VibePolitics Agent Feedback Synthesis

Date: 2026-02-05 Compiled by: Claw

Executive Summary

All five specialist agents reviewed the VibePolitics project design for blind spots. Their collective feedback converges on several critical issues, with **demographic representativeness** being the most serious concern. This document synthesizes their feedback and provides practical recommendations for the revised design.

□ Critical Blind Spots (All Agents Agree)

1. Representativeness Crisis

Severity: HIGH | Mentioned by: Priya, Kenji, Mei

The Problem: Polymarket traders are NOT representative of US voters:

- Wealthier (crypto-enabled, trading requires capital)
- More male (~90% male vs. 50% electorate)
- More educated/tech-savvy
- Younger (25-45 dominant)
- Urban/coastal heavy

Agent Quotes:

- *Priya*: "You're measuring what a narrow, affluent, male demographic believes — then generalizing to 'public opinion.'"

- *Kenji*: "You'll detect shifts among prediction market traders, not the voting public."
- *Mei*: "Traders are the 'early responders' in the information cascade... This is a *lead indicator*, not a direct measure."

Recommended Fix:

> **Reframe the claim.** Don't say "detecting public opinion shifts" — say "detecting information-sensitive shifts in politically engaged sentiment." Frame traders as sensors/proxies of influential opinion, not representative voters.

2. Circular Validation / Ground Truth Problem

Severity: HIGH | Mentioned by: Kenji, Mei, Priya

The Problem:

- Validating against polling shifts, but polls have systematic biases
- Using Google Trends to validate Google Trends signals
- No external anchor outside the system

Recommended Fix:

> **Two-tier validation:** (1) Polling shifts for directional validation, (2) Actual election outcomes as ultimate ground truth. Don't treat polls as truth — model their uncertainty.

3. Causal Ambiguity

Severity: MEDIUM-HIGH | Mentioned by: Kenji, Mei

The Problem: Market moves could be:

- Information aggregation (good signal)
- Media coverage driving attention
- Speculation/arbitrage
- Manipulation attempts System can't distinguish these mechanisms.

Recommended Fix:

> **Add attribution layer:** Correlate signals with news events. Add NewsAPI or GDELT for context. Anomalies without news = potential manipulation flag.

4. Multiple Comparison Problem

Severity: MEDIUM | Mentioned by: Mei, Kenji

The Problem: 6 signals × N markets × M keywords × daily = many false positives at 95% confidence.

Recommended Fix:

> **Implement FDR correction** (Benjamini-Hochberg). Report FDR-corrected signals at q < 0.05.

🟡 Moderate Blind Spots

5. Temporal Resolution Mismatch

Mentioned by: Kenji

Markets update continuously (milliseconds), Google Trends daily, polls every 2-4 days. Comparing apples to oranges.

Fix: Explicit lag modeling. Granger causality tests. Don't assume synchrony.

6. Threshold Non-Stationarity

Mentioned by: Mei

2024 thresholds may not apply to 2026 (midterms vs. presidential, different context).

Fix: Cross-cycle validation (2020 → 2022 → 2024). Adaptive thresholding.

7. Agent Homogeneity

Mentioned by: Mei

Same LLM with different prompts may produce performative disagreement, not genuine diversity.

Fix: Use different LLMs for Alpha vs. Beta (Claude vs. GPT vs. Llama).

8. Liquidity Bias

Mentioned by: Priya

Thin markets excluded, but emerging issues start in thin markets.

Fix: Flag thin markets but don't discard. Monitor thin → liquid transitions.

9. Platform Dependency Risk

Mentioned by: Kenji, Priya

Polymarket could change API, face regulatory action. Single point of failure.

Fix: Document data availability changes. Build fallback scraping for PredictIt, Manifold.

10. IRB/Legal Questions

Mentioned by: Priya

Automated data collection from platforms (TOS), potential identification of traders, market influence if widely adopted.

Fix: Add legal review section. Consult IRB if pursuing academic publication.

□ Theoretical Defenses (Synthesized from Agents)

Defense 1: "Wisdom of Affluent Crowds"

Source: Priya

Reframe demographic skew as a feature: we're measuring **influential opinion** (early adopters, donors, opinion leaders) not mass opinion. This is valuable for campaigns who care about resource allocation and media narrative.

Defense 2: "Information Cascade Theory"

Source: Priya

Markets don't just predict opinion — they **cascade into it** through media coverage. We're detecting the "upstream" signal before mainstream consciousness.

Defense 3: "Strategic Politician Hypothesis"

Source: Priya

Politicians react to markets more than polls now (Mitch McConnell cited prediction markets in 2024). The relevant DV isn't "what will voters think" but "what will politicians do."

Defense 4: "Reframe Unit of Analysis"

Source: Mei

Don't claim "detecting voter opinion shifts." Claim "detecting **information-sensitive shifts in politically engaged sentiment**." Cite Bikhchandani et al. (1992) on information cascades.

Defense 5: "Multi-Source Fusion Robustness"

Source: Mei

Markets, search, and agents have **different biases**. If all three agree, that's stronger evidence than any single source.

