# Millions – Technical Framework & Runbook (v0.1)

Owner: Spencer (PM) · Engineers: Spencer (scripts), Assistant (support) · Target: **Usable Week 1** + iterative upgrades through Week 3

## 0) What we’ve accomplished

* **Game-centric builder in place**: scripts/millions\_build\_game\_view.py (clean rewrite)
  + Builds base games from schedule keys (week + hometm + vistm) → guarantees home\_team/away\_team.
  + Joins **DVOA (Week 1 = 2025 projections)** onto home\_team and away\_team.
  + Pulls current lines with a **robust fallback** hierarchy and **sign-flip backfill**.
  + Produces stable CSV outputs:
    - picks/millions/millions\_roadmap\_game.csv (all weeks)
    - picks/millions/millions\_weekly\_games.csv (per week)
* **Audit in place**: scripts/millions\_audit\_week.py reports missingness and flags expected pre‑Thursday nulls.
* **DVOA projections CSV created**: data/2025\_dvoa\_projections.csv (normalized ingest in builder; alias mapping e.g., WAS→WSH).
* **Spread fallback util**: \_fill\_spreads\_with\_fallbacks(...) (status column for provenance).

## 1) Goals (Season‑long)

1. **Member Dashboard (unbiased)**
   * Quick, readable per‑game cards with: kickoff, CURRENT line, **Circa line** (Thu), line delta badge, **Total/Off/Def/ST DVOA**, basic rest info, EPA (soon), and a notes area.
   * No rankings/ATS recommendations on the member view (kept backend‑only).
2. **Master Dashboard (Spencer only)**
   * Everything in member view **plus** hidden analytics (power ratings, derived edges, consensus trends, pool context).
3. **Tracking & Post‑Mortem**
   * Persist weekly picks (4 exec picks + 1 consensus) and outcomes.
   * Annotate miss reasons (injury note missed, weather, mismatch vs closing, etc.).
4. **Data Lineage & Provenance**
   * Every value in the dashboards traceable to a source field; audits highlight gaps early in week.

## 2) Data Sources & Lineage

### A) Core schedule & contest context (already flowing)

* **File**: picks/survivor/survivor\_roadmap\_expanded.csv (generated by survivor pipeline)
* **Columns used**: week, hometm, vistm, date(\_sch), time(\_sch), home\_or\_away, rest\_days, consensus\_spread, current\_spread, circa\_spread(\_price) (if present), etc.
* **Lineage**: We do **not** modify survivor scripts; Millions builder reads them **read‑only**.

### B) DVOA (Week 1 = projections; post‑W1 = season‑to‑date)

* **Week 1 file**: data/2025\_dvoa\_projections.csv
  + Columns accepted (normalized): team, total\_dvoa, off\_dvoa, def\_dvoa, st\_dvoa (strings with % OK)
  + **Alias map** (example): { 'WAS':'WSH', 'JAC':'JAX', 'NOR':'NO', 'GNB':'GB', 'KAN':'KC', 'SFO':'SF', 'TAM':'TB', 'ARZ':'ARI', 'LVR':'LV', 'SDG':'LAC', 'STL':'LAR' }
* **Post‑W1 file (planned)**: data/2025\_dvoa\_season.csv (same normalized columns)

### C) Market lines

* **Current line**: from survivor expanded CSV (current\_spread, line, spread, with consensus fallback).
* **Opening line**: not yet wired (optional feed).
* **Circa contest line**: ingested **Thursday a.m.** via PDF/CSV parser (planned).

### D) EPA (planned)

* **Source**: nflfastR team‑level EPA per play (offense & defense), season‑to‑date.
* **Join key**: team abbreviation with alias map.

### E) Injuries & Weather (planned)

* **Injuries**: summarized notes per team from reliable public feed (TBD; weekly snapshot Tue/Thu).
* **Weather**: game‑time forecast snapshot (Thu/Sat; indoor tag). Eastern time already available; UI may convert.

