

Project Outline: TrackRecord (working title)

Tagline: "Track the track record of everyone with an opinion"

1. EXECUTIVE SUMMARY

Problem: Commentators, pundits, and experts make countless predictions across media daily. They're rarely held accountable when wrong, leading to a marketplace of ideas polluted with confident but inaccurate voices.

Solution: An AI-powered accountability layer that captures predictions, maps them to Polymarket markets, simulates paper trading positions, and tracks long-term accuracy with real P&L metrics.

Core Innovation: Transform abstract "prediction accuracy" into universal language of profit/loss, making it instantly clear who adds value vs. who's just noise.

Target Launch: 6-month MVP, starting with US political commentators during 2026 election cycle.

2. PRODUCT VISION

Phase 1: Paper Trading Platform (Months 1-6)

- Track 20-50 high-profile pundits automatically
- Map predictions to Polymarket markets
- Simulate paper positions based on stated confidence
- Public leaderboard with P&L tracking
- Basic website + Twitter presence

Phase 2: Scale & Engagement (Months 7-12)

- Expand to 200+ pundits across politics, economics, tech
- Browser extension showing scores inline
- Pundit opt-in verification program
- Mobile app with push alerts
- "Fade or Follow" trading signals

Phase 3: Real Money & Monetization (Year 2)

- Verified pundits can link real Polymarket wallets

- Premium subscription tier with advanced analytics
 - API access for institutional clients
 - Expand beyond Polymarket to other prediction markets
 - Build proprietary resolution system for uncovered predictions
-

3. TECHNICAL ARCHITECTURE

3.1 Core Components

A. Prediction Capture Engine

Input Sources:

- Twitter/X API (real-time monitoring)
- Podcast transcripts (Whisper API + RSS feeds)
- YouTube captions
- News articles & op-eds (web scraping)
- TV transcripts (partnership with transcript services)

NLP Pipeline:

1. Monitor target accounts/shows
2. Identify predictive statements (GPT-4/Claude)
 - Future tense indicators
 - Probability language
 - Conditional statements

3. Extract structured data:

```
{  
  predictor: "name",  
  statement: "full quote",  
  prediction: "structured claim",  
  confidence: "high/medium/low",  
  timestamp: "ISO 8601",
```

```
source_url: "link",  
domain: "politics/economy/tech"  
}
```

4. Store with cryptographic hash

B. Market Matching Engine

Process:

1. Query Polymarket API for active markets
2. Use semantic similarity (embeddings) to match prediction to market
3. Manual review queue for ambiguous matches
4. Store mapping: prediction_id → market_id

Data structure:

```
{  
  prediction_id: "uuid",  
  market_id: "polymarket_id",  
  market_question: "text",  
  match_confidence: 0.0-1.0,  
  match_type: "auto/manual/verified",  
  entry_price: "market price at prediction time",  
  entry_timestamp: "ISO 8601"  
}
```

C. Position Simulator

Logic:

1. Convert verbal confidence → implied probability
2. Standardize position size (\$100 base unit)
3. Calculate shares purchased at entry price
4. Track market price changes (hourly polling)
5. Calculate unrealized P&L for open positions

6. Calculate realized P&L at market resolution

Confidence Mapping:

- "Definitely/will/certain" → 90%+ → \$1,000 position
- "Very likely/expect" → 75-90% → \$500 position
- "Probably/likely" → 60-75% → \$300 position
- "Maybe/could/possible" → 40-60% → \$100 position
- Hedge/qualify → Don't track (too vague)

P&L Calculation:

- Entry: Buy X shares @ \$P
- Exit: Market resolves YES → \$1.00/share, NO → \$0.00/share
- Profit/Loss: (Exit Value - Entry Cost)

D. Scoring & Analytics Engine

Metrics per pundit:

- Total P&L (cumulative)
- Win rate (% of profitable positions)
- ROI (% return on capital deployed)
- Sharpe ratio (risk-adjusted returns)
- Brier score (probabilistic accuracy)
- Prediction volume (activity level)
- Domain breakdown (politics vs economy vs tech)
- vs. Market baseline (outperformance)
- Calibration curve (stated confidence vs actual outcomes)

Leaderboards:

- All-time P&L
- Last 30/90/365 days

- By domain
- Biggest wins/losses
- Most improved
- Contrarian bets (beat consensus)

E. Data Storage

Tech Stack:

- PostgreSQL (main database)
- Redis (caching, real-time updates)
- S3 (raw data backup, prediction snapshots)

Tables:

- pundits (id, name, twitter, affiliation, domains)
- predictions (id, pundit_id, statement, confidence, timestamp, hash)
- markets (id, polymarket_id, question, resolution, close_date)
- positions (id, prediction_id, market_id, entry_price, shares, status)
- resolutions (id, position_id, exit_price, pnl, timestamp)

