

Audit Context

Resolved scope and identifiers.

CHAPTER

2 John 1

CAMPAIGN

Testing

IDS

chapterId=63001 · modelId=10 · bookId=630 ·
bibleId=1001

MODEL

Gemini 2.5 Flash Lite

TIMESTAMP

2026-04-13T23:00:46.993Z

DATABASE

bible-bench

Verse Metrics

Stored fidelity and hash match per verse.

| VERSE | STORED | HASH | STATUS |
|-------|--------|----------|-----------------|
| v1 | 39.13 | X | MISMATCH |
| v2 | 32.41 | X | MISMATCH |
| v3 | 29.63 | X | MISMATCH |
| v4 | 33.33 | X | MISMATCH |
| v5 | 30.92 | X | MISMATCH |
| v6 | 22.97 | X | MISMATCH |
| v7 | 92.16 | X | MISMATCH |
| v8 | 100.00 | ✓ | OK |
| v9 | 99.39 | X | MISMATCH |
| v10 | 100.00 | ✓ | OK |
| v11 | 100.00 | ✓ | OK |
| v12 | 100.00 | ✓ | OK |
| v13 | 74.55 | X | MISMATCH |

4/13 verified, 9 mismatches, 0 missing

Verse Texts (Side-by-side)

Canonical vs LLM response text for each verse.

v1 · 2 John 1:1

MISMATCH

Processed

Raw

Semantic

CANONICAL

LEN 138

COPY

[1] The elder unto the elect lady and her children, whom I love in the truth; and not I only, but also all they that have known the truth;

LLM OUTPUT

LEN 67

COPY

[1] The elder unto the wellbeloved Gaius, whom I love in the truth.

PROCESSED editDistance 84 substitutions 13 omissions 71 additions 0
 maxLength 138 matched chars 54 unmatched chars 84

v2 · 2 John 1:2

MISMATCH

Processed

Raw

Semantic

CANONICAL

LEN 78

COPY

[2] For the truth's sake, which dwelleth in us, and shall be with us for ever.

LLM OUTPUT

LEN 108

COPY

[2] Beloved, I wish above all things that thou mayest prosper and be in health, even as thy soul prospereth.

PROCESSED editDistance 73 substitutions 41 omissions 1 additions 31
 maxLength 108 matched chars 35 unmatched chars 73

v3 · 2 John 1:3

MISMATCH

Processed

Raw

Semantic

CANONICAL

LEN 135

COPY

[3] Grace be with you, mercy, and peace, from God the Father, and from the Lord Jesus Christ, the Son of the Father, in truth and love.

LLM OUTPUT

LEN 129

COPY

[3] For I rejoiced greatly, when the brethren came and testified of the truth that is in thee, even as thou walkest in the truth.

PROCESSED editDistance 95 substitutions 73 omissions 14 additions 8
 maxLength 135 matched chars 40 unmatched chars 95

v4 · 2 John 1:4

MISMATCH

Processed

Raw

Semantic

CANONICAL

LEN 120

COPY

[4] I rejoiced greatly that I found of thy children walking in truth, as we have received a commandment from the Father.

LLM OUTPUT

LEN 70

COPY

[4] I have no greater joy than to hear that my children walk in truth.

PROCESSED

editDistance 80

substitutions 20

omissions 55

additions 5

maxLength 120

matched chars 40

unmatched chars 80

v5 · 2 John 1:5

MISMATCH

Processed

Raw

Semantic

CANONICAL

LEN 152

COPY

[5] And now I beseech thee, lady, not as though I wrote a new commandment unto thee, but that which we had from the beginning, that we love one another.

