흔들 수 있다.

첫째, 에이전트 AI 시대에 광명망 인트라넷의 양립 불가능성이다. 에이전트 AI는 실시간 데이터 피드, 외부 시스템 연동 인터페이스(API), 클라우드 인프라 등 개방된 디지털 환경을 전제한다. 챗봇 시대에는 오픈소스 모델을 다운로드하여 오프라인에서 활용하는 ‘기생적 전략’이 부분적으로 가능했으나, 에이전트 AI는 이러한 폐쇄적 인프라와 구조적으로 양립하기 어렵다. 실제 운영체제 환경에서 AI 에이전트의 업무 자율 수행 능력을 측정하는 OSWorld 벤치마크 성공률이 2024년 12%에서 2026년 초 66%로 1년여 만에 다섯 배 이상 상승한 사실(Stanford HAI 2026, 9)은, 에이전트 AI가 점진적으로 도래할 미래라기보다 이미 측정 가능한 현실에 가까움을 보여준다.

둘째, 생성형 AI의 발전이 외부 정보 유입 경로를 다변화할 가능성이다. AI가 텍스트·음성·이미지를 자유롭게 생성 및 변환하는 환경에서, 기존의 물리적·전자적 정보 차단 방식의 실효성이 약화될 수 있다. 이 경우 체제 유지를 위한 억압 비용과 AI를 통한 경제적·군사적 이득 사이의 균형이 달라질 수 있다.

셋째, AI 활용 양식의 ‘증강(augmentation)’으로의 이동이 폐쇄망에 가하는 추가 압력이다. Anthropic Economic Index Economic Primitives는 Claude 사용에서 증강 방식이 52%, 자동화 방식이 45%였다고 보고한다(Anthropic 2026.1, 5). 별도의 Learning Curves는 6개월 이상 사용자의 과업 성공률이 신규 사용자보다 완전 통제 모형에서 6.4% 높은 상관을 보였다(Anthropic 2026.3, 15). 이 흐름이 시사하는 바는 폐쇄망이 차단하는 것이 단지 모델 접근만이 아니라 사회 전체가 누적해 가는 사용 경험과 메타지식 자체라는 점이다. 챗봇 시대에 일부 가능했던 ‘모델 다운로드 후 격리 활용’이라는 우회 경로는 AI 활용 능력 자체가 사회의 경험 자산으로 누적되는 환경에서는 점차 차선책에 머무를 가능성이 크다.

장기적으로는 세 가지 경로를 상정할 수 있다. 현상유지 시 AI 격차의 확대와 함께 중국 의존이 심화되어 지식종속의 위험이 증대된다. 제한적 개방에서는 DeepSeek 사례가 생성형 AI와 엄격한 표현통제가 기술적으로 양립할 수 있음을 보여주지만(Kerley 2025), 제한적 개방을 장기적으로 유지하는 정치 및 사회적 비용은 별도의 문제다. 구조적 개혁은 글로벌 지식네트워크에 재 연결하는 가장 효과적인 경로이나, 핵무장, 왕조적 체제, 극단적 정보통제라는 북한의 특수성을 고려하면 가장 실현 가능성이 낮은 시나리오다.

여기서 한 가지 부가적 고려가 필요하다. AI는 단지 격차를 벌리는 변수에 그치지 않고, 장기적으로 북한 권위주의 체제의 정치 질서 변화에 새로운 계기를 제공할 잠재 변수가 될 수도 있다. 첫째, 생성형 AI가 정보 차단 비용을 비대칭적으로 끌어올리면, 정권은 통제 강도와 경제적 이득 사이에서 새로운 절충점을 찾도록 강제될 수 있다. 둘째, 단일 외부 행위자(중러)에 대한 지식중속이 심화될수록 자주성 개념에 대한 재정의 압력이 발생할 수 있으며, 이는 다자 매개의 분산 협력 — 이 글에서 검토하는 커먼스 형식 — 으로 자주성을 재구성하는 새로운 노선의 정치적 공간을 열 수 있다. 셋째, AI 기반 행정·경제 운영을 도입하는 과정에서 형성되는 새로운 기술 엘리트 계층은, 기존의 사상·정치 위계와 다른 권력 자원을 축적할 수 있다. 이는 결정론적 예측이 아닌, AI 시대의 정책 환경에서 추적해야 할 잠재 변수로 남겨둘 항목이다.

