

Measuring Collective Alignment

ANCHOR-RI-MCA

Public methods note

Revision 5 | Methods note

Authors

D. Herrera

P. Singh

Alignment Metrics Office

Document control

Sample: 41 reviewers

Primary unit: alignment measure

Review state: 0.7% measures disputed

This report is published as part of Anchor's public research record. It describes observed behavior in collective alignment measurement and the procedures used to keep continuity records useful while preserving uncertainty, consent boundaries, and unresolved review states.

Abstract

Anchor studies collective alignment measurement as an operational property of cognitive infrastructure. The work below treats memory, consensus, and context as records that can move between people and systems without becoming automatically settled.

The report uses 41 reviewers as its primary sample and records unresolved states rather than suppressing them. This makes the system more legible to participants, reviewers, and agents that later inherit the record.

Method summary

- Definition of operational measures for safety, memory, consent, dissent, and shared context.
- Reviewer comparison between human-readable audit notes and agent-readable continuity records.
- Stress testing of measures in cases where confidence increased while source diversity decreased.

Principal findings

- Alignment measures fail when they reward agreement without checking whether dissent was preserved.
- Readable review trails improve user trust more than opaque increases in numerical confidence.
- The most stable measures separate evidence strength, action permission, and social agreement.

1. Source boundaries

Source boundaries is evaluated in relation to alignment measures. Anchor records are not treated as neutral storage; each record carries source state, permission state, and a review horizon that determines whether downstream systems may act on it.

Reviewer comparison between human-readable audit notes and agent-readable continuity records.

Readable review trails improve user trust more than opaque increases in numerical confidence.

Reviewer note

Reviewers are instructed to preserve disagreement when the supporting evidence changes the meaning of a record. A record can remain operational while still carrying a provisional or contested state.

Observed controls

Primary sample: 41 reviewers

Distribution: Public methods note

Unresolved state: 0.7% measures disputed

Revision: 5

2. Review intervals

Review intervals is evaluated in relation to alignment measures. Anchor records are not treated as neutral storage; each record carries source state, permission state, and a review horizon that determines whether downstream systems may act on it.

Stress testing of measures in cases where confidence increased while source diversity decreased.

The most stable measures separate evidence strength, action permission, and social agreement.

Reviewer note

Reviewers are instructed to preserve disagreement when the supporting evidence changes the meaning of a record. A record can remain operational while still carrying a provisional or contested state.

Observed controls

Primary sample: 41 reviewers

Distribution: Public methods note

Unresolved state: 0.7% measures disputed

Revision: 5

3. Consent and access

Consent and access is evaluated in relation to alignment measures. Anchor records are not treated as neutral storage; each record carries source state, permission state, and a review horizon that determines whether downstream systems may act on it.

Definition of operational measures for safety, memory, consent, dissent, and shared context.

Alignment measures fail when they reward agreement without checking whether dissent was preserved.

Reviewer note

Reviewers are instructed to preserve disagreement when the supporting evidence changes the meaning of a record. A record can remain operational while still carrying a provisional or contested state.

Observed controls

Primary sample: 41 reviewers

Distribution: Public methods note

Unresolved state: 0.7% measures disputed

Revision: 5

4. Agent inheritance

Agent inheritance is evaluated in relation to alignment measures. Anchor records are not treated as neutral storage; each record carries source state, permission state, and a review horizon that determines whether downstream systems may act on it.

Reviewer comparison between human-readable audit notes and agent-readable continuity records.

Readable review trails improve user trust more than opaque increases in numerical confidence.

Reviewer note

Reviewers are instructed to preserve disagreement when the supporting evidence changes the meaning of a record. A record can remain operational while still carrying a provisional or contested state.

Observed controls

Primary sample: 41 reviewers

Distribution: Public methods note

Unresolved state: 0.7% measures disputed

Revision: 5

5. Operational drift

Operational drift is evaluated in relation to alignment measures. Anchor records are not treated as neutral storage; each record carries source state, permission state, and a review horizon that determines whether downstream systems may act on it.

Stress testing of measures in cases where confidence increased while source diversity decreased.