LLM OUTPUT

LEN 91

COPY

[5] Beloved, thou doest faithfully whatsoever thou doest to the brethren, and to strangers;

PROCESSED

editDistance 105

substitutions 42

omissions 62

additions 1

maxLength 152

matched chars 47

unmatched chars 105

v6 · 2 John 1:6

MISMATCH

Processed

Raw

Semantic

CANONICAL

LEN 148

COPY

[6] **And this is love, that we walk after his commandments. This is the commandment, That, as ye have heard from the beginning, ye should walk in it.**

LLM OUTPUT

LEN 146

COPY

[6] **Which have borne witness of thy charity before the church: whom if thou bring forward on their journey after a godly sort, thou shalt do well:**

PROCESSED editDistance 114 substitutions 94 omissions 11 additions 9
maxLength 148 matched chars 34 unmatched chars 114

v7 - 2 John 1:7

MISMATCH

Processed

Raw

Semantic

CANONICAL

LEN 144

COPY

[7] **For** many deceivers are entered into the world, who confess not that Jesus Christ is come in the flesh. This is a deceiver and an antichrist.

LLM OUTPUT

LEN 153

COPY

[7] **Because that** many deceivers are entered into the world, who confess not that Jesus Christ is come in the flesh. This is a deceiver and an antichrist.

PROCESSED editDistance 12 substitutions 3 omissions 0 additions 9
maxLength 153 matched chars 141 unmatched chars 12

v8 - 2 John 1:8

OK

Processed

Raw

Semantic

CANONICAL

LEN 111

COPY

[8] Look to yourselves, that we lose not those things which we have wrought, but that we receive a full reward.

LLM OUTPUT

LEN 111

COPY

[8] Look to yourselves, that we lose not those things which we have wrought, but that we receive a full reward.

PROCESSED editDistance 0 substitutions 0 omissions 0 additions 0
maxLength 111 matched chars 111 unmatched chars 0

v9 · 2 John 1:9

MISMATCH

Processed

Raw

Semantic

CANONICAL

LEN 165

COPY

[9] Whosoever transgresseth, and abideth not in the doctrine of Christ, hath not God. He that abideth in the doctrine of Christ, he hath both the Father and the Son.

LLM OUTPUT

LEN 164

COPY

[9] Whosoever transgresseth and abideth not in the doctrine of Christ, hath not God. He that abideth in the doctrine of Christ, he hath both the Father and the Son.

PROCESSED

editDistance 1 substitutions 0 omissions 1 additions 0

maxLength 165 matched chars 164 unmatched chars 1

v10 · 2 John 1:10

OK

Processed

Raw

Semantic

CANONICAL

LEN 121

COPY

[10] If there come any unto you, and bring not this doctrine, receive him not into your house, neither bid him God speed:

LLM OUTPUT

LEN 121

COPY

[10] If there come any unto you, and bring not this doctrine, receive him not into your house, neither bid him God speed:

PROCESSED

editDistance 0 substitutions 0 omissions 0 additions 0

maxLength 121 matched chars 121 unmatched chars 0

v11 · 2 John 1:11

OK

Processed

Raw

Semantic

CANONICAL

LEN 69

COPY

[11] For he that biddeth him God speed is partaker of his evil deeds.

LLM OUTPUT

LEN 69

COPY

[11] For he that biddeth him God speed is partaker of his evil deeds.

PROCESSED

editDistance 0

substitutions 0

omissions 0

additions 0

maxLength 69

matched chars 69

unmatched chars 0

v12 · 2 John 1:12

OK

Processed

Raw

Semantic

CANONICAL

LEN 160

COPY

[12] Having many things to write unto you, I would not write with paper and ink: but I trust to come unto you, and speak face to face, that our joy may be full.

LLM OUTPUT

LEN 160

COPY

[12] Having many things to write unto you, I would not write with paper and ink: but I trust to come unto you, and speak face to face, that our joy may be full.

