



CONVENTION · BA-C-1

Version 1.0 · Stable

The Barkhausen Ladder

The Barkhausen Ladder is a nine-level maturity model (BL-0 through BL-8) describing an entity's readiness to be found, cited, and acted upon by AI assistants and answer engines.

DOCUMENT	BA-C-1
VERSION	1.0 (Stable)
EFFECTIVE	2026
SUBJECTS	Conventions & policy; Measurement & statistics
LICENSE	CC-BY-4.0
AUTHOR	Barkhausen AI

ABSTRACT

The Barkhausen Ladder is a nine-level maturity model (BL-0 through BL-8) describing an entity's readiness to be found, cited, and acted upon by AI assistants and answer engines. Each level is defined by three things: what the level means, the observable criteria that place an entity at it, and why it matters. The levels run from basic search hygiene (BL-0) through rendering and crawler access, structured data, content form, distribution, measurement, algorithmic optimization, and corpus presence to agent-readiness (BL-8). The Ladder is diagnostic, not prescriptive: it states conditions an assessment can verify, not techniques to apply. Levels are cumulative, and an entity is placed at the highest level whose criteria, and all lower levels', it satisfies. This convention uses normative keywords (MUST, SHOULD, MAY), cross-references the metric and sampling conventions (BA-C-2, BA-C-3) that define how visibility itself is measured, and closes with an assessment checklist for each level.

CONTENTS

1	Scope and normative language	4
2	About the name	5
3	BL-0 Search hygiene	5
4	BL-1 Rendering and crawler access	5
5	BL-2 Structured data and entities	6
6	BL-3 Content form	6
7	BL-4 Distribution and sources	7
8	BL-5 Measurement and statistics	7
9	BL-6 Algorithmic optimization (concept level)	8
10	BL-7 Corpus presence (concept level)	8
11	BL-8 Agent-ready	9
12	Using the Ladder	9
13	Conformance	10
14	Limitations	11
15	Appendix: assessment checklist	11

The Barkhausen Ladder is a nine-level maturity model for an entity's readiness to be found, cited, and acted upon by AI assistants and answer engines. It runs from BL-0, basic search hygiene, to BL-8, agent-readiness. Each level is defined by three things: what the level means, the observable criteria that place an entity at it, and why the level matters for AI visibility. The term *entity* is used broadly for whatever is being made visible — an organization, a product, a publication, or a person.

The Ladder is diagnostic rather than prescriptive. It states what an assessment can verify about an entity, not the techniques an entity should apply to advance. The techniques are the province of practitioners, and much of what circulates as technique does not survive measurement (BL-6). Reading the Ladder as a checklist of tricks misuses it.

The levels are cumulative. Each level assumes the ones below it. An entity that has not resolved its crawler access (BL-1) cannot rely on its content form (BL-3), because content a crawler never retrieves cannot be read. An assessment therefore places an entity at the highest level whose observable criteria it satisfies and whose lower levels are also satisfied (see Using the Ladder).

This convention defines the levels only. How visibility itself is measured — the metrics Visibility Probability (VP), Share of Voice (SoV), and Discovery Depth (DD), and the sampling requirements behind them — is specified in the metric and sampling conventions, BA-C-2 and BA-C-3. BL-5 is the level at which an entity adopts those requirements.