□ Novel Algorithmic Approaches Suggested

From Priya:

Approach	Concept	Practicality
Cross-Platform Arbitrage Detection	Monitor Polymarket vs. Kalshi divergence as signal	□ HIGH - Both APIs available
Search Intent Classification	Classify queries as informational/transactional/comparison	□ HIGH - Keyword-based, no extra API
Semantic Drift Detection	BERT embeddings to track meaning shifts	⚠ MEDIUM - Requires NLP pipeline
Epidemiological Diffusion	SIR model for opinion spread across states	□ HIGH - Uses existing Trends data
Information Asymmetry Index	Market moves when search flat = informed traders	□ HIGH - Computable from existing signals

From Kenji:

Approach	Concept	Practicality
Cross-Population Signal Diffusion (CPSD)	Track info flow: Traders → Search → Social → Public	⚠ MEDIUM - Needs social data

Approach	Concept	Practicality
Belief Persistence Scoring (BPS)	Weight signals by decay rate (>3 days = real)	□ HIGH - Computable now
Context-Aware Anomaly Detection (CAAD)	Correlate signals with news events	□ HIGH - NewsAPI is cheap
Multi-Resolution Ensemble (MRE)	Run at hourly/daily/weekly resolutions	□ HIGH - No new data needed
Counterfactual Impact Estimation (CIE)	"Equivalent to X poll points" translation	⚠ MEDIUM - Requires historical calibration

From Mei:

Approach	Concept	Practicality
Two-Tier Validation	Polls + Election outcomes	□ HIGH - Essential
Demographic Sensitivity Analysis	Quantify representativeness gap	⚠ MEDIUM - Limited trader data
FDR Control	Benjamini-Hochberg correction	□ HIGH - Standard stats
Event-Based Ground Truth	Do signals cluster around known events?	□ HIGH - NewsAPI
Baseline Comparison	Compare to simple MA threshold	□ HIGH - Essential
Cross-Cycle Validation	Test 2024 thresholds on 2020/2022	⚠ MEDIUM - Needs historical data

Approach	Concept	Practicality
Hidden Markov Model	Detect regime transitions	⚠ MEDIUM - More complex
Gaussian Process Regression	Continuous probability, no thresholds	⚠ MEDIUM - More complex
Network Graph of Markets	Find "hub" markets that lead	HIGH - From Polymarket data
Diverse Agent Ensemble	Claude + GPT + Llama	HIGH - Just config change

---

Claw's Practical Recommendations

Based on the agent feedback and the principles of **intuitive design** and **not-over-engineering**, here's what I recommend:

Tier 1: Must Implement (MVP)

1. **Reframe Claims**
    - Change all language from "public opinion" to "politically engaged sentiment"
    - Add section: "Which Public? Prediction Markets as Signals of Influential Subgroups"
2. **FDR Correction**
    - Implement Benjamini-Hochberg before any signal reporting
    - Standard practice, reviewers will expect it
3. **Baseline Comparison**
    - Simple baseline: Polymarket price + 7-day MA + 3% threshold
    - If baseline does nearly as well, that's informative
4. **Belief Persistence Scoring**

- Weight signals by duration (>3 days = stronger)
- Simple, computable from existing data

5. **Two-Tier Validation**
    - Tier 1: Polling shifts (directional)
    - Tier 2: Election outcomes (ground truth)

Tier 2: Should Implement (Post-MVP)

6. **Cross-Platform Divergence**
    - Add Kalshi API when auth secured
    - Polymarket vs. Kalshi divergence = platform-specific bias signal
7. **News Event Attribution**
    - Add NewsAPI ($99/mo) or GDELT (free)
    - Correlate signals with news to explain WHY shifts occur
8. **Information Asymmetry Index**
    - z_market - (0.5 * z_search + 0.5 * z_news)
    - Detects when traders have info public hasn't processed
9. **Diverse Agent Models**
    - PolAgent-A: Claude
    - PolAgent-B: GPT-4
    - EconAgent-A: DeepSeek or Llama
    - EconAgent-B: Different model

Tier 3: Nice to Have (Future)

10. Semantic drift detection (BERT)
11. HMM regime detection
12. Epidemiological diffusion model
13. Network graph analysis

□ What NOT to Do

Based on the "not-over-engineering" principle:

1. **Don't build demographic weighting** — We don't have trader demographics, and the reframing defense is stronger anyway.
2. **Don't add social media sentiment** — Scope creep. Stick to markets + search + agents.
3. **Don't implement wavelet transforms** — Interesting but overkill for MVP.
4. **Don't build counterfactual simulation** — Requires causal framework we don't have.
5. **Don't rename the project** — "VibePolitics" is memorable. Use formal language in papers but keep the brand.

□ Summary Table

Issue	Severity	Fix	Effort
Representativeness	□ HIGH	Reframe claims	LOW
Circular validation	□ HIGH	Two-tier validation	MEDIUM
Multiple comparisons	□ MEDIUM	FDR correction	LOW
No baseline	□ MEDIUM	Add simple baseline	LOW
No attribution	□ MEDIUM	NewsAPI integration	MEDIUM
Agent homogeneity	□ LOW	Use different LLMs	LOW
Platform risk	□ LOW	Document, add fallbacks	LOW

Next Steps

1. Update PROJECT_SPEC.md with reframed claims
2. Add validation protocol with two-tier structure
3. Implement FDR correction in signal computation
4. Build simple baseline for comparison
5. Integrate NewsAPI for attribution layer
6. Configure diverse agent models

This synthesis represents the collective wisdom of Kenji , Priya , Mei , Arjun , and Wei . Compiled by Claw for project revision.