III. AI와 북한의 통치, 감시, 사회관리 문제

1. 북한의 AI 감시 및 통제 역량 실태

북한이 지식질서 전환에서 뒤처지고 있는 것은 사실이지만, AI를 활용한 감시 및 통제 영역에서는 일정한 성과를 거두고 있다. 이러한 편향된 활용이 보다 근본적인 지식질서 변동에는 오히려 부정적으로 작용할 가능성이 크지만, AI가 추동하는 북한의 지식 및 거버넌스 변동을 이해하기 위해서는 실제 작동 중인 감시 및 통제 역량을 먼저 짚어 둘 필요가 있다. 북한이 확인 가능한 수준에서 AI를 감시에 활용하고 있는 분야는 안면인식, 생체인식, 컴퓨터 비전이다. 2019년 정보기술성과전람회에서 김일성종합대학이 시연한 ‘담보’는 건물 출입 통제용 얼굴인식 시스템으로, 접근자의 얼굴을 인식해 출입문을 여는 기능이 북한 매체에 소개됐다(Williams & Slavney 2024, 52).

이와 별도로 김일성종합대학 첨단과학기술연구원이 개발한 순안공항용 얼굴인식 시스템은 얼굴 특징과 성별, 나이를 1초 이내에 식별한다고 북한 매체가 주장했으며, 2019년 순안공항에서 사용된 것으로 보도됐다(Williams & Slavney 2024, 57). 이 성능들은 독립시험 결과가 아닌 북한 매체·기업 주장이다. Daily NK의 익명 소식통 보도에 따르면 2021년 함경남도과 평양 등 주요 도시의 기차역에 얼굴인식 장비가 설치됐고, 2024년 당시 평양의 한 역에 설치된 3대는 작동했지만 함경남도 여러 역의 장비는 고장 난 상태였다(Kim Jeong Yoon 2024). 이러한 자동화는 수동 검사 과정의 뇌물·재량을 줄일 수 있지만, 이를 제거했다고 단정하기는 어렵다.

스마트폰 영역에서는 만경대정보기술사가 개발한 진달래 6·7 시리즈(2020)에 심층신경망 기반 지문, 음성, 얼굴, 텍스트 인식 기능이 탑재되었다(Kim 2024a). 압록강기술개발무역회사는 심층신경망 기반 지능형 IP 카메라와 보안 감시 시스템을 개발·판매한다고 북한 매체를 통해 주장해 왔다(Kim 2024a).

국경 감시에서는 중국 기술을 활용한 협력 정황이 보도되고 있다. NK Insider(2025)는 단일 익명 소식통을 인용해 중국제 장비가 북한 국경 초소에 설치되고 양국 지휘부 사이에 관련 정보가 공유된다고 보도했다. 다만 전 국경 배치 완료와 실시간 데이터 공유는 독립적으로 확인되지 않았다.

INSS 전략보고 No. 350(김민정 2025)은 감시 통제에 활용될 수 있는 공개자료에 나타난 네 가지 연구를 검토했다. ① 저조도·저해상도 조건의 안면인식(국가과학원 수학연구소), ② 다중 객체 추적 알고리즘(리과대학 인공지능기술연구소), ③ 모바일·저성능 환경을 겨냥한 경량 음성합성(국가과학원 수학연구소), ④ 지역별 언어 특성을 구별하는 억양 식별(리과대학 인공지능기술연구소)이다. 연구들은 제한된 연산 및 데이터 환경을 겨냥한 경량화 구조에 초점을 둔다.

