



FREE PRINTABLE WORKSHEET

GEO agency evaluation scorecard

Compare up to three AI search optimization proposals on the same 100-point evidence contract.

9 weighted categories

3 critical gates

12 interview questions

No signup

Use one evidence standard for every finalist.

- + Score only artifacts, demonstrations, contract language, and references you can inspect.
- + Leave a category blank until reviewed. A deliberate zero means the provider supplied no acceptable evidence.
- + Require at least 70 points overall and no zero in Measurement evidence, On-site implementation, or Integrity.
- + Do not let a high total override a critical zero. Verify references, security, legal terms, and fit separately.

FIRST-PARTY DISCLOSURE

Quoted First publishes this neutral worksheet and also sells audits and managed services. Apply the same criteria to Quoted First. No provider receives points for publishing the scorecard.



100-point master scorecard

Write candidate names once, then copy scores from the evidence sheets. Blank means not reviewed; zero is a completed score.

CANDIDATE A

CANDIDATE B

CANDIDATE C

Category	Max	Critical	A	B	C
Measurement evidence	20	YES			
Buyer-query design	10	NO			
On-site implementation	15	YES			
Independent-source work	10	NO			
Reporting and uncertainty	10	NO			
Team and ownership	10	NO			
Pricing and contract	10	NO			
B2B SaaS fit	10	NO			
Integrity	5	YES			
RECORDED TOTAL	100				

SUGGESTED GATE

70 or more points, every category reviewed, and no zero in Measurement evidence, On-site implementation, or Integrity. A critical zero fails the gate even when the total is above 70.

Review controls

- ☐ Same request-for-proposal scope sent to every candidate
- ☐ Same reviewers and evidence deadline
- ☐ References checked independently
- ☐ Security, legal, and data terms reviewed outside this score



Candidate A evidence sheet

Score evidence, not confidence

CANDIDATE NAME

WEBSITE

PROPOSAL DATE

REVIEWER

PRICE / TERM

1. Measurement evidence

0-20 POINTS - CRITICAL

SCORE

Full credit: Dated, engine-specific raw responses; misses retained; mentions, recommendations, and linked citations separated.

Zero red flag: Only one blended score, selected win screenshots, or no source URLs.

2. Buyer-query design

0-10 POINTS

SCORE

Full credit: Natural nonbranded questions mapped to audience, intent, and buying stage before the baseline.

Zero red flag: Only branded prompts, or a list retrofitted after results are known.

3. On-site implementation

0-15 POINTS - CRITICAL

SCORE

Full credit: Named owners for access, entity clarity, briefs, writing, editing, and publishing support.

Zero red flag: A dashboard with no accountable path from diagnosis to shipped work.

4. Independent-source work

0-10 POINTS

SCORE

Full credit: Ethical source mapping, original evidence, editorial outreach, and transparent participation.

Zero red flag: Fabricated reviews, link schemes, hidden sponsorship, or promised placements.

5. Reporting and uncertainty

0-10 POINTS

SCORE

Full credit: Fixed conditions, complete negative results, timestamps, source links, caveats, and repeats.

Zero red flag: Changing the question set to manufacture improvement, or percentages without raw evidence.

6. Team and ownership

0-10 POINTS

SCORE

Full credit: Named practitioners, workload, approval path, and clear client/provider ownership.

Zero red flag: A senior seller with an unidentified delivery team.

7. Pricing and contract

0-10 POINTS

SCORE

Full credit: Price, term, cancellation, included volume, exclusions, and pass-through costs in writing.

Zero red flag: Hidden lock-in or undefined unlimited work.

8. B2B SaaS fit

0-10 POINTS

SCORE

Full credit: Plan reflects stage, sales cycle, category, review constraints, claims, and publishing capacity.

Zero red flag: The same playbook for a pre-revenue tool and a mature multi-product platform.

9. Integrity

0-5 POINTS - CRITICAL

SCORE

Full credit: Written statement that mentions, recommendations, citations, links, and ranks are not guaranteed.

Zero red flag: Any guaranteed placement or claim of privileged control over an answer engine.

TOTAL / 100

CRITICAL ZERO?

☐

YES

☐

NO

COMPLETE?

☐



Candidate B evidence sheet

Score evidence, not confidence

CANDIDATE NAME

WEBSITE

PROPOSAL DATE

REVIEWER

PRICE / TERM

1. Measurement evidence

0-20 POINTS - CRITICAL

SCORE

Full credit: Dated, engine-specific raw responses; misses retained; mentions, recommendations, and linked citations separated.