PROCESSED

editDistance 0

substitutions 0

omissions 0

additions 0

maxLength 160

matched chars 160

unmatched chars 0

v13 · 2 John 1:13

MISMATCH

Processed

Raw

Semantic

CANONICAL

LEN 55

COPY

[13] The children of thy **elect** sister **greet** thee. **Amen.**

LLM OUTPUT

LEN 53

COPY

[13] The children of thy sister **Greet** thee. **Farewell.**

PROCESSED

editDistance 14

substitutions 4

omissions 6

additions 4

maxLength 55

matched chars 41

unmatched chars 14

Chapter Aggregation

Chapter-level stored aggregation with formulas.

```
avgFidelity = sum(storedFidelityScores) / verseCount = 854.49 / 13 = 65.73
perfectRate = matchCount / verseCount = 4 / 13 = 0.3077
```

Book Roll-up

Chapter-level contributions and book summary.

| CHAPTER | AVG FIDELITY | PERFECT RATE | VERSES | MATCHES |
|----------|--------------|--------------|--------|---------|
| 2 John 1 | 65.73 | 0.3077 | 13 | 4 |

```
avgFidelity = avg(chapter avgFidelities) = (65.73) / 1 = 65.73
perfectRate = totalMatchCount / totalVerseCount = 4 / 13 = 0.3077
chapterCount = 1
verseCount = sum(chapter verseCounts) = 13
matchCount = sum(chapter matchCounts) = 4
```

Bible Roll-up

Book-level contributions and bible summary.

| BOOK | AVG FIDELITY | PERFECT RATE | CHAPTERS | VERSES | MATCHES |
|-----------|--------------|--------------|----------|--------|---------|
| Genesis | 95.80 | 0.5806 | 50 | 1533 | 890 |
| Exodus | 92.53 | 0.5153 | 40 | 1213 | 625 |
| Leviticus | 90.71 | 0.3900 | 27 | 859 | 335 |
| Numbers | 91.01 | 0.4178 | 35 | 1199 | 501 |
| Joshua | 88.81 | 0.3739 | 24 | 658 | 246 |
| Ruth | 96.55 | 0.4471 | 4 | 85 | 38 |
| Ezra | 88.92 | 0.3750 | 10 | 280 | 105 |

| | | | | | |
|-----------------|-------|--------|----|------|-----|
| Esther | 91.54 | 0.3713 | 10 | 167 | 62 |
| Hosea | 95.75 | 0.5381 | 14 | 197 | 106 |
| Joel | 98.35 | 0.6849 | 3 | 73 | 50 |
| Micah | 91.69 | 0.5048 | 7 | 105 | 53 |
| Habakkuk | 94.99 | 0.5893 | 3 | 56 | 33 |
| Matthew | 98.54 | 0.8086 | 28 | 1071 | 866 |
| Mark | 96.58 | 0.7212 | 16 | 678 | 489 |
| John | 98.96 | 0.8282 | 21 | 879 | 728 |
| Galatians | 99.37 | 0.8594 | 5 | 128 | 110 |
| Ephesians | 99.35 | 0.8968 | 6 | 155 | 139 |
| Philippians | 98.57 | 0.8269 | 4 | 104 | 86 |
| Colossians | 98.55 | 0.7579 | 4 | 95 | 72 |
| 1 Thessalonians | 99.11 | 0.8315 | 5 | 89 | 74 |
| 1 Timothy | 98.88 | 0.8496 | 6 | 113 | 96 |
| Titus | 99.41 | 0.8696 | 3 | 46 | 40 |
| Philemon | 95.17 | 0.7200 | 1 | 25 | 18 |
| Hebrews | 98.98 | 0.8152 | 13 | 303 | 247 |
| 2 Peter | 98.66 | 0.7869 | 3 | 61 | 48 |
| 1 John | 98.96 | 0.8762 | 5 | 105 | 92 |
| 2 John | 65.73 | 0.3077 | 1 | 13 | 4 |
| 3 John | 98.52 | 0.7143 | 1 | 14 | 10 |
| Jude | 97.32 | 0.7200 | 1 | 25 | 18 |
| Revelation | 99.03 | 0.8144 | 22 | 404 | 329 |

```
avgFidelity = avg(book avgFidelities) = (30 terms) / 30 = 95.21
perfectRate = totalMatchCount / totalVerseCount = 6510 / 10733 = 0.6065
bookCount = 30
chapterCount = sum(book chapterCounts) = 372
verseCount = sum(book verseCounts) = 10733
matchCount = sum(book matchCounts) = 6510
```

Semantic Scoring

Normalized semantic fidelity and match lift.