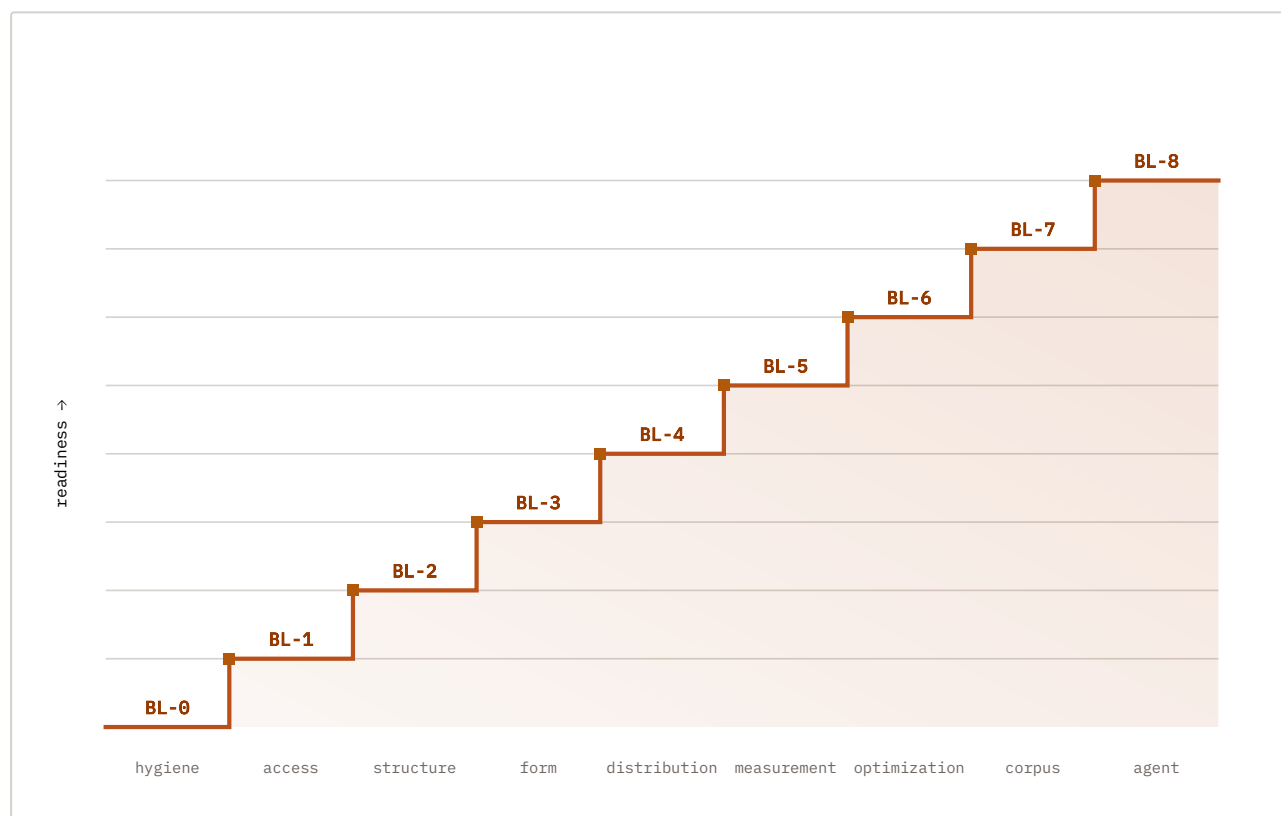


Figure 1. The Barkhausen Ladder: nine cumulative levels of AI-visibility maturity, from search hygiene (BL-0) to agent-readiness (BL-8). Each level assumes the ones below it; an assessment places an entity at the highest level whose criteria — and all lower levels' — are satisfied.

Barkhausen AI · BA-C-1

1 Scope and normative language

This convention applies to any entity seeking to understand its visibility in AI assistants and answer engines, and to any assessment that reports a maturity level for such an entity.

The keywords **MUST**, **MUST NOT**, **SHOULD**, **SHOULD NOT**, and **MAY** are normative and follow the conventions recorded in the conventions process (see /conventions/process/). **MUST** marks a criterion without which a level is not reached; **SHOULD** marks a criterion expected at the level but subject to documented exceptions; **MAY** marks a criterion that strengthens a placement without being required for it.

Observable criteria are conditions a third party can verify from public evidence: the entity's own pages and responses, public crawler and engine documentation, public indices, and repeated sampling of answer engines. A level defined by observable criteria can be assessed without the entity's cooperation and without access to any engine's internals.

Two levels, BL-6 and BL-7, are stated at concept level only. They describe what the level means and why it matters; they do not prescribe methods. This is deliberate: the practices at these levels are the subject of active research and are easily misrepresented as guaranteed outcomes.

2 About the name

The Barkhausen effect, reported by Heinrich Barkhausen in 1919, is the observation that a ferromagnetic material placed in a continuously increasing magnetic field does not magnetize smoothly. It advances in small, discrete jumps as magnetic domains reorient, audible as a crackle when the signal is amplified. AI visibility behaves in a comparable way: continuous effort produces change that arrives in discrete, detectable jumps rather than along a smooth curve. The Ladder names the levels of readiness; the companion notion of a *visibility jump* — a change large enough and sustained enough to be distinguished from sampling noise — is formalized as the Barkhausen Criterion in the sampling convention (BA-C-3). The name records the analogy and carries no other claim.

