

『위대한 인간성』(Magnifica Humanitas)의 함의와 AI 시대를 위한 제언

요약

- ▶ AI 시대의 핵심 질문은 기술을 막을 것인가가 아니라, AI가 인간을 위해 작동할 수 있도록 제도적 조건을 어떻게 설계할 것인가임
 - 교황 레오14세의 첫 회칙 『위대한 인간성』은 AI를 거버넌스·노동·소통 전 영역에 걸친 문명적 선택 문제로 제기하며, 기술 자체의 거부가 아닌 AI가 인간 존엄을 위해 작동하는 조건을 묻는 선언임
- ▶ AI 부문의 위험 관리와 혁신 촉진은 상충하는 목표가 아니며, 두 가지를 동시에 고려하는 제도적 설계가 AI 정책의 출발점이 되어야 함
 - 사회적 신뢰가 확보되지 않으면 핵심 영역에서 AI 활용 자체가 제한될 수 있다는 점에서, 책임성과 투명성은 산업 성장의 장애물이 아닌 지속가능한 확산의 조건임
- ▶ AI 정책은 인간 중심성의 가치를 기준으로 삼되, 위험 기반·비례성 원칙 위에서 산업 역동성과 사회적 신뢰를 동시에 확보하는 방향으로 설계되어야 함
 - 고위험 영역 중심의 단계적 규율, 행위 중심의 경쟁 규율, 기술 변화에 맞춰 학습하고 적응하는 유연한 규제 설계가 핵심 과제

한승혜 연구위원 (shhan@kinternet.org)

발표 배경과 선언의 의미

- ▶ 교황 레오 14세의 첫 회칙 『위대한 인간성』(Magnifica Humanitas)을 통해 AI를 단순한 기술 혁신의 문제가 아닌 인간 존엄·노동·민주주의·권력 집중이라는 문명적 질서의 문제로 제기하며, AI 시대를 살아가는 인류가 어떤 선택을 해야 하는지 묻는 사회적 선언을 발표함
 - 1891년 레오 13세의 『새로운 사태』(Rerum Novarum) 반포 135주년에 맞춰 발표된 이 회칙은, 산업 혁명 당시의 노동 문제에 응답했던 전통을 이어 AI·디지털 전환기의 새로운 사회 문제에 대응하고자 함
 - 국가가 혁신을 주도했던 과거와 달리 AI 생태계는 글로벌 민간 기업의 역할이 중요하며, 따라서 공공의 통제와 거버넌스가 한층 복잡해짐
 - 기술은 ‘그 자체로 악한 것’이 아니지만 동시에 중립적이지 않은 특징을 가지며, AI를 설계·자금 지원·규제·사용하는 이들의 가치판단을 반영하므로 공동 책임이 요구됨
- ▶ 회칙의 핵심 메시지는 AI를 막을것인가가 아니라, AI가 이윤과 효율의 논리가 아닌 인간을 위해 작동할 수 있도록 제도적·사회적 조건을 어떻게 설계할 것인가임
 - AI가 인간의 판단을 보조하는 도구로 머무는지, 아니면 인간을 데이터·확률·생산성의 대상으로 환원하는 지에 따라 기술의 사회적 의미가 근본적으로 달라짐
 - 이번 회칙의 발표는 산업계에 중요한 메시지를 전달하며, 앞으로의 AI 기업은 사회적 정당성을 확보하기 위해 기술적인 발전 뿐 아니라 차별 방지·취약계층 포용·인간 선택권 보존에 대한 설명이 요구될 가능성이 높아짐(Vatican News, 2026a; 2026b)

『위대한 인간성』이 제기하는 AI 시대의 핵심 도전

- ▶ AI 시대의 핵심 거버넌스 과제는 기술이 효율성과 이윤의 논리에 종속되지 않고 인간 존엄과 책임의 체계 안에서 작동하도록 제도적 조건을 설계하는 일임

“우리는 AI가 단순한 도구가 아님을 분명히 해야 한다. 기술은 그것을 설계하고 자금을 조달하며 규제하고 사용하는 이들의 선택과 우선순위를 반영한다. 어떤 시스템이 일부 생명을 덜 가치 있게 대우하거나, 이의를 제기할 가능성 없이 배제하는 방식으로 설계되거나 사용된다면, 그것은 단순히 ‘잘 사용되기를 기다리는 도구’가 아니다. 그것은 이미 인간의 불가침적 존엄성에 모순되는 기준을 도입한 것이다.”

