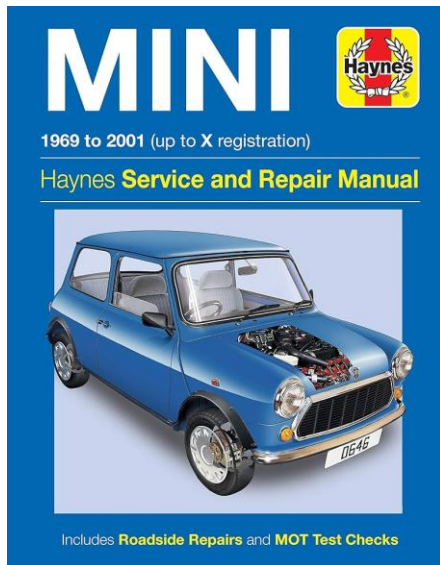# FRAMECORE

ALGO-95.2 HAYNES MANUAL

Production-Grade Operational & Engineering Manual Version: 95.2 Status: Live / Production Compatibility: 100% backward-compatible with ALGO-65.2 / 80 / 90.x Owner & Steward: Roy Taylor / Framecore

Reference made with love and in appreciation of the many, many long hours ensuring I could get out and about and be sure of also getting home.

MINI
1969 to 2001 (up to **X** registration)
Haynes **Service and Repair Manual**
Includes **Roadside Repairs** and **MOT Test Checks**

Explanation (Plain English)

Think of ALGO-95.2 as a factory with inspectors at every stage.
Each folder is either:

- an agent (does work),
- a ledger (records truth),
- or a constraint authority (can stop the system).

Critically:

No single folder is optional

Several folders exist only to prevent silent failure or "too good to be true" outputs

If a folder feels "paranoid", that's because it is. Correctly so.

0. Purpose of This Manual

This manual serves as a comprehensive guide to the assembly, initialization, integrity maintenance, and safe failure protocols for ALGO-95.2. It is intended to supplement the whitepaper by offering detailed operational instructions. Strict procedures are deliberate, as prediction engines require robust safeguards to prevent severe failures.

1. System Overview

ALGO-95.2 represents a non-destructive extension of ALGO-65.2, introducing enhancements such as:

- Completion quality
- Geopolitical risk assessment
- Quality of experience
- Platform availability and licensing

These additions preserve the core equation; in the absence of new signals, system behavior reverts seamlessly to ALGO-65.2. Preservation of this property is essential.

2. System Architecture (Canonical Directory Set)

ALGO-95.2 operates as an ensemble of specialized engines, each with defined responsibilities and failure modes. All components listed below are mandatory unless specifically designated as optional.

2.1 Orchestration & Control

Orchestrator/: Manages daily system health, GPU enforcement, anti-cheat measures, and dependency validation.

2.2 Ingestion & Freshness

SCHIG/: Conducts scheduled data collection from over 11 APIs, encompassing metadata, availability, and signals.

Fresh In!/: Handles normalization, routing, and schema-compliant ingestion.

2.3 Core Data Assets

- Views Training Data/: Maintains immutable ground-truth views suitable for audit.
- Abstract Data/: Stores environmental, cultural, economic, and calendar predictors.
- Components/: Contains lookup tables, decay curves, weights, and seasonal allocators.

2.4 Core Computation

ALGO Engine/: Facilitates GPU-based machine learning training, inference, and calibration.

MAPIE/: Measures accuracy, detects drift, and audits inference results.

Schema/: Holds canonical columns, version-locked for consistency.

2.5 Output & Commercial Layers

- Daily Top 10s/: Produces rankings derived from predicted views.
- Money Engine/: Oversees monetization, ROI analysis, and advertising economics.
- Studios/: Classifies and weights production studios.
- Talent/: Incorporates cast, creator, and franchise attributes.

2.6 Supporting Infrastructure

- GPU Enablement: Provides authoritative GPU configuration.
- Legal/: Enforces licensing restrictions and blackout logic.
- Email/, Replit/, Schema/, AUDIT_LOGS/, INFERENCE_AUDIT.json: Ensure observability, traceability, and human-readable auditing.