Immutability:

- SHA-256 hash of each prediction on capture
- Append-only predictions table
- Periodic Merkle tree snapshots
- Optional: Anchor root hashes to Bitcoin/Ethereum monthly

3.2 External Integrations

Polymarket API:

- Market data (questions, prices, volume)
- Historical price data
- Resolution status
- WebSocket for real-time updates

Social Media:

- Twitter/X API for monitoring
- Post tracking & engagement
- Automated @replies (opt-in)

LLM APIs:

- Claude/GPT-4 for prediction extraction
 - Embedding models for semantic matching
 - Content moderation
-

4. USER INTERFACE

4.1 Web Platform


Homepage:

- Live leaderboard (top 20 pundits)
- Featured predictions (high-stakes, high-confidence)
- Recent resolutions (big wins/losses)
- Search bar (find any pundit)

Pundit Profile Page:

[Profile Header]

Name, photo, affiliation, social links

Overall P&L: +\$12,450 ( +15% this month)


Win Rate: 64% | ROI: +18% | Predictions: 87

[Performance Charts]

- P&L over time (line chart)
- Win/loss by domain (pie chart)
- Calibration curve
- vs. Market comparison


[Open Positions] (live updating)

Market | Entry | Current | Unrealized P&L

"Trump wins PA" | 65¢ | 72¢ | +\$350 

[Closed Positions] (paginated)

Market | Entry | Exit | P&L | Date

"Biden drops out" | 25¢ | \$1.00 | +\$750  | 2024-07-21

[Recent Predictions]

Quote | Confidence | Market | Status

Search & Filter:


- By domain (politics, economics, tech)
- By timeframe (last 30 days, all-time)
- By metric (P&L, win rate, volume)
- By position status (open, closed)

Market Detail Page:

- All pundits who predicted on this market
- Consensus vs contrarians
- Who was right/wrong
- Timeline of predictions

4.2 Browser Extension (Chrome/Firefox)

Features:

- Hover over pundit name on Twitter → popup with stats
- Inline badges:  +\$5k P&L | 68% win rate
- "Track this prediction" button
- Alert when tracked pundit makes new prediction

4.3 Mobile App (Future)

Core flows:

- Browse leaderboards
 - Follow favorite pundits
 - Get push alerts for new predictions
 - Quick stats on the go
-

5. GO-TO-MARKET STRATEGY

5.1 Launch Strategy

Pre-Launch (Month 1-2):

1. Select 20 high-profile pundits (mix of accurate + inaccurate)
2. Backfill 3-6 months of historical predictions
3. Build MVP website + basic tracking
4. Create social media accounts
5. Prepare "launch kit" (graphics, explainers)

Launch (Month 3):

1. Go live with public leaderboard
2. Twitter thread: "We tracked 20 pundits for 6 months. Here's who would've made money..."
3. Individual callouts: "@pundit, you'd be up \$X if you'd bet on your predictions"
4. Submit to HN, Reddit, political forums
5. Pitch to Polymarket for cross-promotion

Growth (Month 4-6):

1. Viral moments: Highlight spectacular failures/successes
2. Add 10-20 new pundits monthly based on community requests
3. Encourage pundits to verify & claim profiles
4. Media outreach: Data journalism partnerships
5. Build API for third-party integrations

5.2 Target Audiences

Primary:

- Polymarket users (want better information sources)
- Politics junkies (chronically online, love drama)
- Traders/investors (looking for alpha)
- Media literate skeptics (tired of BS)

Secondary:

- Journalists (source for stories)
- Fact-checkers (new accountability tool)
- Good-faith pundits (want to prove credibility)
- Academics studying prediction/forecasting

5.3 Content Strategy

Twitter/X:

- Daily updates: "Today's biggest wins/losses"
- Weekly: "Pundit of the week"
- When predictions resolve: "@pundit called it!" or "This aged poorly..."
- Viral formats: Before/after screenshots, hall of shame

Blog/Newsletter:

- Deep dives on methodology
- Case studies: "How we tracked [big prediction]"
- Interviews with accurate predictors
- Industry trends: "Are political pundits getting worse?"