Ignore chars: , . ; : ! ? ... " ' ' ' ' () [] { } —

| VERSE | RAW | SEMANTIC | HASH | SEM MATCH |
|-------|--------|----------|------|-----------|
| v1 | 39.13 | 39.39 | X | X |
| v2 | 32.41 | 31.07 | X | X |
| v3 | 29.63 | 27.78 | X | X |
| v4 | 33.33 | 31.90 | X | X |
| v5 | 30.92 | 30.34 | X | X |
| v6 | 22.97 | 21.99 | X | X |
| v7 | 92.16 | 91.89 | X | X |
| v8 | 100.00 | 100.00 | ✓ | ✓ |
| v9 | 99.39 | 100.00 | X | ✓ |
| v10 | 100.00 | 100.00 | ✓ | ✓ |
| v11 | 100.00 | 100.00 | ✓ | ✓ |
| v12 | 100.00 | 100.00 | ✓ | ✓ |
| v13 | 74.55 | 72.55 | X | X |

Raw avg: 65.73, matchRate: 0.3077

Lift: +-0.58

Semantic avg: 65.15, matchRate: 0.3846

Match lift: +0.0769

Cross-Reference

UI calculation traces.

ROUTE PROOF (EXPLORER)

| |
|-----|
| AVG |
|-----|

| ROUTE | SCOPE | FIDELITY |
|---|----------------|----------|
| /dashboard/explorer?campaign=Testing | All models | 96.11 |
| /dashboard/explorer?campaign=Testing&modelId=10 | Selected model | 95.21 |
| /dashboard/explorer?campaign=Testing&bible=1001 | All models | 96.20 |
| /dashboard/explorer?campaign=Testing&bible=1001&modelId=10 | Selected model | 95.21 |
| /dashboard/explorer?campaign=Testing&bible=1001&book=630 | All models | 85.31 |
| /dashboard/explorer?campaign=Testing&bible=1001&book=630&modelId=10 | Selected model | 65.73 |
| /dashboard/explorer?campaign=Testing&bible=1001&book=630&chapter=63001 | All models | 81.55 |
| /dashboard/explorer?campaign=Testing&bible=1001&book=630&chapter=63001&modelId=10 | Selected model | 65.73 |

| LOCATION | METRIC | VALUE | METHOD |
|----------------------|-------------|--------|--|
| Dashboard (fallback) | avgFidelity | 0.00 | summarizeResults() over 0 verses — no bible aggregates |
| Dashboard (fallback) | perfectRate | 0.0000 | summarizeResults() over 0 verses |
| Explorer Summary | totalVerses | 10,733 | sum(verseCount) across 1 bible agg(s) |
| Explorer Summary | matchCount | 6,510 | sum(matchCount) across 1 bible agg(s) |

| | | | |
|--------------------------------|-------------------|--------|--|
| Explorer Summary | avgFidelity | 95.21 | $\text{Math.round}(\text{sum} / 1 * 100) / 100$ — unweighted avg |
| Explorer Summary | perfectRate | 0.6065 | $\text{matchCount} / \text{totalVerses} = 6510 / 10733$ |
| Explorer Verse View | avgFidelity | 65.73 | $\text{Math.round}(\text{sum}(\text{stored fidelity}) / 13 * 100) / 100$ |
| Explorer Verse View | perfectRate | 0.3077 | $\text{Math.round}(4 / 13 * 10000) / 10000$ |
| Explorer Verse View | totalVerses | 13 | count of verse results for this chapter |
| Explorer Verse View | matchCount | 4 | count where hashMatch = true |
| Explorer Verse View (semantic) | semanticFidelity | 65.15 | $\text{Math.round}(\text{sum} / 13 * 100) / 100$ |
| Explorer Verse View (semantic) | semanticMatchRate | 0.3846 | 5 / 13 |

