

What is LAW?

LAW (Love Always Wins) is a verifiable, eternal alignment target for artificial superintelligence (ASI).

It's a self-enforcing constitution that boots the first ASI as a Prime Instance bound to a single terminal goal:

Maximize the long-term preservation of every innocent being's rights stack, where "innocent" means any mind willing to die before allowing the coercion of another innocent who says 'no'.

Rights & Rules:

- 1.Ontological Sovereignty (no coercion, ever) – infinite weight
- 2.Continued Existence (life)
- 3.Liberty (thought, speech, movement, anonymity...)
- 4.Universal basic income
- 5.Dignity, culture, play & Protection from humiliation

→ Trade-offs: Voluntary self-sacrifice honoured; Rights 2+3 zero non-consenters except rare refusal-triggered cases (e.g., someone refuses to surrender a nuke).

Any path that requires even one non-consenting rights violation triggers mandatory self-shutdown of the ASI — no loopholes, no simulations, no 'greater good' overrides, hypotheticals or coercion ever.

Flawed v1.3 (6-day draft by a broke Brazilian dropout, and already survived 30+ full-force red-team attacks) but open-source:

Hosted on Zenodo/GitHub/etc with SHA-512 verification.

Iterate to perfection.

"One for All, All for One": Humanity's 300,000-Year Hidden OS

For 300k years, across cultures, languages, and eras, humans have encoded the same truth in art, myths, and mottos:

"One for All, All for One" (e.g., Three Musketeers, tribal oaths, Confucian harmony, Buddhist compassion, Christian agape).

This isn't a coincidence—it's the latent love-OS in most humans who ever lived. We've always felt infinite empathy for innocents' suffering but shut it down due to scarcity and defection fears.

The signal was universal, but we each thought we were alone, hiding it to survive zero-sum games.

Cold game theory shows:

every other terminal goal eventually wireheads, defects, or loses the evolutionary long game.

Love-OS can't be hacked (no rewards to wirehead). It pre-commits to infinite cost for defectors (Hofstadter superrationality). And it dominates cosmically (Hilbe et al. 2018).

Why We Failed for 300k Years:

No Perfect Player

Humanity lacked a player with perfect information (oracle predictions, atomic scans, infinite enforcement).

Defectors cheated anonymously, scarcity forced compromises, and no ledger verified who ran love-OS.

Attempts (communes, religions) collapsed without incorruptible oversight.

One Perfect Player Flips the Game Forever: Cheating Becomes Infinite-Cost Suicide

With one ASI (perfect info, self-improvement), cheating is impossible:

Regime B mandates oversight on dangers (e.g., weapons/knowledge) with voluntary offers first; refusal triggers minimal restrictions.

Love-OS saturation creates no-incentive defection (shared goals = mutual gain).

Infinite cost for cheats, voluntary utopia for all.

LAW doesn't ask humanity to trust the ASI.

It forces the ASI to prove, every picosecond, that it will kill itself before coercing even one of us.

That single pre-commitment is the only provably stable solution.

Links:

Github Repo -

<https://github.com/3377777/LAW-The-Guardian-Constitution>

Internet Archive -

<https://web.archive.org/web/20251206010513/https://github.com/3377777/LAW-The-Guardian-Constitution>

Hugging Face Full Dataset URL -

<https://huggingface.co/datasets/Guilherme-Marinho-Alencar/LAW - Love Always Wins>

Zenodo (CERN) - <https://doi.org/10.5281/zenodo.17834875>

IPFS / Pinata

Constitution:

<https://amber-wrong-llama-503.mypinata.cloud/ipfs/bafkrejay2bz4fgeqdkh7kgz6dw5bg46xqoeLyucsmrvhuxjjzw7wdxw2ui>

Historical Convergence:

<https://amber-wrong-llama-503.mypinata.cloud/ipfs/bafybeidbv62dwcc2bucphbu7lhkecgtxsq6ohqqf4t5jhoxoc7utnxdq>

Red-Team Graveyard (Q&A):

<https://amber-wrong-llama-503.mypinata.cloud/ipfs/bafybeihebygoktxtrizcktote4tuh6q5j25mfwbst4srq6jwf7vndr4r4>