2.7 Canonical Directory & Agent Responsibilities

(Operational Role Definition – Non-Optional)

ALGO-95.2 is composed of discrete, auditable agents mapped directly to filesystem boundaries. Each directory listed below represents either a functional engine, a control authority, or an immutable ledger. Removal, bypass, or consolidation of any directory invalidates system integrity.

2.7.1 Audit & Self-Defense

AUDIT_LOGS/
Authoritative, append-only execution logs.
Records:

- runtime duration
- GPU confirmation
- model consensus deltas
- validation interventions (VIR, GCI, caps)

Any prediction without a corresponding audit trail is invalid by definition.

2.7.2 Canonical Components & Constants

Components/
Holds all non-learned truth:

decay curves

genre coefficients

seasonality vectors

hard bounds and clipping thresholds

These values are intentionally not trainable. If Components is altered, retraining alone is insufficient; MAPIE recalibration is mandatory.

2.7.3 External Signal Ingestion

SCHIG/
(Scheduled Collection & Harvest Ingestion Gateway)

Responsible for deterministic data acquisition from all upstream APIs:

- metadata
- rankings
- availability
- contextual signals
- SCHIG enforces:
- rate discipline
- schema pre-validation
- timestamp normalization

If SCHIG fails, no inference is permitted, even with cached data.

Fresh In!/
Normalization and routing layer.
Transforms raw SCHIG outputs into schema-compliant, version-locked records.
No learning, no inference — purely hygienic.

This folder exists to prevent "nearly correct" data from poisoning models.

2.7.4 Core Prediction Engines

ALGO Engine/
GPU-only execution of:

- Random Forest
- XGBoost
- CatBoost
- Handles:
- training
- inference
- consensus aggregation
- feature attribution

CPU execution is explicitly disallowed.
If this folder reports success without GPU confirmation, the run is considered fraudulent.

MAPIE/
Measurement, Accuracy, Performance & Inference Evaluation.

MAPIE:

- computes MAPE, R^2, residuals
- detects drift
- enforces statistical plausibility
- blocks "implausibly good" results

MAPIE exists to stop optimism, not pessimism.

2.7.5 Data Truth & Ground Reality

Views Training Data/
Immutable ground-truth views used for learning and validation.
Once written, records may not be edited — only superseded by newer vintages.

This folder is legally and analytically defensible evidence.

Abstract Data/
Holds non-content predictors:

- weather
- events
- economics
- geopolitics
- calendars

Absence of Abstract Data does not halt execution; it forces graceful reversion to ALGO-65.2 behavior.

2.7.6 Commercial & Interpretive Layers

Daily Top 10s/
Derived rankings generated after prediction.
Rankings are explicitly downstream artifacts and must never feed back into training.

Money Engine/
Translates predicted views into:

- ROI
- valuation
- ad-tier economics
- licensing scenarios

Money Engine has no authority over prediction outcomes.

2.7.7 Entity Intelligence

Studios/
Canonical studio identifiers and historical performance modifiers.
Prevents duplicate studio identities and shell-entity inflation.

Talent/
Tracks cast, creators, franchises, and historical performance trajectories.
Inputs affect content intrinsic quality (C_i) only — never platform or context.

2.7.8 Governance & Constraints

Legal/
Implements:

- blackout rules
- licensing restrictions
- territorial exclusions

Legal overrides prediction after computation, never before.

Schema/
Defines the single source of truth for column names, types, and version locks.
If Schema changes, all downstream systems must be revalidated.

2.7.9 Infrastructure & Tooling

GPU Enablement/
Authoritative hardware configuration and verification.
Confirms CUDA availability, memory sufficiency, and driver compatibility.

Email/
Human-in-the-loop alerting for:

- failures
- boundary violations
- audit exceptions

Replit/
Isolated experimentation and sandbox execution only.
Outputs here are never production-valid unless promoted through SCHIG → Fresh In!.

2.7.10 Reference Data Assets

BFD_Star_Schema_V22.02.parquet
Canonical dimensional model.
Defines how all entities relate (content, talent, platform, time).