3 BL-0 Search hygiene

An entity at BL-0 satisfies the baseline conditions of conventional search indexing. This is the floor of the Ladder: an entity that is not reliably indexable by general search engines is not reliably retrievable by the systems that sit on top of them.

An entity at BL-0 **MUST** be indexable — its important pages return successful responses, are not excluded by robots directives or noindex meta tags, and declare canonical URLs that resolve consistently. It **MUST** expose a valid, reachable XML sitemap whose entries resolve, and **MUST** present coherent page metadata (title, description, and a language declaration). It **SHOULD** maintain redirect health, with no chains or loops, carry no broken internal links, and meet baseline performance and stability thresholds such as Core Web Vitals. Duplicate and near-duplicate URLs **SHOULD** be consolidated by canonicalization.

The level matters because answer engines that ground their responses in live retrieval draw from conventional search indices, or from their own crawls of the same web. Basic indexation failures — a page excluded by a stray directive, a sitemap that returns an error, a canonical tag that points to the wrong URL — remove content from that pool before any AI-specific consideration applies. BL-0 is unglamorous and non-negotiable.

4 BL-1 Rendering and crawler access

An entity at BL-1 has content that is actually retrievable, in full, by the crawlers that feed answer engines, and has made a deliberate decision about which crawler classes it admits.

An entity at BL-1 MUST serve its primary content in server-rendered or otherwise complete HTML, such that the content is present in the initial response without client-side JavaScript execution. Public analyses indicate that AI crawlers largely do not execute JavaScript [3]; content that exists only after client-side rendering is, in practice, invisible to them. The entity MUST operate a robots.txt policy that distinguishes crawler classes — training crawlers, retrieval and index crawlers, and user-triggered fetchers — rather than a single undifferentiated rule, and its policy MUST match its intent. A common failure is blocking a retrieval crawler while meaning to block only a training crawler, which removes the entity from answer citations. The entity SHOULD support timely discovery through instant-indexing submission mechanisms such as IndexNow where its target engines consume them. An entity MAY publish an llms.txt file as a low-cost hygiene item; llms.txt is a proposal for exposing curated content to language models [4], and no major engine has publicly confirmed consuming it during answer generation, so its presence MUST NOT be treated as an efficacy guarantee.

The level matters because both retrieval and corpus construction begin with a fetch. If the fetch returns an empty JavaScript shell, or a misconfigured directive turns a retrieval crawler away, everything above this level is unreachable. BL-1 is where a large share of otherwise capable entities silently fail.

5 BL-2 Structured data and entities

An entity at BL-2 is machine-legible as a consistent, disambiguated entity, and its key facts are marked up while also existing in the visible text of the page.

An entity at BL-2 MUST express its key facts as schema.org structured data (JSON-LD), using the types appropriate to what it is. Its entity representation MUST be consistent across pages and across languages — the same name, identifiers, and relationships — and SHOULD assert disambiguating links, for example sameAs references to authoritative profiles and presence in public knowledge graphs. Critically, every fact expressed in markup MUST also be present in the visible, rendered text of the page. Structured data is necessary but not sufficient: it aids disambiguation and traditional search, but current answer engines read the page, not the markup, and a fact that exists only in JSON-LD is not reliably available to them.

The level matters because an answer engine must decide which real-world entity a page is about before it can attribute a fact to that entity. Inconsistent names across pages or languages, or the absence of any knowledge-graph anchor, produce entity confusion: the entity's facts are attributed to a similarly named other, or to no one. Markup accelerates disambiguation; visible text carries the facts. An entity that has one without the other is not at BL-2.

6 BL-3 Content form

An entity at BL-3 writes its content in the forms that answer engines can excerpt, cite, and synthesize.

Content at BL-3 SHOULD be organized into self-contained, excerpt-safe passages, each passage answering one specific question completely, so that it remains correct and useful when quoted in isolation. It SHOULD support claims with statistics, direct quotations, and citations to sources. A controlled experiment over a benchmark of roughly 10,000 queries found that adding quotations, statistics, and cited sources measurably raised a passage's inclusion in generated answers — the strongest methods improving a position-adjusted visibility metric by up to roughly 40% — while a traditional keyword-stuffing tactic produced no comparable gain [1]. Content SHOULD use comparison, FAQ, and list structures where the material genuinely enumerates, and SHOULD cover the sub-queries a topic fans out into rather than a single phrasing. Each page SHOULD maintain single-language purity, because

mixed-language pages score poorly against the language filters used in both retrieval indexing and corpus construction.