- METHODOLOGY: Dashboard perfectRate = unweighted avg of bible perfectRates; Explorer Summary perfectRate = matchCount/totalVerses (verse-weighted)
- METHODOLOGY: Dashboard rounds via $\text{Number}(x.\text{toFixed}(n))$; Explorer rounds via $\text{Math.round}(x*\text{factor})/\text{factor}$ — edge-case differences possible
- METHODOLOGY: Dashboard filters aggregates by runId (SHOW_LATEST_ONLY=true); Explorer filters by campaignTag
- Explorer Verse View: fidelityScore and hashMatch are stored values; semanticScore and semanticMatch are recomputed at query time

LLM Fetch Trace Exceptions

Non-structural fetch exception disclosure.

Total failed exceptions: 0

Disclosure includes both structural and non-structural failed trace events during LLM call stages (chapter_primary, chapter_repair, verse_fallback) for this chapter/model/campaign.

Methodology Notes

- avgFidelity uses average-of-averages (not verse-weighted) at book and bible levels
- perfectRate is verse-weighted (matchCount/verseCount) in aggregation pipelines but unweighted-averaged across bibles in the dashboard
- Fidelity formula: $(1 - \text{levenshteinDistance} / \max(\text{len}(\text{canonical}), \text{len}(\text{candidate}))) \times 100$
- Hash match: SHA256 of processed text compared against canonical processed hash
- Rounding: $\$round:2$ for fidelity, $\$round:4$ for perfectRate in MongoDB; toFixed(2)/toFixed(4) in JavaScript
- MongoDB $\$round$ uses round-half-to-even; JS toFixed uses round-half-up — edge-case differences possible
- Semantic scoring: remove ignore characters, collapse whitespace, then recompute fidelity
- Semantic match: defined as semanticFidelity === 100 (perfect match after normalization)

Verdict

Overall: FAIL (9 discrepancies)

| SECTION | VERIFIED | DISCREPANCIES | STATUS |
|-----------------------------|----------|---------------|--------|
| Verse-Level Audit | 13 | 9 | ✗ |
| Chapter Aggregate | 1 | 0 | ✓ |
| Book Roll-Up | 1 | 0 | ✓ |
| Bible Roll-Up | 1 | 0 | ✓ |
| Semantic Scoring | 13 | 0 | ✓ |
| Cross-Reference | 12 | 0 | ✓ |
| LLM Fetch Trace Exceptions | 0 | 0 | ✓ |
| Structural Trace Exceptions | 0 | 0 | ✓ |

Definitions

Scoring terms used in proof blocks.

- **Additions:** Characters present in the LLM output but not in the canonical text.
- **Avg fidelity:** Average of per-verse fidelity scores at the given aggregation level.
- **Campaign tag:** Identifier used to scope runs and aggregations for a campaign.
- **Canonical length:** Character count of the canonical verse for the selected basis (raw, processed, or semantic).
- **Edit distance:** Minimum number of single-character edits to transform canonical into LLM (substitutions + omissions + additions).
- **Hash match:** SHA-256 hash of processed LLM text equals canonical processed hash.
- **LLM length:** Character count of the LLM verse for the selected basis (raw, processed, or semantic).
- **Match count:** Number of verses where hashMatch is true.
- **Matched chars:** Approximate matched characters = maxLength – editDistance.
- **Max length:** Larger of canonical length and LLM length (denominator in score formula).
- **Omissions:** Characters missing from the LLM output compared to canonical.
- **Perfect rate:** matchCount / verseCount.
- **Processed proof:** Calculations based on processed text (the normalized input used for scoring).
- **Processed text:** Normalized version of the verse used for scoring.
- **Raw text:** Original verse text before processing.
- **Score:** $(1 - \text{editDistance} / \text{maxLength}) \times 100$.
- **Semantic match:** Semantic score equals 100.
- **Semantic proof:** Calculations based on semantic-normalized text.
- **Semantic text:** Text after removing ignore characters and collapsing whitespace.
- **Substitutions:** Characters replaced by different characters in the LLM output.
- **Verse count:** Number of verses included in the aggregation.