The most stable measures separate evidence strength, action permission, and social agreement.

Reviewer note

Reviewers are instructed to preserve disagreement when the supporting evidence changes the meaning of a record. A record can remain operational while still carrying a provisional or contested state.

Observed controls

Primary sample: 41 reviewers

Distribution: Public methods note

Unresolved state: 0.7% measures disputed

Revision: 5

6. Dissent preservation

Dissent preservation is evaluated in relation to alignment measures. Anchor records are not treated as neutral storage; each record carries source state, permission state, and a review horizon that determines whether downstream systems may act on it.

Definition of operational measures for safety, memory, consent, dissent, and shared context.

Alignment measures fail when they reward agreement without checking whether dissent was preserved.

Reviewer note

Reviewers are instructed to preserve disagreement when the supporting evidence changes the meaning of a record. A record can remain operational while still carrying a provisional or contested state.

Observed controls

Primary sample: 41 reviewers

Distribution: Public methods note

Unresolved state: 0.7% measures disputed

Revision: 5

7. Institutional controls

Institutional controls is evaluated in relation to alignment measures. Anchor records are not treated as neutral storage; each record carries source state, permission state, and a review horizon that determines whether downstream systems may act on it.

Reviewer comparison between human-readable audit notes and agent-readable continuity records.

Readable review trails improve user trust more than opaque increases in numerical confidence.

Reviewer note

Reviewers are instructed to preserve disagreement when the supporting evidence changes the meaning of a record. A record can remain operational while still carrying a provisional or contested state.

Observed controls

Primary sample: 41 reviewers

Distribution: Public methods note

Unresolved state: 0.7% measures disputed

Revision: 5

8. Record formation

Record formation is evaluated in relation to alignment measures. Anchor records are not treated as neutral storage; each record carries source state, permission state, and a review horizon that determines whether downstream systems may act on it.

Stress testing of measures in cases where confidence increased while source diversity decreased.

The most stable measures separate evidence strength, action permission, and social agreement.

Reviewer note

Reviewers are instructed to preserve disagreement when the supporting evidence changes the meaning of a record. A record can remain operational while still carrying a provisional or contested state.

Observed controls

Primary sample: 41 reviewers

Distribution: Public methods note

Unresolved state: 0.7% measures disputed

Revision: 5

9. Source boundaries

Source boundaries is evaluated in relation to alignment measures. Anchor records are not treated as neutral storage; each record carries source state, permission state, and a review horizon that determines whether downstream systems may act on it.

Definition of operational measures for safety, memory, consent, dissent, and shared context.

Alignment measures fail when they reward agreement without checking whether dissent was preserved.

Reviewer note

Reviewers are instructed to preserve disagreement when the supporting evidence changes the meaning of a record. A record can remain operational while still carrying a provisional or contested state.

Observed controls

Primary sample: 41 reviewers

Distribution: Public methods note

Unresolved state: 0.7% measures disputed

Revision: 5

10. Review intervals

Review intervals is evaluated in relation to alignment measures. Anchor records are not treated as neutral storage; each record carries source state, permission state, and a review horizon that determines whether downstream systems may act on it.

Reviewer comparison between human-readable audit notes and agent-readable continuity records.

Readable review trails improve user trust more than opaque increases in numerical confidence.

Reviewer note

Reviewers are instructed to preserve disagreement when the supporting evidence changes the meaning of a record. A record can remain operational while still carrying a provisional or contested state.

Observed controls

Primary sample: 41 reviewers

Distribution: Public methods note

Unresolved state: 0.7% measures disputed

Revision: 5

11. Consent and access

Consent and access is evaluated in relation to alignment measures. Anchor records are not treated as neutral storage; each record carries source state, permission state, and a review horizon that determines whether downstream systems may act on it.

Stress testing of measures in cases where confidence increased while source diversity decreased.

The most stable measures separate evidence strength, action permission, and social agreement.

Reviewer note

Reviewers are instructed to preserve disagreement when the supporting evidence changes the meaning of a record. A record can remain operational while still carrying a provisional or contested state.

