Creative Intelligence Framework (CIF) — Technical Whitepaper v1.2

Status: Draft for review (methods-first; code-free)
Date: 2025-11-03 (America/Chicago)
Supersedes: v1.1.1 (+ Addendum A clarifications)
Primary Author: Keith Hetrick (Architecture, Methods, Governance)
Contributors: Keith Hetrick — Maintainer
Affiliation: Bellweather Studios
Contact: keith@bellweatherstudios.com (mailto:keith@bellweatherstudios.com)
Doc ID: **CIF-TW-v1.2**

Confidentiality: Internal draft; redistribution only with approval.

Abstract

We present **CIF v1.2**, an auditable framework for evaluating songs using one client-facing KPI, the **Hit Confidence Index (HCI)**. HCI is presently computed from a six-axis **Audio Echo Engine (AEE)** whose core model is the **Historical Echo Model (HEM)**. Lyrics are **advisory-only** in v1.2. The release formalizes parameter discipline and bounds; codifies a **Dual-Engine** future with a **Lyric Echo Engine (LEE)**; fixes governance via the **40/40 Balance Rule** (equal policy weight to audio/lyric upon LEE graduation); and adds stability-first validation (bootstrap CIs, independence checks), a Host-level **Suggestion/Optimization Layer**, Trend-hydrated platform lanes (Radio US canonical; Spotify advisory), fail-ledger observability, and strict provenance. **HCI math is unchanged** (audio-only KPI; lyrics advisory).

Keywords: Music Information Retrieval (MIR), historical echo, composite index, bootstrap CIs, independence (VIF), TTC, LUFS, governance, provenance.

Document Control

- **Version**: v1.2 (draft)
- **Change Summary**: Integrates Addendum A; defines Radio US as canonical KPI lane; Spotify advisory; Host-only advisory invariant; Trend Snapshot advisory data; axis weight manifests; constants table; golden-run tolerance; policy switch registry; FAQ; accessibility checklist.
- **Reviewers**: [list]
- **Approvals**: [list]
- **Effective Date**: [date]
- **Superseded Docs**: CIF v1.1.1 (PDF) + Addendum A

Global Acronyms, Symbols & Terms (one-page reference)

Engines & Models

AEE — Audio Echo Engine (active KPI domain)

LEE — Lyric Echo Engine (advisory in v1.2; score-contributing post-graduation)

HEM — Historical Echo Model (inside AEE)

HLM — Historical Lyric Model (inside LEE)

Language Intelligence (Lyric) — domain label for the lyric/linguistic layer within CIF (served by LEE)

Composites & KPI

AER — Audio Echo Resonance (AEE composite)

LER — Lyric Echo Resonance (LEE composite; future KPI input)

HCI — Hit Confidence Index (public KPI)

EACM — Equal-Axis Composite Mean (engine 6-axis mean)

Variables & Ops

x (raw feature), μ,σ (profile norms); $z=(x-\mu)/\sigma$; logistic map $s=\sigma(\gamma z)\in[0,1]$; $\gamma\in[0.8,1.2]$

$A_k\in[0,1]$ (axis k); $EACM^* = (1/6)\sum A_k\in[0,1]$

$\beta\in[0,1]$ (cross-engine policy weight); c_audio,c_lyric (domain caps)

HEM geometry: cosine distance; σ via median 8-NN
Diagnostics: correlation matrix; VIF; condition number κ
DBI — Decadal Balance Index (diagnostic; report-only)
TTC — Time-to-Chorus (s); Chorus-lift — Δ short-term LUFS (dB)

Governance & Provenance
Canonical KPI lane: **Radio US**; advisory lanes: Radio, Spotify
40/40 Balance Rule (β=0.5 post-LEE graduation)
Whitening (Lenient): OFF unless κ>50 or max VIF>7
Seeds: {42, 314, 2718}; float64; BLAS threads pinned
Cards: Profile / Run / Environment; Provenance footer on all figs/tables

1. Introduction & Scope

CIF measures **balanced creative fit** via six interpretable axes aggregated into **HCI**. v1.2 clarifies methods, governance, and reproducibility, establishes **Radio US** as the canonical evaluation lane, and introduces a Host-level, deterministic **Suggestion/Optimization** system that is lane- and Trend-aware yet **read-only** with respect to scoring. Lyrics remain advisory-only; HCI math is unchanged.

