

SiteScope

AI-Powered Data Center Site Screening

Subconscious Hackathon 2026 | Project Report

Executive Summary

SiteScope is an AI research agent that compresses data center site selection screening from weeks of manual consultant work into a single API call taking minutes. Built on the Subconscious platform using the TIM-CLAUDE engine, the agent systematically researches candidate markets across seven dimensions — power availability, community sentiment, tax incentives, natural hazards, connectivity, recent development activity, and regulatory landscape — and produces a structured, ranked market dossier with actionable signal ratings.

The project targets the information bottleneck behind what is arguably the largest infrastructure buildout since the interstate highway system: the global data center expansion required to support AI and cloud computing. With \$3 trillion in investment needed by 2030 and \$98 billion in projects blocked or delayed in a single quarter of 2025 alone, the gap between available capital and deployable infrastructure is defined not by money, but by the inability to research and evaluate candidate markets fast enough.

Market Context: Why This Matters Now

The Scale of the Buildout

The data center sector is experiencing an infrastructure investment supercycle. According to JLL's 2026 Global Data Center Outlook, roughly 100 GW of new capacity is expected to come online between 2026 and 2030, representing \$1.2 trillion in real estate asset value creation alone. Tenants will likely spend an additional \$1–2 trillion fitting out space with IT equipment. The global data center market is projected to reach \$300 billion in 2026, up from \$270 billion in 2025.

This growth is driven primarily by AI workloads. With over \$500 billion spent on data centers in 2025 alone and projections of \$5–8 trillion in AI infrastructure spending through 2030, hyperscale operators are racing to secure capacity. But securing capacity requires sites, and sites require research.

The Bottleneck: Site Selection

Site selection has fundamentally changed. Where developers once optimized for cheap land and basic connectivity, the gating factors are now power availability, community acceptance, and regulatory alignment. Utility interconnection timelines of 4–10 years are now common in primary markets, fundamentally reshaping how early site selection must begin.

The research burden is enormous and fragmented. A typical site selection process involves:

- Screening 30+ candidate markets across multiple criteria simultaneously
- Commissioning power studies, hiring brokers, retaining environmental consultants, and engaging utility representatives — often in parallel with poor information sharing
- Tracking community sentiment across local news, council meeting minutes, and public records
- Monitoring rapidly changing state tax incentive programs and regulatory postures
- Assessing natural hazard exposure, water availability, and climate suitability for cooling

This process typically takes 2–4 weeks for preliminary screening alone, before any on-the-ground engagement begins. Teams rely on major consulting firms — Cushman & Wakefield, CBRE, JLL, Colliers, Arup, AlixPartners — all of which maintain dedicated data center site selection practices charging premium fees.

The \$98 Billion Wall

Community opposition has emerged as the single most consequential risk factor in data center development. According to Data Center Watch, \$98 billion in planned data center investment was blocked or delayed in Q2 2025 alone — more than all previous quarters since 2023 combined. Twenty-five projects were canceled due to local opposition that year, quadrupling from the prior year.

Anti-data-center groups are now active in at least 24 US states. Opposition is bipartisan, locally organized, and increasingly effective. Concerns center on water usage, energy costs, noise, and environmental impact. High-profile cancellations in Virginia, Indiana, Georgia, and Arizona demonstrate that even well-capitalized projects with secured land can fail at the community engagement stage.

A February 2026 analysis by Sightline Climate estimates that 30–50% of data center capacity expected to come online in 2026 may not be delivered on schedule, reflecting constraints that now extend well beyond power availability into permitting and community acceptance.

The Consulting Segment

Data center consulting services represent the fastest-growing segment within the broader \$300B+ market. Skills shortages in data center design, networking architecture, and site planning are pushing organizations to rely on specialist consultants for feasibility studies, standards alignment, vendor selection, and risk management. This creates both the market opportunity and the cost structure that SiteScope targets.

What SiteScope Does

SiteScope replaces the first two weeks of the site selection screening process. It does not replace the consultant; it replaces the manual research phase that precedes on-the-ground engagement. The positioning is deliberate: the screening layer before the relationship layer.

Input

The user provides their requirements in natural language:

- Capacity (MW)
- Target timeline to energization
- Geographic region or preferences
- Workload type and density (e.g., AI training at 40–50kW/rack)
- Priority weighting (power availability vs. incentives vs. community receptiveness)
- Cooling preferences and water dependency tolerance

Research Process

The agent systematically researches each candidate market across seven dimensions, using Subconscious's built-in platform tools (`web_search`, `news_search`, `google_search`, `company_search`) to chain 20–30 searches within a single inference:

- Power & Grid: utility capacity, interconnection queue status, grid expansion plans, time-to-energize estimates
- Community Sentiment: opposition groups, moratoriums, zoning battles, canceled projects, supportive officials
- Tax & Incentives: state equipment tax exemptions, property abatements, economic development programs
- Natural Hazards & Climate: flood zones, seismic risk, wildfire exposure, hurricane corridors, ambient cooling advantage, water availability
- Connectivity: internet exchange proximity, carrier-neutral facilities, fiber route diversity
- Recent Development Activity: active operators, project announcements, campus developments underway
- Regulatory & Political Landscape: pending legislation, zoning changes, utility rate restructuring, political posture

