

# Picks & Shovels Stock Oracle - Full Project Description

## Executive Summary

Build an autonomous, always-on “Oracle” that ingests every public hint—RFPs, contracts, news, blogs, forums, social chatter, even TV transcripts—about AI data-center expansions, power projects, and water infrastructure. It applies NLP, relation extraction, custom classifiers (e.g., Legal-BERT), rule-based filters, and ML-driven scoring to surface under-the-radar “picks & shovels” companies whose growth curves, ancillary offerings (racks, HVAC, power gear, water treatment), and contracts with major AI players qualify them as pre-IPO opportunities. The system powers a three-tier subscription service (web dashboard, API access, enterprise alerts) and uses email/Slack notifications for high-value signals. You’ll phase the build as an MVP ingesting SAM.gov and news feeds, then layer in state portals, social media, advanced NLP, and scoring, while leveraging SBIR/DOE grants and venture funding.

## Background & Rationale

**Explosive Compute Demand:** Data-center electricity use is set to more than double to 945 TWh by 2030, outpacing Japan’s total consumption today, driven largely by AI training and inference workloads.

**Picks & Shovels Phenomenon:** In gold rushes, the greatest profits went to those selling picks and shovels. Similarly, during the AI infrastructure boom, suppliers of racks, cooling, power distribution, and water-treatment stand to win regardless of which hyperscaler strikes “gold”.

## Project Objectives

1. Comprehensive Signal Capture
2. Automated NLP & Filtering
3. Criteria Matching (growth curves, ancillary supply, contract detection)
4. Scoring & Ranking
5. Real-Time Alerts & Dashboard
6. Phased Funding & Grants

## Technical Architecture

- Data Sources: SAM.gov API, state portals, NewsAPI.org, Twitter API v2, forums, blogs, TV transcripts.
- Storage: S3/GCS (raw), PostgreSQL (metadata), Neo4j (relationships).
- NLP Pipeline: spaCy (tokenization, NER), Legal-BERT fine-tuning, rule-based and ML classification.
- Signal Detection: Match against growth curves, ancillary supply taxonomy, contract detection.
- Scoring: Composite "Oracle Score" for filtering.
- Dashboard: React/Next.js frontend; email/Slack alerts.
- API: RESTful with FastAPI, JWT authentication.

## Implementation Roadmap

- Phase 1 – MVP (0–3 months)
- Phase 2 – State & Social Integration (3–6 months)
- Phase 3 – Advanced NLP (6–9 months)
- Phase 4 – Alerts & Public API (9–12 months)
- Phase 5 – Scale & Harden (12–18 months)

## **Go-to-Market & Monetization**

- Launch Beta pilots.
- Conference presentations.
- Channel partnerships with consultants.
- Subscription Pricing: Basic (\$25K), Professional (\$60K), Enterprise (\$150K) annual tiers.

## **Regulatory & Compliance**

- GDPR/CCPA Compliance
- Export Control Classification
- SOC 2 Type II Security Standards

## **Funding & Grants**

- DOE RFI partnerships and response.
- SBIR/STTR grant applications (DOE, NSF, DOD).
- Series A venture capital post-MVP.

## **Conclusion**

This “Picks & Shovels” Stock Oracle unites cutting-edge NLP, vast public data, and rigorous scoring to spot under-the-radar AI infrastructure suppliers. Phased as an MVP, with commercial APIs, state/federal signals, and strategic VC funding—it will become the definitive oracle for AI infra investors.