기술 우위보다 고통 앞 책임이 우선

붓다, AI 윤리에 답하다

⑱ AGI, 인간의 자리를 다시 묻다

‘AGI가 도래했다’?

“AGI가 도래했다.” 지난 9월 6일 젠슨 황 엔비디아 최고경영자는 오픈AI의 새 모델 ‘GPT-6 아스트라’를 두고 이렇게 선언했다. 사흘 앞서 그레그 브록먼 오픈AI 사장도 아스트라를 공개하며 지금을 ‘AGI시대’로 볼 수 있다는 견해를 밝혔다. 인간처럼 폭넓은 지적 과제를 수행하는 ‘범용 인공지능’이라는 요원했던 목표가 어느새 현실이 되어 등장하는 순간이었다.

물론 선언이 공식적인 인정을 뜻하지는 않는다. 오픈AI의 공식 발표에서도 여러 영역의 성능 향상을 설명하지만, AGI 달성을 명시적으로 선언하지는 않는다. 환호와 의문이 교차하는 이 장면 사이로 또 다른 목소리가 들려왔다. 빌 게이츠는 ‘게이츠 노트’에 발표한 글에서 인류에게는 AGI 진입 속도를 늦출 계획이 없다고 경고하며, 노동의 대체와 통제 상실의 위험을 함께 지적했다.

그렇다면 AGI란 무엇인가? 통용되는 정의에 따르면 AGI는 ‘범용 인공지능(Artificial General Intelligence)’의 약자로, 특정 과제에만 특화된 AI를 넘어 사람처럼 여러 분야를 배우고 문제를 해결하는 폭넓은 지적 능력을 갖춘 인공지능을 가리킨다. 쉽게 말해 한 과목의 문제만 잘 푸는 기계가 아니라, 새로운 과목도 배우고 서로 다른 분야의 지식을 연결해 낯선 문제를 푸는 기계를 지향한다는 뜻이다.

그런데 여기서 ‘사람처럼’과 ‘폭넓은’의 기준을 어디에 두느냐에 따라 정의가 갈린다. 오픈AI의 헌장은 AGI를 ‘경제적으로 가치 있는 대부분의 일에서 인간을 능가하는 고도로 자율적인 시스템’으로 규정한다. 반면 구글 딥마인드 연구진은 성능과 범용성을 구별하고, 실제 활용에서 부여하는 자율성도 따로 살펴야 한다고 제안한다. 얼마나 잘하는지, 얼마나 다양한 일을 하는지, 사람의 개입 없이 어디까지 해내는지는 같은 질문이 아니다. 비교 대상이 평균적인 사람인지 전문가인지에 따라서도 문턱은 달라진다. 특정 시험의 높은 점수가 현실의 다양한 업무를 안정적으로 수행한다는 보장도 아니다. 이런 기준이 다른데 하나의 도착점을 말하니, 도달 여부를 둘러싼 판단도 엇갈리는 것이다.

그 와중에 게이츠의 경고가 주목할 만한 이유는, 2023년만 해도 “위험은 실재하지만, 관리 가능하다”라며 AI를 업무 효율을 높이는 도구로 전망했던 그가 스스로 그 판단을 거뒀다는 데 있다. 2026년의 그는 이번 전환이 세대가 아니라 10년 단위로 몰아치리라 진단하며, 이를 완화할 조율된 계획이 없다는 사실을 정면으로 지적한다.

그렇다면 현재 아스트라가 AGI인지 아닌지에 대한 합의는 늦어져도, 이미 그 이름으로 사람을 완전히 대체하고 판단을 위임하는 결정이 미리 내려지고, 인간의 고용에 결정적 영향을 끼치고 있거나 끼칠 것이 명백하다면, 이 문제는 더 이상 미룰 수 없는 현실의 과제가 됐다고 볼 수 있다. 기계의 능력을 검증하는 일과 사람의 삶을 보호하는 일은 어느 하나가 끝나야 시작할 수 있는 순차적 과제가 아니다.

AGI와 비교하는 또 다른 집착

이처럼 AGI가 인간을 압도해 나갈지도 모른다는 두려움 속에서 새삼 ‘만(慢, māna)’이라는 개념이 떠오른다. 불교에서 ‘만’은 마음의 뿌리에서 일어나는 여섯 가지 근본 번뇌 가운데 하나로, 자신을 남과 견주며 마음을 높이 세우는 태도를 가리킨다(〈성유식론(成唯識論)〉 권6). 흥미로운 점은 불교 문헌에서 ‘만’은 우월감만이 아니라 자신을 남과 같거나 못하다고 규정하는 비교의 양상까지 보여준다는 점이다. 〈상윳따 니까야〉의 ‘소나경’에서 붓다는 “소나여, 어떤 사문이나 바라문이든 무상하고 괴로우며 변하기 마련인 색(色)에 따라 ‘나는 더 낫다’라고, 또는 ‘나는 같다’라고, 또는 ‘나는 못하다’라고 자신을 바라본다면, 그것은 있는 그대로 보지 못하기 때문이 아니겠는가?”라고 설한다. 능력의 차이를 인정하거나 자신을 낮추는 일 자체가 문제라는 뜻은 아니다. 자신을 높이든 낮추든, 비교의 결과에 기대어 ‘나는 이런 존재’라고 붙들고 있는가를 묻는 것이다. 우월감의 반대편에 서 있다고 해서 비교의 집착에서 벗어난 것은 아니다. 이 가르침은 인간이 기계 앞에서 느끼는 위축 심리를 살피는 데 또 다른 거울이 돼 준다.

