Verdict

Whitepaper

An independent watchdog for prediction market resolution quality

https://ver.watch

Prepared for researchers, institutional traders, and platform developers

Version 1.0 | June 2026

Document Scope

This document describes the current Verdict product, its scoring methodology, platform architecture, and known limitations. It is a technical whitepaper and does not constitute legal, financial, investment, or trading advice.

The methodology applies to active, non-closed Polymarket markets with non-zero trading volume that are retrieved through the public Polymarket Gamma API and processed under the current Verdict ruleset. Scores are intended to make resolution-rule risk more visible before settlement; they are not a prediction of market direction, profitability, or final adjudication.

◇

Table of Contents

Contents

1. Abstract

Prediction markets are increasingly used as public instruments for aggregating expectations about political, financial, technological, cultural, and sporting events. Their value depends not only on liquidity and pricing efficiency, but also on the quality of market resolution. A market can produce an apparently precise probability while its written criteria remain ambiguous, underspecified, dependent on discretionary interpretation, or vulnerable to manipulation at the moment of settlement.

Verdict is an independent platform that measures this form of resolution risk for active Polymarket markets by applying a deterministic scoring methodology to public data from the Polymarket Gamma API. It evaluates written resolution criteria across six weighted dimensions: Time Clarity, Resolution Source, Outcome Definition, Evidence Standard, Edge Case Handling, and Conflict of Interest. Each market receives a score from 0 to 100 and is classified as Low, Medium, High, or Critical risk.

The Verdict Score is not a prediction about the underlying event. It does not guarantee correct settlement, forecast future price movement, or determine whether a trader should take a position. It is a structured estimate of how much resolution-rule ambiguity is visible before the market resolves. This whitepaper describes the problem Verdict addresses, the formal scoring framework, the current system architecture, the community layer, the data refresh policy, known limitations, and future work.

CORE CLAIM

For prediction markets to function as reliable information infrastructure, the rules that convert real-world events into market outcomes must be measurable, inspectable, and reproducible.

2. Introduction - The Problem with Prediction Market Resolution

A prediction market converts beliefs about an event into tradable prices. In principle, the market price can be interpreted as a public probability signal. In practice, that signal depends on the contract underneath it. A market must define what event is being measured, what evidence will be accepted, when determination occurs, and how conflicting or delayed information is handled. If those terms are weak, the price may reflect both event probability and uncertainty about how the rules will be applied.

Resolution problems often emerge only at settlement. A question may appear simple while omitting a timezone, failing to identify a decisive source, or using qualifiers such as “confirmed,” “major,” or “credible” without definition. A corporate event may occur before a deadline but be disclosed later. A sports event may be postponed. An official data release may be revised. In each case, participants can hold reasonable but incompatible interpretations of the same market text.

The consequences extend beyond individual losses. Ambiguous resolution criteria can reduce trader confidence, distort prices, weaken the research value of market data, and create reputational risk for market operators. If large positions are at stake, weak evidence standards may also create incentives to influence public narratives around resolution. For markets that aim to function as information infrastructure, contract clarity is therefore not a cosmetic feature; it is a core condition of integrity.

Verdict addresses this problem by evaluating market text before resolution. It does not attempt to decide whether the correct outcome is Yes or No. It asks whether the market is written in a way that permits a consistent, evidence-based decision. This prior assessment helps separate price risk from rule risk and makes resolution quality visible while the market is still active.

3. What is Verdict?

Verdict is an independent watchdog for prediction market resolution quality. It monitors active Polymarket markets and scores the written rules that govern settlement. Its purpose is to surface markets that may be ambiguous, underspecified, or vulnerable to discretionary interpretation before traders, researchers, or the public rely on the resulting price signal.

