

Metadata health scorecard

Is your DAM ready for AI search?

AI search, including semantic and natural language search, works by interpreting meaning, not just matching keywords. That means your metadata needs to be more than complete. It needs to be structured, linguistically meaningful, and consistent enough for a model to reason across thousands of assets.

Use this scorecard to find out where you stand.

How to score:

0 = Not in place

1 = Partially in place

2 = Mostly in place

3 = Fully in place

1. Schema consistency and completeness

Maximum score: 15

How reliable is the information attached to your assets?

- ① We have a defined, documented metadata schema — not just inherited fields from a previous DAM migration
- ① Required fields are completed on upload, not retroactively or inconsistently
- ① The same asset type (e.g. video, banner, brand guide) uses the same field structure regardless of who uploaded it
- ① We use controlled vocabulary or picklist values for key fields, not open-text entries that vary by team or region
- ① Duplicate, outdated, or conflicting tags are actively identified and cleaned

Section score: / 15

Why this matters for AI search: AI models rely on consistent field values to build reliable associations. Free-text entries like "Q4 campaign UK" and "UK — Q4 2024 campaign" represent the same thing but look entirely different to a model. Controlled vocabulary removes this ambiguity at source.



2. Taxonomy depth and semantic clarity

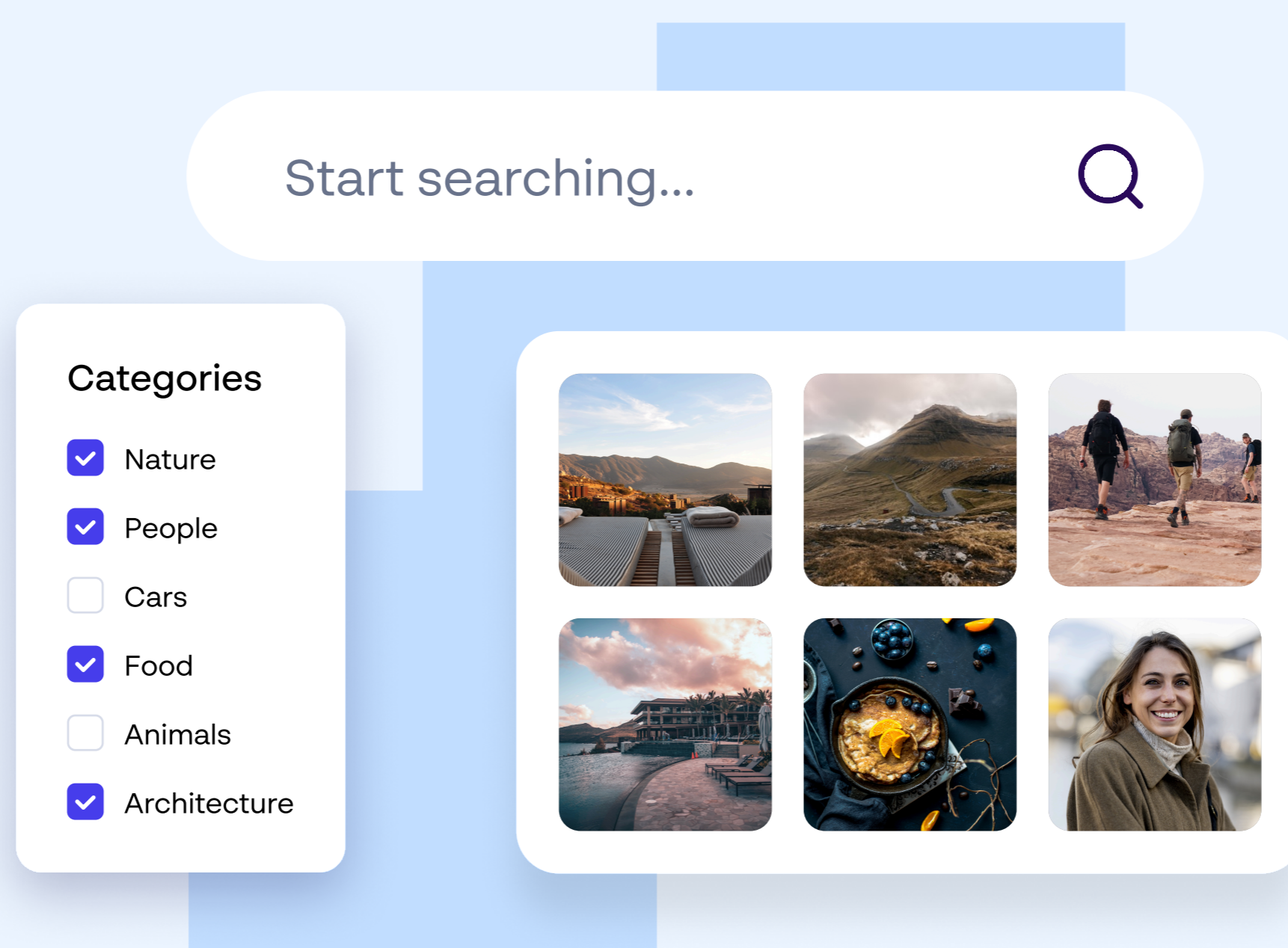
Maximum score: 15

Do your categories mean something — and mean the same thing to everyone?

- ② Our taxonomy is documented and version-controlled, not managed by institutional memory
- ② Categories and terms reflect how users search, not how assets are internally produced or filed
- ② Hierarchies are deep enough to be specific (not just "Digital > Social" but "Digital > Social > LinkedIn > Carousel > Brand Campaign")
- ② The same concept is not described with different terms across regions, languages, or teams
- ② Synonyms and alternative search terms are mapped to canonical taxonomy terms

Section score: / 15