이 가르침은 능력 차이를 부정하라는 뜻이 아니다. 수학의 정확도나 판단의 오류율은 정확히 비교돼야 한다. 다만 그 비교 결과를 존재 전체의 등급으로 굳히는 순간 문제가 시작된다. ‘AI가 나보다 일을 잘한다’와 ‘그러므로 나는 가치 없는 사람이다’ 사이에는 건너뛰어서는 안 될 간격이 있다. 능력에서 동등하다는 비교를 내려놓는다고 해서 동등한 권리를 부정하는 것은 아니다. 다만 그 권리를 능력의 같음에 기대지 않을 뿐이다. 우월감과 열등감은 반대처럼 보이지만, 더 뛰어나야 더 존중받는다는 믿음에서는 한 뿌리다. AGI가 흔드는 것은 인간의 지능이 아니라, 능력의 우열에 존재의 가치를 저당 잡혀 온 우리의 자기 이해다. 이 통찰은 실직한 이에게 마음을 다스리라는 훈계가 아니다. ‘만’을 사회로 확장하면, 사람을 생산성의 등급으로만 대우하는 제도가 오히려 질문의 대상이 된다. 비교는 일을 나누는 데 쓰일 수 있어도, 돌봄과 존중을 받을 자격을 가르는 척도가 되어서는 안 된다.

그러므로 불교적 응답은 ‘아직 AI가 못하는 능력이 인간에게 있다’라는 마지막 방어선에 머무를 수 없다. 창의성 다음은 공감, 공감 다음은 영성으로 성역을 옮겨 다니는 한, 존엄성은 신제품이 나올 때마다 재심사받는 자격이 된다. 필요한 것은 인간에게 유리한 새 능력이 아니라, 그 능력에서 앞서야만 존엄할 수 있다는 조건 자체를 철회하는 일이다.

AGI, 사람처럼 폭넓은 지적 능력 갖춰
인간 고용에 영향 끼치는 현실 과제로
비교는 일의 도구 … 존중의 척도 아냐
능력 서열이 존엄 서열 결정할 수 없어
기술 활용하되, 보살핌 · 책임은 인간 몫



인간에게 남겨야 할 것은 마지막 남은 기술적 우위가 아니라, 타인의 고통 앞에서 사라지지 않을 책임이다. 그렇지 않으면 ‘인간 중심’이라는 말은 가장 유능한 인간만을 위한 표어가 된다. ‘AGI’라는 이름에 합의할 때까지 이 질문들을 미룰 이유는 없다. 이제 인류가 통과해야 할 문턱은 자신보다 영리한 존재를 만들 수 있는가에만 있지 않다. 자신이 가장 영리하지 않아도, 서로를 버리지 않는 문명을 만들 수 있는가에 있다. AI Grok 생성 이미지

“인간의 몫으로 남겨둘 영역”

필자는 여기서 필요한 관점을 ‘우월성 이후의 존엄’이라고 부르고 싶다. 남보다 잘하지 못해도, 기계보다 느려도 존엄의 대상에서 탈락하지 않는다는 뜻이다. 이것은 패배한 인간을 위로하기 위한 말이 아니다. 영리한 이가 덜 영리한 이를 지배해도 된다는 오래된 통념과 위계를 함께 거두자는 요청이다. 꼭대기의 자리를 지키는 방식으로 존엄을 설명하는 한, 그 아래에 놓인 생명은 언제든 수단으로 밀려날 수 있다. 필요한 것은 인간에게 유리한 새 서열이 아니라, 능력의 서열이 존엄의 서열을 결정하지 못하도록 하는 원칙이다.