The level matters because answer engines decompose a question into sub-questions, retrieve passages for each, and compose an answer from the passages they judge most self-sufficient and best supported. A passage that only makes sense in the context of its surrounding page, or that asserts without evidence, is a weaker candidate for inclusion. BL-3 is about writing for extraction, not for a single reader scrolling a page.

7 BL-4 Distribution and sources

An entity at BL-4 is present in the third-party and community sources that answer engines cite, not only on its own domain.

An entity at BL-4 SHOULD have a substantive, accurate presence in the external sources that answer engines draw on disproportionately — community platforms, reputable third-party publications, and well-maintained directories relevant to its domain. It SHOULD accumulate authority signals in the form of citations and links from independent, established sources. Public analyses observe that presence on certain third-party platforms is associated with higher AI-citation frequency; one analysis of roughly 75,000 brands reported an association between a brand’s presence on a large video platform and its visibility in AI answers [8]. Such observational findings establish association, not causation, and the specific platforms that engines favor shift over time.

The level matters because, when an answer engine grounds a claim, it frequently cites a third party rather than the entity’s own site, and a self-published assertion carries less evidential weight than the same fact reported independently. An entity confined to its own domain depends entirely on that domain being retrieved and trusted. Distribution across the sources the engines already favor widens the number of paths by which the entity’s facts can enter an answer.

8 BL-5 Measurement and statistics

An entity at BL-5 measures its own AI visibility as a statistical quantity — a distribution estimated from repeated sampling — rather than inferring it from individual answers.

An entity at BL-5 MUST estimate its visibility from repeated sampling and report it with uncertainty. Concretely, it MUST express visibility as Visibility Probability (VP) — the probability that the entity is mentioned in an answer to a given query, estimated as a proportion $\hat{p}$ with a stated confidence interval (for a proportion, a 95% interval requires roughly $n \approx 1.96^2 p(1-p)/E^2$ observations for margin of error E) — and SHOULD track Share of Voice (SoV) relative to comparators and Discovery Depth (DD), the constraint depth at which it enters a recommendation set. It MUST monitor change over time and evaluate any claimed improvement against the four conditions of the Barkhausen Criterion (BA-C-3): the jump must be statistically significant against proper intervals, sustained across consecutive windows, not attributable to a flagged engine change, and assessed with multiplicity control — not a single favorable sample. The metric definitions and sampling requirements are specified in BA-C-2 and BA-C-3.

The level matters because a single answer from an assistant is one draw from a distribution that varies with phrasing, session, time, and the engine’s own nondeterminism. A screenshot demonstrates that an outcome is possible, not that it is probable. An entity that treats a favorable screenshot as evidence of visibility, or an unfavorable one as evidence of its loss, is measuring noise. An entity is not at BL-5 until it measures visibility as a distribution with confidence intervals, tracks it over time, and applies a change

threshold. Everything below BL-5 can be done blind; from BL-5 on, an entity can tell whether its actions changed anything.

9 BL-6 Algorithmic optimization (concept level)

An entity at BL-6 treats content optimization as a measured, iterated loop rather than a fixed list of tactics: candidate changes are proposed, evaluated against a visibility signal, and kept when they measurably work. This level is defined at concept level only; this convention does not prescribe methods.

An assessment places an entity at BL-6 when the entity's optimization is evidence-driven and iterative rather than heuristic-driven and static. The distinguishing marker is that the entity can say which of its changes moved a measured metric and by how much, rather than citing a list of practices believed to help. Public research supports the distinction: an evaluation that compared a set of hand-written optimization heuristics against an iterative, measured optimization procedure found that the iterative procedure surpassed the fixed heuristics [2]. No specific method is endorsed here.

The level matters because the field is full of confidently asserted practices that do not survive measurement. An entity that applies a fixed checklist cannot distinguish the changes that helped from those that did nothing or hurt, and it inherits whichever folk heuristics are in circulation. Anchoring optimization to BL-5 measurement replaces belief with evidence. Because the loop's methods are contested and easily oversold, this convention states only what the level is and why it matters, not how to run the loop.