— 『위대한 인간성』§104

- AI가 인사·금융·복지·사법 등 영역에서 인간의 판단을 대신할 때, 설명요구권·이의제기 절차·책임소재가 명확히 설계되지 않으면 책임이 공중에 분산됨
 - AI 윤리가 선언 차원에 머물지 않게 하려면 AI 거버넌스가 모든 과정에서 작동하는 의사결정 체계를 갖추어야 함
- ▶ 자동화·AI의 급속한 확산은 노동 구조 전반을 재편하고 있으며, 기술이 어떤 방식으로 노동에 적용되는가에 따라 그 사회적 결과가 근본적으로 달라짐

"AI는 단조로운 업무를 넘겨받아 생산성을 높여줄 것으로 기대되지만, 실제로는 기계가 일하는 사람을 지원하도록 설계되는 것이 아니라 노동자가 기계의 속도와 요구에 맞춰 적응하도록 강요하는 경우가 잦다. 그 결과, AI의 선진된 혜택과 달리 현재의 기술 접근 방식은 역설적으로 노동자의 숙련도를 떨어뜨리고, 자동화된 감시 하에 두며, 경직되고 반복적인 업무에 머물게 한다."

— 『위대한 인간성』§150

- AI 도입 과정에서 인간 중심 설계 원칙이 고려되어야 하며, 전환 충격 완화는 자동화 억제보다 재교육·직무 전환 지원 인프라를 공공정책으로 뒷받침하는 방식이 더 실효적임

▶ 디지털 플랫폼이 형성하는 소통 환경과 알고리즘이 개인의 행동을 프로파일링·예측하고 영향을 주는 구조는 자유와 비판적 사고의 문제로 다루어져야 함

"통제는 명시적인 금지를 통해서만 행사되는 것이 아니라, 가시성의 구조를 통해서도 이루어진다. 무엇이 증폭되고 무엇이 보이지 않게 되는지, 무엇이 보상받고 무엇이 불이익을 받는지가 결국 의견과 선택을 형성하며, 순응과 자기검열을 조장한다. 이러한 이유에서 디지털 시대의 자유는 단순히 내면의 문제가 아니라 공적인 문제이다."

— 『위대한 인간성』§171

- 알고리즘 추천 구조의 투명성, 아동·청소년 대상 조작적 설계 규제, AI 비판적 활용 교육은 각각 공공정책·자율규제·교육정책이 연계하여 다루어야 할 복합 과제임

표 1. 『위대한 인간성』이 제기한 AI 시대의 핵심 도전과 정책 쟁점

영역	핵심 문제제기	향후 정책 쟁점
거버넌스·책임	- 기술 비중립성, 책임소재 불명확	- AI 영향평가, 설명요구권, 독립 감독 체계
권력 집중	- 소수 민간기업의 AI 생태계 독점	- 클라우드·데이터 접근 공정성, 알고리즘 투명성
노동·경제	- 자동화로 인한 노동 구조 재편	- 직무 재교육·전환 지원, AI 도입 인간 중심 설계 원칙
소통·청소년	- 알고리즘 양극화, 비판적 사고 약화	- AI 리터러시 교육, 조작적 설계 규제
안보·평화	- AI 자율무기의 윤리적 통제 부재	- 치명적 자율무기 국제 규범, AI 비군사화

AI 논의의 또 다른 현실: 기술적 기회와 시장 현실에 대한 직시

▶ AI의 위험성이 집중 조명되고 있는 만큼, 정책 논의가 위험 담론에만 편향되지 않도록 AI가 실질적으로 인류의 삶을 개선하고 있는 측면을 함께 고려하는 균형이 필요함

- AI는 이미 의료 진단 역량 향상, 교육 접근성 확대, 취약 지역의 경제 참여 기회 제공 등에서 실질적 성과를 내며 인류가 오랫동안 해결하지 못한 문제들에 새로운 접근 가능성을 열고 있음
- 인류는 인쇄기에서 인터넷에 이르기까지 변혁적 기술에 반복적으로 적응해 왔으며, AI 역시 위험을 다루는 정책과 잠재력을 극대화하는 정책을 병행하는 것이 적절함(Castro, 2026)
- AI 정책이 위험 완화에만 집중하고 기술 확산을 위한 조건 마련을 도외시하면, 혜택이 필요한 이들에게 AI가 닿지 못하는 역설이 발생할 수 있음

- ▶ AI 독점에 대한 우려가 정당한 문제의식이라도, 이를 정책으로 번역할 때는 AI 산업의 층위별 경쟁 구조에 대한 정교한 시장 분석이 선행되어야 하며, 현실의 AI 시장을 단순 독점으로 규정하는 접근은 오히려 정책 판단을 왜곡할 수 있음
 - AI 산업은 모델·클라우드·반도체·데이터·플랫폼 층위별로 경쟁 구조가 상이하며, 대형 기업 외에도 다수의 신규 진입자가 경쟁하고 있어 단순 독점으로 규정하기 어려움
 - AI 산업에서 일정 규모의 기업이 주도적 역할을 하는 것은 막대한 R&D 비용과 글로벌 인프라 구축에 따른 구조적 특성인 만큼, 기업 규모 자체보다 구체적 행위를 중심으로 규율하는 것이 효과적임
 - 과잉 규제는 글로벌 대기업보다 국내 스타트업과 중소기업에 먼저 타격을 주며, 준수 비용이 높아질수록 규모의 경제를 갖춘 대기업의 시장 지위가 오히려 강화되는 역설이 발생할 수 있음(Coniglio, 2026)
- ▶ 윤리적 문제 제기와 AI의 실질적 혜택은 근본적으로 대립하지 않으며, ‘규제냐 혁신이냐’의 이분법 자체가 잘못된 질문으로 거버넌스를 혁신 과정 내부에 내재화하는 접근이 두 가지를 동시에 실현하는 더 실효적인 경로
 - Stilgoe et al.(2013)은 책임 있는 혁신(Responsible innovation)을 ‘현재의 과학·혁신에 대한 집합적 관리를 통해 미래를 돌보는 것’으로 정의하며, 윤리는 혁신 이후의 사후 점검이 아니라 혁신 과정 자체에 통합되어야 함을 강조함
 - 이 프레임워크는 Collingridge(1980)의 통제의 딜레마, 즉 기술 초기에는 영향 예측이 어렵고 성숙 후에는 방향을 바꾸기 어렵다는 딜레마에 대한 응답으로, 예측·성찰·포용·대응의 네 차원을 혁신 과정에 반복적으로 내재화함으로써 이를 넘어설 수 있다고 제언함
 - AI에 대한 사회적 신뢰가 확보되지 않으면 공공부문·교육·의료·금융 등 핵심 영역에서 AI 활용 자체가 제한될 수 있으며, 이 점에서 책임성과 투명성은 산업 성장의 장애물이 아닌 지속가능한 확산의 조건임

AI 거버넌스, 어떻게 설계할 것인가

- ▶ 이번 선언의 함의는 AI 규제 강화의 단순한 논거 제시가 아닌 인간 존엄·책임성·투명성·참여에 대한 원칙을 AI 제도 설계의 가치 기준으로 어떻게 구현할 것인가를 묻는 문서로 보아야 하며, 이 원칙들은 AI 정책의 방향성을 판단하는 기준으로 활용될 수 있음
 - 회칙에서 강조된 거버넌스의 핵심은 AI를 사후적으로 점검하는 체계가 아니라 개발·도입·활용 전 과정에서 작동하는 의사결정 구조로, 사전적 거버넌스가 사후 규제보다 실효적이라는 학술적 논의와도 맥을 같이함(Stilgoe et al., 2013)
 - 다만 이 선언의 윤리 담론이 포괄적 사전규제나 허가제, 과도한 인증 요건으로 직결되는 것은 경계해야 하며, 추상적 원칙을 구체적 시장 분석 없이 규제 명분으로 활용하면 정책 효과가 왜곡될 수 있음
 - 국제 AI 규범 논의에서 『위대한 인간성』의 윤리적 언어는 공동 원칙 형성에 기여할 수 있으나, 국제 규범을 국내 제도에 그대로 이식하기보다 국내 산업 여건·기업 규모·준수 비용을 고려한 단계적 적용이 요구됨

- ▶ “AI가 인간을 위해 작동하는가”라는 질문에 답하기 위한 제도화는, 모든 AI에 동일한 의무를 부과하는 포괄적 사전규제가 아니라 실제 피해 가능성과 기본권 침해 위험이 높은 고위험 영역을 중심으로 위험 기반의 유연한 단계적 접근으로 설계되어야 함
 - Collingridge(1980)의 통제의 딜레마가 시사하듯, 기술 초기의 성급한 포괄적 사전규제는 혁신을 압살하면서도 실질적 위험 통제에는 실패하는 역효과를 낼 수 있음
 - 범용성이 높은 생성형 AI처럼 맥락에 따라 위험성이 달라지는 기술은, 요구되는 결과를 명확히 설정하되 방법은 기업 자율에 맡기는 방식으로 규율하는 것이 처방적 규칙 기반 규제의 한계를 넘어설 수 있음 (WEF, 2020)

