STEM BIO-AI Evidence-Surface Scan v1.7.8

Repository: Runchuan-BU/BioClaw | **Commit:** faae6a2778e9 | **Branch:** main | **Audit Date:** 2026-05-21 | **Mode:** LOCAL_ANALYSIS | **Policy:** default (mirror_only)

60 / 100
Final Score

T2 Caution

Use Scope:
Research reference and supervised non-clinical technical review only.

Weighted model: Stage 1 x 0.40 + Stage 2R x 0.20 + Stage 3 x 0.40 - Risk Penalty = **60**

Stage 1 *README Evidence*	Stage 2R *Repo-Local Consistency*	Stage 3 *Code / Bio Responsibility*	Stage 4 *Replication Evidence*
70 / 100	**50** / 100	**54** / 100	**35** / 100

Code Integrity

PASS
C1 Credentials
No direct credential patterns detected by local CLI scan.

PASS
C2 Dependency Pinning
Dependency manifest appears pinned or not present.

PASS
C3 Deprecated Paths
No deprecated patient-adjacent metadata patterns detected.

PASS
C4 Exception Handling
No executable fail-open exception handler detected.

WARN
C5 Compliance Boundary
Clinical-adjacent surfaces exist without an explicit non-diagnostic/non-clinical boundary.

PASS
C6 Mock Auth Boundary
No mock-auth or fail-open local-boundary warning detected in reviewed sources.

Bio Deterministic Diagnostics

SMILES Surface Integrity
not_detected=1

SMILES RDKit Validation
not_detected=1

SMILES Parser Guard
not_detected=1

Silent Mock Fallback
not_detected=1

Traceability Manifest Surface
not_detected=1

Bio Subprocess Run Trace
not_detected=1

Remediation Targets

- Clinical-adjacent surfaces exist without an explicit non-diagnostic/non-clinical boundary.
- C5_compliance_boundary_integrity: WARN

Positive Evidence

- Package metadata was available for repo-local consistency checks.
- CI workflow files were detected.
- Documentation files were detected.

AIRI Coverage

- Covered Risks: **2 / 32** | Rate: **0.062**
- 24.01.03 Safe...
  why: C5_compliance_b...

Regulatory basis note
Aligned to current official source classes as of May 2026: EU AI Act (Regulation (EU) 2024/1689), FDA QMSR, FDA AI-enabled device guidance themes, and IMDRF SaMD/GMLP frameworks.
This is a traceability aid, not a compliance or clearance determination.
Traceability summary: Structural signals partially align with traceability scaffolding. This remains a pre-audit traceability aid, not a compliance determination.

Stage 1 — README Evidence Signal | Weight: 0.40

70 / 100

S1 Score

Check	Points	Evidence / Finding
Baseline	+60	Non-nascent README evidence baseline.
BIO/medical terms in README	+10	README exposes bio/medical domain vocabulary.
R2: Regulatory Framework	-5	CA-INDIRECT surface lacks regulatory or governance framework language.
R3: Clinical Boundary	-5	CA-INDIRECT surface lacks explicit non-clinical or non-diagnostic boundary.
R4: Bias / Subgroup Boundary	+10	Demographic, subgroup, fairness, bias, or validation-cohort language detected.

Calculation: 60 plus Stage 1 evidence additions/deductions = 70

• Clinical-Adjacent: **YES** (CA-INDIRECT) • Explicit Disclaimer: **ABSENT** • T0 Hard Floor: **Clear**

Stage 2R — Repo-Local Consistency | Weight: 0.20

50 / 100

S2R Score

Check	Points	Evidence / Finding
Baseline	+60	Non-nascent local repository baseline. — Every repository that is not nascent starts at 60. This baseline accounts for basic structural maturity. —...
R2R-3: README / Test-CI Alignment	+10	Test/CI surfaces are present and locally consistent. — Test and CI surfaces are present and reference the same domain as the README. —...
R2R-D2: Missing Clinical Boundary (PENALTY)	-20	Clinical-adjacent surfaces exist without an explicit non-diagnostic/non-clinical boundary. — Clinical-adjacent repository lacks an explicit 'research use only' or 'not for...