## 3) Pipeline Architecture

Survivor generator (existing)  
 └─ picks/survivor/survivor\_roadmap\_expanded.csv  
 └─ millions\_build\_game\_view.py  
 ├─ Load BASE games from (week,hometm,vistm)  
 ├─ Load DVOA (projection if week==1 else season) → join to BASE (home/away)  
 ├─ Attach team rows (HOME/AWAY) for spreads/rest/inj/weather  
 ├─ Fill spreads (current/open/closing/circa) with backfills → compute deltas  
 └─ Write outputs (all weeks + weekly)  
 └─ millions\_audit\_week.py (optional)  
 └─ diagnostics markdown

**Outputs:**

* picks/millions/millions\_roadmap\_game.csv — all games (season)
* picks/millions/millions\_weekly\_games.csv — filter by --week
* picks/millions/diagnostics/millions\_week\_audit.md — weekly audit

## 4) Schemas (key columns)

### Millions weekly file (subset shown)

* Keys: week, game\_num, home\_team, away\_team, matchup, kickoff\_local, venue
* Lines: open\_spread\_home/away, current\_spread\_home/away, closing\_spread\_home/away, circa\_spread\_home/away, circa\_spread\_price\_home/away
* Line values: line\_value\_current\_home/away, line\_value\_closing\_home/away, line\_value\_home/away (alias to current)
* DVOA: total\_dvoa\_home/away, off\_dvoa\_home/away, def\_dvoa\_home/away, st\_dvoa\_home/away
* EPA (placeholders): off\_epa\_per\_play\_home/away, def\_epa\_per\_play\_home/away, epa\_diff\_home/away
* Context: rest\_days\_home/away, rest\_days\_diff, injuries\_key\_home/away, weather\_notes
* **Provenance** (new): current\_spread\_status (primary/secondary/flip/fallback0)

## 5) Member vs Master dashboards

### Member (unbiased)

* Show: kickoff, CURRENT line, Circa line (— pre‑Thu), delta badge, Total/Off/Def/ST DVOA, rest diff, EPA rows when available, notes placeholder.
* Hide: power ratings, ATS ranks, any nudging signals.

### Master (Spencer only)

* Show everything in member view **plus**: power ratings & gaps, model edges, confidence bands, historical notes & miss tagging, pool context (common picks %), trend charts.

## 6) Operational Cadence (weekly)

**Mon–Tue**

* Refresh survivor expanded CSV.
* Run builder → audit. Investigate missing CURRENT lines, DVOA alias hits.
* Spot‑check 3–5 games (see checklist below).

**Wed**

* Integrate EPA (refresh) and injuries/early weather notes.

**Thu morning**

* Ingest **Circa contest lines** from PDF/CSV. Rebuild → audit → member HTML export.
* Send member links for review; collect notes & exec picks.

**Fri–Sat**

* Update injuries & weather snapshot; re‑export member HTML if needed (no Circa change).

**Sun night / Mon a.m.**

* Post‑mortem: add closing vs Circa deltas and notes; update tracking sheet.

## 7) Spot‑check checklist (per‑game)

1. **Teams & week**: away @ home matches known schedule; week correct.
2. **Kickoff**: kickoff\_local (Eastern) looks right; indoors flagged in notes if applicable (later).
3. **Lines**:
   * Current present; current\_spread\_status == current:primary ideally. If secondary/flip/fallback0, investigate.
   * Circa — until Thu; after Thu, ensure deltas compute.
4. **DVOA**: Total/Off/Def/ST non‑null; magnitudes plausible (roughly within ±40).
5. **Rest**: rest\_days\_diff sign matches expectations (home vs away rest days).
6. **EPA (when added)**: values non‑crazy (≈ −0.3..+0.3 per play), offense vs defense directions make sense.

## 8) Two‑week roadmap (detailed)

### Tonight (get to green for Week 1 member preview)

* ✅ Builder + audit green; DVOA W1 projections wired; current spread fallback ✅
* Add **member HTML exporter** (simple cards; hide deltas pre‑Thu).  
  *Command:* python -m scripts.millions\_export\_member\_html --week 1 --out picks/millions/member\_week\_1.html
* Add **team alias registry** (see §9) directly in builder util.
* Label audit’s expected‑null fields as “(expected pre‑Thu)”.