The current product focuses on Polymarket because its public Gamma API provides accessible market data and because its markets span many categories and event types. Verdict uses only publicly available data and has no financial relationship with Polymarket or any prediction market platform. It is not a trading venue, broker, market maker, arbitrator, oracle, or enforcement authority. It does not hold funds, place trades, resolve outcomes, or compel platforms to revise settlements.

Verdict's function is informational. It produces a reproducible assessment of resolution-rule quality using public market text. Traders can use the score as one input in procedural due diligence. Researchers can use the score to filter market data before drawing conclusions from prices. Platform developers can use the methodology as a drafting benchmark for improving market contracts.

OPERATING PRINCIPLES

- **Deterministic.** The same documented ruleset applied to the same market should produce the same score.
- **Transparent.** Rules, weights, penalties, and edge cases are published and inspectable.
- **Independent.** Verdict uses public data and has no financial relationship with prediction market platforms.
- **Conservative.** When the methodology is uncertain, it flags potential risk rather than assuming the criteria are harmless.

4. The Resolution Risk Framework

Verdict defines *resolution risk* as the risk that a market's written criteria are insufficient to determine the intended outcome in a consistent, timely, and evidence-based manner. This definition is intentionally narrow. It does not cover liquidity risk, price volatility, platform solvency, oracle failure, trader strategy, regulatory risk, or counterparty risk. It focuses on the text-level quality of the contract that determines settlement.

The framework identifies six recurring failure modes. Time ambiguity occurs when the relevant deadline, timezone, or distinction between event time and announcement time is unclear. Source ambiguity occurs when no authoritative source or source hierarchy is named. Outcome ambiguity occurs when the Yes condition is not clearly defined. Evidence ambiguity occurs when acceptable and excluded evidence are not specified. Edge-case ambiguity occurs when delays, revisions, cancellations, postponements, disputes, or late reports are not addressed. Conflict-of-interest risk occurs when the resolution authority may be structurally exposed to the outcome.

These categories make resolution quality comparable across markets. A sports market may be clear about time but weak on postponed events. A corporate market may cite official filings but fail to specify whether the transaction date or disclosure date controls. A crypto market may reference on-chain evidence but omit the relevant chain, contract, address, or block time. By scoring these patterns systematically, Verdict allows users to identify hidden procedural risk that price and volume alone do not reveal.

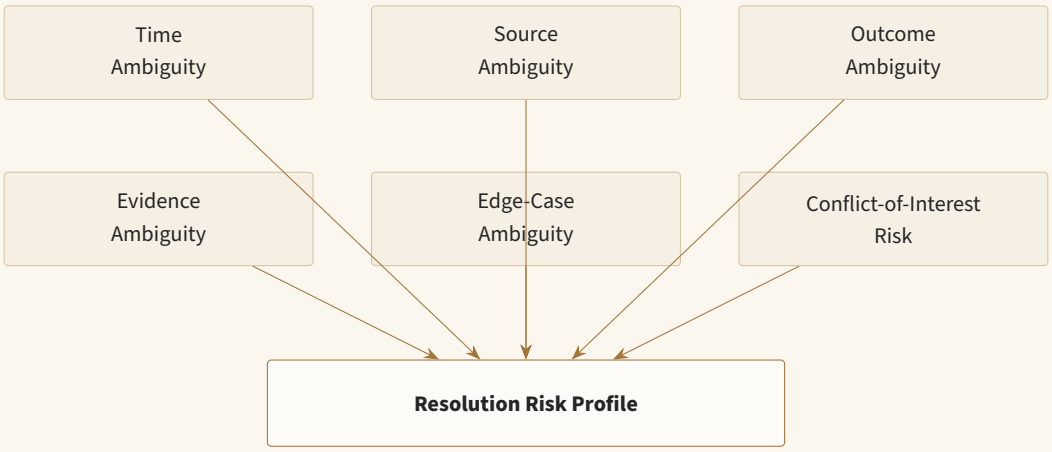

Figure 1: Verdict's six failure-mode categories for resolution quality.