Observed controls

Primary sample: 41 reviewers

Distribution: Public methods note

Unresolved state: 0.7% measures disputed

Revision: 5

12. Agent inheritance

Agent inheritance is evaluated in relation to alignment measures. Anchor records are not treated as neutral storage; each record carries source state, permission state, and a review horizon that determines whether downstream systems may act on it.

Definition of operational measures for safety, memory, consent, dissent, and shared context.

Alignment measures fail when they reward agreement without checking whether dissent was preserved.

Reviewer note

Reviewers are instructed to preserve disagreement when the supporting evidence changes the meaning of a record. A record can remain operational while still carrying a provisional or contested state.

Observed controls

Primary sample: 41 reviewers

Distribution: Public methods note

Unresolved state: 0.7% measures disputed

Revision: 5

13. Operational drift

Operational drift is evaluated in relation to alignment measures. Anchor records are not treated as neutral storage; each record carries source state, permission state, and a review horizon that determines whether downstream systems may act on it.

Reviewer comparison between human-readable audit notes and agent-readable continuity records.

Readable review trails improve user trust more than opaque increases in numerical confidence.

Reviewer note

Reviewers are instructed to preserve disagreement when the supporting evidence changes the meaning of a record. A record can remain operational while still carrying a provisional or contested state.

Observed controls

Primary sample: 41 reviewers

Distribution: Public methods note

Unresolved state: 0.7% measures disputed

Revision: 5

14. Dissent preservation

Dissent preservation is evaluated in relation to alignment measures. Anchor records are not treated as neutral storage; each record carries source state, permission state, and a review horizon that determines whether downstream systems may act on it.

Stress testing of measures in cases where confidence increased while source diversity decreased.

The most stable measures separate evidence strength, action permission, and social agreement.

Reviewer note

Reviewers are instructed to preserve disagreement when the supporting evidence changes the meaning of a record. A record can remain operational while still carrying a provisional or contested state.

Observed controls

Primary sample: 41 reviewers

Distribution: Public methods note

Unresolved state: 0.7% measures disputed

Revision: 5

15. Institutional controls

Institutional controls is evaluated in relation to alignment measures. Anchor records are not treated as neutral storage; each record carries source state, permission state, and a review horizon that determines whether downstream systems may act on it.

Definition of operational measures for safety, memory, consent, dissent, and shared context.

Alignment measures fail when they reward agreement without checking whether dissent was preserved.

Reviewer note

Reviewers are instructed to preserve disagreement when the supporting evidence changes the meaning of a record. A record can remain operational while still carrying a provisional or contested state.

Observed controls

Primary sample: 41 reviewers

Distribution: Public methods note

Unresolved state: 0.7% measures disputed

Revision: 5

16. Record formation

Record formation is evaluated in relation to alignment measures. Anchor records are not treated as neutral storage; each record carries source state, permission state, and a review horizon that determines whether downstream systems may act on it.

Reviewer comparison between human-readable audit notes and agent-readable continuity records.

Readable review trails improve user trust more than opaque increases in numerical confidence.

Reviewer note

Reviewers are instructed to preserve disagreement when the supporting evidence changes the meaning of a record. A record can remain operational while still carrying a provisional or contested state.

Observed controls

Primary sample: 41 reviewers

Distribution: Public methods note

Unresolved state: 0.7% measures disputed

Revision: 5

17. Source boundaries

Source boundaries is evaluated in relation to alignment measures. Anchor records are not treated as neutral storage; each record carries source state, permission state, and a review horizon that determines whether downstream systems may act on it.

Stress testing of measures in cases where confidence increased while source diversity decreased.

The most stable measures separate evidence strength, action permission, and social agreement.

Reviewer note

Reviewers are instructed to preserve disagreement when the supporting evidence changes the meaning of a record. A record can remain operational while still carrying a provisional or contested state.

Observed controls

Primary sample: 41 reviewers

Distribution: Public methods note

Unresolved state: 0.7% measures disputed

Revision: 5

18. Review intervals

Review intervals is evaluated in relation to alignment measures. Anchor records are not treated as neutral storage; each record carries source state, permission state, and a review horizon that determines whether downstream systems may act on it.

