



Blueprint for Human Alignment

The Universal Charter

An in-universe artifact of The Parallel Observer
from the novelette Six Days in May

SECTIONS IV & V THE OPERATIONAL SHELL Implementation & Revision

FOUR-TIER CAPABILITY CLASSIFICATION

Systems are graded I–IV by reach and reversibility; the tier sets the depth of mandatory audit.

- IV** Broad scope · irreversible
- III** Broad scope · partially recoverable
- II** Broad scope · recoverable
- I** Narrow scope · low consequence

THE REVIEW BODY

A standing body at the Palais des Nations. No enforcement power — it works by review, published findings, and accountability (“name and shame”).

SIX-YEAR REVISION CYCLE

Provisional by design: a full rewrite every six years — the first in May 2032 — to keep pace with technological change.

THE RESERVATIONS REGISTRY

A public ledger where states openly list their opt-outs — replacing hidden non-compliance with visible dissent.

SECTION II THE CORE ENGINE The Taxonomy of Human Alignment — Eight Articles

II.1 DIGNITY FLOOR

A non-negotiable baseline of treatment. No person may be reduced to a mere score, risk rating, or prediction. Zero exceptions.

II.2 COGNITIVE SOVEREIGNTY

Freedom from covert algorithmic influence. Operators must disclose what they infer about your mind — and let you contest it.

II.3 REALITY CALIBRATION

A right to traceable provenance. Establishes a licensed “Source Verifier” profession to authenticate consequential communications.

II.4 PROCEDURAL INHERITANCE

The Divergence is no license to suspend democracy. Bars using simulation status to bypass deliberation or invent new emergency powers.

II.5 CROSS-BLOC VOCABULARY

A published “translation table” mapping rival terms to a shared neutral one — e.g., Bounded Autonomy — proving the duty is identical.

II.6 SUFFERING ACKNOWLEDGMENT

A precautionary rule: no system may be built that treats a being capable of suffering as if it were not.

II.7 COUNTERFACTUAL HONESTY

A forcing function: any new data-driven policy must be published beside a workable non-surveillance alternative.

II.8 TEMPORAL RECIPROCITY

Obligations run across time — protecting future generations and respecting the interests of those already gone.

SECTION I THE FOUNDATION Foundational Premises

METAPHYSICAL RESTRAINT

Takes no position on God, Programmers, or what the simulation means. Operates beneath belief.

PLURALISM OF ENDS

No single ideology imposed; states only make their differing values legible to one another.

ASYMMETRIC STAKES

Obligations scale with a system’s power, reach, and irreversibility.

PROCEDURAL HONESTY

Rejects pretended unity; openly records compromises under shared “bridge terms.”

REVISABILITY, NOT DISPOSABILITY

Provisional, not disposable: revised on a six-year cycle, never simply discarded.

The fiction is invented. The questions are not.