

Data Documentation: Pokémon TCG Tournament Matchup Data and Formally Verified Game-Theoretic Analysis

This document describes the structure, provenance, and interpretation of all files included in the dataset.

1. Data Source and Collection

All empirical data were scraped from Trainer Hill (<https://www.trainerhill.com/meta?game=PTCG>) on February 19, 2026. The query filters were:

- Game: Pokémon Trading Card Game (PTCG)
- Tournament size: 50+ players
- Date range: January 29 – February 19, 2026
- Platforms: all
- Tie counting: 1/3 win (i.e., WR% = $(W + T/3) / (W + L + T)$)

Trainer Hill aggregates results from officially sanctioned Pokémon TCG tournaments. The data represent thousands of matches across dozens of events during the Stellar Crown – Prismatic Evolutions (SVI-PFL) format.

2. File Descriptions

2.1 trainerhill_matchups_2026-02-19.md

Format: Markdown with tables

Size: ~12 KB (325 lines)

This is the primary empirical data file containing:

- Meta shares (overall): popularity of each of the 14 archetypes across all tournaments
- Top-8 shares: popularity among top-8 finishers only

- Full 14×14 matchup matrix: for each ordered pair of decks (row = player, column = opponent), the file reports the win rate percentage and the raw win-loss-tie record
- Condensed 6×6 matrix: a subset of the top 6 decks for quick reference
- Sample sizes: total match counts for key matchups (e.g., Dragapult mirror: 2,845 games)

The 14 Archetypes

#	Archetype	Meta Share (%)
1	Dragapult Dusknoir	15.5
2	Gholdengo Lunatone	9.9
3	Grimmsnarl Froslass	5.1
4	Mega Absol Box	5.0
5	Other (aggregated)	4.8
6	Gardevoir	4.6
7	Charizard Noctowl	4.3
8	Gardevoir Jellicent	4.2
9	Dragapult Charizard	3.5
10	Raging Bolt Ogerpon	3.4
11	Alakazam Dudunsparce	3.2
12	N's Zoroark	3.1
13	Ceruleedge	2.8
14	Kangaskhan Bouffalant	2.5

Reading the Matchup Matrix

Each archetype section contains a table with columns:

- vs — the opponent archetype
- WR% — win rate as a percentage, computed as $(W + T/3) / (W + L + T)$
- W-L-T — raw win, loss, and tie counts

The matrix is not perfectly symmetric because $WR_A_vs_B + WR_B_vs_A \approx 100\%$ but not exactly, due to the tie-counting formula.

2.2 bootstrap_nash_results.txt

Format: Plain text

Size: ~4 KB (90 lines)

Contains the results of a 10,000-iteration bootstrap stability analysis of the Nash equilibrium. For each iteration, the 14×14 matchup matrix was resampled (with replacement from original W-L-T counts), and a new Nash equilibrium was computed. The file reports:

- Support frequency: how often each deck appears in the Nash support across 10,000 resamples
- Weight statistics: mean, median, standard deviation, and Wilson 95% confidence intervals for each deck's Nash weight
- Game value: the expected payoff at equilibrium ($47.96\% \pm \text{CI}$)
- Core support decks appear in >96% of resampled equilibria, confirming qualitative stability

2.3 symmetric_nash_results.txt

Format: Plain text

Size: ~4 KB (77 lines)

Contains the Nash equilibrium of the symmetrized constant-sum game. The symmetrization transforms the matchup matrix M into M' where $M'_{ij} = (M_{ij} + (1000 - M_{ji})) / 2$, yielding a zero-sum game. The file reports:

- Maximum deviation from perfect symmetry: $\max|M_{ij} + M_{ji} - 1000| = 73$ permil
- Symmetric Nash strategy: 7-deck support with exact rational weights
- Game value: exactly 500 permil (as expected for a symmetric zero-sum game)
- Dragapult Dusknoir receives 0% weight despite 15.5% meta share (the "popularity paradox")

2.4 predictive_validation.txt

Format: Plain text

Size: ~4 KB (68 lines)

Contains a predictive validation test of the replicator dynamics model. Single-step replicator dynamics on the full 14-deck game were computed using meta shares as initial population frequencies and the matchup matrix as the payoff matrix. Predictions (grow/shrink/stable) were compared against actual Trainer Hill trend arrows observed between February 19 and 20, 2026. The file reports:

- Per-deck fitness values and predicted direction
- Observed trend arrows from Trainer Hill
- Match/mismatch classification: 2 out of 3 testable predictions confirmed

2.5 native Decide audit.txt

Format: Plain text

Size: ~7 KB (137 lines)

An audit of all native Decide proof closures in the Lean 4 artifact. native Decide is a Lean tactic that trusts the compiler's evaluation rather than constructing kernel-checkable proof terms. This file catalogues every use, enabling reviewers to assess the trust boundary. The file reports:

- File-by-file listing of every native Decide invocation
- What each invocation proves (e.g., matrix entry verification, Nash optimality check)
- Total count and classification by proof category

3. Units and Conventions

- Win rates in the Markdown file are percentages (0–100). Win rates in the Lean formalization use permil (0–1000) for exact integer arithmetic.
- Meta shares are percentages of total tournament entries.
- Ties are counted as 1/3 of a win per Trainer Hill convention.
- Nash equilibrium weights are reported as percentages (summing to 100%).
- The "Other" category aggregates all archetypes outside the top 14.

4. Relationship to the Lean 4 Artifact

The empirical data in this dataset are encoded in the Lean 4 proof artifact (available separately) in two canonical representations:

- RealMetagame.lean — matchup win rates as a function over the Deck inductive type (permil)
- NashEquilibrium.lean — matchup data as an Array (Array Nat) for computational proofs

A machine-checked proof (MatrixConsistency.lean) establishes that both representations agree entry-by-entry. All game-theoretic results in the accompanying paper trace to this single empirical source.

5. Reproducibility

To reproduce the computed results from the raw data:

- Nash equilibria: use any linear programming solver (e.g., `scipy.optimize.linprog`) on the 14×14 matrix
- Bootstrap analysis: resample W-L-T counts with replacement, recompute Nash per iteration
- Replicator dynamics: apply the standard replicator equation $\frac{dx_i}{dt} = x_i(f_i - \bar{f})$ with the matchup matrix as payoff

Python scripts for all three computations are included in the accompanying code artifact.

6. Citation

If you use this dataset, please cite:

A. F. Ramos and T. Soria, "From Rules to Nash Equilibria: Formally Verified Game-Theoretic Analysis of a Competitive Trading Card Game," submitted to IEEE Transactions on Games, 2026.