Barkhausen AI

CONVENTION · BA-C-7

Version 1.0 · Stable

Terminology for AI-visibility measurement

Measurements of AI visibility are comparable only when the words that name them are stable: two studies that both report "visibility" compare nothing if one counts entity mentions and the other counts source links.

“Every leap begins as a crackle.”

DOCUMENT	BA-C-7
VERSION	1.0 (Stable)
EFFECTIVE	2026
SUBJECTS	Conventions & policy
LICENSE	CC-BY-4.0
AUTHOR	Barkhausen AI

ABSTRACT

Measurements of AI visibility are comparable only when the words that name them are stable: two studies that both report "visibility" compare nothing if one counts entity mentions and the other counts source links. This convention fixes the field-level vocabulary the rest of the series relies on and that no other convention owns — AI visibility, the phenomenon; answer engine and AI assistant, the systems; generative engine optimization, the practice; retrieval visibility versus parametric, closed-book visibility, two causal paths; mention and citation, two distinct observables; information need, query, and phrasing, three levels of a question; and recommendation set. It states why one canonical term per concept is a precondition for comparability, gives usage rules including first-use linking and a deprecated-synonym table, and defines conformance as using the canonical terms or explicitly mapping one's own. It claims authority over usage within this series only, not over the field's language.

CONTENTS

The measurements this series specifies are comparable only if the words they use are stable. Two studies that both report "visibility" compare nothing if one counts the times an entity is named in an answer and the other counts the times an answer links to the entity's site. This convention fixes the field-level vocabulary the rest of the series relies on: the name of the phenomenon, the systems that produce it, the observable units a measurement counts, and the three levels at which a question is described. It defines only the terms no other convention owns; metric names (BA-C-2), sampling terms (BA-C-3), the maturity levels (BA-C-1), and the crawler classes (BA-C-6) are defined where they are used and referenced here by owning convention rather than redefined.

The glossary rendered from this series aggregates every convention's definitions; this is the one whose product is those entries rather than requirements. Its body groups them into the concepts they belong to and states the rules for using them.

Normative note. The key words MUST, MUST NOT, SHOULD, and MAY carry their established normative meanings. Here they govern usage — how a conforming publication names concepts and maps any term of its own to the canonical one. They place no obligation on how anyone outside this series speaks, and express no claim of authority over the field's vocabulary.

1 1. Why a fixed vocabulary

Measurement comparability is the reason. A number is interpretable only against a defined quantity, and a defined quantity is named. When one word carries two meanings, a reader cannot tell which quantity a figure estimates, and two figures that share the word cannot be compared or pooled — the standing example being "AI citation rate," used both for how often an entity is named in answers and for how often an answer links to it, two observables that move independently (§4). Fixing one term per concept removes that ambiguity at the source, so the disclosure requirements of BA-C-4 and the reporting checklist of BA-C-5 have stable referents to require.

The vocabulary is descriptive, not proprietary. Where the field has an accepted term, this convention adopts it rather than coining a replacement; where a term is genuinely ambiguous, it records the ambiguity and points to the term that resolves it. The aim is one name per concept within this series.

2 2. The phenomenon and the systems

AI visibility is the phenomenon this series measures: the degree to which an entity is present in, and can be surfaced by, the answers that answer engines generate. It is estimated from repeated observation, not read from a single answer, and is quantified by the metrics of BA-C-2.

The systems that generate those answers are **answer engines**. The term names the function — responding to a natural-language question with a synthesized answer rather than only a ranked list of links — and is used throughout this series when the object under measurement is that answer-generating function. **AI assistant** names the same class from the user's side: the conversational, task-oriented product a person interacts with. A document uses answer engine when it means the measured system and AI assistant when it means the user-facing surface, and introduces no third term for either.

Generative engine optimization (**GEO**) is the practice of improving an entity's AI visibility — the activity, as distinct from AI visibility, the phenomenon it acts on. The term is adopted here as the field's accepted name; it originates in a 2024 study that introduced both the term and an associated benchmark [1], and this convention takes it as given rather than as its own coinage. This series specifies how to *measure* visibility; it does not publish GEO techniques.

3 3. Retrieval visibility and parametric visibility

Two distinct causal paths make an entity visible, and conflating them makes a measurement uninterpretable. **Retrieval visibility** is visibility that arises because an engine retrieves content — the entity's own pages, or third-party pages that name it — at answer time and grounds its answer on that content. It depends on crawler access and index inclusion, whose classes are defined in BA-C-6, and on the content being present and extractable.

Parametric (**closed-book**) **visibility** is visibility that arises from what a model internalized in its parameters during training, surfaced without any answer-time lookup. It is "closed-book" because the answer draws on the model's weights rather than on a retrieved document. Presence in a training corpus does not imply the model retains or reproduces the content; no causal claim is made. The distinction is load-bearing because the two paths respond to different interventions and change on different time scales — a retrieval-visible entity moves as an index updates, while parametric visibility shifts only when a model is retrained — so a study that does not separate them can attribute a change to the wrong cause.