10 BL-7 Corpus presence (concept level)

An entity at BL-7 is present in the openly crawlable web that feeds the training corpora behind AI models — not only in the live index that feeds retrieval. This level is defined at concept level only.

An assessment places an entity at BL-7 when its content is present, in a durable and crawlable form, in the open web that supplies public training corpora. The shared upstream for most open corpora is Common Crawl: the highest-quality open datasets are derived from it, and it is the common raw material from which many others are filtered [5]. Presence in that upstream is governed largely by the link graph; public analysis of Common Crawl describes its crawl selection as governed substantially by domain-level link centrality — harmonic centrality in the host-level web graph — which influences whether, and how deeply, a domain is crawled [6]. An entity with little inbound linking from established domains may not enter the corpus pipeline at all. This level is stated conceptually and prescribes no technique.

Presence in a training corpus does not imply the model retains or reproduces the content; no causal claim is made. BL-7 describes an upstream condition — being present in the crawlable web that feeds corpora — that is itself observable. Whether a given fact is thereby memorized or surfaced by a deployed model is a separate, downstream question this convention does not assert.

The level matters because retrieval visibility (BL-1 through BL-4) and parametric, or closed-book, visibility are different channels on different time scales. Retrieval visibility can change within days; corpus presence is slow, tied to model training cycles, and hard for a competitor to reverse once established. An entity invisible to the corpus pipeline forgoes the slower channel entirely. The distinction between the two channels is why the Ladder separates BL-4 from BL-7.

11 BL-8 Agent-ready

An entity at BL-8 exposes machine-actionable endpoints so that AI agents can invoke it — query its data, or act on it — rather than merely mention it in prose.

An entity at BL-8 exposes structured, queryable interfaces to its data and functions so that an agent can discover and call them programmatically. An assessment places an entity at BL-8 only when at least one machine-actionable interface to the entity's data or functions is publicly discoverable and documented — this condition **MUST** hold for the placement. Which protocol implements the interface is open: the entity **MAY** expose it through any emerging agent protocol, and no specific one is required, because the protocol landscape is still forming. The Model Context Protocol is one public, open example of such a specification for connecting AI systems to external tools and data [7]. The mark of a BL-8 entity is that an agent can obtain a current, structured answer from it directly, and where applicable take an action, without a human intermediary parsing a web page.

The level matters because the lower levels concern being found and cited in text, while BL-8 concerns being usable. As assistants move from answering questions to completing tasks, the entities that expose reliable, structured, invokable interfaces can participate in agent workflows — being acted upon, not only described. This is the top of the Ladder because it presupposes the rest: an agent must be able to find, disambiguate, and trust an entity before invoking it.

12 Using the Ladder

An assessment places an entity at a single level: the highest level whose observable criteria the entity satisfies and for which every lower level is also satisfied. The Ladder is cumulative because the levels depend on one another — content form (BL-3) does not help if crawlers cannot retrieve the content (BL-1), and measurement (BL-5) presupposes that there is retrievable, well-formed content to measure. An entity that satisfies BL-3 but fails BL-1 is placed at BL-0, with the specific failure noted; the placement records the binding constraint, not the highest isolated achievement.

LEVEL	NAME	THE QUESTION IT ANSWERS	ASSESSABLE FROM PUBLIC EVIDENCE
BL-0	Search hygiene	Is the entity reliably indexable at all?	Yes
BL-1	Rendering and crawler access	Can AI crawlers actually retrieve the content?	Yes
BL-2	Structured data and entities	Is the entity legible as a consistent, disambiguated entity?	Yes
BL-3	Content form	Is the content written to be excerpted and cited?	Yes
BL-4	Distribution and sources	Is the entity present in the sources engines cite?	Yes
BL-5	Measurement and statistics	Does the entity measure visibility as a distribution?	Inferred
BL-6	Algorithmic optimization	Is optimization an iterated, measured loop?	Inferred
BL-7	Corpus presence	Is the entity in the corpus-feeding open web?	Partly, via public indices
BL-8	Agent-ready	Can agents invoke the entity, not just mention it?	Yes, endpoints are public

A placement is a snapshot against observable criteria at a point in time, not a score or a ranking. Two entities at the same level may differ in every other respect. Levels BL-0 through BL-4 are assessable entirely from public evidence — the entity’s pages, public crawler and engine documentation, public indices — and are the levels a sector-wide survey can score at scale. BL-5 and above concern what the entity does (measurement, iterated optimization) and which channel it has entered (corpus presence, agent interfaces), which an external assessment infers from observable proxies and states with appropriate uncertainty.