Non-goals: sales forecasts; legal/licensing judgments; auto-mixing or lyric rewriting in the scoring path.

2. Background & Prior Art

Aligned with MIR practice: per-feature normalization; logistic mappings to [0,1]; convex axis construction; arithmetic consolidation for interpretability (geometric mean used only in sensitivity). Emphasis on reproducibility (seeds, dtype, corpora stability) and fairness guardrails (caps, independence). Prior work on structure detection (TTC) and loudness (LUFS) informs segmentation features.

3. System Overview

3.1 End-to-End Pipeline

Automator → Preprocess (standardize, winsorize, logistic) → Axis construction (**AEE**; **LEE** advisory) → Engine composites (AER/LER) → Caps → **Host** Consolidation (HCI; Radio lane) → **Host** Suggestion/Optimization (Radio + Spotify; Trend-hydrated) → Artifacts (figs/tables/cards).

3.2 Music Advisor Logic — Present & Modular Future

Present (v1.2): AEE → AER → HCI (audio-only KPI). LEE outputs advisory analyses; any lyric-like proxies in AEE are flagged and tapered post-graduation.
Future: Host consolidates AER and LER by policy β with domain caps.

HCI = β·min(EACM_audio,c_audio) + (1−β)·min(EACM_lyric,c_lyric)
Today β=1.0; post-graduation default β=0.5 (**40/40**).

3.3 Component Placement & Responsibilities

- **Layer 0 — Ingest & QA**: Automator extracts tempo, key, TTC(+conf), structure, loudness (integrated & ST-LUFS), lyrics(+conf); emits QA flags.
- **Layer 1 — Standardization & Mapping**: z-scores; winsor $|z|>4$; logistic $s=\sigma(\gamma z)$; validators can set NA; axes renormalize locally.
- **Layer 2A — AEE**: HEM; six audio axes → AER.
- **Layer 2B — LEE**: HLM; six lyric axes (advisory) → LER.
- **Layer 3 — Host**: domain caps; consolidation (HCI, Radio lane); **Suggestion/Optimization** (Radio+Spotify; Trend Snapshot; Baseline vs Curation).
- **Layer 4 — Governance & Validation**: β, caps, whitening triggers; bootstrap CIs; independence; DBI.
- **Layer 5 — Artifacts**: KPI, axis summaries, advisory, Cards, figs/tables with provenance.

> **Architectural Invariant — Host-Only Advisory > Engines (AEE/LEE) are measurement systems only.** They **do not** read Trend data, platform lanes, or user preferences, and they **do not** emit suggestions. Their sole output is AER/LER with axis-level provenance.
> **The Host (Music Advisor) alone executes all suggestion/optimization** after caps & consolidation (HCI). Host advisory may read lanes/Trend/QA/preferences **but has no feedback path** into features, axes, AER/LER, caps, β, or HCI.

3.4 Axis Contract (engine-neutral interface)

Input: FeatureSet + profile ID + model ver + corpus hash + seeds.
Output: six axis scores, `axis_components`, EACM*, `quality` flags, full `provenance`. Missing features are **excluded** with **local renormalization**. Host never multiplies axes or HCI.

3.5 Governance: Lanes, β, Caps, Overrides

- **Canonical KPI lane**: **Radio US**.
- **Advisory lanes**: Radio US + Spotify (duplicates auto-hidden unless advice differs).
- **β**: 1.0 today; **40/40 (β=0.5)** post-LEE graduation.
- **Caps**: c_audio=c_lyric=0.58 defaults (sensitivity ±0.02 in Appendix).
- **Overrides**: Profile-scoped; require Profile Card & Change-Log entries.