Zero red flag: Only one blended score, selected win screenshots, or no source URLs.

2. Buyer-query design

0-10 POINTS

SCORE

Full credit: Natural nonbranded questions mapped to audience, intent, and buying stage before the baseline.

Zero red flag: Only branded prompts, or a list retrofitted after results are known.

3. On-site implementation

0-15 POINTS - CRITICAL

SCORE

Full credit: Named owners for access, entity clarity, briefs, writing, editing, and publishing support.

Zero red flag: A dashboard with no accountable path from diagnosis to shipped work.

4. Independent-source work

0-10 POINTS

SCORE

Full credit: Ethical source mapping, original evidence, editorial outreach, and transparent participation.

Zero red flag: Fabricated reviews, link schemes, hidden sponsorship, or promised placements.

5. Reporting and uncertainty

0-10 POINTS

SCORE

Full credit: Fixed conditions, complete negative results, timestamps, source links, caveats, and repeats.

Zero red flag: Changing the question set to manufacture improvement, or percentages without raw evidence.

6. Team and ownership

0-10 POINTS

SCORE

Full credit: Named practitioners, workload, approval path, and clear client/provider ownership.

Zero red flag: A senior seller with an unidentified delivery team.

7. Pricing and contract

0-10 POINTS

SCORE

Full credit: Price, term, cancellation, included volume, exclusions, and pass-through costs in writing.

Zero red flag: Hidden lock-in or undefined unlimited work.

8. B2B SaaS fit

0-10 POINTS

SCORE

Full credit: Plan reflects stage, sales cycle, category, review constraints, claims, and publishing capacity.

Zero red flag: The same playbook for a pre-revenue tool and a mature multi-product platform.

9. Integrity

0-5 POINTS - CRITICAL

SCORE

Full credit: Written statement that mentions, recommendations, citations, links, and ranks are not guaranteed.

Zero red flag: Any guaranteed placement or claim of privileged control over an answer engine.

TOTAL / 100

CRITICAL ZERO?

☐

YES

☐

NO

COMPLETE?

☐



Candidate C evidence sheet

Score evidence, not confidence

CANDIDATE NAME

WEBSITE

PROPOSAL DATE

REVIEWER

PRICE / TERM

1. Measurement evidence

0-20 POINTS - CRITICAL

SCORE

Full credit: Dated, engine-specific raw responses; misses retained; mentions, recommendations, and linked citations separated.

Zero red flag: Only one blended score, selected win screenshots, or no source URLs.

2. Buyer-query design

0-10 POINTS

SCORE

Full credit: Natural nonbranded questions mapped to audience, intent, and buying stage before the baseline.

Zero red flag: Only branded prompts, or a list retrofitted after results are known.

3. On-site implementation

0-15 POINTS - CRITICAL

SCORE

Full credit: Named owners for access, entity clarity, briefs, writing, editing, and publishing support.

Zero red flag: A dashboard with no accountable path from diagnosis to shipped work.

4. Independent-source work

0-10 POINTS

SCORE

Full credit: Ethical source mapping, original evidence, editorial outreach, and transparent participation.

Zero red flag: Fabricated reviews, link schemes, hidden sponsorship, or promised placements.

5. Reporting and uncertainty

0-10 POINTS

SCORE

Full credit: Fixed conditions, complete negative results, timestamps, source links, caveats, and repeats.

Zero red flag: Changing the question set to manufacture improvement, or percentages without raw evidence.

6. Team and ownership

0-10 POINTS

SCORE

Full credit: Named practitioners, workload, approval path, and clear client/provider ownership.

Zero red flag: A senior seller with an unidentified delivery team.

7. Pricing and contract

0-10 POINTS

SCORE

Full credit: Price, term, cancellation, included volume, exclusions, and pass-through costs in writing.

Zero red flag: Hidden lock-in or undefined unlimited work.

8. B2B SaaS fit

0-10 POINTS

SCORE

Full credit: Plan reflects stage, sales cycle, category, review constraints, claims, and publishing capacity.

Zero red flag: The same playbook for a pre-revenue tool and a mature multi-product platform.

9. Integrity

0-5 POINTS - CRITICAL

SCORE

Full credit: Written statement that mentions, recommendations, citations, links, and ranks are not guaranteed.