The value of a level is diagnostic: it identifies the binding constraint. An entity assessed at BL-1 learns that its next return on effort is in rendering and crawler access, not in content or measurement, however capable those might already be. The Ladder orders the constraints so that effort is spent where it is currently limited.

13 Conformance

A Ladder placement is reported in conformance with this convention when it states the level assigned and the basis on which it was assigned, so that a reader can check the placement against the criteria. A conforming placement **MUST** state the assessment date, because the observable criteria and the engine behavior behind them are perishable (see Limitations). It **MUST** state the evidence basis for each level it claims to have checked — the public evidence examined and, for BL-5 and above, whether the level was assessed directly or inferred from observable proxies. It **MUST** apply the highest-satisfied-level rule: the entity is placed at the highest level whose criteria, and all lower levels’, are satisfied, and where a lower level fails, the binding constraint is named rather than the highest isolated level reached.

A conforming placement **SHOULD** be cited in a form that carries the convention and the version under which it was made — for example, “assessed at BL-3 under BA-C-1 v1.0, 2026-07”. Because the criteria are

versioned, a placement is interpretable only against the version it names.

14 Limitations

The Ladder describes readiness, not outcome. A high placement indicates that the conditions for visibility are in place; it does not guarantee that any given engine will mention or cite the entity. Visibility is measured (BL-5), never inferred from the level.

The observable criteria reference a moving ecosystem. Crawler behavior, engine grounding, structured-data consumption, and agent protocols change without notice; criteria that discriminate today may not next year. This convention is versioned for that reason, and specific criteria should be read against the reference material current at assessment time. Criteria and results tied to particular engines should be assumed perishable.

Levels are ordinal, not interval. The distance from BL-0 to BL-1 is not comparable to the distance from BL-4 to BL-5, and a level is not a quantity to be averaged casually across unlike entities. Sector aggregates such as “a sector averages BL-2” are useful summaries, but they rest on this ordinal caveat.

BL-6 and BL-7 are stated at concept level and are the least externally verifiable. An external assessment infers them from proxies and should state that inference’s uncertainty; a claim to be at BL-7 is a claim about an upstream condition, subject to the honesty boundary stated in that level.

Cumulative placement can understate an entity. An entity with advanced content and measurement but a single BL-1 misconfiguration is placed low. That is correct as a diagnosis of the binding constraint, but it should not be read as a judgment of the entity’s overall sophistication.

The Ladder is one framework among possible ones. It orders constraints in the sequence that has held for current answer engines; a different engine architecture could reorder them. The framework earns its place by being useful and falsifiable, not by being the only way to describe the terrain.

15 Appendix: assessment checklist

This appendix distills the criteria stated above into a per-level checklist. Each row lists the level, the criteria an assessment checks — restated from the level’s own text, adding no new requirement — and how each is verified from public evidence. A sector-wide survey scores BL-0 through BL-4 from public evidence directly; BL-5 and BL-6 are inferred from observable proxies, BL-7 is partly assessable via public web indices, and BL-8 is assessable from public endpoints, consistent with the assessability column in Using the Ladder. Normative keywords in the table carry the strength each level assigns them; BL-6 and BL-7 are stated at concept level and have no MUST or SHOULD criteria to check.