2. 중국·러시아로부터의 기술 유입과 ‘기술 종속의 딜레마’

이 감시 역량은 중국 기술에 크게 의존한다. Williams & Slavney(2024, 58)는 압록강기술개발무역회사가 판매하는 DT-20 안면인식 장치가 중국 Shenzhen Wit Easy 제품으로 보인다고 추적했다. 기술 유입의 경로는 다음과 같이 구분할 수 있다. ① 감시 카메라 및 스마트폰 부품 등 상업적 하드웨어 수입 — ASIAPRESS의 내부 취재협력자 보도에 따르면 중국제 카메라는 2024년부터 설치됐고 2025년 8월 말 이후 확대 정황이 나타났다(Ishimaru & Kang 2025). ② 중국 국경 인접 대학과의 학술 공동연구 — 2017~2023년 잠재적 AI 관련 논문 161편 가운데 중국 기관 참여 논문이 약 70편이었다(Kim 2024b). ③ NK Insider가 보도한 국경 AI 경보 시스템 — 단일 익명 소식통에 근거해 독립 확인이 필요하다. ④ DeepSeek R1 같은 중국산 오픈소스 AI 모델의 활용 가능성 — 북한이 실제로 다운로드 및 활용했다는 직접 증거는 아직 확인되지 않았다. ⑤ 러시아 파견 — 2025년 7월 김일성대 인사가 AI 연구자 파견이 진행 중이라고 밝혔다고 NK News가 보도했으나 실제 도착과 규모는 독립적으로 확인되지 않았다(Kim 2025).

이처럼 중국과 러시아가 가장 큰 기술 전파의 허브가 되고 있는데, 여기서 하나의 역설이 나타난다. 감시기술 전반에서 중국 혹은 러시아가 독점적인 공급자가 된다면, 북한이 이들의 기술 네트워크를 통해 AI 감시 역량을 확보할수록, 체제 유지의 기술적 기반 자체가 이들 두 국가에 종속될 위험이 커진다.

이 종속이 안고 있는 정치적 리스크 — 데이터 주권의 약화, 공급 중단 리스크, 핵심 기술 비공개로 인한 자주적 발전 경로의 제약 — 는 앞서 지식종속을 논의하며 이미 정리한 바 있다. 다만 그 영역에서의 종속이 지식 생산의 장기적 혁신 역량을 잠식하는 간접적 위험이라면, 본 절에서의 종속은 보다 직접적이다 — 안면인식, 국경 경보 및 생체정보 처리 등 체제 유지의 핵심 도구 자체가 외부 행위자의 전략적 판단에 좌우될 수 있다는 점에서다. 즉, AI 권위주의를 추구하는 행위 자체가 외부 종속을 심화시키는 ‘기술 종속의 딜레마’가, 감시·통제 영역에서는 보다 직접적인 형태로 드러나는 셈이다.