Partnerships:

- Polymarket: Official integration, co-marketing
- Podcasts: Provide guest stats, sponsor segments
- News orgs: Licensing data for chyrons/articles

6. MONETIZATION

6.1 Revenue Streams

Freemium Model:

Free Tier:

- Browse leaderboards
- View pundit profiles (limited history)
- Basic stats (P&L, win rate)
- See top predictions

Premium (\$9.99/month or \$99/year):

- Full prediction history
- Advanced analytics (Sharpe, calibration, etc.)
- Custom alerts ("Tell me when X makes a prediction")
- Export data (CSV, API)
- Ad-free experience
- Early access to new features

Pro Tier (\$49/month - for traders/institutions):

- API access (10k requests/month)
- Real-time WebSocket feeds
- Custom filters & queries
- Institutional reporting
- White-label options

Additional Revenue:

- Polymarket affiliate (rev share on referred users)
- Data licensing (sell aggregated data to hedge funds, media)
- Verified badges for pundits (\$99/year)
- Sponsored predictions (brands can sponsor market creation)

6.2 Unit Economics (Year 1 projections)**Costs:**

- Infrastructure: \$500/month (AWS, databases, APIs)
- LLM APIs: \$1,000/month (prediction extraction)
- Data feeds: \$300/month (news APIs, transcripts)

- Development: 1-2 FTEs (founders sweat equity initially)

Revenue (Conservative):

- Month 6: 100 free users, 10 premium (\$100/month)
- Month 12: 5,000 free users, 200 premium (\$2,000/month)
- Year 2: 50,000 free, 2,000 premium (\$20k/month)

Path to profitability: 12-18 months with lean team

7. KEY RISKS & MITIGATION

7.1 Technical Risks

Risk: NLP fails to accurately capture predictions **Mitigation:**

- Human review queue for high-profile predictions
- Community flagging system
- Pundits can contest/verify their own predictions

Risk: Market matching is unreliable **Mitigation:**

- Conservative matching (only track high-confidence matches)
- Manual review for ambiguous cases
- Allow pundits to approve/reject matches

Risk: Polymarket API goes down or changes **Mitigation:**

- Cache all historical data locally
- Build adapters for multiple prediction markets
- Fallback to manual resolution if needed

7.2 Legal Risks

Risk: Defamation claims from pundits with bad track records **Mitigation:**

- Only track public statements with links to sources
- Use objective metrics (market data, not opinions)
- Clear disclaimer: "Simulated positions, not financial advice"
- Allow pundits to provide context/corrections

Risk: Copyright issues with quotes/transcripts **Mitigation:**

- Keep quotes minimal (fair use)
- Link to original sources
- Partner with transcript services for licensing

Risk: Financial regulations (are we giving investment advice?) **Mitigation:**

- Clear disclaimers everywhere
- Educational content, not recommendations
- Consult lawyers on terms of service

7.3 Business Risks

Risk: Pundits stop making predictions when tracked **Mitigation:**

- Start with historical tracking (can't avoid past)
- Frame as credibility boost for accurate predictors
- Most won't change behavior (ego, job requirements)

Risk: Polymarket loses liquidity or shuts down **Mitigation:**

- Diversify across prediction markets (Kalshi, Metaculus)
- Build proprietary resolution for non-market predictions
- Community voting as backup

Risk: Competitors copy the idea **Mitigation:**

- First-mover advantage, build brand early
- Network effects (more pundits = more value)
- Proprietary data (years of tracked predictions)
- Community moat (engaged users)

8. SUCCESS METRICS

8.1 Phase 1 (Months 1-6)

Product:

- 20+ pundits tracked
- 500+ predictions captured
- 100+ resolved positions

- <24hr lag from prediction to capture

Engagement:

- 1,000 monthly active users
- 50+ premium subscribers
- 10,000+ Twitter followers
- One viral moment (100k+ impressions)

Technical:

- 95%+ uptime
- <5% prediction capture error rate
- <10% market matching error rate

8.2 Phase 2 (Months 7-12)

Product:

- 200+ pundits tracked
- 5,000+ predictions captured
- Browser extension launched
- API in beta

Engagement:

- 10,000 monthly active users
- 200+ premium subscribers
- 5 verified pundits (claimed profiles)
- Featured in 3+ major media outlets

Revenue:

- \$2,000 MRR
- \$10k in annual contracts (API/data licensing)

8.3 Phase 3 (Year 2)

Product:

- 500+ pundits tracked
- Real money tracking for 10+ verified pundits

- Mobile app launched
- Multi-market support (beyond Polymarket)

Engagement:

- 100,000 monthly active users
- 2,000+ premium subscribers
- 50+ verified pundits
- Mentioned on major TV/podcast at least once

Revenue:

- \$20,000 MRR
 - Profitable (ramen profitable minimum)
-

9. TECHNICAL ROADMAP**Month 1-2: Foundation**

- Set up infrastructure (AWS, databases)
- Build prediction capture pipeline
- Integrate Polymarket API
- Create basic data models
- Backfill 20 pundits, 6 months history

Month 3-4: MVP Launch

- Build web interface (leaderboard, profiles)
- Implement position simulator
- Create scoring algorithms
- Launch public beta
- Set up social media presence

Month 5-6: Iteration & Growth

- Add more pundits based on demand
- Improve NLP accuracy
- Build manual review tools

