ArtMetrics — Technical Whitepaper

ArtMetrics Foundation

June 28, 2026

CONTENTS

ArtMetrics — Technical Whitepaper

A governance-risk scoring and longitudinal-trajectory engine for arts & culture nonprofits.

Abstract

ArtMetrics is a nonprofit governance-risk scoring engine and public transparency tool. It ingests IRS Form 990 financials and state charity filings to produce objective, data-driven scorecards for arts & culture nonprofits, expressed as an overall 0–100 score composed of financial resilience, governance risk, and program scale. On top of the cross-sectional score, a deterministic longitudinal **trajectory engine** classifies each organization's multi-year history into one of twelve organizational archetypes, and an LLM-based narrative layer reads each organization's own program description to surface mission and impact. This paper documents the data pipeline, the full scoring formulas, a worked example, the trajectory engine, the qualitative overlay, the officer network, and the system's limitations.

The guiding principle throughout is **lens, not shield**: the platform exists to provide context, not to punish. Design choices — such as the neutral governance default for organizations with no compensation signal — follow from that commitment.

1. Motivation

Nonprofit governance is opaque. Funders, boards, and organizations themselves lack objective, comparable signals about financial resilience and governance health, even though the source material is public. IRS Form 990 filings and state charity registries contain the relevant facts, but they are distributed across agencies, formats, and filing years, and are not structured for comparison or for longitudinal analysis.

ArtMetrics addresses this gap by transforming raw public filings into structured, comparable, and contextual scorecards. The aim is to help grantmakers and boards see clearly and ask better questions, and to give every organization — strong, healthy, or in need of support — a fair and contextual hearing. The product is deliberately constructive and non-punitive: data provides context, not a verdict.

2. Data Pipeline & Sources

ArtMetrics draws on three categories of public data.

IRS Form 990 (via ProPublica Nonprofit Explorer). The core financial record for each organization: total revenue, total expenses, aggregate compensation, and program-service descriptions. These power the financial resilience and program scale components and provide the aggregate compensation figure used in the governance signal.

IRS TEOS XML, Part VII. The Tax-Exempt Organization Search bulk XML exposes per-officer detail: name, title, hours per week, and compensation. ArtMetrics parses this into a structured officer dataset of **44,630 records across 9,526 organizations**. This per-officer granularity underpins both the officer-network graph and finer governance context.

State attorney-general charity registries. Three states are covered in depth: New Mexico (COROS), Oregon (DOJ CT-12), and Colorado (SOS). State data supplies control-concentration signals that supplement the federal compensation signal where available.

A separate qualitative pass reads IRS 990 **Part III** (Statement of Program Service Accomplishments) directly from raw TEOS XML; this feeds the narrative overlay described in Section 7.

3. Scoring Model

3.1 OVERALL SCORE

The overall score is a weighted combination on a 0–100 scale:

```
Overall = 0.40 × Financial resilience
        + 0.35 × Governance risk
        + 0.25 × Program scale
```

Scores map to four descriptive tiers:

Tier	Range
Strong	70+
Healthy	55–69
At-Risk	35–54
Priority Review	< 35

The lowest tier is named **Priority Review** rather than “Critical” by design: it marks where contextual attention is most warranted, not a judgment.

3.2 FINANCIAL RESILIENCE SUB-SCORE (0–100)

Financial resilience combines three factors plus a deficit-spending penalty.

Earned-income share / diversification (0–40 pts). Rewards revenue diversification through earned income:

Earned-income share	Points
≥ 50%	40
≥ 30%	30
≥ 10%	15
otherwise	5

Reserve strength (0–40 pts). Months of operating expenses held in net assets:

Reserve (months)	Points
≥ 18	40
≥ 12	30
≥ 6	15
> 0	5
otherwise	0

Solvency (0–20 pts). Based on the liabilities-to-assets ratio. A ratio of **≤ 0.2** is the strongest position; **0.5** is treated as a warning threshold.

Deficit-spending penalty (–10 pts). Applied when revenue is less than expenses, reflecting that an organization is drawing down rather than building resilience in the period.

3.3 GOVERNANCE RISK

The primary governance signal is the **compensation-to-revenue ratio**, with the following reference thresholds:

Compensation-to-revenue	Reading
10%	Healthy ceiling
20%	Warning
30%	Critical

Where state attorney-general data exists (NM, OR, CO), **control-concentration** signals are added to the governance component. Section 6 documents the neutral default applied when no compensation signal is present.

3.4 PROGRAM SCALE

Program scale reflects operating-revenue magnitude relative to peers, combined with the program-expense ratio (the share of expenses going to program services).

4. Worked Example: Flickinger Center for Performing Arts

The Flickinger Center for Performing Arts (New Mexico) illustrates how the components combine.

Metric	Value
Overall score	**61 — Healthy**
Financial sub-score	60
Governance sub-score	35
Program sub-score	100
Revenue	$634,764
Expenses	$682,373
Net assets	$1,137,377 (~20 months of reserves)
Program-expense ratio	94%
Historical filings	16 years

Applying the 40/35/25 weighting:

```
Overall = 0.40 × 60 + 0.35 × 35 + 0.25 × 100
        = 24.0 + 12.25 + 25.0
        = 61.25  →  61 (Healthy)
```

The financial sub-score reflects a well-reserved organization: roughly 20 months of operating expenses held in net assets places it in the strongest reserve band. The program sub-score of 100, paired with a 94% program-expense ratio, shows an organization with the overwhelming majority of its spending reaching mission.

At the same time, the governance sub-score of 35 surfaces a signal worth contextual review. This is precisely the value of a multi-component model: a strong-program, strong-reserve organization still earns a Healthy tier overall, while the scorecard transparently shows *where* a closer, supportive look may be warranted. The number is an invitation to context, not a conclusion. Sixteen years of historical filings also make Flickinger a candidate for the longitudinal analysis described next.