IV. 북한의 AI 지식-거버넌스 시나리오: 교육, 연구, 의료

본 절은 앞의 두 절에서 검토한 북한의 구조적 후행과 통치 및 감시 차원의 분석 위에서, AI 시대의 지식-거버넌스 변동이 교육, 연구, 의료 세 도메인에서 북한과 관련해 단·중기에 어떻게 전개될 수 있는지를 도메인별로 살펴본다. 이 세 영역은 국가의 생산력 관리와 노동력 재생산에 직결되는 사회적 거버넌스의 핵심이며, 군사·경제 영역과 가장 깔끔하게 분리되는 지식 및 거버넌스 고유의 영토이다. 북한 헌법은 2026년 3월 23일 최고인민회의에서 개정됐고, 영토 조항 신설과 통일 관련 조항 삭제를 통해 두 국가 노선이 간접적으로 반영됐다(차두현·한기범 2026). 직접적 남북 양자 협력이 차단된 현재의 환경에서 협력 경로는 ① 다자기구(UNICEF, UNESCO, WHO, WFP 등 UN 체계의 기구, Gavi·Global Fund 같은 국제 파트너십, IAU·WMO 등 국제 학술·전문기구), ② 글로벌 사우스 채널(가나, 나이지리아, 인도 등에서 검증된 저자원 AI 시범 사례의 한반도 적용), ③ 제3국 학술·시민 네트워크(Royal Society·SISSA·오슬로대·Eugene Bell Foundation 등 비미·비중 양자 트랙) 가운데 하나 또는 둘 이상의 조합을 경유한 우회 형태로 설계될 필요가 있다. 한국의 글로벌 AI 포지션과 대북 협력 전략의 종합적 검토는 이 글의 범위를 넘어서며, 협력 가능성과 한계의 종합은 마지막 절에서 이어 다룬다.

1. 교육·인적 역량 분야

북한 교육의 현재 모습은 ‘제도는 있으나 측정이 어려운’ 상태로 요약할 수 있다. 2012년 법제화된 12년제 의무교육이 2017년 전국 시행되었으나(NK News 2017), UNESCO UIS 국가 프로필에는 학습성과와 교육의 질에 관한 지표가 광범위하게 비어 있다(UNESCO UIS 2025). 이러한 조건에서 측정 가능한 협력의 출발점은 데이터 수집 그 자체가 될 수밖에 없으며, UNICEF DPRK가 평양 중앙통계국과 공동 수행한 2017 MICS는 그 선례로 참고할 만하다(UNICEF DPRK 2018). 글로벌 사우스 AI 교육 사례 중 광명망 환경에 이식 가능성을 검토할 수 있는 사례는 Rising Academies의 WhatsApp 기반 수학 AI 튜터 Rori다. 약 500명을 대상으로 한 예비평가에서 효과크기 0.36이 보고됐고, 기본 휴대전화와 저 대역폭 네트워크에서 작동했다(Henkel et al. 2024). 세계은행 연구는 나이지리아 9개교에서 교사의 지원 아래 GPT-4 기반 Microsoft Copilot을 활용한 6주 방과 후 수업을 무작위 평가했으며, 0.31 SD, 즉, 통상적인 학습량 약 2년분에 해당하는 효과를 보고했다(De Simone et al. 2025). 이는 교사 지원형 프로그램의 맥락적 근거이며, 챗봇 단독 효과나 광명망 내 한국어 SLM의 직접적인 효과 상한으로 해석할 수 없다. 한국 측에서는 NIA 정보접근센터가 2025년 세계은행 소개 기준 46개국 61개소에서 운영된 경험을 참고할 수 있다(NIA n.d.; World Bank 2025). UNICEF·UNESCO를 매개로 비정치 STEM·문화 콘텐츠에 한정한다면 사상교양과의 충돌을 줄일 가능성이 있다. 다만 광명망 폐쇄성과 콘텐츠 검열을 감안하면, 협력은 비대화형 자동 채점·STEM·문화 등 비교적 협소한 기능에서부터 단계적으로 출발하는 것이 현실적이다.

2. 지식·연구·학술 분야

앞서 살펴본 인지-구현 괴리에도 불구하고 북한은 2025년 7월 김일성종합대학 정보과학부를 인공지능학부로 격상하면서 AI를 학제적 우선순위로 끌어올렸다(전명훈 2026). 다만 2017~2023년 북한 국내지를 제외하고 식별된 잠재적 AI 관련 논문 161편 가운데 중국 기관 참여 논문이 약 70편이어서, AI 분야 자체에서의 다변화 여지는 크지 않은 것으로 보인다(Kim 2024b). 그러나 이중용도(dual-use) 우려가 상대적으로 적은 비 AI 학술 영역에서는 안보리 제재기에도 작동해 온 다자 채널이 분야별로 형성되어 있어, 한국이 우회 협력의 진입점으로 활용할 수 있는 여지가 적지 않다.

