

STRATEGIC NOTE · EXTERNAL PUBLIC AUDIT

# SEO & GEO Audit Urssaf.fr

Decide quickly, correct precisely and measure honestly — without GSC, Analytics, CMS or any other client connection.

---

| CLIENT | DATE            | METHOD         | MODE                                                                 |
|--------|-----------------|----------------|----------------------------------------------------------------------|
| Urssaf | August 21, 2026 | SEO/GEO V3.1.0 | Public measurements +<br>disclosed proxies;<br>no business estimates |

NAVIGATION

# Contents

|    |                                                      |    |
|----|------------------------------------------------------|----|
| 01 | Verdict in 30 seconds                                | 3  |
| 02 | Scope and method                                     | 4  |
| 03 | Audit coverage                                       | 5  |
| 04 | Reading F/V/O/E/M                                    | 6  |
| 05 | What already works                                   | 7  |
| 06 | Decision priorities                                  | 8  |
| 07 | Technical foundations — P0/P1 issues                 | 9  |
| 08 | Architecture and semantics                           | 10 |
| 09 | Performance: favorable field data, heavy lab results | 11 |
| 10 | Entities, structured data and provenance             | 12 |
| 11 | Content and metadata                                 | 13 |
| 12 | Public demand and visibility competition             | 14 |
| 13 | AI visibility: actual status                         | 15 |
| 14 | Blind spots and confirmation protocol                | 16 |
| 15 | 0–90 day roadmap                                     | 17 |
| 16 | Measurement and success criteria                     | 18 |
| 17 | Evidence register — summary                          | 19 |
| 18 | Sources and glossary                                 | 20 |

## VERDICT

## Verdict in 30 seconds

Urssaf.fr has a solid public foundation: HTTP 200 responses, clean canonicals, complete sitemaps, substantial content and legal references. The main risk is not a visible SEO failure, but a heavy template, incomplete machine semantics and the absence of first-party measurement.

**SOLID****Apparent indexability**

199/199 unique URLs returning 200, self-referencing canonical and present in the sitemap.

**TO STRENGTHEN****Foundations**

Three URLs with spaces, 302 chains, dominant megamenu and non-responsive images.

**TO STRUCTURE****Entity & provenance**

11/12 templates without JSON-LD; visible dates but little machine-readable data.

**NOT MEASURED****Business & AI**

No GSC/GA4 or fresh GEO panel: no visibility or return rate published.

### The figures guiding the decision

**839****Sitemap URLs**

823 FR + 16 EN, 0 duplicates

**3****URLs with spaces**

priority normalization fix

**≈7 MB****Homepage in the lab**

single PageSpeed run, to repeat

**11/12****No rendered JSON-LD**

one minimal NewsArticle on news

**Main obstacle**

The site contains a large amount of reliable information, but the template and machine signals do not sufficiently prioritize this value. Without GSC/analytics measurement, the team cannot then prioritize by impact or prove gains.

SCOPE

# Scope and method

## What was observed

- public crawl respecting robots.txt: 200 responses, 0.8 s delay, 200-URL cap;
- analysis of 199 unique HTML pages, links, metadata, Hn, images, dates, sources and canonicals;
- sitemap index + 823 French URLs + 16 English URLs;
- 12 templates in the rendered DOM, including 4 mobile checks;
- public redirects, TLS, headers, robots, llms.txt and security.txt;
- public PageSpeed: 28-day CrUX and one Lighthouse run per device;
- 40 public queries across nine intent families;
- primary Google, OpenAI, Anthropic and Perplexity documentation.

## What was never used

- no GSC, GA4, Bing Webmaster, CMS, CRM or Urssaf account;
- no credentials, secrets, client session cookies or personal data;
- no forms, logins, messages, publication or external changes;
- no traffic, volume, conversion, ranking or revenue estimates;
- no AI citation rate inferred from a contaminated session.

## Evidence rule

| Status       | Use in this report                                            | Example                                       |
|--------------|---------------------------------------------------------------|-----------------------------------------------|
| OBSERVED     | Direct measurement, denominator and capture retained.         | 199/199 self-referencing canonicals.          |
| PROXY        | Useful public signal, limitations written next to the figure. | Single Lighthouse run, public search results. |
| NOT MEASURED | No zero score; consequence and confirmation procedure.        | GSC CTR, GA4 conversion, AI citations.        |

### Coverage decision

Representative external audit accepted: the main critical templates are covered and unsampled families are declared. No sample metric is mechanically generalized to the 839 URLs.

SAMPLE

# Audit coverage

23.7 %  
of sitemap sampled  
199 unique URLs out of 839

12  
rendered templates  
DOM, Hn, JSON-LD, landmarks

4  
mobile templates  
390 x 844, no root overflow

## Template matrix

| Family                                                   | HTML | Rendered       | Mobile | Assessment                    |
|----------------------------------------------------------|------|----------------|--------|-------------------------------|
| Homepage                                                 | Yes  | Yes            | Yes    | Complete for the public scope |
| Employer / Self-employed / Individual hubs               | Yes  | 3              | No     | Good coverage of major hubs   |
| Regulatory articles                                      | Many | 2              | 2      | Two editorial depths          |
| Services and tools                                       | Many | 2              | 1      | Cesu, first hire, tools       |
| News                                                     | List | List + article | No     | NewsArticle checked           |
| Institutional                                            | Yes  | Yes            | No     | "Discover Urssaf" page        |
| Employee, artist-author, seafarer, third-party declarant | No   | No             | No     | NOT SAMPLED                   |

Known bias

The crawl follows the megamenu discovery order: Employer and Self-employed are overrepresented. The complete graph makes all collected URLs reachable in one click, which proves neither real UX depth nor the absence of orphan pages.

## Expected checks

21/23 checks have observed or proxy evidence. Two intentionally remain unmeasured: **geo.panel** and **measurement.first\_party**.

F / V / O / E / M

# Reading F/V/O/E/M

No overall average: each dimension answers a different question. Bars are published only when coverage exceeds the methodological threshold.



### Why F is not higher

Seven open technical findings account for the deductions: URLs with spaces, 302s, template H1, megamenu, lab, images and hreflang.

### Why V and O remain empty

The number of findings or a public search replaces neither homogeneous AI repetitions nor validated business components.

### Interpreting E

38.8 does not grade Urssaf's teams. It describes the plan's status at handover: proposed owners and acceptance criteria are present, but no deadlines are imposed without client governance.

## TO PRESERVE

## What already works

### Crawlable foundation

- 199 unique URLs returning 200;
- 199 titles and descriptions present;
- 199 self-referencing canonicals;
- 199 URLs present in the sitemap;
- French language declared across the sample.

### Structured sitemaps

- 839 unique URLs;
- 0 duplicates, 0 HTTP URLs;
- 839 lastmod dates;
- index separating French and English;
- robots.txt points to the index.

### Substantial content

- median near 985 words in the main content;
- no exact duplicates of the main content;
- visible dates on 166/199 pages;
- frequently asked questions and related content already present on many pages.

### Verifiable authority

- