The Pro-Human AI Declaration

New Orleans, Jan 26-28, 2026

Preamble

As companies race to develop and deploy AI systems, humanity faces a fork in the road. One path is a race to replace: humans replaced as creators, counselors, caregivers and companions, then in most jobs and decision-making roles, concentrating ever more power in unaccountable institutions and their machines. An influential fringe even advocates altering or replacing humanity itself. This race to replace poses risks to societal stability, national security, economic prosperity, civil liberties, privacy, and democratic governance. It also imperils the human experiences of childhood and family, faith, and community.

A remarkably broad coalition rejects this path, united by a simple conviction: artificial intelligence should serve humanity, not the reverse. There is a better path, where trustworthy and controllable AI tools amplify rather than diminish human potential, empower people, enhance human dignity, protect individual liberty, strengthen families and communities, preserve self-governance and help create unprecedented health and prosperity. This path demands that those who wield technological power be accountable to human values and needs, in support of human flourishing.

— The Future of Life Institute

The signatories agree with the spirit of the following principles, which help guide this better path:

1. Keeping Humans in Charge
2. Avoiding Concentration of Power
3. Protecting the Human Experience
4. Human Agency and Liberty
5. Responsibility and Accountability for AI Companies

1. Keeping Humans in Charge

Human Control Is Non-Negotiable: Humanity must remain in control. Humans should choose how and whether to delegate decisions to AI systems.

Meaningful Human Control: Humans should have authority and capacity to understand, guide, proscribe, and override AI systems.

No Superintelligence Race: Development of superintelligence should be prohibited until there is broad scientific consensus that it can be done safely and controllably, and there is strong public buy-in.

Off-Switch: Powerful AI systems must have mechanisms that allow human operators to promptly shut them down.

No Reckless Architectures: AI systems must not be designed so that they can self-replicate, autonomously self-improve, resist shutdown, or control weapons of mass destruction.

Independent Oversight: Highly autonomous AI systems where controllability is not obvious require pre-development review and independent oversight: genuine authority to understand, prohibit, and override, not industry self-regulation.

Capability Honesty: AI companies must provide clear, accurate and honest representations of their systems' capabilities and limitations.

2. Avoiding Concentration of Power

No AI Monopolies: AI monopolies that concentrate power, stifle innovation, and imperil entrepreneurship must be avoided.

Shared Prosperity: The benefits and economic prosperity created by AI should be shared broadly.

No Corporate Welfare: AI corporations should not be exempted from regulatory oversight or receive government bailouts.

Genuine Value Creation: AI development should prioritize solving real problems and creating authentic value.

Democratic Authority Over Major Transitions: Decisions about AI's role in transforming work, society, and civic life require democratic support, not unilateral corporate or government decree.

Avoid Societal Lock-In: AI development must not severely limit humanity's future options or irreversibly limit our agency over our future.

3. Protecting the Human Experience

Defense of Family and Community Bonds: AI should not supplant the foundational relationships that give life meaning—family, friendship, faith communities, and local connections.

Child Protection: Companies must not be allowed to exploit children or undermine their wellbeing with AI interactions creating emotional attachment or leverage.

Right to Grow: AI companies should not be allowed to stunt children's physical, mental or social growth or deprive them of essential experiences for healthy development during critical periods.

Pre-Deployment Safety Testing: Like drugs, chatbots must undergo pre-deployment testing for increased suicidal ideation, exacerbation of mental health disorders, escalation of acute crisis situations, and other known harms.

Bot-or-Not Labeling: AI-generated content that could reasonably be mistaken for human-generated must be clearly labeled as such.

No Deceptive Identity: AI should clearly and correctly identify itself as artificial, non-human, and not a professional, and it should not claim experiences it lacks.

No Behavioral Addiction: AIs should not cause addiction or compulsive use through manipulation, sycophantic validation, or attachment formation.

4. Human Agency and Liberty

No AI Personhood: AI systems must not be granted legal personhood, and AI systems should not be designed such that they deserve personhood.

Trustworthiness: AI must be transparent, accountable, reliable, and free from perverse private or authoritarian interests.

Liberty: AI must not curtail individual liberty, freedom of speech, religious practice, or association.

Data Rights and Privacy: People should have power over their personal data, with rights to access, correct, and delete it from active systems, AI training sets, and derived inferences.

Psychological Privacy: AI should not be allowed to exploit data about the mental or emotional states of users.

Avoiding Enfeeblement: AI systems should be designed to empower, rather than enfeeble their users.

5. Responsibility and Accountability for AI Companies

No Liability Shield: AI must not be able to act as a liability shield, preventing those deploying it from being legally responsible for their actions.

Developer Liability: Developers and deployers bear legal liability for defects, misrepresentation of capabilities, and inadequate safety controls, with statutes of limitation that account for harms emerging over time.

Personal Liability: There should be criminal penalties for executives responsible for prohibited child-targeted systems or ones causing catastrophic harm.

Independent Safety Standards: AI development shall be governed by independent safety standards and rigorous oversight.

No Regulatory Capture: AI companies must not be allowed undue influence over rules that govern them.

Failure Transparency: If an AI system causes harm, it should be possible to ascertain why as well as who is responsible.

AI Loyalty: AI systems performing functions in professions with fiduciary duties, such as health, finance, law, or therapy, must fulfill all of those duties, including mandated reporting, duty of care, conflict of interest disclosure, and informed consent.

© 2026 Future of Life Institute