EXTRACTION-READY CONTENT · FIELD GUIDE

Why Some Competitors Are Easier to Reuse and Cite

A passage-level diagnostic for editors, SEO/AEO teams, analysts and reviewers. Diagnose the evidence unit before rewriting the page.

> **Core rule**
>
> Do not optimize for extraction until you can name the exact passage, exact failure, evidence for that failure and smallest corrective change.

Updated 16 August 2026. This field guide does not claim that clear passages guarantee crawling, ranking, retrieval, answer inclusion, reuse or citation.

1. The passage identity card

Extraction-ready content preserves meaning when a useful passage is separated from surrounding prose. Six things must still travel with it:

Field	Question
Subject	Who or what is the claim about?
Claim	What exactly is asserted or recommended?
Conditions	Which scope, qualifier, exception or prerequisite changes the meaning?
Evidence	What supports the claim at the strength asserted?
Source	Where did the support come from, and what role does that source play?
Time	When does the claim or observation apply?

Evidence object ladder

Domain → Page → Section → Passage → Claim → Evidence → Source → Observation

Use the broadest object only when the evidence supports the broader inference. A citation of one passage under one query is not evidence that the whole domain is easier to cite.

Stage boundary

Crawling/access → indexing/eligibility → retrieval/reranking → passage selection/extraction → synthesis → attribution/citation. A content-side repair cannot substitute for an upstream access or retrieval test.

2. Ten failure classes

Route the first material failure. Secondary issues matter only if they change the next action.

#	Class	Diagnostic question
01	Scope	Are the question, market, product and evidence unit controlled?
02	Access / retrieval	Was the page eligible and retrieved under the observation?
03	Passage boundary	Can the useful evidence unit be isolated without changing meaning?
04	Referent / claim	Are actor, object, quantifier and assertion explicit?
05	Conditions	Does the qualifier or exception travel with the claim?
06	Relationship	Is the relation among entities stated rather than implied?
07	Evidence / provenance	Can a reviewer reconstruct the support and source role?
08	Temporal	Is a material date, version or window present?
09	Structure	Does presentation hide an otherwise valid relationship?
10	Causal inference	Is the team attributing a downstream result to an edit without a controlled test?

> **Stop rule**
>
> Once the passage is independently understandable and appropriately evidence-bound, stop rewriting it for hypothetical machine preferences. Move downstream.

3. The canonical seven-step method

1. **Define the exact question or claim.** Lock decision question, market/locale, product scope and observation window.
2. **Isolate the smallest meaningful passage.** Test one definition, rule, comparison or evidence-bearing claim instead of the whole page.
3. **Make subject, claim and relationship explicit.** Repair hidden actors, pronouns, category ambiguity and unbounded claims.
4. **Preserve conditions and necessary context.** Carry only the qualifier or exception required to prevent a different interpretation.
5. **Bind evidence, source, provenance and time.** Make the support reconstructable and time-bounded.
6. **Test in isolation; fix only the failing condition.** Do not reward extra length or formatting when the passage already passes.
7. **Observe retrieval, reuse and citation separately.** Use repeated controlled observations; downstream reuse never erases an upstream failure.

4. Controlled competitor audit

Hold the question, market, product scope, time window and evidence unit constant. Compare causes you can inspect; do not copy the winner.

Control	Compare	Do not infer
Subject / relationship	Explicit actor, object and relation	That entity density caused citation
Conditions	Whether qualifiers travel with the claim	That shorter text is better
Evidence / provenance	Support, locator, source role, date	That authority can replace evidence
Structure	Whether semantics expose an already valid relation	That formatting is a universal citation lever
Observed reuse	Repeated matched observations	That one citation proves causation

Current evidence changes the playbook

- Google says no tiny-chunk strategy or special generative-AI schema is required for its search features.
- A 2026 Ahrefs matched intervention found no major citation uplift from adding schema alone.
- A controlled 2026 competitive GEO study found topical relevance and position stronger than formatting-only edits in its testbed.
- AgentGEO found targeted failure diagnosis and repair could outperform generic rewriting while modifying much less content in its benchmark.

These findings have different evidence roles and scopes. None is treated as a universal platform law.

5. Worked example and decision scenarios

Starting passage: “The integration reduces reporting time by 40% and is suitable for multi-location teams.” Elsewhere, the page says the result came from a six-week, three-location pilot using one reporting workflow.

Check	Diagnosis	Action
Subject / claim	Clear enough	Preserve
Conditions	Pilot scope detached - first material failure	Carry the pilot boundary with the claim
Evidence / time	Locator and six-week window remote	Bind them to the evidence unit
Structure	Not the first failure	Do not redesign the section

> **Repaired passage**
>
> In a six-week, three-location pilot using one reporting workflow, the integration reduced reporting time by 40%; treat the result as pilot evidence, not as a universal performance claim.

Six symptoms, six next tests

Symptom	Next-best test
Competitor cited, your page not cited	Confirm matched retrieval before rewriting
Citation drops a qualifier	Check boundary, conditions and absorption/fidelity
Strong ranking, weak citation	Separate search authority from answer support
Schema added, no change	Ask whether schema was the relevant intervention
One rewrite, one new citation	Repeat matched observations before causal inference
Content checks pass, no reuse	Stop rewriting; move to access, retrieval, authority and observation

6. Measurement and handoff

Keep four outcomes separate so one metric cannot overwrite the diagnosis.

Evidence object	Question
Content validity	Did the passage identity card and isolation test pass?
Retrieval / selection	Was the page or passage surfaced under the controlled query and context?
Reuse / citation	Was the source cited, and did the answer absorb the supporting content?
Business outcome	Did exposure create qualified traffic, trust, pipeline or a better decision?

Workbook handoff rules

- Set one START decision question and one market / locale before activating repeated records.
- Every repeated ASSESS row retains the governed scope.
- Wrong-case or unsupported controlled values fail closed.
- Blank START scope cannot become VALID through blank = blank equality.
- SHARE remains blank when scope is blank and cannot open while an activated record is invalid.
- Test row 1, row 2 and a later repeated row before accepting formula fill-down.

Browser diagnostic

The web diagnostic routes one passage to the first content-side failure. It deliberately returns no numeric extraction score. The workbook manages repeated evidence and controlled handoff.

7. Review checklist and source boundaries

Before you approve a passage revision, confirm:

- The exact question, market and evidence unit are explicit.
- The first material failure was recorded before editing.
- Subject, claim, conditions, evidence, source and time travel together where material.
- Entity relationships are explicit enough for the decision; keywords are not being used as a substitute for meaning.
- The smallest correction was made; unaffected content was preserved.
- Retrieval, citation and business outcomes will be observed separately.
- One observation is not being converted into a universal platform rule.
- Reopening conditions are named: new claim, source, market, page structure, date/version or material platform change.

Source hierarchy

Role	Use
Primary platform documentation	Public access rules, feature definitions and explicit guidance
Controlled / academic research	Bounded experiments, failure models and measurement frameworks
Vendor observational research	Current patterns and hypotheses, not automatic causal proof
Competitor editorial guidance	Market advice and terminology, not authority about undisclosed algorithms

> **Final rule**
>
> Diagnose the exact evidence unit, repair the first material failure, preserve the rest, then observe downstream systems without turning one result into a universal law.

Companion article: https://www.serpifai.com/insights/extraction-ready-content/
Workbook: /assets/downloads/insights/extraction-ready-content/serpifai-extraction-readiness-passage-assessment-v1-1-0.xlsx