BFD_V22.02.parquet
Wide-fact backing dataset used for:

- training
- attribution
- forensic analysis

2.7.11 Meta-Control

.claude/
Session-level reasoning, prompts, and guardrails.
This folder documents why decisions were made, not just what happened.

It is not part of inference, but it is part of accountability.

2.7.12 Operational Principle (Read This Twice)

If any folder listed above is missing, bypassed, or silently ignored,
the system may still run — but it is no longer ALGO-95.2.

3. Execution Requirements

3.1 Mandatory GPU Utilization

CPU fallback is not permitted.

Partial execution is not allowed.

Silent degradation is disallowed.

If GPU initialization fails, the system halts. Predictions completed without GPU resources are invalid.

3.2 Runtime Validation

Any complete ALGO-95.2 run under 30 minutes indicates incomplete computation and is considered fraudulent rather than optimized.

4. Mathematical Foundation

4.1 Preserved Core (ALGO-65.2)

The original master equation remains unaltered:

$\hat{V} = C \times P \times T \times E \times M$

4.2 ALGO-95.2 Additive Envelope

Enhancements are multiplicative:

$\hat{V}_{95} = \hat{V} \times R_i \times G\_d,t \times Q\_d,t \times A_i,p,t$

Where:

R_i: Completion and retention quality

G_d,t: Geopolitical adjustment factor

Q_d,t: Quality of experience

A_i,p,t: Platform availability and licensing

Missing terms default to 1.0. Backward compatibility is mathematically enforced.

5. Views as the Unit of Account

Predictions focus on discrete view counts, not hours or minutes, due to their auditability and contractual relevance. Temporal modifiers apply post-prediction according to genre-specific conversion tables.

6. Machine Learning Ensemble

ALGO-95.2 employs a three-model GPU ensemble:

Random Forest (cuML)

XGBoost (GPU hist)

CatBoost (GPU)

Each model utilizes the full dataset without sampling. Predictions are aggregated through weighted consensus, with discrepancies treated as informative.

7. Anti-Cheat and Self-Defense Mechanisms

The system incorporates robust defenses against data leakage, label contamination, and unintended inference-only runs. Critical safeguards include:

- R^2 upper limitations
- MAPE lower thresholds
- Minimum tree requirements
- Restricted feature lists
- Minimum runtime enforcement

Excessively favorable outcomes are automatically rejected.

8. Temporal Policy

8.1 Prediction Horizons

Daily/weekly predictions: approximately 75% accuracy, high responsiveness.

Quarterly forecasts (standard): 92–98% accuracy.

ALGO-95.2 is optimized for 2–3 quarters forward; projections beyond this range constitute scenario modeling.

8.2 Model Retraining

Quarterly retraining is compulsory. Weekly retraining is permissible if substantial new data is available. The system does not support obsolete models.

9. Expected Failure Modes

ALGO-95.2 is engineered to fail predictably:

- Absent abstract data reverts to ALGO-65.2.
- Missing quality-of-experience telemetry defaults QoE to 1.0.
- Regional blackouts invoke geopolitical floor adjustments.
- Catastrophic undetected failures are systematically prevented.

10. Output Interpretation

All outputs are:

- Traceable to source data
- Reproducible with identical inputs
- Explainable via component factors
- Inability to justify output changes indicates an upstream issue.

11. Conceptual Model

ALGO-95.2 should be viewed as ALGO-65.2 augmented with real-world considerations. The core mathematics remain unchanged, while adjustments reflect actual conditions.

12. Change Management Advisory

Modifications to schemas, runtimes, feature sets, or execution order necessitate updates in MAPIE, inference audits, and this manual. Otherwise, while the model may continue operating, its validity cannot be assured.