5. The Neutral-50 Governance Design Choice

Of the 13,278 scored organizations, **9,526** have at least one year of IRS 990 Part VII compensation data. The remaining **3,752** are all-volunteer organizations or filers of the 990-N (postcard), which carries no compensation information.

For these organizations there is no compensation-to-revenue signal to compute. A naive system might score the absence as zero or as a red flag. ArtMetrics instead applies a **neutral default of 50** to the governance component in these cases. This is a deliberate, principled choice: it avoids penalizing an organization for the *absence of data*, and it explicitly does not constitute a finding. An all-volunteer organization with no paid staff has no compensation ratio to evaluate, and the model treats that honestly rather than punitively — consistent with the lens-not-shield philosophy.

6. Longitudinal Trajectory Engine: Twelve Organizational Thumbprints

Beyond the cross-sectional score, ArtMetrics classifies each organization's financial history into one of twelve **organizational thumbprints**, computed **deterministically** from financial history (no model inference). Classification runs over 3-, 5-, and 10-year horizons. An organization needs **at least 3 years of data**; otherwise its trajectory is reported as **"Insufficient Data."** When no extreme divergence is detected, the default classification is **"Steady State."**

Each archetype pairs a diagnosis with a constructive **path forward / remedy** — the engine is built to suggest where an organization can go next, not merely to label where it has been.

6.1 ACUTE THUMBPRINTS (3-YEAR HORIZON)

- **Acute Resource Divergence** — overall score down > 10 while compensation-to-revenue percentage is up > 2.

- **Governance Lag** — overall score up > 15 while governance is flat or down.
- **Programmatic Contraction** — score stable or up, but program share down > 10.
- **Acute Stabilization** *(positive)* — score up > 10 and governance up > 5 from a sub-30 base; a recovery signature.

6.2 STRUCTURAL THUMBPRINTS (5-YEAR HORIZON)

- **Structural Deficit** — 5-year average financial score < 20 despite a program score > 40.
- **Compensation Escalation Cycle** — compensation-to-revenue up > 5 points and now above 12%.
- **Compliance Decay** — governance down > 15 while overall is roughly stable.
- **Mission Drift** — 5-year score volatility > 40, characteristic of chasing restricted grants rather than sustaining a stable core.

6.3 GENERATIONAL THUMBPRINTS (10-YEAR HORIZON)

- **Founder’s Trap** — a roughly flat decade ($|\Delta| < 10$) with compensation-to-revenue < 3%, indicating founder sweat-equity and an attendant succession risk.
- **Endowment Ossification** — score up > 20 but program share down > 20, a capital-hoarding signature.
- **Institutional Maturation** *(positive, “gold standard”)* — score up > 10, governance steady or up, and governance > 50: a decade of concurrent growth in both health and oversight.
- **The Hollow Shell** — score down > 20 and program down > 15, indicating a long structural decline.

Because every threshold is explicit and the computation is deterministic, a trajectory classification is fully reproducible and explainable from an organization’s own filing history. The constructive-remedy framing ensures that even the more difficult archetypes (e.g., The Hollow Shell, Structural Deficit) are presented alongside a path forward rather than as a verdict.

7. Qualitative Narrative Overlay (LLM on Part III)

Financial scores describe an organization's structure, not its mission. To balance this, ArtMetrics includes an AI/LLM parsing engine that reads IRS 990 **Part III** (Statement of Program Service Accomplishments) from raw TEOS XML and surfaces each organization's **primary discipline, core audience, and community impact**. The intent is explicit: so that the platform *celebrates mission work, not just measures finances.*

Organizations can **"Claim"** their profile and contribute this narrative directly, supplementing the parsed text with their own voice. The narrative overlay sits alongside the quantitative scorecard rather than feeding into it, keeping the objective score and the qualitative story distinct and legible.

8. Officer Network

The per-officer Part VII dataset (44,630 records across 9,526 organizations) is linked into a searchable **officer network** spanning **119,000+ historical officer records**. Because officers are linked organization-to-organization, the network reveals shared leadership across the sector. This supports the control-concentration governance signals (where state AG data is available) and powers the enterprise officer-network graph visualization.

9. Limitations

ArtMetrics is built to be defensible, which requires stating its boundaries plainly.

- **Data latency.** IRS Form 990 data is inherently lagged. Depending on the filing and processing cycle, the most recent available data for an organization may be anywhere from roughly **6 months to 3 years** old. Scores reflect the most recent *available* filings, not necessarily the current moment.
- **Control-concentration coverage.** Control-concentration signals depend on state attorney-general data, which is currently in depth only for **New Mexico, Oregon, and Colorado.** Organizations outside these states are scored without that supplementary signal.

- **Descriptive, not predictive.** The scores and trajectory thumbprints are descriptive characterizations of past and present filings. They are not forecasts and should not be read as predictions of future outcomes.

These limitations are reasons to treat ArtMetrics as a lens that adds context to human judgment, not a replacement for it.

10. Roadmap

ArtMetrics is structured as a layered product: a **public free tier** (search plus profiles with a three-view teaser), **enterprise analytical tools** (peer comparison, officer-network graph, per-organization visualizations), and a **funder-facing API** on the roadmap. The platform is built in part around AlleyCorp's **Nonprofit ENG(INE)** program.

The near-term direction moves ArtMetrics from a tool *about* nonprofits toward a tool *for* them. Organizations can already claim a profile and add a narrative overlay; the next milestone is **organization self-service dashboards**, giving leaders direct visibility into their own scorecard and trajectory and the constructive path forward their data implies.