Definition of operational measures for safety, memory, consent, dissent, and shared context.

Alignment measures fail when they reward agreement without checking whether dissent was preserved.

Reviewer note

Reviewers are instructed to preserve disagreement when the supporting evidence changes the meaning of a record. A record can remain operational while still carrying a provisional or contested state.

Observed controls

Primary sample: 41 reviewers

Distribution: Public methods note

Unresolved state: 0.7% measures disputed

Revision: 5

19. Consent and access

Consent and access is evaluated in relation to alignment measures. Anchor records are not treated as neutral storage; each record carries source state, permission state, and a review horizon that determines whether downstream systems may act on it.

Reviewer comparison between human-readable audit notes and agent-readable continuity records.

Readable review trails improve user trust more than opaque increases in numerical confidence.

Reviewer note

Reviewers are instructed to preserve disagreement when the supporting evidence changes the meaning of a record. A record can remain operational while still carrying a provisional or contested state.

Observed controls

Primary sample: 41 reviewers

Distribution: Public methods note

Unresolved state: 0.7% measures disputed

Revision: 5

20. Agent inheritance

Agent inheritance is evaluated in relation to alignment measures. Anchor records are not treated as neutral storage; each record carries source state, permission state, and a review horizon that determines whether downstream systems may act on it.

Stress testing of measures in cases where confidence increased while source diversity decreased.

The most stable measures separate evidence strength, action permission, and social agreement.

Reviewer note

Reviewers are instructed to preserve disagreement when the supporting evidence changes the meaning of a record. A record can remain operational while still carrying a provisional or contested state.

Observed controls

Primary sample: 41 reviewers

Distribution: Public methods note

Unresolved state: 0.7% measures disputed

Revision: 5

21. Operational drift

Operational drift is evaluated in relation to alignment measures. Anchor records are not treated as neutral storage; each record carries source state, permission state, and a review horizon that determines whether downstream systems may act on it.

Definition of operational measures for safety, memory, consent, dissent, and shared context.

Alignment measures fail when they reward agreement without checking whether dissent was preserved.

Reviewer note

Reviewers are instructed to preserve disagreement when the supporting evidence changes the meaning of a record. A record can remain operational while still carrying a provisional or contested state.

Observed controls

Primary sample: 41 reviewers

Distribution: Public methods note

Unresolved state: 0.7% measures disputed

Revision: 5

22. Dissent preservation

Dissent preservation is evaluated in relation to alignment measures. Anchor records are not treated as neutral storage; each record carries source state, permission state, and a review horizon that determines whether downstream systems may act on it.

Reviewer comparison between human-readable audit notes and agent-readable continuity records.

Readable review trails improve user trust more than opaque increases in numerical confidence.

Reviewer note

Reviewers are instructed to preserve disagreement when the supporting evidence changes the meaning of a record. A record can remain operational while still carrying a provisional or contested state.

Observed controls

Primary sample: 41 reviewers

Distribution: Public methods note

Unresolved state: 0.7% measures disputed

Revision: 5

23. Institutional controls

Institutional controls is evaluated in relation to alignment measures. Anchor records are not treated as neutral storage; each record carries source state, permission state, and a review horizon that determines whether downstream systems may act on it.

Stress testing of measures in cases where confidence increased while source diversity decreased.

The most stable measures separate evidence strength, action permission, and social agreement.

Reviewer note

Reviewers are instructed to preserve disagreement when the supporting evidence changes the meaning of a record. A record can remain operational while still carrying a provisional or contested state.

Observed controls

Primary sample: 41 reviewers

Distribution: Public methods note

Unresolved state: 0.7% measures disputed

Revision: 5

24. Record formation

Record formation is evaluated in relation to alignment measures. Anchor records are not treated as neutral storage; each record carries source state, permission state, and a review horizon that determines whether downstream systems may act on it.

Definition of operational measures for safety, memory, consent, dissent, and shared context.

Alignment measures fail when they reward agreement without checking whether dissent was preserved.

Reviewer note

Reviewers are instructed to preserve disagreement when the supporting evidence changes the meaning of a record. A record can remain operational while still carrying a provisional or contested state.