5. Scoring Methodology

Verdict assigns each market a maximum score of 100 points. Time Clarity, Resolution Source, and Outcome Definition each contribute up to 20 points. Evidence Standard and Edge Case Handling each contribute up to 15 points. Conflict of Interest contributes up to 10 points. The weighting reflects the central role of time, source, and outcome language in settlement, while still recognizing that evidence rules and edge-case procedures are essential for difficult markets.

The model is additive. Points are awarded for explicit strengths and penalties are applied for specific risk indicators. This structure avoids an all-or-nothing judgment. A market can receive credit for naming an official source while losing points for failing to define late reporting. The resulting score is therefore a profile of textual strengths and weaknesses rather than a single opaque label.

TOTAL SCORE FORMULA

$$S_{total}(m) = S_{time} + S_{source} + S_{outcome} + S_{evidence} + S_{edge} + S_{conflict}$$

$$0 \leq S_{total}(m) \leq 100$$

where m denotes a market and each S_i denotes the scored contribution of one dimension.

Table 1: Six-dimension scoring model.

DIMENSION	MAX POINTS	PRIMARY QUESTION
Time Clarity	20	Does the market specify when resolution occurs and which temporal reference applies?
Resolution Source	20	Does the market name authoritative sources and define conflict handling?
Outcome Definition	20	Does the market define the Yes/No outcome clearly and concisely?
Evidence Standard	15	Does the market define accepted and excluded evidence types?
Edge Case Handling	15	Does the market address delays, revisions, cancellations, postponements, and late reports?
Conflict of Interest	10	Is the resolution authority structurally independent from the outcome?

5.1 Dimension 01 - Time Clarity

Time Clarity evaluates whether the market specifies when resolution occurs and which temporal reference governs the outcome. It awards points for an end date, temporal terms such as before or by, a stated timezone such as ET or UTC, and language distinguishing event time from announcement, report, or disclosure time. It applies a penalty when time-sensitive language appears without a timezone.

This dimension matters because many events occur, are reported, and are officially disclosed at different times. Without a clear temporal rule, participants may disagree over whether the relevant fact happened within the market window. Time clarity is especially important for corporate filings, legal deadlines, sports schedules, elections, and fast-moving geopolitical events.

TIME CLARITY RULES

+6	End date is specified.
+5	Question or description references temporal terms such as before, by, after, between, until, or an explicit date.
+5	Description includes a timezone such as ET, UTC, GMT, EST, or EDT.
+4	Description distinguishes event time from announcement, report, or disclosure time.
-5	Time-sensitive language is present but no timezone is mentioned.

5.2 Dimension 02 - Resolution Source

Resolution Source evaluates whether the market identifies authoritative evidence. It awards points for a source URL or named authority, for official or institutional references such as government agencies, SEC filings, FIFA, the Federal Reserve, or on-chain data, and for a hierarchy that resolves conflicts between sources. It penalizes reliance on undefined phrases such as “credible reporting,” “reliable sources,” or “substantial evidence.”

The objective is not to require a single type of source for every market. Different markets need different evidence. The issue is whether the contract tells users which sources count and what happens when sources disagree. Without that hierarchy, settlement can depend on discretion rather than rule application.

RESOLUTION SOURCE RULES

+10	A resolution source URL or named authority is provided.
+5	Official or institutional sources are referenced, such as government, SEC, FIFA, Federal Reserve, or on-chain data.
+5	A source hierarchy for conflict resolution is defined.
-5	Undefined "credible reporting," "reliable sources," or "substantial evidence" is used.

5.3 Dimension 03 - Outcome Definition

Outcome Definition evaluates whether the question and description define what must happen for a Yes resolution. It awards points for concise questions, binary Yes/No outcomes, and explicit Yes criteria. It penalizes ambiguous qualifiers such as significant, major, reportedly, effectively, or confirmed when those terms are not defined.