- Launch premium tier
- Implement basic API

Month 7-9: Scale

- Browser extension (Chrome/Firefox)
- Advanced analytics dashboard
- Verified pundit program
- Alert system (email/push)
- Multi-domain expansion (economics, tech)

Month 10-12: Monetization

- Full API launch
 - Data licensing program
 - Institutional partnerships
 - Mobile app (iOS/Android)
 - Real money tracking pilot
-

10. TEAM & RESOURCES

Initial Team (Month 1-6)

- **1x Full-stack Engineer/Founder:** Platform development, infrastructure
- **1x ML/NLP Engineer:** Prediction extraction, market matching
- **0.5x Designer:** UI/UX, branding (contract/part-time)
- **0.5x Content/Community:** Social media, content, user engagement (contract)

Expanded Team (Month 7-12)

- Add: 1x Backend Engineer (scaling, API)
- Add: 1x Mobile Developer (apps)
- Add: 0.5x Data Analyst (insights, reporting)

Technology Stack

Backend:

- Python (FastAPI for API, Celery for async tasks)

- PostgreSQL (main database)
- Redis (caching, queues)
- Docker (containerization)

Frontend:

- React/Next.js (web app)
- Tailwind CSS (styling)
- Recharts (data visualization)

Infrastructure:

- AWS (EC2, RDS, S3, Lambda)
- CloudFlare (CDN, DDoS protection)
- Vercel (frontend hosting)

ML/AI:

- OpenAI/Anthropic APIs (prediction extraction)
- Hugging Face (embeddings, matching)
- Custom fine-tuned models (future)

Monitoring:

- Sentry (error tracking)
- Grafana/Prometheus (metrics)
- PostHog (product analytics)

11. COMPETITIVE LANDSCAPE

Direct Competitors

None exist exactly, but adjacent:

- **Metaculus:** Crowd forecasting, not pundit tracking
- **Good Judgment Open:** Structured forecasting tournaments
- **PolitiFact/FactCheck:** Past claims, not future predictions

Indirect Competitors

- Twitter Community Notes (fact-checking)

- Prediction markets themselves (Polymarket, Kalshi)
- Media criticism sites (no quantitative tracking)

Competitive Advantages

1. **First mover:** No one tracking pundits systematically with P&L
 2. **Universal metric:** P&L beats abstract accuracy scores
 3. **Automated:** No manual fact-checking, scales infinitely
 4. **Real-time:** Track predictions as they happen
 5. **Market integration:** Piggyback on Polymarket's infrastructure
-

12. LONG-TERM VISION (3-5 years)

Become the IMDB/Rotten Tomatoes for predictions:

- Every pundit has a profile and score
- Media orgs display scores in chyrons
- Standard metric: "64% TrackRecord"
- Browser extension used by millions
- Premium data product for institutions

Expand beyond politics:

- Corporate predictions (analyst earnings calls)
- Sports (commentator predictions)
- Tech (product launch predictions, startup success)
- Science (research outcome predictions)

Build the resolution layer:

- Own prediction market for non-Polymarket topics
- Community-driven resolution
- Become infrastructure for accountability economy

Exit scenarios:

- Acquisition by Polymarket (vertical integration)
- Acquisition by Twitter/X (native feature)

- Acquisition by Bloomberg/Financial Times (data product)
 - Stay independent (profitable SaaS)
-

13. NEXT STEPS

Immediate Actions (Week 1)

1. ☒ Finalize project outline
2. ☐ Register domain (TrackRecord.com, ThePunditIndex.com, TrackRecord.io?)
3. ☐ Set up GitHub repo, project management
4. ☐ Design database schema
5. ☐ List target pundits (first 20)

Sprint 1 (Weeks 2-4)

1. ☐ Build prediction capture MVP (Twitter only)
2. ☐ Integrate Polymarket API
3. ☐ Create basic position simulator
4. ☐ Test on 5 pundits, 3 months history

Sprint 2 (Weeks 5-8)

1. ☐ Build web frontend (leaderboard + profiles)
2. ☐ Implement scoring algorithms
3. ☐ Backfill 20 pundits
4. ☐ Internal testing & refinement

Sprint 3 (Weeks 9-12)

1. ☐ Public beta launch
 2. ☐ Marketing push (Twitter, HN, Reddit)
 3. ☐ Gather feedback, iterate
 4. ☐ Add more pundits
-

14. OPEN QUESTIONS

1. **Name?** TrackRecord, TrackRecord, PredictorIndex, AccountabilityLayer?
2. **Start with paper or push hard for real money opt-ins?**
3. **Which 20 pundits first?** (Mix of good/bad, left/right, mainstream/indie)
4. **Monetize early or grow free users first?**
5. **Build in public or stealth until launch?**
6. **Solo founder or find co-founder?**
7. **Bootstrap or raise pre-seed?**
8. **Which domain first?** (Politics easiest, but economics has \$ incentive)