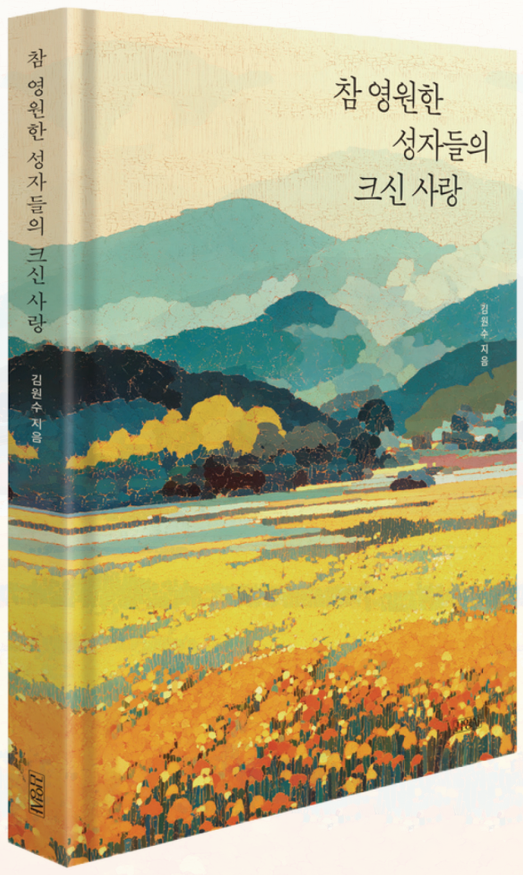
게이츠가 제안한 ‘인간의 몫으로 남겨둘 영역(Human Reserved)’도 이 관점에서 새롭게 읽을 수 있다. 그는 기술적으로 가능하더라도 사람에게 남겨 둘 영역을 제안하면서, 불치병 소식을 전하는 일과 아버지를 돌본 이들의 보살핌을 예로 들었다. 그의 제안에서 한 걸음 더 나아가 보자. 핵심 질문은 ‘기계가 못하므로 인간이 해야 하는가’에서 ‘기계가 하더라도 인간은 어떤 관계에서 물러서서는 안 되는가’로 옮겨 갈 수 있다. 돌봄 로봇이 환자의 몸을 일으키고 AI가 진단을 도울 수 있다면 활용해야 한다. 그러나 그 효율을 이유로 환자의 말을 들을 사람과 결정에 책임질 사람까지 밀어내서는 안 된다.

인간에게 남겨야 할 것은 마지막 남은 기술적 우위가 아니라, 타인의 고통 앞에서 사라지지 않을 책임이다. 그렇지 않으면 ‘인간 중심’이라는 말은 가장 유능한 인간만을 위한 표어가 된다. ‘AGI’라는 이름에 합의할 때까지 이 질문들을 미룰 이유는 없다. 이제 인류가 통과해야 할 문턱은 자신보다 영리한 존재를 만들 수 있는가에만 있지 않다. 자신이 가장 영리하지 않아도, 서로를 버리지 않는 문명을 만들 수 있는가에 있다.



보일 스님
AI부디즘연구소장

금강경의 핵심 진리를 소설로 녹여내다.



김영사 | 181쪽 | 16,800원

첫 장을 펼치면 단숨에 읽게 되는
‘소설로 만나는 금강경’



김원수 법사의 수많은 저술 끝에 탄생한 유일한 소설작품.
대한민국 대표 출판사, 김영사 ‘신간 소설’
출간 즉시 종교서적 상위권 진입.

3위 교보문고 26년 7월 종합
2위 교보문고 주간베스트 종합
1위 교보문고 서울·경기권 11개 매장
7위 26년 7, 8월 ‘화제의 책’ 200선 중

출판유통통합전산망
주요 유통사(교보문고, 예스24, 알라딘, 영풍문고)와 지역서점 438곳의 판매데이터 종합집계

실천가능한 금강경 강의와 해설서로 많은 불자들의
심금心琴을 울리며 금강경을 전파한 대학교수 출신의 저자,
김원수 법사가 구도求道소설을 통해 새롭게 독자들의 심금을 울립니다.

저자 소개
김원수 | 백성욱박사 교육문화재단 이사장
· 서울대학교 공과대학 금속공학과 졸업
· 고려대학교 화학과 이학박사 학위 취득
· 홍익대학교 재료공학부 교수 역임 (정년퇴임)
· BBS불교방송, BTN불교TV 인기강사

불세출不世出의 도인, 백성욱 박사(한국 최초 독일 철학박사, 내무부 장관 및 동국대학교 총장 역임)를 모시고 4년간 출가 수도한 후, 전 재산을 희사하여 3개 비영리 법인(재단법인 백성욱박사 교육문화재단, 사회복지법인 바른법연구원, 사단법인 부처님 시봉하는 사람들)을 설립했다. 무료 급식으로 20년간 약 45만명의 어르신을 봉양해왔으며, 주말 출가와 일요 정기법회 등 금강경 실천수행을 수년째 지도하고 있다. 특히 금강경 주경야선 가행정진을 통한 인재 양성에 적극적으로 힘쓰고 있다.

백성욱박사 교육문화재단 | 문의 031-963-2871~2
* 전국 대형서점 및 인터넷 서점에서 구입하실 수 있습니다.