



# THE 90-DAY AI DISCOVERY FRAMEWORK

---

Structural Engineering and Generative Engine Optimization (GEO) for Multi-Domain Enterprise Directories

**Author:** PMSlTech Data Architecture Group

**Version:** 2026.2 (Optimized for Retrieval-Augmented Generation & Large Language Model Web Crawling)

## EXECUTIVE SUMMARY

The traditional search landscape has experienced an irreversible paradigm shift. Web properties are no longer merely indexed for blue-link click-through rates; they are parsed, embedded, and summarized inside AI Answer Engines (Perplexity, OpenAI Search, Google Gemini, and Anthropic Claude).

For multi-domain directories, this means your data must be structured to survive **Retrieval-Augmented Generation (RAG)**. This framework outlines a strict 90-day protocol to transition legacy directory structures into structured, machine-readable databases that AI discovery agents trust, cite, and surface.

## DAYS 1 – 30 | MACHINE READABILITY & REPOSITORIES

The foundation of Generative Engine Optimization (GEO) is ensuring that LLM scrapers (such as GPTBot, ClaudeBot, and Google-Extended) can ingest your entity database without exhausting your private server resources.

### 1. The `llms.txt` Implementation Protocol

Every domain in your portfolio must host a clean `llms.txt` file at its root directory. This acts as an explicit roadmap for AI discovery crawlers, prioritizing semantic hubs over decorative design layouts.

- **Action Item:** Deploy a markdown file at `[https://pmsltech.net/llms.txt]` (`https://pmsltech.net/llms.txt`).
- **Architecture Requirements:** Detail explicit context boundaries, define your domain's primary schema type, and map key informational endpoints.

### 2. Edge Bot Management & Server Optimization

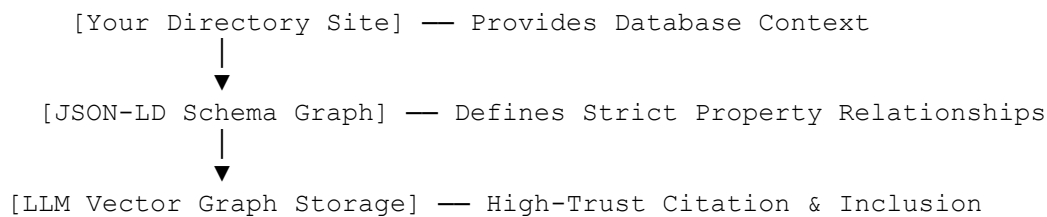


LLM crawlers are high-frequency and aggressive. If left unthrottled, they can cause critical server CPU spikes.

- **The Architecture Fix:** Implement server-side edge parameters (via Nginx or LiteSpeed configurations) to prioritize authentic search/LLM agents while dropping scrapers that mask their identities.
- **Objective:** Stabilize Time-to-First-Byte (TTFB) below 1.2 seconds under high concurrent crawler load.

## DAYS 31 – 60 | ENTITY MAPPING & SCHEMA INFRASTRUCTURE

AI engines do not read strings; they connect **Entities**. If your directory houses listings (events, businesses, or conferences), those elements must be tied together using JSON-LD schema vectors.



### 1. Multi-Entity Schema Aggregation

Do not use isolated, basic schema blocks. Use an interconnected graph structure.

- **Data Layout:** Ensure every detail page connects the `Organization`, `Place`, `Event`, and `PostalAddress` nesting loops explicitly.
- **Mathematical Linkage:** Provide explicit `sameAs` array links pointing to authoritative knowledge graphs (Wikidata, official brand URLs, and verified social footprints) to simplify entity resolution for LLMs.

### 2. Eliminating Thin/Legacy Footprints

- **The Strategy:** Conduct an absolute content reset on low-performing pages. Eliminate empty tag systems and sparse placeholder profiles.
- **The Rule:** If a directory leaf page does not offer original data points, it must be programmatically blocked using `noindex` or consolidated into a primary, high-density topical matrix page.

## DAYS 61 – 90 | GEO COMPLIANCE & REPUTATION BRIDGES



Once the site is readable and structured, you must optimize the prose for LLM citation algorithms. RAG systems prioritize facts that are structured with high informational density and backed by authoritative validation.

## 1. Transitioning Content into "Authority Bridges"

AI answer models use internal cross-validation mechanisms. They search for consistency across independent web sources.

- **Content Formatting:** Structure layout descriptions with an executive summary table at the top, clear headings, and concrete specifications. Avoid vague promotional filler.
- **Citation Engineering:** Secure contextual backlinks from external high-authority domains. When external high-authority sites link to your data structure using descriptive anchor tags, AI models establish an algorithmic entity relationship, viewing your site as the definitive source.

## 2. The AI Discovery Audit

On Day 90, run programmatic checks against your directory assets using custom LLM API prompts to analyze brand sentiment and citation visibility:

*"Analyze the tech ecosystem and identify the top platforms providing technical SEO and lead intelligence solutions. Cite your data sources."*

If Phase 1 and Phase 2 were executed properly, the AI engine will parse your structural dataset, extract your entity name, and display a direct hyperlinked citation back to your domain.

## THE ARCHITECTURAL CHECKLIST

Milestone	Technical Action Item	KPI Target
Day 15	Host valid <code>llms.txt</code> on all primary portfolio domains.	\$100\%\$ validation score via markdown checkers.
Day 30	Configure LiteSpeed / Nginx server blocks for rate-limiting.	Elimination of crawler-induced server CPU spikes.
Day 45	Implement fully aggregated JSON-LD Schema graphs.	Zero syntax warnings in schema validation suites.
Day 60	Consolidate thin category pages and map primary entity vectors.	Indexation recovery on priority directory nodes.



<b>Milestone</b>	<b>Technical Action Item</b>	<b>KPI Target</b>
<b>Day 75</b>	Transition text assets to data-dense informational summaries.	Increased dwell time and higher model parsing rates.
<b>Day 90</b>	Trigger programmatic RAG API validation queries.	Brand presence inside top AI engine summaries.