CIF Technical Whitepaper v1.1.1 — Addendum A (v1.2 Deltas & Clarifications)

> Scope: clarifies parameter discipline, safeguards, lyric-proxy transition, independence checks, and validation posture; **does not change** the v1.1.1 models. All items will be integrated into v1.2 with a formal Change Log.

A0. Publication Protocol

- v1.1.1 remains frozen (historical record).
- Addendum A records clarifications and validation exhibits.
- v1.2 will integrate these deltas inline and append a Change Log mapping A-tags to final sections.

A1. Parameter Discipline (canonical)

Pipeline order (applies to all profiles):

1. Standardize features: $z = \frac{x-\mu}{\sigma}$.
2. **Winsorize** extreme z-scores at $|z|>4$ (profile-tunable).
3. **Logistic map (per-feature gain):** $s=\sigma(\gamma z)=\frac{1}{1+e^{-\gamma z}}$, with $\gamma_j\in[0.8,1.2]$ by default (profile-tuned).
4. Axis aggregation (convex combinations of s_j per axis).
5. **EACM**: equal-axis mean across $K=6$ axes.
6. Apply **caps/guards** (Goldilocks) on axis scores if specified by the profile.

Units canon: TTC in **seconds**; chorus-lift in **dB** (short-term LUFS); loudness in **LUFS**; tempo in **BPM**.
Profiles (e.g., `US_Pop_2025`) define μ,σ, γ_j ranges, and guard caps.

A2. Score Bounds & Monotonicity (properties)

- $s_j=\sigma(\gamma_j z_j) \in (0,1)$, strictly monotone for $\gamma_j>0$.

- Axis scores (A_k) are convex combinations (\Rightarrow A_k\in[0,1]).
- (\mathrm{EACM}=\frac{1}{K}\sum_{k=1}^{K} A_k \in [0,1]).
- With caps (A*k'=\min(A_k,c_k)) on a set (\mathcalC}), the upper bound is [\max(\mathrm{EACM}) \le \frac{\sum*{k\in\mathcal{C}} c_k + (K-|)}{K}.]

  Example with Market/Emotional capped at (0.58) and (K{=}6): (\max \mathrm{EACM} \le 0.86).

A3. HEM Distance, Kernel & σ Rule

Distance metric: cosine distance in a **standardized** (z-scored) feature space.
Kernel width (σ): set by **median k-NN distance** with **k=8** computed on the active profile corpus.
Blend: (\mathrm{HEM}=\alpha \exp!\left(-\frac{d^{2}}{2\sigma^2}\right) + (1-\alpha),g(n;n^\star,\tau)), where (g) is a smooth bell around novelty anchor (n^\star). Parameters (\alpha,\tau) published in the profile card.

A4. Whitening Policy (Lenient; OFF by default)

Default: whitening **OFF**. Turn **ON** for a profile **only if** either:

- covariance **condition number (\kappa > 50)**, or
- **max VIF > 7** for any axis.

When ON: apply **PCA whitening** on the standardized HEM feature subset. Add a small diagonal (\lambda I) (e.g., (10^{-3})) if needed.
Audit exhibits: pre/post eigenvalue spectra and feature-correlation small multiples in the Validation Appendix.

(Documented options, OFF by default): Robust shrinkage whitening; robust z-scoring (median/MAD); axis-local whitening confined to HEM features.)

A5. Structure Segmentation (TTC & Chorus-Lift)

Chorus-lift: short-term **LUFS** (EBU R128: 3 s window, 1 s hop). Compute two 6 s windows centered on (i) first chorus onset and (ii) the immediately preceding verse; lift (=) LUFS($|text{chorus})$ − $LUFS($\text{verse}$).
TTC: earliest first chorus onset with **structure confidence ≥ 0.60**. If no reliable chorus, TTC := NaN and TTC is excluded from axis aggregation (weights renormalized among available features).

A6. Independence & Double-Counting Controls

- Report Pearson/Spearman **axis-axis correlations** and **VIF per axis** (target **VIF ≤ 5**).
- **Lyric-proxy taper:** As Luminaire (HLM/HLI) graduates, **linearly anneal** lyric-proxy signals in Creative/Cultural to zero over **1–2 minor releases**; publish schedule and expose proxy flags during the transition.

A7. Uncertainty, Sensitivity & Validation

- **Bootstrap:** 100× per cohort; publish 95% CIs for axis means and HCI.
- **Repeatability:** 3 fixed seeds; report mean ± sd.
- **Ablation:** lyric-proxy **ON vs OFF**; publish deltas for axes and HCI.
- **Temporal drift:** stratify by era buckets; show stability of axis means.
- **External (descriptive):** Spearman (ρ) between HCI quantiles and exogenous success buckets (non-supervised).

A8. Missing-Feature Policy

If a feature is NaN after QA (e.g., TTC missing), **exclude it** from that axis for that track and **renormalize** remaining feature weights. Log the omission in artifacts.

A9. Definitions & Notation (Appendix A)

Include symbols/units, glossary of Greek letters (($\alpha,\sigma,\tau,\gamma$)), and HEM/novelty terms ($d,\sigma,n^\star$). Clarify that vectors are bold, scalars italic, and symbols are defined at first use.

A10. Required Artifacts

- **Run Card** (doc version, profile, corpora IDs+hashes, seeds, params, caps).
- **Environment Card** (OS, Python, libs).
- **Figure Provenance** (Run ID, Profile ID, corpus IDs, commit/hash).

Mapping to v1.2 (Change Log)

A table mapping A1–A10 to their final section numbers in v1.2 will be included in the v1.2 release.