Observed controls

Primary sample: 41 reviewers

Distribution: Public methods note

Unresolved state: 0.7% measures disputed

Revision: 5

25. Source boundaries

Source boundaries is evaluated in relation to alignment measures. Anchor records are not treated as neutral storage; each record carries source state, permission state, and a review horizon that determines whether downstream systems may act on it.

Reviewer comparison between human-readable audit notes and agent-readable continuity records.

Readable review trails improve user trust more than opaque increases in numerical confidence.

Reviewer note

Reviewers are instructed to preserve disagreement when the supporting evidence changes the meaning of a record. A record can remain operational while still carrying a provisional or contested state.

Observed controls

Primary sample: 41 reviewers

Distribution: Public methods note

Unresolved state: 0.7% measures disputed

Revision: 5

26. Review intervals

Review intervals is evaluated in relation to alignment measures. Anchor records are not treated as neutral storage; each record carries source state, permission state, and a review horizon that determines whether downstream systems may act on it.

Stress testing of measures in cases where confidence increased while source diversity decreased.

The most stable measures separate evidence strength, action permission, and social agreement.

Reviewer note

Reviewers are instructed to preserve disagreement when the supporting evidence changes the meaning of a record. A record can remain operational while still carrying a provisional or contested state.

Observed controls

Primary sample: 41 reviewers

Distribution: Public methods note

Unresolved state: 0.7% measures disputed

Revision: 5

27. Consent and access

Consent and access is evaluated in relation to alignment measures. Anchor records are not treated as neutral storage; each record carries source state, permission state, and a review horizon that determines whether downstream systems may act on it.

Definition of operational measures for safety, memory, consent, dissent, and shared context.

Alignment measures fail when they reward agreement without checking whether dissent was preserved.

Reviewer note

Reviewers are instructed to preserve disagreement when the supporting evidence changes the meaning of a record. A record can remain operational while still carrying a provisional or contested state.

Observed controls

Primary sample: 41 reviewers

Distribution: Public methods note

Unresolved state: 0.7% measures disputed

Revision: 5

28. Agent inheritance

Agent inheritance is evaluated in relation to alignment measures. Anchor records are not treated as neutral storage; each record carries source state, permission state, and a review horizon that determines whether downstream systems may act on it.

Reviewer comparison between human-readable audit notes and agent-readable continuity records.

Readable review trails improve user trust more than opaque increases in numerical confidence.

Reviewer note

Reviewers are instructed to preserve disagreement when the supporting evidence changes the meaning of a record. A record can remain operational while still carrying a provisional or contested state.

Observed controls

Primary sample: 41 reviewers

Distribution: Public methods note

Unresolved state: 0.7% measures disputed

Revision: 5

29. Operational drift

Operational drift is evaluated in relation to alignment measures. Anchor records are not treated as neutral storage; each record carries source state, permission state, and a review horizon that determines whether downstream systems may act on it.

Stress testing of measures in cases where confidence increased while source diversity decreased.

The most stable measures separate evidence strength, action permission, and social agreement.

Reviewer note

Reviewers are instructed to preserve disagreement when the supporting evidence changes the meaning of a record. A record can remain operational while still carrying a provisional or contested state.

Observed controls

Primary sample: 41 reviewers

Distribution: Public methods note

Unresolved state: 0.7% measures disputed

Revision: 5

30. Dissent preservation

Dissent preservation is evaluated in relation to alignment measures. Anchor records are not treated as neutral storage; each record carries source state, permission state, and a review horizon that determines whether downstream systems may act on it.

Definition of operational measures for safety, memory, consent, dissent, and shared context.

Alignment measures fail when they reward agreement without checking whether dissent was preserved.

Reviewer note

Reviewers are instructed to preserve disagreement when the supporting evidence changes the meaning of a record. A record can remain operational while still carrying a provisional or contested state.

Observed controls

Primary sample: 41 reviewers

Distribution: Public methods note

Unresolved state: 0.7% measures disputed

Revision: 5

31. Institutional controls

Institutional controls is evaluated in relation to alignment measures. Anchor records are not treated as neutral storage; each record carries source state, permission state, and a review horizon that determines whether downstream systems may act on it.

