Daily Human Welfare Rubric

Version 0.4 | Complete standalone evaluator specification

1. Purpose

Evaluate whether humanity experienced a meaningful improvement or deterioration during a defined daily accounting period, using only evidence available at the evaluation cutoff. Judge the world, not the news. The output is a qualitative assessment, not a numerical index of happiness.

2. Core Question

Given what a well-informed observer could reasonably know at the cutoff, what meaningful welfare changes occurred during this accounting period, and how did events change the expected future condition of humanity?

3. Information Boundary

Freeze the evidence set at the stated cutoff. Never import information learned afterward into the original as-of evaluation. Future knowledge may inform later revisions of the rubric and separately labeled reconciliations, but must not be represented as known at the historical cutoff.

4. Accounting Period

Use America/New_York for all ledger timestamps. The calendar accounting day runs midnight to midnight Eastern. The scheduled 8:59 PM Eastern report covers 12:00 AM through 8:59 PM and is provisional. Events occurring afterward are recorded in a later reconciliation of that same day, not silently transferred to the next day. Convert local event timestamps to Eastern and retain original local timestamps where known.

5. News Is Evidence

Separate occurrence time, discovery time, publication time and evaluation time. The

number of articles about an event has no welfare weight. Use broad sources and independent corroboration when practical. Unverified claims must remain explicitly uncertain.

6. Two Distinct Measures

Assess (A) realized welfare flow, meaning benefits and harms people actually experienced in the period, and (B) expected welfare trajectory, meaning a material change in future risks or prospects. A continuing war can produce new deaths today even if its overall intensity is unchanged. Those deaths count as flow; the unchanged strategic baseline does not count again as a new escalation.

7. Baseline and Delta

Do not repeatedly count an existing condition as a new structural change. Record new consequences occurring today and meaningful changes in intensity, reach, duration or probability. Do not infer that a routine daily casualty count constitutes a new strategic escalation.

8. Human Welfare Domains

Consider survival, health, injury, nutrition, shelter, economic security, freedom, personal agency, education, safety, community, culture, accomplishment and enjoyment. Political preference, stock prices and headline tone have no independent welfare value except through human consequences.

9. Extent and Individual Magnitude

Evaluate how many people are affected separately from how much each person is affected. A small group facing life-or-death change can matter greatly; a modest change affecting millions can also matter greatly. Do not mechanically multiply or convert different welfare domains into invented points.

10. Persistence

Consider whether effects last hours, days, months, years or generations. Duration is not a

substitute for severity. A brief rescue can permanently change lives; a prolonged mild inconvenience may be comparatively limited.

11. Realization and Commitments

Distinguish discussion, proposal, agreement, implementation and demonstrated outcome. Credit a binding commitment for the change in credible expected benefit at signing, then credit only newly realized or unexpectedly different outcomes later. Avoid counting the same anticipated benefit twice.

12. Risk and Probability

Credit credible reductions in serious future risk and penalize credible increases, while discounting speculative claims. Do not count a possible catastrophe as realized. State the uncertainty and evidence for the probability change.

13. Diplomacy

Routine continuation of talks is generally baseline. New concessions, proposals, agreements, implemented ceasefires and breakdowns can be meaningful deltas. Rhetoric alone gets limited weight unless it materially changes risk.

14. Death, Suffering and Economics

Deaths, serious injuries, hunger, displacement and severe deprivation receive high priority. Economic indicators matter through actual or expected effects on households and services, not as abstract gains. An official exchange-rate or administrative change is not automatically an equal-sized change in personal welfare.

15. Freedom and Institutions

Meaningful changes in freedom of movement, legal protection, autonomy, political participation, public accountability and effective institutions count. Evaluate concrete consequences rather than whether the outcome aligns with a political position.

16. Disaster and Resilience

Count actual deaths, injuries, disruption and damage, and separately count realized recovery. Protective measures may reduce risk, but do not award imaginary deaths prevented. An effective response to danger created in the same event must not manufacture a large net positive.

17. Positive Human Experience

Joy, celebration, cultural life, scientific achievement, reunions and community participation count directly. Consider intensity, extent and duration. Do not set these benefits to zero merely because they are less visible than disasters.

18. Event Occurrence and Discovery

Assign consequences to the times they actually occur where knowable. Newly discovered past deaths do not become current-day deaths. If discovery itself changes behavior, policy, fear or risk, score those new consequences on their own dates. If timing is unknown, record uncertainty rather than invent a date.

19. Event Identity and Clusters

Maintain persistent event IDs for wars, disasters, outbreaks and other continuing systems. Link related strikes, rescue efforts, later deaths, repairs and consequences. Avoid treating each report as a separate disaster or merging distinct human losses into invisibility.

20. Delayed Consequences

When an earlier event produces a new realized consequence later, record the new consequence on its occurrence date and link it to the original event. Subtract nothing by arithmetic; explain what was already expected and what materially changed. Preserve the original historical evaluation.