먼저, 자연과학 분야의 작동 사례부터 짚을 수 있다. 김일성대 소속 Chung-Sik Sin, Hyon-Sok Choe, Jin-U Rim은 분수계 미분방정식 분야의 2023년 논문을 Mathematical Methods in the Applied Sciences에 게재했다(Sin, Choe & Rim 2023). 국가과학원 레이저연구소의 김광현(Kim Kwang-Hyon)은 2015~2019년 TWAS Young Affiliate로 활동했다(TWAS n.d.). 2019년 3월에는 이탈리아 외무부 승인 하에 SISSA(Trieste) ↔ 김일성대 정식 교환협정(분야: 계산신경과학; 매개자 Alessandro Treves)이 체결되었다(Abbott 2019). 가장 두드러진 사례는 영국 케임브리지·임페리얼 팀이 DPRK 지진청 및 국가과학원과 공동 진행한 백두산 화산 공동연구로, Royal Society MoU와 AAAS 사이언스 디플로머시가 16건의 영·미·UN 수출허가를 매개하여 Ri Kyong-Song 외 공저로 Science Advances에 게재되었다(Ri et al. 2016).

공중보건 영역에서도 학자 단위 협력이 이어지고 있다. Eugene Bell Foundation·Partners In Health·Stanford·평양보건성이 함께 진행한 다제내성결핵(MDR-TB) 치료 보고가 PLOS Medicine 2016에 실렸고(Seung et al. 2016), 김일성대 평양의학대학의 Ri Hak Chol은 오슬로대 연구진과 The Lancet Planetary Health 2019에 “Drivers of Health Transition in the DPRK”를 공저했다(Ri et al. 2019).

인문 분야에서는 거래말큰사전 남북공동편찬위원회가 2005~2015년 25차 회의를 통해 약 12만 8천 개 표제어의 정의 및 발음에 합의했고(거래말큰사전 남북공동편찬위원회 2005-2015), 개성 만월대 남북 공동발굴(2007~2018)이 단속적으로 이어지며 약 16,500점의 유물이 출토되었다. 한편 만월대가 포함된 「개성 역사유적지구」는 2013년 UNESCO 세계유산에 등재되었다(UNESCO World Heritage Centre 2013).

한국이 활용할 수 있는 우회 설계는 이러한 작동 트랙 위에서 구체화될 수 있다. 한국연구재단(NRF) 글로벌 사우스 협력 사업이나 TWAS-COMSTECH-ICTP-AP(베이징, 2018년 이후) 펠로우십에 한국이 편당하는 좌석을 북한 또는 북한-제3국 공저자에게 우선 배분하는 간접 설계, BGE-M3(Chen et al. 2024) 같은 다국어 임베딩 모델로 북한 공개 학술 코퍼스를 한국어 코퍼스와 동일 의미 공간에 매핑하는 기술적 통로 등을 검토할 만하다. 민간 AI 기업의 비영리 과학 프로그램도 잠재적 자원으로 검토할 수 있다.

Anthropic의 AI for Science Program은 자격을 갖춘 연구자에게 API 크레딧을 제공하는 프로그램이다(Anthropic 2025). 북한 연구자의 실제 이용 가능성은 미국의 제재·수출통제 검토 없이는 단정할 수 없다. 특정 도메인(생명과학, 기후, 천문 및 작물 유전체 등)의 다자 학술

컨소시엄에 한국이 펀딩을 매개해 북한 또는 북한-제3국 공저자가 참여하는 구조를 검토할 수 있지만, 협력 가능 영역은 UN 안보리 1718위원회의 dual-use 해석을 감안할 때 순수수학, 이론물리, 언어학, 고대사, 해양·기후, 공중보건 등 AI를 방법론적 도구로만 결합하는 분야에 한정될 가능성이 크며, 사상교양·정치와 결합되는 콘텐츠는 접근이 매우 제한된다.