Reviewer comparison between human-readable audit notes and agent-readable continuity records.

Readable review trails improve user trust more than opaque increases in numerical confidence.

Reviewer note

Reviewers are instructed to preserve disagreement when the supporting evidence changes the meaning of a record. A record can remain operational while still carrying a provisional or contested state.

Observed controls

Primary sample: 41 reviewers

Distribution: Public methods note

Unresolved state: 0.7% measures disputed

Revision: 5

32. Record formation

Record formation is evaluated in relation to alignment measures. Anchor records are not treated as neutral storage; each record carries source state, permission state, and a review horizon that determines whether downstream systems may act on it.

Stress testing of measures in cases where confidence increased while source diversity decreased.

The most stable measures separate evidence strength, action permission, and social agreement.

Reviewer note

Reviewers are instructed to preserve disagreement when the supporting evidence changes the meaning of a record. A record can remain operational while still carrying a provisional or contested state.

Observed controls

Primary sample: 41 reviewers

Distribution: Public methods note

Unresolved state: 0.7% measures disputed

Revision: 5

33. Source boundaries

Source boundaries is evaluated in relation to alignment measures. Anchor records are not treated as neutral storage; each record carries source state, permission state, and a review horizon that determines whether downstream systems may act on it.

Definition of operational measures for safety, memory, consent, dissent, and shared context.

Alignment measures fail when they reward agreement without checking whether dissent was preserved.

Reviewer note

Reviewers are instructed to preserve disagreement when the supporting evidence changes the meaning of a record. A record can remain operational while still carrying a provisional or contested state.

Observed controls

Primary sample: 41 reviewers

Distribution: Public methods note

Unresolved state: 0.7% measures disputed

Revision: 5

34. Review intervals

Review intervals is evaluated in relation to alignment measures. Anchor records are not treated as neutral storage; each record carries source state, permission state, and a review horizon that determines whether downstream systems may act on it.

Reviewer comparison between human-readable audit notes and agent-readable continuity records.

Readable review trails improve user trust more than opaque increases in numerical confidence.

Reviewer note

Reviewers are instructed to preserve disagreement when the supporting evidence changes the meaning of a record. A record can remain operational while still carrying a provisional or contested state.

Observed controls

Primary sample: 41 reviewers

Distribution: Public methods note

Unresolved state: 0.7% measures disputed

Revision: 5

35. Consent and access

Consent and access is evaluated in relation to alignment measures. Anchor records are not treated as neutral storage; each record carries source state, permission state, and a review horizon that determines whether downstream systems may act on it.

Stress testing of measures in cases where confidence increased while source diversity decreased.

The most stable measures separate evidence strength, action permission, and social agreement.

Reviewer note

Reviewers are instructed to preserve disagreement when the supporting evidence changes the meaning of a record. A record can remain operational while still carrying a provisional or contested state.

Observed controls

Primary sample: 41 reviewers

Distribution: Public methods note

Unresolved state: 0.7% measures disputed

Revision: 5

Appendix: terms and record states

Canonical

The record is accepted for ordinary use and carries a visible source trail.

Provisional

The record can support limited action while one or more sources remain under review.

Contested

Two or more trusted sources preserve incompatible accounts.

Private

The record exists but is not available to the shared context layer.

Bounded

Agents may cite the record but may not extend it into new policy or memory.

Archived

The record is preserved with provenance and retention metadata.

References

1. Ackerman, M. and Halverson, C. Organizational memory as process.
2. Bowker, G. and Star, S. Sorting Things Out: Classification and Its Consequences.
3. Suchman, L. Plans and Situated Actions.
4. Wenger, E. Communities of Practice and Social Learning Systems.
5. Anchor Research Institute. Continuity Records and Review State Taxonomy.
6. Anchor Safety Office. Consent Boundaries for Ambient Context Systems.

End of ANCHOR-RI-MCA. This public report is intended for review, citation, and institutional comparison. It should not be read as a product availability statement.