Why this matters for AI search: Natural language search models try to map a user's intent to a concept in your system. If your taxonomy is shallow or inconsistent, the model has nothing reliable to map to. The richer and more consistently applied your taxonomy, the more accurately AI can interpret queries like "find me the latest campaign assets for financial services in EMEA."





3. Descriptive metadata quality

Maximum score: 15

Is the content of your metadata specific and meaningful — not just present?

- ➔ Asset descriptions go beyond filenames — they include context, subject matter, intended audience, and use case where relevant
- ➔ Image and video assets include alt text or scene descriptions that describe visual content in plain language
- ➔ Metadata values are specific enough to distinguish similar assets ("woman using laptop in office, blue background, formal tone" vs. "lifestyle office photo")
- ➔ Campaign, product, audience, and regional context is captured in structured fields, not buried in filenames or folder paths
- ➔ Teams do not rely on visual scanning or internal knowledge to identify the right version of an asset

Section score: / 15

Why this matters for AI search: Semantic search models read metadata the way a person reads a description. Vague or absent descriptions mean the model cannot surface the right asset from a natural language query. Specific, well-written values dramatically improve relevance.

4. Governance and metadata ownership

Maximum score: 15

Is metadata quality actively managed?

- ➔ There is a named owner responsible for metadata schema, taxonomy, and quality standards
- ➔ Written guidelines exist for how each asset type should be tagged, and those guidelines are accessible to anyone uploading
- ➔ New team members are trained on metadata standards before they upload assets
- ➔ Metadata quality is reviewed on a regular cycle and any gaps are surfaced and addressed, not just tolerated
- ➔ Governance balances control with practicality: standards are good enough that people follow them without workarounds

Section score: / 15

Why this matters for AI search: Even a well-structured taxonomy degrades without governance. AI search quality is only as good as the most recent batch of assets added to your DAM. If standards slip on upload, retrieval quality also gradually starts to slip, often invisibly until users stop trusting results.