3.6 Gates & Advisory (Two-Mode)

Placement & Isolation. All suggestion/optimization runs in the Host *after* HCI is finalized. Advisory is lane-/Trend-aware and deterministic, and is **strictly read-only** with respect to scoring: **no adjustments** to features, axes, AER/LER, caps, β, or HCI are permitted.
Two-Mode Advisory. *Evaluation* renders baseline market-norm suggestions (Radio canonical; Spotify advisory). *Curation* (opt-in) may re-order/hide advice per preference; safety items are pinned; `preferences_used` is recorded.

4. Methods (formal)

4.1 Parameter Discipline & Units

Standardization: z=(x−μ)/σ; winsor |z|>4. Logistic mapping: s=1/(1+e^{−γz}), γ∈[0.8,1.2]. Units canon: tempo(BPM), TTC(s), Chorus-lift(dB via 6s ST-LUFS), loudness (integrated LUFS).

4.2 Axis Construction (per engine)

Axis A*k* = ∑{j∈F_k} w*{k,j}·s*{k,j}, w*{k,j}≥0, ∑ w*{k,j}=1.
Invalid features are excluded; remaining weights renormalize to 1 for that track. Axes ∈[0,1].

AEE axes (today): Market, Sonic, Emotional, Historical/Echo, Cultural, Creative.
LEE axes (future): Lyric Craft, Hook Architecture, Narrative Specificity, Prosody Alignment, Thematic Resonance, Linguistic Flow. *(Prosody uses alignment metadata only; no spectral imports.)*

4.3 Engine Composites, Caps, Bounds

EACM* = (1/6)∑ A_k; domain caps post-aggregation: ~E* = min(EACM*, c_*).
Bounds/Monotonicity: increasing any A_k never decreases EACM* or HCI after caps; single-axis change Δ bounded in KPI by β·Δ/6 (audio) and (1−β)·Δ/6 (lyric).

4.4 Historical Echo Model (HEM)

Cosine geometry in standardized space; kernel width σ via median 8-NN within the anchored 40-year corpus. **Whitening** OFF unless κ>50 or max VIF>7 (Lenient); if triggered, Appendix D audit prints pre/post spectra and corr matrices.

4.5 Segmentation Features (AEE)

TTC: first chorus onset with detector conf ≥ 0.60; else NA.
Chorus-lift: Δ of 6s ST-LUFS windows (chorus vs verse). Validators may set NA, leading to local axis renormalization.

4.6 Consolidation (Host)

HCI = β·min(EACM_audio, c_audio) + (1−β)·min(EACM_lyric, c_lyric).
Today: β=1.0 ⇒ HCI = min(EACM_audio,c_audio).
Sensitivity (Appendix): geometric mean reported for completeness (not KPI).

4.7 Missing-Feature Policy

Invalid features → NA; exclude and renormalize within the affected axis. Max HCI impact bounded by β/6 (audio) and (1−β)/6 (lyric), further limited by caps.

5. Automator (Data Extract Layer)

Extractor, not a model. Produces FeatureRecord: tempo, key, TTC(+conf), structure, loudness (integrated & ST-LUFS), lyrics(+conf), provenance. **Goldilocks Safeguard** for lyrics; may set lyric domain NA (advisory may still render with low-confidence labels). QA taxonomy: BLOCK/WARN/INFO; Run Card logs acceptance.