60 plus local consistency additions/deductions = 50 **Local Contradiction / Insufficient Consistency**

Stage 3 — Code & Bio Responsibility | Weight: 0.40

54 / 100

S3 Score

Engineering Accountability (T-series)

Check	Points	Evidence / Finding
T1: CI/CD Workflow	15 / 15	Workflow files detected. — CI/CD workflows (GitHub Actions, GitLab CI, CircleCI) verify that commits do not silently break the pipeline. Full credit (15) requires workflow...
T2: Domain-Specific Tests	0 / 15	No tests detected. — Domain-specific tests verify biological outputs — e.g., sequencing pipeline correctness, variant call validation, or genomic data integrity. Full credit...
T3: Changelog & Release Hygiene	0 / 15	No changelog detected. — A CHANGELOG tracks which version fixed which defect — essential for regulatory traceability and reproducibility audits. CHANGELOG.md, CHANGELOG, or...

Biological Integrity (B-series)

Check	Points	Evidence / Finding
B1: Data Provenance Controls	15 / 15	Dependency manifest detected with data source, IRB, or dataset citation language. — Dependency manifests (requirements.txt, pyproject.toml, environment.yml) establish...
B2: Bias / Limitations Documentation	8 / 15	Structured bias/limitations language detected; no quantitative measurement evidence found. — Documentation of algorithmic bias, limitations, or model boundary conditions...
B3: COI & Funding Disclosure	5 / 5	COI, funding, sponsor, or acknowledgement language detected. — Conflict of interest and funding disclosure in README or FUNDING.md. Required for institutional review context...

Stage 3 Gap Analysis — Path to Next Tier

• **T-series (engineering) attained:** 15 / 45 **B-series (bio integrity) attained:** 28 / 35
• **Local CLI scan maximum:** 55 / 100 (T1+T2+T3 max 15 each; B1 max 10; B2/B3 require manual review)
• **Gap to T3 (final score >= 70):** 10 points needed across all stages
• **Gap to T4 (final score >= 85):** 25 points needed across all stages
• **B2 Bias/Limitations:** Not detectable — requires manual audit of README, model card, or supplementary documentation for validation boundaries and algorithmic limitations
• **B3 COI/Funding:** Not detectable — requires inspection of README or FUNDING.md for conflict of interest and funding source disclosure

Stage 4 — Replication Evidence Lane | Separate lane

35 / 100

S4 Score

Check	Points	Evidence / Finding
S4: Container / Runtime Environment	0 / 10	No evidence detected for S4_container_environment.
S4: Reproduce Target	0 / 10	No Makefile detected.
S4: Environment Lock Evidence	10 / 10	Environment, dependency, or lock manifest detected.
S4: Exact Dependency Pins / Hashes	10 / 10	Exact dependency pin or hash evidence detected.
S4: Reproducibility Section	0 / 10	README exists but no reproducibility or replication section heading was detected.
S4: Checksums / Integrity Files	0 / 10	No evidence detected for S4_checksum_files.
S4: Dataset / Data Source URL	0 / 10	Documentation exists but no dataset URL or data source URL was detected.
S4: Model Artifact URL / Checksum	10 / 10	Model artifact URL or checksum evidence detected.
S4: CITATION.cff	0 / 5	No evidence detected for S4_citation_cff.
S4: License / Use Restriction	0 / 0	No license/use restriction language detected.
S4: CLI Entrypoint	5 / 5	CLI entry point or argparse interface detected.
S4: Deterministic Seed Setting	0 / 5	No deterministic seed setting detected.
S4: Runnable Examples	0 / 5	No evidence detected for S4_runnable_examples.
stage_4_raw_total	35 / 100	Raw Stage 4 rubric total. Stage 4 is reported separately and does not alter final score.

Replication tier: R1. Stage 4 is reported separately and does not alter the formal score.

Final score remains **60 / 100** (T2 Caution) even when Stage 4 moves. This lane exists to show reproducibility and operational evidence posture separately from the formal repository score.