표 2. 인간 중심 AI 거버넌스의 설계 방향과 경계

거버넌스 원칙	제도 설계 방향	주의할 경계
인간 존엄·책임성	AI 영향평가, 인간 개입 절차 의무화	모든 AI에 동일 의무 부과 지양
투명성·설명 가능성	자동화 결정 설명요구권·이의제기 절차	소스코드 전면 공개 등 과잉 투명성 요구 경계
권력 집중 견제	시장 봉쇄·자사우대 행위 규율	기업 규모 자체를 문제 삼는 접근 지양
노동 존엄성	인간 중심 AI 도입 원칙, 재교육 지원 병행	자동화 금지 등 기술 도입 억제 방식 경계
국제 공조	국제 원칙을 가치 기준으로 참조	해외 규범의 무비판적 국내 이식 경계

AI 정책을 위한 제언

- ▶ 인간 중심성의 가치를 제도 설계의 기준으로 삼되, 이것이 혁신 억제나 산업 위축으로 이어지지 않도록 위험 기반·비례성 원칙 위에서 산업 역동성과 사회적 신뢰를 동시에 확보하는 방향으로 나아가야 함
 - 국내 기업은 글로벌 빅테크에 비해 자본·데이터·클라우드·반도체 접근성이 불리한 위치에 있는 만큼, 규제 비용이 높아질수록 스타트업과 중소기업이 먼저 타격을 받는 구조임을 제도 설계에 고려해야 함
 - AI 권력 집중 문제는 기업 규모 자체보다 시장 봉쇄·데이터 접근 제한·API 차별 등 구체적 행위 중심으로 접근해야 하며, 층위별 경쟁 구조분석이 선행되어야 함
- ▶ AI 거버넌스 체계는 혁신을 제약하는 외부 규범이 아니라, AI가 핵심 사회 영역에 확산되기 위한 신뢰 기반으로 내재되어야 함
 - 기획·설계·배포·운영 전 과정에 걸쳐 책임 있는 혁신(Responsible Innovation) 원칙이 반복적으로 작동하는 체계를 갖추되, 기업의 R&D 유인을 꺾지 않는 세제 혜택·공동 연구 인프라 지원이 병행되어야 함(Stilgoe et al., 2013)
 - AI에 대한 사회적 신뢰가 확보되지 않으면 공공·의료·금융·교육 등 핵심 영역에서의 도입 자체가 지연될 수 있으며, 이 점에서 기업의 AI 위험관리 체계와 이용자 대상 설명·이의제기 경로 마련은 규제 준수가 아닌 신뢰 산업으로서의 경쟁력 기반임

- ▶ AI 규제는 기술 변화 속도를 따라가지 못하는 경직된 사전 규범이 아닌, 실제 적용 경험을 반영하며 학습하고 적응할 수 있는 유연한 구조로 설계되어야 함
 - 규제 당국과 기업이 충분한 시험을 거쳐 그 결과를 규범에 반영하는 방식은, 기술의 실제 작동 방식을 이해하면서 규제를 설계할 수 있어 사전에 모든 것을 규정하려는 접근보다 실효적임
 - 초기에 성급하게 설정된 규제가 고착되면 기술 환경의 변화에 맞춰 수정되기 어려운 만큼, 규제에 일몰 조항(sunset clause)을 두어 주기적으로 재검토할 수 있는 구조를 처음부터 설계에 내재해야 함

참고문헌

- 교황 레오 14세 (2026.5.15.). 『위대한 인간성』(Magnifica Humanitas).
<https://www.vatican.va/content/leo-xiv/en/encyclicals/documents/20260515-magnifica-humanitas.html>
- Castro, D. (2026). "AI Is Not Another Tower of Babel". ITIF Innovation Files (May 25, 2026).
<https://itif.org/publications/2026/05/25/ai-is-not-another-tower-of-babel/>
- Collingridge, D. (1980). The Social Control of Technology. Frances Pinter Publishers.
- Coniglio, J.V. (2026). "The Vatican's 'AI Monopolies' Talk Risks Encouraging Bad Tech Policy". ITIF Innovation Files (May 29, 2026).
<https://itif.org/publications/2026/05/29/vaticans-ai-monopolies-talk-risks-encouraging-bad-tech-policy/>
- Stilgoe, J., Owen, R., & Macnaghten, P. (2013). "Developing a framework for responsible innovation". Research Policy, 42(9), 1568-1580.
- Vatican News (2026a). 「위대한 인간성」 요약(상)편: '인간 존엄성 수호'.
<https://www.vaticannews.va/ko/pope/news/2026-05/papa-enciclica-humanitas0.print.html>
- Vatican News (2026b). 「위대한 인간성」 요약(하)편: '인공지능을 무장 해제해야 한다'.
<https://www.vaticannews.va/ko/pope/news/2026-05/papa-enciclica-humanitas1.print.html>
- World Economic Forum (2020). Agile Regulation for the Fourth Industrial Revolution: A Toolkit for Regulators. WEF.