6. Language Intelligence (Lyric) — LEE (status & graduation)

Status: advisory-only; analyses do not change HCI. Lyric-like proxies in AEE are flagged; taper to 0 on graduation to prevent double-counting.
Six Lyric Axes (parity): Lyric Craft; Hook Architecture; Narrative Specificity; Prosody Alignment; Thematic Resonance; Linguistic Flow.
Graduation Gate: repeatability; bootstrap CI tightness; independence (corr≤~0.6; VIF≤5 target, ≤7 cap vs AEE); temporal stability; placebo (lyric shuffle collapses to baseline); face validity.
Post-graduation: LER contributes by β with caps; 40/40 default.

7. Validation & Uncertainty

Fig 7.1 (Stability): track-level bootstrap (B=100) 95% CI ribbons for HCI (Radio lane). Seeds {42,314,2718}.
Fig 7.2 (Independence): axis corr heatmap + VIF bars (goal ≤5; cap 7). Whitening audit in Appendix if triggers fire.

8. Ethics, Data Governance, IP

Provenance: corpus hashes; anchored_40yr window on Profile Card. Lyric compliance: permissions, PII redaction, limited display. Advisory-not-predictive. Methods/parameterizations proprietary; **no code** in doc.

9. Results — Canonical Profile (US_Pop_2025)

KPI lane: **Radio US** (canonical). **Advisory lanes**: Radio + Spotify; duplicates auto-hidden unless lane advice differs.

> **Advisory is post-score and Host-only.** HCI is computed first (Radio canonical), then lane-specific checklists are rendered. **Suggestions never modify HCI.**

Table 9.1: axis means ± sd (over seeds); **HCI** mean ± sd; 95% bootstrap CI; cap-utilization %.
Worked Example (Track Alpha, anonymized): standardize → map → axes → EACM → cap → **HCI (Radio)**; then lane-specific checklists (Radio, Spotify).

10. Discussion & Limitations

Segmentation brittleness (TTC/lift), profile drift risk, proxy leakage risk pre-LEE graduation, compensability trade-offs (arithmetic KPI vs geometric sensitivity), interpretability governance (caps, β fixed by policy).

11. Conclusion & Roadmap

HCI remains stable and auditable. v1.2 prepares a dual-engine future with explicit governance and Host-only advisory. Near-term: LEE shadow mode; placebo tests; independence audits; proxy taper; Trend Snapshot cadence; Radio canonical lane review windows.

12. Change Log (v1.1.1 → v1.2)

Parameter discipline; bounds/monotonicity; HEM σ-rule; Lenient whitening; segmentation acceptance; consolidation math; missing-feature bound; modular Axis Contract; governance (β/caps/overrides, lanes); Host-only advisory invariant; validation exhibits; DBI diagnostic; Cards & Provenance; constants table; policy switch registry; FAQ.

Axis Weight Manifests (main text mini-tables)

> *Note:* populate weights from source repo manifests; rows must sum to 1. Proxy taper flagged where applicable.

AEE — Market Axis

Feature (mapped)	Weight	Units / Window	Proxy-taper note
[placeholder]	0.00	[e.g., BPM band]	–
…	…	…	…
Sum	**1.00**		

AEE — Sonic Axis

Feature (mapped)	Weight	Units / Window	Proxy-taper note
[placeholder]	0.00	[e.g., spectral tilt]	–
…	…	…	…
Sum	**1.00**		

AEE — Emotional Axis

Feature (mapped)	Weight	Units / Window	Proxy-taper note
[placeholder]	0.00	[e.g., valence proxy]	–
…	…	…	…
Sum	**1.00**		

AEE — Historical/Echo Axis

Feature (mapped)	Weight	Units / Window	Proxy-taper note
[placeholder]	0.00	[e.g., HEM proximity]	–
…	…	…	…
Sum	**1.00**		

AEE — Cultural Axis

Feature (mapped)	Weight	Units / Window	Proxy-taper note
[placeholder]	0.00	[e.g., rhythm archetype]	–
…	…	…	…
Sum	**1.00**		

AEE — Creative Axis

Feature (mapped)	Weight	Units / Window	Proxy-taper note
[placeholder]	0.00	[TTC, ST-LUFS lift]	–
…	…	…	…
Sum	**1.00**		

> **LEE axis tables** mirror the above (Lyric Craft, Hook Architecture, Narrative Specificity, Prosody Alignment, Thematic Resonance, Linguistic Flow). Populate once HLM components are frozen. Prosody uses alignment metadata only.

