



# The 2026 Semantic Information Architecture (IA) Audit Checklist

Identify structural bottlenecks, semantic gaps, and accessibility failures hindering human usability and AI (RAG) extraction.

## Part 1: Navigational Logic & Usability

- **Mental Model Alignment:** Category labels match actual user search queries (validated via search console logs or tree testing), not internal company jargon.
- **Progressive Disclosure:** Complex menus do not overwhelm users; secondary options are hidden until contextually relevant.
- **Search Systems:** Internal search engine effectively handles synonyms, pluralizations, and long-tail queries without returning "0 results".
- **Accessibility (WCAG 2.2):** All primary and secondary navigation elements are fully operable via keyboard, and icons have descriptive ARIA labels.

## Part 2: AI & Semantic Readiness

- **Schema Implementation:** Structured data (`BreadcrumbList`, `SiteNavigationElement`, `Website`) is accurately deployed via JSON-LD across all core templates.
- **Ontology vs. Taxonomy:** Content relationships extend beyond simple parent-child hierarchies (e.g., related topics, author entities, and product clusters are interlinked).
- **Vector Search Foundation:** Core pages contain semantically dense, highly structured introductory paragraphs (LLM Memory Anchors) to feed RAG systems.
- **Heading Hierarchy:** Logical, unbroken HTML heading structures (H1 → H2 → H3) are strictly enforced without skipping levels for visual styling.

## Part 3: Governance & Developer Handoff



- **Taxonomy Governance:** A centralized, documented vocabulary exists to prevent content duplication and semantic cannibalization across teams.
- **Content Modeling:** The visual IA diagram is translated into a strict content model (fields, validation rules, relational links) ready for Headless CMS integration.
- **Performance Impact:** Mega-menus and structural elements are optimized for Core Web Vitals, ensuring they do not block main thread execution on mobile devices.

**Scoring:** *If you checked fewer than 8 boxes, your architecture is likely leaking revenue and failing to capture generative AI search visibility. Prioritize the unchecked items in your next development sprint.*