Concise and binary language reduces interpretive load, but concision alone is not enough. A short question can still be ambiguous if the description does not define the triggering condition. Verdict therefore evaluates both the question and the supporting description.

OUTCOME DEFINITION RULES

+8	Question is under 180 characters.
+6	Outcomes are binary Yes/No.
+6	Description explicitly defines what must happen for Yes resolution.
-5	Ambiguous qualifiers such as significant, major, reportedly, effectively, or confirmed are used.

5.4 Dimension 04 - Evidence Standard

Evidence Standard evaluates whether the market states what evidence counts and what evidence does not count. It awards points for accepted evidence, excluded evidence, and named evidence types such as filings, on-chain data, official statements, or published reports. It penalizes undefined dependence on confirmation or credible sources.

This dimension is distinct from Resolution Source. A market may name a source while failing to define what evidence from that source is sufficient. Conversely, a market may mention evidence categories without a hierarchy. Strong evidence standards reduce the room for post-hoc interpretation at settlement.

EVIDENCE STANDARD RULES

+5	Description specifies what evidence counts toward resolution.
+5	Description specifies what evidence does not count.
+5	Named evidence types are referenced, such as filings, on-chain data, official statements, or published reports.
-5	Resolution depends on undefined confirmation or credible sources.

5.5 Dimension 05 - Edge Case Handling

Edge Case Handling evaluates whether the market anticipates common settlement complications. It awards points for addressing delays, revisions, corrections, cancellations, postponements, disputes, and events reported after the deadline. It penalizes time-sensitive markets that do not address late reporting or delayed disclosure.

Edge cases matter because markets often resolve under imperfect conditions. Official data can be revised, events can be postponed, and reports can arrive after the relevant window. A market that does not specify how these situations are treated may appear clear until the edge case becomes decisive.

EDGE CASE HANDLING RULES

+5	Description addresses delays.
+5	Description addresses revisions, corrections, cancellations, postponements, or disputes.
+5	Description explains what happens if events are reported after the deadline.
-5	Market is time-sensitive but does not address late reporting or delayed disclosure.

5.6 Dimension 06 - Conflict of Interest

Conflict of Interest evaluates whether the resolution authority is structurally independent from the outcome. It awards points when the authority has no direct financial stake and when the market does not involve the platform operator or creator as a party. It imposes a severe penalty when the creator or operator is also the resolution authority and has a financial stake.

The rule is intentionally limited to observable structural conflicts. It does not infer hidden incentives. Its purpose is to recognize that precise rules can still lose credibility if the entity interpreting them is directly exposed to the result.

CONFLICT OF INTEREST RULES

+5	Resolution authority has no direct financial stake in the outcome.
+5	Market does not involve the platform operator or market creator as a party.
-10	Market creator or operator is also the resolution authority and has financial stake.

6. Risk Level Classification

After dimension scores are calculated, the total is mapped to four risk levels: Low, Medium, High, and Critical. These categories summarize the quality of the written resolution criteria. They do not represent expected profit, market direction, or the probability that the underlying event will occur.

The thresholds are designed for triage. A Low score does not guarantee dispute-free settlement, and a Critical score does not prove that the market will be misresolved. The classification indicates how much further review may be appropriate before a user relies on the market as a trading instrument or informational signal.

Table 2: Risk level thresholds.

RISK LEVEL	SCORE RANGE	INTERPRETATION
LOW	80–100	Criteria are comparatively well specified; residual risk may remain.
MEDIUM	60–79	Material drafting weaknesses exist and should be reviewed.
HIGH	40–59	Significant ambiguity may affect settlement expectations.
CRITICAL	0–39	Criteria are weak, incomplete, or highly discretionary.

RISK LEVEL CLASSIFICATION FUNCTION

$$R(S) = \begin{cases} \text{LOW}, & 80 \le S \le 100, \\ \text{MEDIUM}, & 60 \le S \le 79, \\ \text{HIGH}, & 40 \le S \le 59, \\ \text{CRITICAL}, & 0 \le S \le 39. \end{cases}$$

Here S is the total score across all six dimensions.