3. 의료 및 보건 분야

남북 직접 협력이 차단된 환경에서도 의료 및 보건 분야에서 다자 채널은 일정 수준 작동하고 있는 것으로 보인다. 2024년 7월 UNICEF 지원으로 400만 회분 이상의 백신이 북한에 반입됐으며, 이 가운데 약 200만 회분은 따라잡기 캠페인용, 나머지는 정기 접종용이었다. 같은 해 9월 UNICEF·Gavi·WHO와 북한 보건성은 80만 명 이상의 아동과 12만 명의 임산부를 목표로 전국 210개 군에서 캠페인을 시작했다(UNICEF 2024a; 2024b). 한국 측에는 KOICA-Stop TB Partnership 협약(2024)과 KM-BERT(Kim et al. 2022)·KMedBERT·서울대병원 임상 LLM(KMLE 86%) 등의 한국어 의학 LLM 자원이 갖추어져 있어, 한국이 기술과 자금을 공급하고 WHO 평양사무소(2001년 이래)·UNICEF·Gavi·Global Fund가 북한 내 배치를 담당하는 분업 모델이 실재 인프라 위에서 검토될 수 있다. Global Fund는 2010년 이래 북한 TB·말라리아 사업에 누적 1억 달러 이상을 집행했으며, 2019년 9월 이사회가 4,170만 달러를 승인했다(Global Fund 2020; NK News 2019).

V. 지식 및 거버넌스 부문에서의 남북 협력 가능성과 한계

: AI 커먼스 형성과 거버넌스 변화의 단·중기 전망

세 도메인의 협력 경로는 공통적으로 두 가지 기능 — AI 시대 고립이 초래할 장기 비용의 시그널링과 중러 이외 대안 자원의 가시화 — 을 수행하지만, 그 누적 효과에 시선을 옮기면 단순한 사업 결과 이상의 함의를 발견할 수 있다. 앞 절에서 도메인별로 살펴본 KM-BERT 같은 한국어 의료 텍스트 모델, 다국어 임베딩 모델·다자 학술 펀딩 좌석, 저자원 의료 AI 도구와 그 학습·검증 데이터는 단발적 산출물이 아닌 공유 가능한 자원의 풀을 단계적으로 축적해 가는 흐름으로 읽을 수 있다. 이렇게 다수의 행위자가 함께 만들고 함께 관리하는 공유 자원·공간을 Last(2017)는 시장 논리(이윤)와 국가 논리(통제) 어느 쪽으로도 환원되지 않는 제3의 조직 형식, 커먼스(common)로

개념화한 바 있다. 그에 따르면 인터넷 및 AI 같은 분산 디지털 인프라는 이 커먼스가 물질적으로 가능해지는 매개가 된다. 그의 분석은 글로벌 스케일을 대상으로 했지만, 분단으로 남북 사이의 공유 자원이 가장 결핍되어 있는 — ‘커먼스 갭(common gap)’이 가장 응축된 — 한반도에도 부분적으로 적용해 볼 수 있다.

이 커먼스 축적이 단·중기에 북한 내부 거버넌스에 미칠 수 있는 미세 변화는 앞서 정보통제 비용의 변화를 전망하며 잠재 변수로 식별한 세 가지 — 자주성 개념의 재정의 압력, 정보통제 비용의 비대칭적 상승, 신기술 엘리트 계층의 형성 — 에 비추어 채널별로 정리할 수 있다. 의료 채널에서 진단 및 데이터 표준화와 한국어 의료 텍스트 모델의 운용은 WHO를 통한 북한의 선택적 수용으로 이어질 수 있다. 이는 자력갱생 서사를 폐기하는 것이 아닌 수정 및 확장하는 방향으로 — 단일 외부 의존을 분산시키면서 다자 커먼스 위에서 자주성을 재정의하는 한 형태로 — 작동할 가능성이 있다. 교육 채널에서 광명망 내 한국어 SLM·자동 채점 코퍼스의 누적은 학습자별 데이터 흐름을 만들어 사상 일률성의 운영 비용을 비대칭적으로 상승시킬 수 있다. 학술 채널에서 다국어 임베딩 모델·다자 펠로우십 협력의 누적은 김일성대 인공지능학부·국가과학원 신진 연구진층의 형성을 가속해 기능적 신기술 엘리트의 층위를 두텁게 할 여지가 있다. 다만 이러한 효과는 점진적이며, 단·중기에는 미시적 변화의 누적에 머무를 가능성이 크다. 본격적 체제 변동 시나리오는 이 글의 범위를 넘는 별도의 과제다.