### Next 3–4 days

* **Circa parser**: PDF to CSV for contest lines (data/circa\_week\_{w}.csv). Auto‑merge in builder.
* **EPA join**: nflfastR team‑level EPA per play (off & def), S2D; add to builder and audit.
* **Injuries**: initial text feed ingest; attach injuries\_key\_home/away summaries (Tue/Thu snapshots).
* **Weather**: stadium + forecast snapshot; add weather\_notes.

### Week 2–3

* **Master dashboard**: private analytics (power gaps, edges, pool pop tracking). First cut as HTML export.
* **Picks tracking store**: picks/millions/picks\_{year}.csv and a small CLI to append week’s choices & freeform notes.
* **Regression fixes**: stricter schema validation (types/ranges); alias expansion as needed.

## 9) Team alias registry (proposed default)

Use in DVOA/EPA joins and any external source merges.

{  
 'WAS':'WSH', 'JAC':'JAX', 'NOR':'NO', 'TAM':'TB', 'SFO':'SF', 'ARZ':'ARI',  
 'GNB':'GB', 'KAN':'KC', 'NWE':'NE', 'SDG':'LAC', 'STL':'LAR', 'OAK':'LV',  
 # common nicknames to abbreviations (if encountered)  
 'Washington':'WSH','Jacksonville':'JAX','New Orleans':'NO','Tampa Bay':'TB',  
 'San Francisco':'SF','Arizona':'ARI','Green Bay':'GB','Kansas City':'KC',  
 'New England':'NE','San Diego':'LAC','St. Louis':'LAR','Oakland':'LV'  
}

## 10) Commands cheat‑sheet (PowerShell)

# 0) Activate venv (if not already)  
.\venv\Scripts\Activate.ps1  
  
# 1) Build Millions (Week 1 uses projections)  
python -m scripts.millions\_build\_game\_view --week 1  
  
# 2) Audit Week 1  
python -m scripts.millions\_audit\_week --week 1 --write-report  
  
# 3) Export Member HTML (first cut)  
python -m scripts.millions\_export\_member\_html --week 1 --out picks/millions/member\_week\_1.html

## 11) Testing & QA

* **Unit sanity**: after each rebuild, run the audit and verify: current\_spread\_\* ≤ 5% missing (target 0%); DVOA 0% missing; Circa 100% missing pre‑Thu only.
* **Schema checks**: verify all expected columns exist (see §4). Add assert guards before export.
* **Visual diff**: open the member HTML; scan 3–5 random cards for plausibility.

## 12) Known gaps & risks

* Circa lines not automated yet (manual until PDF parser lands).
* EPA/injuries/weather not yet merged (placeholders in file + audit).
* External source alias drift (add to registry upon first miss).
* Survivor changes upstream (columns renamed) — builder now defensive but we’ll monitor audits.

## 13) Handover guidance (if someone else continues)

* Read this doc top‑to‑bottom.
* First run the **three commands** in §10 to reproduce Week 1 outputs.
* Use §7 checklist to spot‑check 3–5 games.
* For any missingness > 0% in DVOA/current lines, check alias map (§9) and spread fallbacks.
* Open issues:
  1. Implement Circa parser.
  2. Add nflfastR EPA join.
  3. Add injuries/weather feeds.
  4. Build Master dashboard exporter.

## 14) Appendix: Member dashboard fields (initial)

* matchup, kickoff\_local, current\_spread\_home/away, circa\_spread\_home/away (— pre‑Thu), line\_value\_home/away (only show when Circa present), total\_dvoa\_\*, off\_dvoa\_\*, def\_dvoa\_\*, st\_dvoa\_\*, rest\_days\_diff, current\_spread\_status, notes (UI only).

*End v0.1 – update as we wire EPA, injuries, weather, Circa parser.*