Numerical Constants & Switches (main text)

Constant / Policy	Value	Where used		
Caps (audio, lyric)	0.58	§3.5, §4.3 (post-aggregation)		
TTC acceptance conf	0.60	§4.5 validators		
Winsor cutoff		z	> 4	§4.1
Seeds	{42,314,2718}	§7, Cards		
Bootstrap B	100	§7.1		
Whitening triggers	κ>50 or VIF>7	§4.4, App D		
Golden-run tolerance		ΔHCl	≤ 0.002 per track	HOWTO_REPRODU

Policy Switch Registry (one-pager)

Switch	Default	Owner	Change path	Evidence required
Canonical lane	Radio US	Governance	Profile override + Change-Log	DBI ≥ 0.85; decade stability
β policy	1.0 (today)	Governance	Graduation decision	LEE gate: repeatability, CIs, independence, placebo
Caps	0.58	Governance	Profile override	Sensitivity ±0.02 review
Whitening	OFF	Methods	Triggered by κ/VIF	Appendix D audit
Trend cadence	periodic	Data	Snapshot approval	Delta thresholds met

Trend Layer (Host advisory data)

Advisory-only; hydrates lane targets for Suggestion/Optimization. Precedence: Manual curated snapshot → Datahub export → fallback. Stamped with `norms_timestamp`, `source_name`, `source_version`, `commit`. Guards: delta thresholds; conflict bounds; provenance; no auto-flip of canonical lane.

Host Suggestion/Optimization Layer

Deterministic rule graph: **Axis deficit → Intent → Playbook → Actions**. Inputs: final axes, QA/validators, Trend Snapshot, platform lane. Outputs: top-K (default 5) actions per lane, evidence links (measurement → target), counterfactual nudge

bands with confidence tags. **Baseline vs Curation** views; safety items pinned; `preferences_used` recorded. **Duplicates hidden** between Radio & Spotify unless advice differs.

Fail-Ledger & RAG

Failure taxonomy (`AUDIO.INTEGRITY.*` , `EXTRACT.TTC.*` , `EXTRACT.LYRICS.*` , `ENGINE.*.OWNERSHIP` , `NORMS.*` , `SUGGEST.ROUTER.*` , etc.) with severity (BLOCK/WARN/INFO), reason, recovery, exposure. RAG banner on Run Card; SLI/SLO observability for TTC accepted %, lyric accepted %, independence OK %, CI width.

Ethics & Compliance (lyric handling box)

Permissible ingestion; storage redaction; display limits (advisory snippets only); no redistribution of full texts; access controls; audit trails.

##xs HOWTO_REPRODUCE

Seeds/dtype/threads; commands to render Cards, figs, tables; golden-run diff with **|ΔHCI| ≤ 0.002** acceptance; expected hashes.

FAQ (short)

Why Radio as canonical? Slower drift, decadal continuity—closest to 40-year archetypes.
Do preferences affect HCI? No—preferences re-rank advisory only.
What if lyrics are missing? Lyric domain NA; HCI renormalizes to audio; advisory may still render with low-confidence labels.
Why caps=0.58? Fairness guardrail; see sensitivity ±0.02.

What's Trend vs Corpus? Trend hydrates advisory lane targets; engines remain anchored to historical corpus.
Can advice change the score? No—Host-only, post-score, read-only.

> *Side note (margin callout, optional):* 🔒 **Host-Only Advisory (Do Not Bypass)** — Engines: measure only → AER/LER; Host: caps→HCI→advisory; advisory reads lanes+Trend+QA; **never writes to scoring**; no feedback into HCI.