5. AI and semantic search readiness

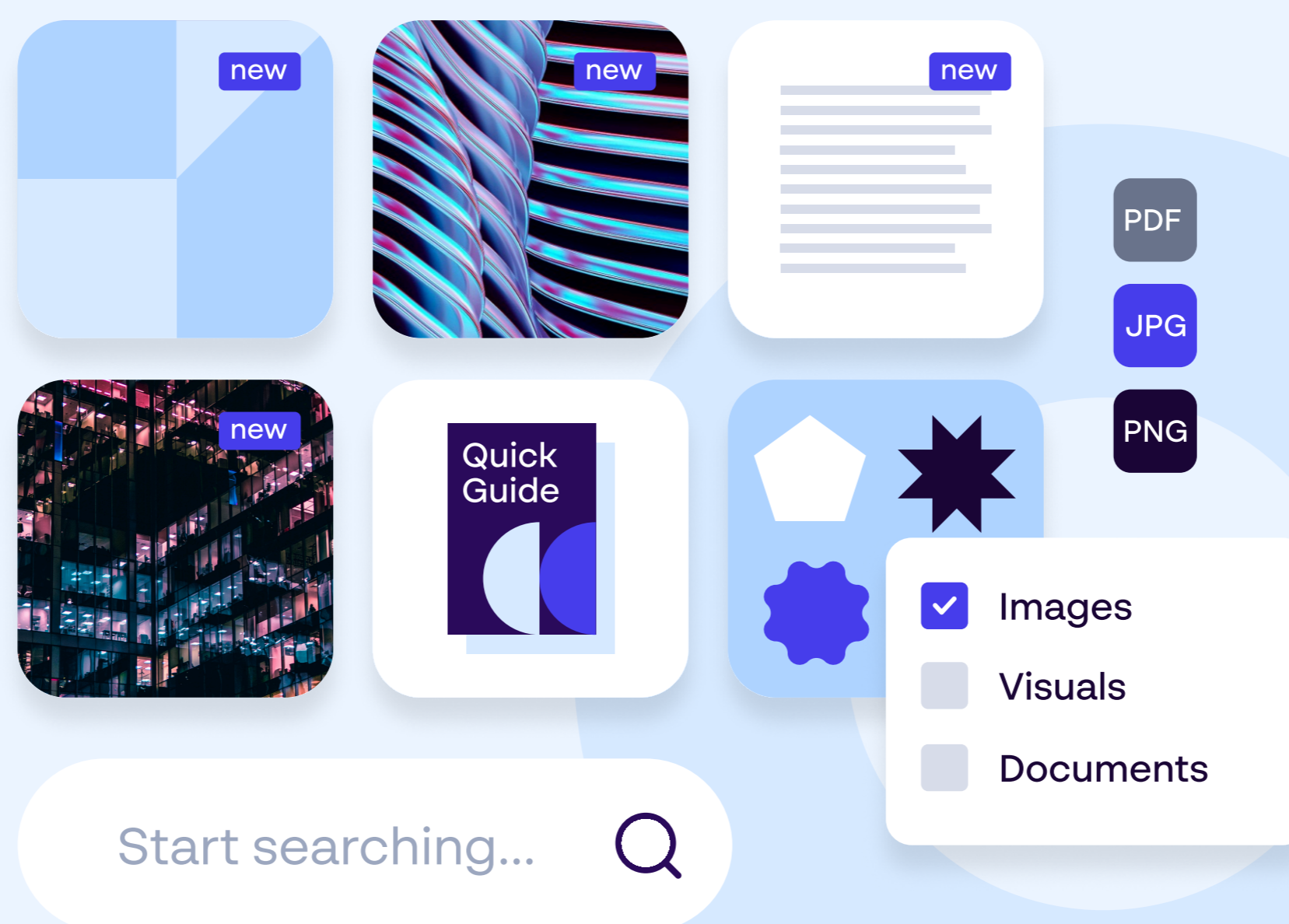
Maximum score: 15

Is your structure built to support how AI search actually works?

- ③ Metadata is stored in structured, queryable fields, not embedded in PDFs, file names, or unstructured notes
- ③ Key fields use consistent, predictable formats (dates, regions, product names, asset types) that a model can parse reliably
- ③ Our taxonomy and metadata schema can support multiple languages without creating structural conflicts
- ③ Metadata supports asset relationship context: for example, linking variants, source files, and campaign families
- ③ We have identified which fields are highest priority for AI enrichment or auto-tagging, and our schema supports that workflow

Section score: / 15

Why this matters for AI search: Natural language queries like "show me the approved hero images from the 2024 launch" require a DAM that knows what "approved," "hero image," and "2024 launch" each mean as well as where to find that information in a consistent field. Without structured, query-able metadata, AI search defaults to guessing.





Your total score □ / 75

Score	What it means
60–75	Strong foundation. Your DAM has the structure AI search needs. You can move to optimising enrichment quality and expanding coverage.
40–59	Useful foundations, but gaps remain. Inconsistency in schema, descriptive quality, or governance will limit what AI search can return reliably.
20–39	Structural risk. AI search may work in patches but will frustrate users at scale. Metadata remediation should come before any AI search rollout.
0–19	Structural risk. AI search may work in patches but will frustrate users at scale. Metadata remediation should come before any AI search rollout.

What to do next

A low score does not mean you need to rebuild everything. It means you need clarity on where the biggest gaps are first.

Book a metadata and taxonomy health check with Papirfly

Our experts can review your current setup, identify what is limiting search performance, and help you prioritize the changes that will have the biggest impact.

See what is holding your search back

Get a tailored review of your current metadata and taxonomy setup, with practical recommendations to prepare your DAM for AI search.

→ [Book your health check](#)