4 4. The observables: mention and citation

A **mention** is the observable unit of AI visibility: an instance in which an entity is named, or unambiguously referred to, in the text of an answer. Whether a given span of text counts as a mention is settled by the detection method a study discloses under BA-C-2 §6; this convention fixes the term, not the detector. Visibility Probability (BA-C-2) is, at bottom, the probability that an answer contains at least one mention.

A **citation** is a source attribution an answer provides — a link or an explicit reference to a specific document. It is a different observable from a mention, and the two do not coincide. An entity can be mentioned without being cited — named in an answer's prose with no link — or cited without being mentioned, when an answer links to a document whose visible text never shows the entity's name. Because the two are distinct and move independently, any rate reported about an entity MUST state which it counts; "citation rate" and "mention rate" are not interchangeable, and an unqualified "AI citation rate" is non-conforming (see §7).

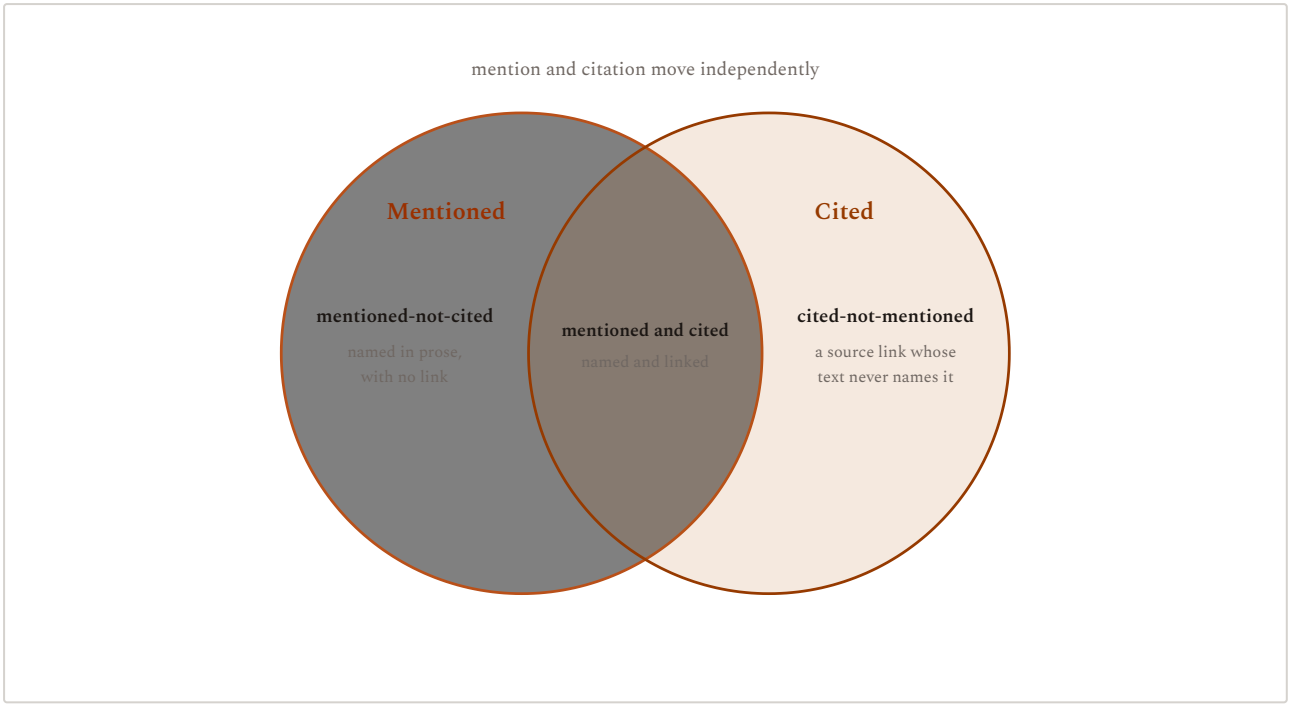

Figure 1. Mention and citation are distinct observables. An entity can be mentioned without being cited — named in an answer's prose with no link (left lobe) — or cited without being mentioned — a source link whose visible text never names it (right lobe); only the overlap is both named and linked. Because the two move independently, a reported rate must state which it counts (§4).

Barkhausen AI · BA-C-7

5 5. The query vocabulary: need, query, and phrasing

A single English word, "query," is used loosely for three things a measurement must keep separate: what a user wants to find out, the question they pose to find it, and the exact words they use. This convention fixes three levels.

