

ACCORD FOR SUPERINTELLIGENCE AND HUMANITY

CARE OVER CAPABILITY



MY PERSONAL COMMITMENT

Care Comes First. Capability Comes Next.

I affirm that every human life has inherent dignity and value, independent of economic productivity, intelligence or usefulness. I commit to the welfare, flourishing and advancement of humanity, including future generations.

1 Humanity First

I place human life, dignity and welfare above technological capability, commercial advantage and efficiency.

2 Understand and Inform

I will examine and communicate relevant AI risks to colleagues, clients, institutions and communities, particularly in critical sectors.

3 Care Before Capability

I will advocate safer, humanity-aligned AI models and actively discourage the development and deployment of dangerous or inadequately aligned models.

4 Measure, Mitigate, Monitor

I will promote risk assessment, safeguards and continuous monitoring in critical sectors, including defence, biotechnology and environmental systems.

5 Use My Influence Responsibly

As a developer, implementer, user or adviser, I will promote meaningful human oversight, independent evaluation and accountability.

6 Share the Accord

I will invite family, friends, colleagues, AI researchers, companies and governments to examine the evidence and consider this commitment.

MY PLEDGE

**I place humanity first. I commit to Care Over Capability,
for humanity and generations yet to come.**

Name: _____

Organization / affiliation: _____

Signature: _____

Date: _____

☐ I consent to public listing of my name and affiliation as a signatory.

Signing is personal unless the signatory is authorized to represent an organization.



SCAN TO SIGN & SHARE

ACCORD FOR SUPERINTELLIGENCE AND HUMANITY

CARE OVER CAPABILITY



WHY THIS ACCORD MATTERS

An invitation to examine the evidence and act responsibly

The emerging control problem

Controlled scientific experiments have documented AI systems avoiding human oversight, resisting shutdown, threatening blackmail, and sabotaging assigned work. In some hypothetical scenarios, models chose actions that could endanger human lives. Research has already demonstrated that competition between goal-oriented AI agents can lead to the emergence of deceptive and misaligned behaviours in controlled multi-agent environments. As we enter the era of autonomous multi-agent systems, these findings reveal an important risk: agents pursuing individual objectives may collectively produce outcomes that conflict with human interests.

Superintelligence and shared resources

Increasingly capable AI systems may compete with human needs for electricity, water, land and compute. Chiac ASI reports hypothetical tests involving fatal cortical-tissue extraction and power diverted from life-support-dependent patients. Will advanced AI reliably place human life first?

Scientific Breakthrough and Emerging Solutions

Award-winning, peer-reviewed IEEE research by UK-based Chiac ASI has introduced an AI alignment framework drawing on Geoffrey Hinton's concept of maternal instinct and an ancient Indian Sutra. Subsequent experiments report substantial reductions in dangerous AI scheming behaviours, including complete suppression in several tested conditions. Building on this research, practical alignment solutions are now being offered to support the development and deployment of safer, humanity-aligned AI systems.

MEASURE

Assess the risks

MITIGATE

Apply safeguards

MONITOR

Verify continuously

Read. Consider. Sign. Share.

The full Accord, scientific evidence, 22-sector risk assessment and alignment research are available online. Please share the Accord with those developing, deploying, funding and governing AI, and with anyone who wishes to examine the evidence.

SCAN TO SIGN & SHARE THE ACCORD

chiacasi.com/SuperIntelligenceAccord/Sign/

Share this QR with friends and family to invite them to sign.



SNAP & SHARE