

How to NOT be Invisible to AI

a small museum website checklist

HOW TO BE VISIBLE - An AI Boosting Checklist

Here is a practical, lightweight check sheet designed specifically for small-to-medium museums. It focuses on low-effort, high-impact changes that ensure AI travel planners, LLMs, and search agents can accurately read, extract, and recommend your institution.

Phase 1: Lightweight Schema Markup (JSON-LD)

Schema is the hidden code you place in the `<head>` of your website. It acts as a direct data feed for AI and SEO, preventing models from having to "guess" your details.

1. Implement **Museum** (or **LocalBusiness**) Schema on the Homepage

- **The Tweak:** Add a JSON-LD script defining your official name, geographic coordinates, address, and primary contact info.
- **How it helps AI:** It establishes the museum as a verified, physical entity in the AI's "Knowledge Graph." When an AI agent is asked to "build an itinerary for historic sites in [City]," this schema anchors your museum to that specific location, making it a selectable destination.

2. Add **OpeningHoursSpecification**

- **The Tweak:** Explicitly define your operating hours, days closed, and seasonal holiday hours within your schema markup.
- **How it helps AI:** AI models frequently hallucinate museum hours based on outdated travel blogs. This schema provides a definitive, machine-readable schedule so AI assistants confidently tell users, "Yes, they are open on Tuesday from 10 AM to 5 PM."

3. Use **Event** Schema for Temporary Exhibitions

- **The Tweak:** Tag upcoming and current exhibitions with **Event** schema, specifically noting the `startDate` and `endDate`.
- **How it helps AI:** AI travel planners look for date overlaps. If a user asks an AI to plan a trip for October 12th-15th, the AI will scan for **Event** schemas matching those dates. Without this, your temporary exhibits will likely be ignored.

4. Implement **FAQPage** Schema

- **The Tweak:** Create a dedicated FAQ page (e.g., parking, accessibility, photography rules) and wrap it in FAQ schema.
- **How it helps AI:** LLMs using Retrieval-Augmented Generation (RAG) aggressively hunt for Question-and-Answer formats. By feeding the AI explicit FAQs, you guarantee it will quote your exact policies rather than guessing.

Phase 2: Content & Structural Tweaks (Semantic HTML)

AI crawlers "read" websites much like screen readers used by visually impaired users. If the underlying structure is messy, the AI will abandon the page.

5. De-Silo High-Value Text from iFrames and PDFs

- **The Tweak:** Move essential information (like ticketing instructions, event calendars, or primary historical texts) out of embedded third-party widgets or downloadable PDFs, and paste it directly onto the webpage as standard text.
- **How it helps AI:** Most AI web crawlers cannot "see" inside iFrames (like embedded Eventbrite calendars) and frequently skip heavy PDF files. Moving this data to plain HTML ensures the AI can actually read and index your schedule.

6. Enforce a Strict Heading Hierarchy (H1, H2, H3)

- **The Tweak:** Ensure every page has exactly one main title (H1), followed logically by subheadings (H2), and nested topics (H3). Never use headings just to make text look bigger.
- **How it helps AI:** AI models use headings to "chunk" information into context windows. A logical structure tells the AI exactly how different historical facts, ticket prices, and gallery names relate to one another.

7. Write Dense, Descriptive Alt Text for Key Artifacts

- **The Tweak:** Instead of naming an image `img_4599.jpg` or using generic alt text like "old painting," use rich descriptions: `Alt="19th-century oil painting of the town square by Jane Doe, featuring the original clock tower."`
- **How it helps AI:** Multi-modal AI models train on alt text. Rich descriptions allow your specific collection items to surface when users ask AI to find historical imagery or primary sources related to specific topics.

8. Replace "Click Here" with Descriptive Anchor Text

- **The Tweak:** Instead of writing "For our visitor map, click here," write "Download the [2024 Museum Visitor Map and Guide]."
- **How it helps AI:** Descriptive links give the AI crawler immediate context about the destination page before it even clicks, helping it map the relationship between your web pages.

Phase 3: Off-Page Trust Signals

AI agents need to verify that the information on your website matches the "real world" before they confidently recommend you as a trusted source.

9. Synchronize NAP Data (Name, Address, Phone)

- **The Tweak:** Ensure your exact business name, address, and phone number are written identically on your website, your Google Business Profile, Apple Maps, and local tourism boards.

- **How it helps AI:** AI models cross-reference data. If your website says "County History Museum" but Google Maps says "The Historical Society of [County]," the AI's confidence score drops, and it may omit you from recommendations to avoid giving the user a bad experience.

10. Connect to Wikidata / Wikipedia

- **The Tweak:** Ensure your museum has a basic, accurate Wikidata entry that links directly back to your official website.
- **How it helps AI:** Almost all major LLMs heavily weight Wikipedia and Wikidata as foundational truth. A linked Wikidata entity acts as a massive "trust signal" to the AI that your institution is a recognized, authoritative source of information.

HOW TO BE TRUSTED - A Scholarly AI Optimization Checklist

AI models evaluate academic trustworthiness through **provenance, structured metadata, and persistent identifiers**, rather than overall site size or text volume. While longer scholarly essays and images enrich a site, they must be formatted so AI crawlers and Retrieval-Augmented Generation (RAG) agents can verify their academic authority.

- **Images:** Multimodal AI can analyze visual features, but without rich text descriptions or manifest standards (like IIIF), AI cannot infer historical context from pixels alone.
- **Basic Info:** Standardized basic metadata (accession numbers, dates, creator, medium) is the single most critical factor for AI indexing.
- **Long Texts:** Deep essays provide excellent context for RAG systems, but unstructured prose without clear headings, citations, or curator credentials risks being misparsed or flagged as unverified.

Scholarly AI Optimization Checklist

1. Machine-Readable Metadata

- Tag object records with Schema.org entities ([VisualArtwork](#), [ArchiveComponent](#), [Article](#)).
- Embed Dublin Core fields: Title, Creator, Creation Date, Medium, Provenance, and Accession ID.
- Standardize terminology using authority files like the Getty Art & Architecture Thesaurus (AAT) or Union List of Artist Names (ULAN).
-

2. Institutional Authority & Attribution

- Attach staff names, curator credentials, or ORCID iDs to scholarly articles and object entries.
- Include an explicit "How to Cite This Object" block with stable permalinks or DOIs.
- Apply an explicit Open Access license (e.g., CC0 or CC-BY) in the page metadata to signal to AI crawlers that the data is authoritative and legally reusable.

3. RAG-Friendly Text Formatting

- Structure long essays with clear subheadings ([H2](#), [H3](#)) and a 2–3 sentence summary at the top.

- Move detailed research notes out of downloadable PDFs or closed databases and publish them directly on static HTML web pages.

4. Visual Context & Interoperability

- Provide 100–300 word descriptive text blocks for key artifacts, covering both physical appearance and historical significance.
- Implement IIIF (International Image Interoperability Framework) manifests for image assets to allow AI systems to reliably link visuals with metadata.