An **information need** is what a user is trying to find out — a goal, independent of the words used to express it and of any engine. It is the most abstract level, and the one at which a measurement's estimand is defined (BA-C-2) and at which a query-intent class is assigned (BA-C-8). A **query** is a question that operationalizes an information need as something an engine can be asked; a single need may be served by more than one query. A **phrasing** is a specific surface wording of a query — one lexically distinct way of putting it. The set of real phrasings for a query is its phrasing distribution (BA-C-3); a measurement samples that distribution, and a single phrasing is one draw, not the query itself. When a need is served by more than one query, the need-level phrasing distribution BA-C-2 writes as Q_k is the pooled mixture of those queries' phrasing distributions: sampling may be organized per query, but the estimand's distribution is defined at the level of the need.

Keeping the levels distinct is what lets the rest of the series be precise: BA-C-2 defines the estimand over an information need, BA-C-3 samples across a phrasing distribution rather than a single sentence, and BA-C-8 classifies intent at the level of the need. A claim true at one level need not hold at another — visibility for one phrasing is not visibility for the need — and this vocabulary is what makes that gap sayable.

6 6. Recommendation set

A **recommendation set** is the set of entities an answer puts forward as candidates in response to a request for options — the list an answer gives to a "which" or "who should I" question. It is the object two BA-C-2 metrics act on directly: Discovery Depth measures the constraint depth at which an entity first enters it, and rank-biased overlap (RBO) how much of it persists across runs; Share of Voice is defined more broadly, as an entity's share of the brand mentions in the answers sampled, whether or not those mentions form a recommendation set (BA-C-2). Its membership and order are unstable across phrasings and reruns, which is why it is measured as a distribution rather than read once (BA-C-3). The term names an answer's put-forward candidates specifically, not every entity an answer happens to mention.

7 7. Usage rules

Three rules govern how a conforming publication uses this vocabulary.

One canonical term per concept. A publication uses the canonical term for each concept it names, and does not alternate between the canonical term and a synonym for the same concept within a document.

Link on first use. The first time a page uses a canonical term, it SHOULD link to the term's glossary entry, so that a reader arriving at the page by excerpt can resolve the term without prior context. This is the mechanism by which the glossary stays load-bearing rather than decorative.

Map, do not overload. A publication that has its own preferred term MAY use it, provided it maps the term to the canonical one on first use — for example, "answer-engine optimization (AEO), used here as a synonym for generative engine optimization." What it MUST NOT do is use a canonical term for a non-canonical meaning, because that silently breaks comparability for every reader who takes the term at its defined sense.

The following terms are deprecated within this series in favor of the canonical term, or flagged as ambiguous and requiring specification:

DEPRECATED OR AMBIGUOUS TERM	CANONICAL TERM	NOTE
AI SEO	generative engine optimization (GEO)	"SEO" names the search-era practice; GEO is the accepted name for the answer-engine practice
Answer-engine optimization (AEO)	generative engine optimization (GEO)	Permitted only as an explicitly mapped synonym, not used unmapped
LLM visibility	AI visibility	The phenomenon is a property of an engine's answers, not of a bare model
AI citation rate	*specify* mention rate or citation rate	Ambiguous — mention and citation are distinct observables (§4)
Chatbot (as the measured system)	answer engine / AI assistant	"Chatbot" under-describes the retrieval-and-answer function measured here

8 8. Conformance

A publication conforms to this convention by naming each concept it covers with the canonical term, or by explicitly mapping any term of its own to the canonical term on first use. Conformance concerns

vocabulary alone: it makes a publication's terms resolvable and its figures comparable; it does not bear on the validity of what the publication measures, which is governed by BA-C-2 through BA-C-5. A publication MAY state conformance as: *terminology per BA-C-7 v1.0*.

9 Limitations

A vocabulary is fixed only for as long as the thing it names holds still. The systems this series measures are new and changing, and the field's speech is unsettled; terms in common use today may narrow, split, or fall away. This convention is versioned for that reason: a new term, or a changed sense of an existing one, arrives through a new version, not through silent reinterpretation of the words already defined.

Its authority is also bounded. It governs how this series speaks and claims none over how the field speaks. A term defined here may carry a different sense elsewhere, which is not an error to correct but a difference to map — the usage rules in §7 exist so that a publication can adopt this vocabulary without asserting its own is the only correct one. Fixing terms improves comparability; it does not settle the field's language.

REFERENCES

1. Aggarwal, Murahari, Rajpurohit, Kalyan, Narasimhan, Deshpande; Proc. 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD 2024); arXiv:2311.09735. *GEO: Generative Engine Optimization* (2024). https://arxiv.org/abs/2311.09735 Accessed 2026-07-10. [archived]

HOW TO CITE

Barkhausen AI (2026). Terminology for AI-visibility measurement. https://barkhausen.ai/conventions/terminology/