Output

The agent produces a structured JSON dossier validated against a typed schema, containing:

- Executive summary with top recommendation
- 3–5 ranked candidate markets, each with signal ratings (favorable/mixed/constrained, supportive/mixed/hostile, etc.) across all dimensions
- Detailed narrative per market covering power, community, incentives, hazards, connectivity, and recent activity
- Key risks identified per market
- Recommended next steps for each market (what to investigate next with human engagement)
- Markets to avoid with specific reasons
- Methodology note disclosing sources and limitations

Business Impact

Before SiteScope	After SiteScope
2–4 weeks of manual screening	2–3 minutes per screening run
Multiple consultants working in parallel	One API call with structured output
\$50K–\$150K in preliminary research fees	Pennies per run in API token costs
30+ candidate markets narrowed to 5 manually	Ranked shortlist with signal ratings in one pass
Information often stale by delivery date	Re-run anytime as the landscape shifts

The core value proposition is not cost savings alone — it is speed. In a market where utility interconnection timelines stretch to 4–10 years and community opposition can kill a project overnight, the ability to screen and re-screen markets in minutes rather than weeks represents a material competitive advantage. Development teams can:

- React to regulatory changes (moratoriums, new incentive programs) within hours
- Run multiple scenario analyses with different priority weightings in a single afternoon
- Screen emerging secondary and tertiary markets that traditional consultants may not cover
- Maintain always-current market intelligence rather than relying on quarterly reports
- Focus expensive human relationship capital only on markets that pass initial viability screening

Why Subconscious

This project demonstrates a use case that plays directly to the Subconscious platform’s core strengths:

Multi-hop tool use: Site screening requires chaining 20–30 searches across fragmented sources in a single inference. TIM’s recursive decomposition naturally breaks “screen 5 markets across 7 criteria” into subtasks, and TIMRUN’s context pruning ensures market #5 receives the same research depth as market #1 — no context rot.

Structured output: The answerFormat schema guarantees the agent returns typed JSON with signal enums, not unstructured text. This enables direct integration into dashboards, CRM workflows, and automated reporting pipelines.

Single API call: No multi-agent framework, no manual context engineering. One client.run() call with instructions and tools produces the complete dossier. Subconscious handles orchestration, tool execution, and context management.

Platform tools: The built-in web_search, news_search, google_search, and company_search tools map directly to the information sources site selectors rely on — industry reports, local news coverage, government sites, and company filings.

Technical Architecture

SiteScope is implemented as a two-service web application: a FastAPI backend orchestrating the Subconscious agent, and a React frontend rendering the structured dossier.

The backend exposes a simple REST API: POST /api/screen triggers an agent run and returns a run ID; GET /api/screen/{run_id} polls for completion status. The agent call uses the TIM-GPT engine with four platform tools and a structured answerFormat schema. Output is validated server-side against a Pydantic SiteScreeningDossier model before being returned to the client.

The frontend implements a three-state view machine (input → loading → dossier) with 2-second polling during the loading phase. The dossier view renders a signal matrix with color-coded dots (green/amber/red) for quick comparison across markets, with detailed drill-down into each market's assessment. The design uses a custom paper-feel palette with Fraunces (display), Instrument Sans (body), and JetBrains Mono (labels) typography — no external component library.

Two hardcoded mock dossiers support offline development and live demo scenarios without requiring agent API calls.

Limitations & Honest Scope

SiteScope is a screening tool, not a final recommendation engine. Important limitations include:

- **Power availability:** General grid capacity signals are researchable from public sources, but specific MW available at a given substation requires direct utility engagement.
- **Community sentiment:** News coverage captures public signals but not private negotiations, backroom deals, or relationships with key officials.
- **Gated data sources:** Proprietary utility GIS layers, ISO interconnection queue databases behind paywalls, and detailed FEMA flood mapping tools are not accessible via web search.
- **Parcel-level detail:** Specific land availability, title, easements, and Phase I environmental assessments require broker and consultant engagement.

These limitations are by design, not oversight. The agent explicitly discloses them in its methodology note and frames its output as a screening layer that identifies where to focus human engagement — not a replacement for that engagement.

Future Extensions

- Custom function tools for ISO interconnection queue data, FEMA flood maps, and state incentive databases
- Interactive map-based frontend for geographic visualization of candidate markets
- Saved screening history with alert monitoring for regulatory changes in tracked markets
- Webhook-based async processing for long-running deep research runs
- Multi-user collaboration with shared screening pipelines for development teams

Conclusion

The data center industry is in the early stages of the largest infrastructure buildout in a generation, constrained not by capital but by the ability to research and evaluate sites fast enough. Community opposition, power scarcity, and regulatory complexity have transformed site selection from a real estate exercise into a multi-dimensional intelligence problem.

SiteScope demonstrates that AI agents — specifically, Subconscious's TIM-GPT engine with multi-hop tool use and structured output — can compress the research synthesis layer of this problem from weeks to minutes. The agent doesn't replace human judgment or relationships; it ensures that when those expensive human resources are deployed, they're pointed at markets that are actually viable.

In a market where the most important approval may no longer come from investors or utilities, but from the communities asked to live alongside data centers, SiteScope helps you find the communities that will say yes.