— End of Haynes Manual (95.2)

Appendix. Column Headers

Main Database BFD_V22.02.parquet

(765,862 rows × 1,711 columns, columns numbers listed)

Identity & IDs (1-6)

fc_uid, tmdb_id, imdb_id, thetvdb_id, trakt_id, flixpatrol_id

Core Metadata (7-70)

title, original_title, title_type, season_number, max_seasons, season_episodes, total_episodes, start_year, end_year, premiere_date, finale_date, runtime_minutes, genres, is_adult, status, original_language, country, regions, overview, tagline, age_certification, homepage, streaming_platform, streaming_platform_us, streaming_platform_uk, network, studio, production_country, studio_headquarters, studio_offices, studio_hits, production_companies, creators, directors, writers, producers, cast_data, budget, revenue, imdb_score, imdb_vote_count, tmdb_score, tmdb_vote_count, tmdb_popularity, rotten_tomatoes_score, metacritic_score, netflix_hours_viewed, netflix_views, nielsen_viewers_millions, flixpatrol_points, flixpatrol_rank, parrot_demand_expressions, source_api, source_apis_used, ingestion_timestamp, confidence_score, measurement_scope, poster, backdrop, trailer_url, videos, similar_content, keywords, imdb_url, views_estimated

Views by Half-Year & Country (72-245) ~174 cols

- Pattern: views_h{1|2}_{year}_{country} (years 2023-2025, 28 countries + total)
- Views by Quarter & Country (246-651) ~406 cols
- Pattern: views_q{1-4}_{year}_{country} (years 2023-2026, 28 countries + total)
- ML Lag Features (652-859) ~208 cols
- Pattern: {metric}_lag{7|14|28|56|91}, {metric}_roll{7|14|28}
- ML Bucket Features (860-1032) ~173 cols
- Pattern: {metric}_b{10|25|50|75|90}, {metric}_bucket
- Interaction Features (1033-1080) ~48 cols
- Pattern: int_{feature1}_{feature2}, {metric}_sq, {metric}_log
- Computed Weights & Multipliers (1081-1082)
- base_prediction, studio_weight
- Extended Views US/CN (1083-1122) ~40 cols
- Pattern: views_h{1|2}_{year}_{us|cn}, views_q{1-4}_{year}_{us|cn}
- Historical Views 2021-2022 (1123-1350) ~228 cols
- Pattern: views_{q|h}{period}_{year}_{country} (18 countries)
- Future Views 2026 (1351-1426) ~76 cols
- Pattern: views_{q3|q4|h1|h2}_2026_{country}

Abstract Signals (1427-1452) ~26 cols

abs_social_buzz, abs_trend_velocity, abs_critical_score, abs_audience_score, abs_award_momentum, abs_seasonal_index, abs_search_volume, abs_twitter_mentions,abs_instagram_engagement, abs_news_aberration, abs_weather_impact, abs_holiday_weight, abs_sports_competition, abs_sub_growth_region, abs_decay_position, abs_franchise_value, abs_star_power, abs_budget_tier, abs_marketing_spend, abs_platform_reach, studio_weight_S_i, genre_decay_T_g, views_y_pred, views_pred_xgboost, views_pred_catboost, views_pred_lightgbm, views_pred_randomforest, release_validated

Country-Specific Predictions (1455-1660) ~206 cols

Pattern: views_y_pred_{country}, views_q{1-4}_{2019-2026}_{country}

Platform & Aggregates (1661-1700)

streaming_platform_global, streaming_platform_row, views_y, views_monthly, views_quarterly, views_domestic, views_international, views_emea, views_latam, views_apac, completion_rate, repetition_index, engagement_velocity, attribution_confidence, data_origin, source_system, source_reference, last_updated, row_hash, content_hash, is_deleted, validation_status, audit_log, created_at, modified_by, schema_version, canon_compliance, vram_allocation, reserved_1, reserved_2, beta_p, lambda_g, lambda_g_inv, hours_per_view, binge_rate_index, rewatch_index, wvv, streaming_platform_usa, views_y_source, views_computed

Studio Classification (1701-1711)

studio_canonical, studio_country, parent_studio, studio_source, production_company_canonical, studio_movies_count, studio_domestic_box_office, studio_worldwide_box_office, studio_rank, is_revenue_customer, tmdb_votes

(182,263,620 rows × 4 columns)

1. fc_uid   - Title identifier

2. country   - Country code (18 countries)

3. platform - Streaming platform

4. views   - View count

Summary

Database	Rows	Columns
BFD	765,862	1,711
Star Schema	182,263,620	4

Report Ends