7. Platform Architecture and Features

Verdict is implemented as a serverless web application that evaluates the resolution quality of prediction markets using a transparent, rule-based methodology. The platform separates deterministic scoring from community interaction and provides tools for exploration, critique, and integration.

7.1 System Architecture

Market data is retrieved from the Polymarket Gamma API, normalized, processed through deterministic scoring logic, and rendered through the Verdict interface. Community interactions–such as comments, risk votes, watchlists, and dispute submissions–are stored separately to preserve the integrity of the score. This separation ensures reproducibility, transparency, and independence while enabling real-time exploration and discussion.

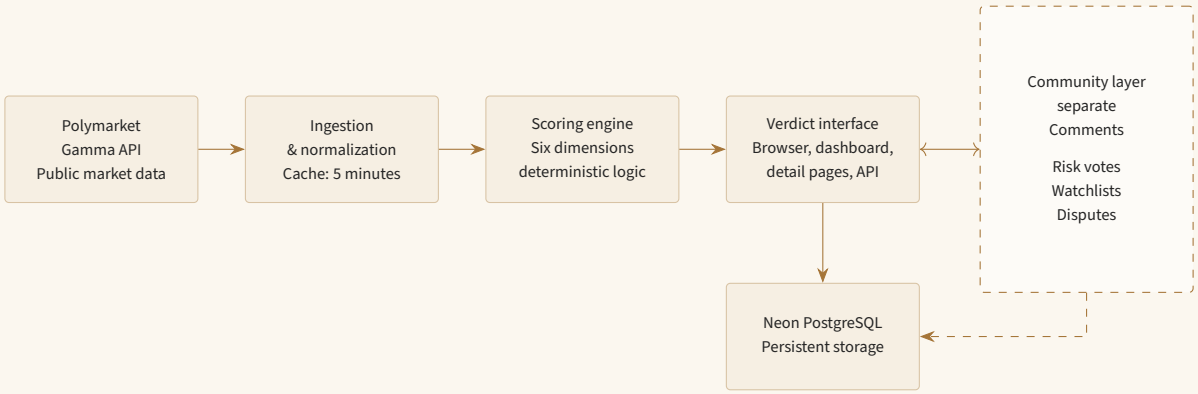

Figure 2: Current Verdict architecture. Deterministic scoring and community input are stored and interpreted as separate signals.

7.2 Market Browser and Dashboard

The Market Browser allows users to search, filter, and sort active Polymarket markets by risk level, category, trading volume, and resolution score. Categories include Politics, Crypto, Sports, Business, Tech & Science, Pop Culture, World, and Other. The Dashboard provides aggregate insights, including risk distribution histograms, score distribution charts, and a filterable market table. These views are intended for users who need system-level insight: researchers assessing market data quality, traders screening for procedural risk, and platform developers identifying repeated drafting weaknesses.

7.3 Market Detail Page

Each Market Detail Page presents a per-market breakdown of all six scoring dimensions, resolution criteria analysis, trading data, community risk vote, and discussion thread. The objective is explainability. Users should be able to see why a score was assigned rather than relying on a number without context. The page also supports methodological critique: a user may disagree with how a phrase is classified or whether a description adequately handles late reporting. Because the rules are public, these disagreements can be framed as proposals for rule improvement rather than arguments with an opaque model.

7.4 Community, AI, Disputes, and Embeds

Each market detail page includes an AI-powered Q&A section where users can ask natural language questions about the market's resolution criteria. This feature, powered through the Vercel AI Gateway, is designed to assist reading and interpretation. It does not replace the deterministic score or change the published methodology.

The Dispute System provides a structured form for reporting markets whose resolution appears to diverge significantly from written criteria. Anonymized cases may be published in the public interest. The feature is for documentation and public record purposes only. The Embed API allows third-party sites to display market-level risk information through an iframe.