LEVEL	CRITERIA TO CHECK (DISTILLED FROM THE LEVEL)	HOW IT IS VERIFIED FROM PUBLIC EVIDENCE
BL-0 Search hygiene	Important pages return successful responses, are not excluded by robots or noindex, and declare canonical URLs that resolve consistently (MUST); a valid, reachable XML sitemap whose entries resolve (MUST); coherent page metadata — title, description, and language declaration (MUST).	Fetch key pages and the declared sitemap; inspect status codes, robots.txt and noindex directives, rel=canonical targets, and the title, description, and language attributes.
BL-1 Rendering and crawler access	Primary content is present in the initial HTML response without client-side JavaScript (MUST); robots.txt distinguishes crawler classes — training, retrieval, user-triggered — and matches stated intent (MUST).	Fetch each page with a non-JS client and diff the primary content against the rendered page; read robots.txt and compare the per-class rules against the entity's stated access intent.
BL-2 Structured data and entities	Key facts are expressed as schema.org JSON-LD (MUST); the entity representation is consistent across pages and languages (MUST); every marked-up fact also appears in the visible text (MUST).	Extract JSON-LD from key pages; check name, identifier, and relationship consistency across pages and languages; diff each marked-up fact against the visible, rendered text.
BL-3 Content form	Passages are self-contained and excerpt-safe (SHOULD); claims carry statistics, quotations, and citations (SHOULD); enumerable material uses comparison, FAQ, or list structures (SHOULD); each page keeps single-language purity (SHOULD).	Read sampled passages for standalone correctness and supporting evidence; inspect page structure and per-page language.
BL-4 Distribution and sources	Substantive, accurate presence in the third-party and community sources engines cite (SHOULD); independent authority signals — citations and links from established sources (SHOULD).	Search the external sources engines draw on for the entity; check independent citations and inbound links from established domains.
BL-5 Measurement and statistics	Visibility estimated from repeated sampling and reported as VP with a confidence interval (MUST); change monitored over time and evaluated against the Barkhausen Criterion (MUST); SoV and DD tracked (SHOULD).	Inferred from the entity's published measurement disclosures; a survey infers from observable proxies and states the uncertainty.
BL-6 Algorithmic optimization	Optimization is an iterated, measured loop, and the entity can attribute a measured metric's movement to specific changes (concept level).	Inferred from proxies; not directly assessable from public evidence.
BL-7 Corpus presence	Content is present, in a durable and crawlable form, in the open web that feeds public training corpora (concept level).	Partly assessable via public web indices and link-graph presence; otherwise inferred, with the uncertainty stated.
BL-8 Agent-ready	At least one machine-actionable interface to the entity's data or functions is publicly discoverable and documented (MUST for placement); the implementing protocol is open (MAY).	Discover and fetch the entity's documented machine-actionable interface, and confirm an agent could call it programmatically.

REFERENCES

1. Aggarwal et al., KDD 2024. *GEO: Generative Engine Optimization* (2024). <https://arxiv.org/abs/2311.09735> Accessed 2026-07-08. [archived]

2. Bagga, Farias, Korkotashvili, Peng, Wu; arXiv:2511.20867. *E-GEO: A Testbed for Generative Engine Optimization in E-Commerce* (2025). <https://arxiv.org/abs/2511.20867> Accessed 2026-07-08. [archived]
3. Vercel engineering (with Merj). *The rise of the AI crawler* (2025). <https://vercel.com/blog/the-rise-of-the-ai-crawler> Accessed 2026-07-08. [archived]
4. llmstxt.org (llms.txt proposal). *The /llms.txt file* (2024). <https://llmstxt.org> Accessed 2026-07-08. [archived]
5. Penedo et al., NeurIPS 2024 Datasets and Benchmarks Track. *The FineWeb Datasets: Decanting the Web for the Finest Text Data at Scale* (2024). <https://arxiv.org/abs/2406.17557> Accessed 2026-07-08. [archived]
6. Baack; Proc. 2024 ACM Conference on Fairness, Accountability, and Transparency (FAccT '24), pp. 2199–2208. *A Critical Analysis of the Largest Source for Generative AI Training Data: Common Crawl* (2024). <https://doi.org/10.1145/3630106.3659033> Accessed 2026-07-08.
7. Model Context Protocol (open specification). *Model Context Protocol* (2024). <https://modelcontextprotocol.io> Accessed 2026-07-08.
8. Ahrefs (study of ~75,000 brands). *Top Brand Visibility Factors in ChatGPT, AI Mode, and AI Overviews (75k Brands Studied)* (2026). <https://ahrefs.com/blog/ai-brand-visibility-correlations/> Accessed 2026-07-08. [archived]

HOW TO CITE

Barkhausen AI (2026). The Barkhausen Ladder. <https://barkhausen.ai/conventions/barkhausen-ladder/>

Published under the Creative Commons Attribution 4.0 International (CC-BY-4.0).



Barkhausen AI · Est. MMXXVI · *Every leap begins as a crackle.*