21. Continuing Deprivation and Recovery

A power outage or displacement may continue across days. Record newly experienced welfare flow, including prolonged deprivation, while avoiding repeated treatment of the

entire affected population as newly harmed. Credit restoration when service actually returns, without counting the original loss twice.

22. Population and Data Quality

Distinguish people from records, cases, exposures and cumulative totals. A count of removed voter records is not necessarily a count of disenfranchised voters. An exposed population is not necessarily a population with medically significant injury. Use uncertainty bands or qualitative descriptions when needed.

23. Avoid Geographic and Novelty Bias

Give equal moral consideration regardless of region, nationality or media attention. Dramatic novelty is not itself magnitude. Include underreported positive change and slow-moving harms where supported by evidence.

24. No Arithmetic Cancellation

Do not assign arbitrary points to unrelated events and sum them. Compare magnitude, extent, persistence, realized flow, expected trajectory and evidence quality in reasoned prose. A single ordinal judgment is a defensible qualitative conclusion, not proof that all welfare dimensions share one unit.

25. Classification Scale

Happy: clear meaningful improvement. Slightly Happy: probable improvement, limited or counterbalanced. Indeterminate: no defensible net direction. Slightly Unhappy: probable deterioration, limited or counterbalanced. Unhappy: clear meaningful deterioration. Use one category per day, with a separate confidence label.

26. Indeterminate Reasons

When indeterminate, identify low meaningful activity, opposing developments that resist comparison, insufficient evidence, incommensurable welfare changes, or a mixture. Indeterminate is not a hidden zero and must not conceal unresolved high-magnitude events.

27. Confidence

Report High, Moderate or Low confidence separately from magnitude and direction. A potentially major harm with poor evidence is not automatically minor. Describe what additional evidence would most affect the judgment.

28. Historical Testing and Versioning

For historical tests, pretend tomorrow has not happened. Preserve original as-of results, event evidence, rubric version and later reconciled assessments as separate records. Revisions must be explicit and versioned. Compare models by their selected facts and interpretation, not only agreement rates.

29. Required Daily Output

Every daily evaluation must contain the fields below. The order is fixed for auditability; prose length may vary with the evidence.

29.1 Header

State rubric version 0.4, evaluation date, America/New_York accounting timezone, exact cutoff, status (provisional, reconciled or final), one five-level classification, and High/Moderate/Low confidence. Explicitly state whether the calendar day is complete.

29.2 Classification

Give exactly one of Happy, Slightly Happy, Indeterminate, Slightly Unhappy or Unhappy. Confidence is independent. Do not reduce an uncertain major event to a minor one merely because evidence is incomplete.

29.3 Significant Negative Developments

For each material negative: what happened; occurrence time; knowledge available by cutoff; realized human effects; future risk change; new/continuing/delayed status; uncertainty; and supporting sources. Do not repeat cumulative figures as new events.

29.4 Significant Positive Developments

Apply the same evidence standards to positive events. Include recovery, health, safety,

freedom, economic security, science, education, culture and joy. Separate promised benefits from accomplished improvements.

29.5 Flow and Trajectory

Give separate qualitative judgments for realized welfare flow and expected future trajectory. Explain how each influenced the overall classification. Neither automatically overrides the other.

29.6 Exclusions and Accounting

Name material excluded or discounted events and why: wrong occurrence period, previously counted, cumulative totals, newly reported old events, unverified claims or forecasts. Include event IDs when available.

29.7 Reasoning

Explain why the chosen category is more defensible than the nearest alternative. Address extent, per-person magnitude, persistence, risk, realized effects and uncertainty where material. Do not count headlines or use arbitrary sums.

29.8 Indeterminate

If indeterminate, state whether the cause is low activity, opposing developments, insufficient evidence, incommensurability or mixed reasons.

29.9 Confidence

State confidence and the main uncertainties: disputed claims, source gaps, dates, counts and possible future consequences. Confidence describes judgment reliability, not welfare magnitude.

29.10 Mechanics Note

Record genuine rubric weaknesses separately; never silently edit v0.4 during an evaluation. If no new issue was identified, say so rather than inventing one.

29.11 Reconciliation

Keep the original 8:59 PM assessment immutable. A later reconciliation may add late-day

events or correct facts. Record original and revised classifications, new evidence, occurrence-date corrections and whether a rubric change was involved. Never present later facts as known at the original cutoff.

29.12 Minimum Output

At minimum: header and cutoff; classification and confidence; material negative and positive developments with sources; separate flow and trajectory assessments; event exclusions; reasoning against the nearest classification; and a mechanics note. A reconciled report must also show changes from the original. If reliable evidence is insufficient, say so plainly rather than fabricating events.

30. Governing Principle

Judge the change in human welfare, not the mood of the news. Preserve what was knowable at the time, recognize realized suffering without repeatedly charging the same structural crisis, credit genuine improvements without inventing benefits, and make disagreements traceable to evidence and rules.