Table 3: Technical stack.

LAYER	TECHNOLOGY / COMPONENT	FUNCTION
Frontend	Next.js 16 App Router, React 19, Tailwind CSS v4, shadcn/ui	User interface, market browser, dashboard, and detail pages.
Backend	Next.js Server Actions and API Routes	Scoring workflows, API endpoints, and application logic.
Database	Neon PostgreSQL serverless, Drizzle ORM	Storage for users, sessions, votes, comments, and community records.
Authentication	Better Auth, Google OAuth, Ethereum SIWE	Account creation, login, wallet authentication, and session management.
AI	Vercel AI Gateway	Natural language Q&A about resolution criteria.
Deployment	Vercel	Application hosting and serverless deployment.
Market Data	Polymarket Gamma API	Public REST source for active market data.

8. Community and Governance Layer

Verdict's community layer allows registered users to post comments, cast risk votes, and maintain watchlists while preserving the integrity of deterministic scores. Users can sign in through Google OAuth or Ethereum wallet authentication using EIP-4361 SIWE signatures. The database schema includes users, sessions, accounts, verifications, market_comment, comment_vote, and risk_vote tables, supporting identity, persistence, comments, and market-level assessments.

Community input is deliberately separated from the deterministic score. A community vote may show that users perceive a market as riskier or safer than the rules-based score indicates, but it should not silently alter the score. Verdict therefore presents two signals: a reproducible textual assessment and a user-generated risk assessment.

The governance function is procedural. Users can identify markets requiring attention, discuss edge cases, submit disputes, and critique the methodology. Future versions may formalize methodology proposals, but any changes should be versioned so that historical scores remain interpretable. This approach preserves the deterministic core of Verdict while allowing the framework to improve through public scrutiny.

COMMUNITY PURPOSE

The community layer is designed to surface experience, context, and scrutiny around markets without compromising the reproducible core of Verdict's scoring methodology.

9. Data Sources and Refresh Policy

Verdict uses the public Polymarket Gamma API as its current data source. The API requires no authentication. Verdict covers active, non-closed markets with non-zero trading volume, up to 500 markets per refresh. Market data is cached for five minutes to balance freshness with operational efficiency.

The platform is not designed as a tick-level market data service. It monitors resolution criteria and market-level rule risk. If a market's description changes, the score may update after the next refresh. Because historical score tracking is not yet implemented, displayed scores should be treated as the current observed state of market criteria rather than a permanent record of prior language.

Reliance on public data supports independence and reproducibility, but it also limits coverage to fields exposed by the API and visible market text. If relevant settlement guidance exists outside the public criteria, Verdict cannot evaluate it unless it is incorporated into publicly available market data.

10. Limitations and Future Work

10.1 Current Limitations

Verdict's scores are heuristic estimates based on written resolution criteria only. They do not assess whether a market will actually be resolved correctly. A high-scoring market can still produce a dispute under unusual facts, and a low-scoring market may resolve cleanly if the underlying event is simple and uncontested.

The current rules can also be affected by verbose but semantically weak criteria. A sophisticated operator could add language that triggers scoring credit without improving practical interpretability. Future versions should incorporate semantic analysis that evaluates whether criteria are meaningful, not merely present.

Verdict currently lacks historical score tracking. Scores reflect the present state of market criteria, which operators can modify. Historical snapshots would allow users to see how rules changed over time and would support research into whether low-scoring markets produce more disputes. The platform currently covers Polymarket only and has no enforcement power. Support for additional prediction market platforms is planned, but cross-platform scoring will require adapter logic and possibly platform-specific rule extensions.

10.2 Future Work Priorities

Future development should prioritize historical score tracking, methodology versioning, semantic NLP analysis, and cross-platform coverage. Historical tracking would create timestamped evidence of criteria changes. Versioning would allow researchers to reproduce prior scores. Semantic analysis would reduce superficial keyword gaming. Cross-platform support would permit comparison of market design practices across operators.