한편 커먼스의 축적은 그 기능적 관리 필요성 — 데이터 표준, 접근 권한, 품질, 윤리 — 을 동반하므로, 이를 조정하는 남북 및 다자의 거버넌스 협의체가 점진적으로 등장할 가능성도 함께 검토할 수 있다. Last(2017)가 정리한 Ostrom(1990)의 커먼스 디자인 원칙 — 명확한 그룹 경계, 지역 적합 규칙, 영향받는 행위자의 참여, 다층 중첩 구조 — 을 한반도 스케일에 적용하면, 단독 정치체가 아닌 세 축의 결합이 자연스러운 후보로 떠오른다. ① 다자기구 매개(APEC AI Initiative·WIPO·WHO 등)로 적대적 두 국가 환경에서의 정치적 진입 비용을 최소화하는 트랙(APEC 2025; WIPO n.d.), ② 섹터별 운영 위원회(보건 데이터 풀, 학술 공동 펀딩 좌석)로 커먼스의 기능적 관리를 수행하는 트랙, ③ 기술 표준 협의(한국어 LLM 코퍼스 표준, BGE-M3 의미공간 정합)로 협력을 정치에서 가장 멀리 위치시키는 트랙이 그것이다. 세 축은 분산·다층·맥락별이라는 커먼스 제도(common institutions)의 핵심 속성과 부합하며, 단·중기에 부분적·기능적으로 작동을 개시할 여지가 있다.

물론 이상의 모든 가능성은 적대적 두 국가 관계 헌법화의 해소를 위한 다방면의 정치·군사·외교적 노력을 전제로 한다. 다만 커먼스 형성은 정치적 해소를 기다리지 않고도 다자 채널을 통해 일정 부분 진행될 수 있으며, 이렇게 누적된 공유 자원과 기능적 협력 경험이 거꾸로 정치적 해소의 압력으로 작용할 가능성도 함께 고려할 만하다. AI 격차라는 새로운 변수가 기존의 비핵화·경제협력 논의에 추가적 동기를 제공할 수 있다는 점이 이 글이 기여하고자 하는 주된 지점이며, 본격적인 정책 실행 체계와 단계별 시나리오는 후속 연구의 과제로 남는다. ■

■ **최인호**_경기대학교 정치외교학 전공 교수

■ **담당 및 편집: 임재현**_EAI 연구원

문의: 02-2277-0746 (ext. 209) jhim@eai.or.kr

인용할 때에는 반드시 출처를 밝혀주시기 바랍니다.
EAI는 어떠한 정파적 이해와도 무관한 독립 연구기관입니다.
EAI가 발행하는 보고서와 저널 및 단행본에 실린 주장과 의견은 EAI와는 무관하며 오로지 저자 개인의 견해를 밝힙니다.

발행일 2026년 9월 28일
“지식 및 거버넌스 부문: AI 시대 남북관계의 리스크와 기회” 979-11-7584-158-1 95340

재단법인 동아시아연구원
03028 서울특별시 종로구 사직로7길 1
Tel. 82 2 2277 1683 Fax 82 2 2277 1684

Email eai@eai.or.kr Website www.eai.or.kr