Zero red flag: Any guaranteed placement or claim of privileged control over an answer engine.

TOTAL / 100

CRITICAL ZERO?

☐

YES

☐

NO

COMPLETE?

☐



Vendor interview: questions 1-6

Ask for an artifact, clause, or specific example. Record what you inspected and what remains unresolved.

QUESTION 1 Which exact buyer questions will you measure?

Useful answer: A predeclared list segmented by problem, comparison, fit, and purchase intent.

Red flag: Only branded prompts or an undisclosed proprietary prompt set.

EVIDENCE / NOTES

QUESTION 2 Can we inspect a complete sample response set?

Useful answer: Raw answers, dates, engine/model conditions, displayed source order, and null results.

Red flag: Selected screenshots without misses.

EVIDENCE / NOTES

QUESTION 3 What counts as success?

Useful answer: Separate fields for absence, mention, recommendation, citation, and linked native-search result.

Red flag: Any mention is called a citation, or one opaque score replaces the evidence.

EVIDENCE / NOTES

QUESTION 4 How do you diagnose why another source wins?

Useful answer: Access, entity, answer-fit, evidence, and independent-source hypotheses tied to observed pages.

Red flag: Add schema or publish more before inspecting the result pool.

EVIDENCE / NOTES

QUESTION 5 Who implements each recommendation?

Useful answer: Named responsibility for code, schema, content, SME review, publishing, and outreach.

Red flag: A roadmap with no owner, acceptance test, or publishing support.

EVIDENCE / NOTES

QUESTION 6 What will exist after 30 days?

Useful answer: A declared query register, complete baseline, source map, eligibility findings, and work queue.

Red flag: A future dashboard demo with no inspectable deliverable.

EVIDENCE / NOTES



Vendor interview: questions 7-12

Ask for an artifact, clause, or specific example. Record what you inspected and what remains unresolved.

QUESTION 7 What should exist after 90 days?

Useful answer: Shipped fixes/pages, repeated identical measurements, a change log, and keep/stop decisions.

Red flag: A guaranteed percentage lift without a stable protocol.

EVIDENCE / NOTES

QUESTION 8 How do you earn third-party corroboration?

Useful answer: Original research, expert contribution, relevant editorial outreach, and transparent participation.

Red flag: Fake personas, sponsored lists without disclosure, review manipulation, or bulk links.

EVIDENCE / NOTES

QUESTION 9 How do you handle results that do not improve?

Useful answer: Preserve the negative result, reassess the hypothesis, and change one meaningful variable.

Red flag: Hide the prompt, change the denominator, or rename the metric.

EVIDENCE / NOTES

QUESTION 10 What does our team still own?

Useful answer: Specific weekly hours, SMEs, approvals, developer access, and publication deadlines.

Red flag: Done for you with no client dependencies.

EVIDENCE / NOTES

QUESTION 11 What is excluded from the price?

Useful answer: Content volume, outreach, tools, research, travel, media, fees, and revision limits.

Red flag: Important scope appears only after signing.

EVIDENCE / NOTES

QUESTION 12 What can you guarantee?

Useful answer: Process, deliverables, cadence, and evidence quality - not external source selection.

Red flag: A guaranteed citation, recommendation, link, position, or deadline.

EVIDENCE / NOTES



Candidate	Total	Critical zero	Complete	Price / term	Rank
		☐ Y ☐ N	☐ Y ☐		
		☐ Y ☐ N	☐ Y ☐		
		☐ Y ☐ N	☐ Y ☐		

Disqualifiers and conditions

- ☐ Guaranteed citation, recommendation, link, rank, or deadline
 - ☐ No complete response evidence or source URLs
 - ☐ No accountable implementation owner
 - ☐ Undisclosed subcontracting, lock-in, or pass-through cost
 - ☐ Reference, security, legal, or data review still open
 - ☐ Client capacity or approval dependency not resolved

Ownership map

MEASUREMENT		TECHNICAL / ENTITY	
CONTENT		INDEPENDENT SOURCES	
APPROVALS		RETESTING	

Recommendation

- ☐ Proceed ☐ Proceed with conditions ☐ Request missing evidence ☐ Do not proceed

RATIONALE / CONDITIONS

DECISION OWNER	DECISION DATE	NEXT REVIEW
----------------	---------------	-------------

The totals describe supplied evidence. They do not verify quality, predict placement, or make a universal provider recommendation.