The dispute system can also evolve into a structured public archive containing anonymized case summaries, evidence categories, and resolution outcomes. The community layer may support formal feedback on weights, penalties, and edge-case categories. These additions should preserve the deterministic score of Verdict while improving the framework's empirical grounding.

11. Conclusion

Prediction markets depend on written rules that can be applied consistently when events occur, fail to occur, or become contested. Without clear timing, source hierarchy, outcome definitions, evidence standards, edge-case handling, and conflict safeguards, a market price may conceal significant rule risk. That risk is not visible from price and volume alone.

Verdict provides a structured method for identifying such risk in active Polymarket markets. Its deterministic framework does not decide outcomes, predict profitability, or enforce settlements. It evaluates whether market criteria are written in a way that supports reliable resolution. The platform combines this framework with browsing, dashboards, detail pages, community input, natural language Q&A, dispute reporting, and embed features.

For researchers, Verdict can help filter the quality of market-derived signals. For institutional traders, it can support procedural due diligence. For platform developers, it can serve as a drafting checklist and benchmark. As prediction markets become more visible, resolution quality should be treated as market infrastructure: measurable, inspectable, and open to public scrutiny.

12. Appendix - Full Scoring Ruleset Reference Table

The table below summarizes the deterministic scoring rules currently used by Verdict. Positive values add credit when a market satisfies a rule. Negative values apply penalties when risk indicators are present.

Table 4: Full scoring ruleset reference.

DIMENSION	RULE	POINTS
Time Clarity	End date is specified.	+6
Time Clarity	Question or description references before, by, after, between, until, or an explicit date.	+5
Time Clarity	Description includes a timezone such as ET, UTC, GMT, EST, or EDT.	+5
Time Clarity	Description distinguishes event time from announcement, report, or disclosure time.	+4
Time Clarity	Time-sensitive language is present but no timezone is mentioned.	-5
Resolution Source	A resolution source URL or named authority is provided.	+10
Resolution Source	Official or institutional sources are referenced, such as government, SEC, FIFA, Federal Reserve, or on-chain data.	+5
Resolution Source	A source hierarchy for conflict resolution is defined.	+5
Resolution Source	Undefined credible reporting, reliable sources, or substantial evidence are used.	-5
Outcome Definition	Question is under 180 characters.	+8
Outcome Definition	Outcomes are binary Yes/No.	+6
Outcome Definition	Description explicitly defines what must happen for Yes resolution.	+6
Outcome Definition	Ambiguous qualifiers are used, including significant, major, reportedly, effectively, or confirmed.	-5
Evidence Standard	Description specifies what evidence counts toward resolution.	+5
Evidence Standard	Description specifies what evidence does not count.	+5
Evidence Standard	Named evidence types are referenced, such as filings, on-chain data, official statements, or published reports.	+5
Evidence Standard	Resolution depends on undefined confirmation or credible sources.	-5
Edge Case Handling	Description addresses delays.	+5
Edge Case Handling	Description addresses revisions, corrections, cancellations, postponements, or disputes.	+5
Edge Case Handling	Description explains what happens if events are reported after the deadline.	+5
Edge Case Handling	Market is time-sensitive but does not address late reporting or delayed disclosure.	-5

DIMENSION	RULE	POINTS
Conflict of Interest	Resolution authority has no direct financial stake in the outcome.	+5
Conflict of Interest	Market does not involve the platform operator or market creator as a party.	+5
Conflict of Interest	Market creator or operator is also the resolution authority and has financial stake.	-10

12.1 Methodology Versioning Note

Any future change to weights, penalties, source parsing, semantic analysis, platform coverage, or historical tracking should be released under a named methodology version so that historical scores remain reproducible and interpretable.

◇

End of Document