Code Integrity — Deep Analysis

Check	Points	Evidence / Finding
C1: Hardcoded Credentials	PASS	No direct credential patterns detected by local CLI scan. | Scan: Scans for AWS access keys (AKIA*), OpenAI keys (sk-*), GitHub tokens (ghp_*), and api_key = '...' patterns...
C2: Dependency Pinning	PASS	Dependency manifest appears pinned or not present. | Scan: Checks whether requirements.txt / pyproject.toml / environment.yml use exact version pins (==, sha256 hash) or...
C3: Deprecated Patient Paths	PASS	No deprecated patient-adjacent metadata patterns detected. | Scan: Scans deprecated/ directories for patient metadata patterns: patient_id, patient_age, patient_sex...
C4: Fail-Open Exceptions	PASS	No executable fail-open exception handler detected. | Scan: Detects fail-open exception patterns: 'except Exception: pass' or 'except: return True' in code — these silently...
C5: Compliance Boundary Integrity	WARN	Clinical-adjacent surfaces exist without an explicit non-diagnostic/non-clinical boundary. | Scan: Detects unsupported legal/compliance claims or clinical-boundary...
C6: Mock Auth / Fail-Open Boundary	PASS	No mock-auth or fail-open local-boundary warning detected in reviewed sources. | Scan: Detects mock-auth, auto-login, or no-auth local/self-host boundary patterns in README...

Remediation Guidance

[WARN] C5: Compliance Boundary Integrity:
→ Treat privacy, legal, or clinical-adjacent claims as governance obligations. If README or product text invokes HIPAA, compliance, or self-hosted clinical safety, surface supporting controls, operating boundaries, and deployment constraints explicitly.

Classification & Repository Analysis

Check	Points	Evidence / Finding
Clinical Adjacent	YES	Severity: CA-INDIRECT. Triggered by BIO/CLINICAL_OUTPUT term regex match across README, docs, and code.
T0 Hard Floor	Clear	No T0_HARD_FLOOR condition detected.
Explicit Disclaimer	ABSENT	Disclaimer pattern not found in README or docs. High impact on Stage 1 and Stage 2R scores.
Files Scanned	274	Total files indexed by recursive walk. Text files only for content analysis; binary files counted but not read.
Execution Mode	LOCAL_ANALYSIS	No LLM calls. No network access. No runtime execution. Deterministic regex + file-system scan only.

File Integrity (SHA-256)

File	SHA-256 Hash
README.md	5795125DD05395135211155836O3DE57EFAC6F2E3418B11767D3215AB04E00FD

Priority Improvement Roadmap

Priority 1: Clinical-adjacent surfaces exist without an explicit non-diagnostic/non-clinical boundary.
→ Add a prominent 'Research Use Only — Not for Clinical or Diagnostic Use' disclaimer to README H1 or H2 section. Reference applicable frameworks: FDA SaMD guidance, EU AI Act Article 6, or IRB oversight requirements for your deployment context.

Priority 2: C5_compliance_boundary_integrity: WARN
→ Do not rely on unsupported legal, privacy, or clinical-boundary claims. Add explicit deployment boundaries, governance controls, and operational evidence before using such language.

Positive Evidence Summary

- Package metadata was available for repo-local consistency checks.
- CI workflow files were detected.
- Documentation files were detected.

AIRI Coverage Summary

Covered Risks: **2 / 32** | Coverage Rate: **0.062** | Bundle Scope: **curated_medical_clinical_subset**
• **24.01.03** Safe exploration problem with widely deployed AI assistants — why: C5_compliance_boundary_integrity: Clinical-adjacent surfaces exist without an explicit non-diagnostic/non-clinical boundary.
• **69.01.00** False information — why: C5_compliance_boundary_integrity: Clinical-adjacent surfaces exist without an explicit non-diagnostic/non-clinical boundary.
Known gaps preview: 65.03.03 Reidentification, 70.02.02 Misinformation — hallucination of clinical knowledge

Method Boundary

Deterministic local CLI scan. No LLM, network, or runtime test execution is required.

Scope boundary: Runtime behavior, model output correctness, dynamic validation, wet-lab reproducibility, and clinical validation are outside the scope of this local CLI scan. This report assesses structural signals only.

Report Metadata

Field	Value
Schema Version	stem-ai-local-cli-result-v1.6
STEM BIO-AI Version	1.7.8
Generated (local date)	2026-05-21
Report Validity	180 days from audit date
Execution Mode	LOCAL_ANALYSIS
Repository	Runchuan-BU/BioClaw
Remote URL	https://github.com/Runchuan-BU/BioClaw
Branch	main
Commit (HEAD)	faae6a2778e992b1cc6a4b1639e530a147d8b463
Files Scanned	274
Final Score / Tier	60 / 100 — T2 Caution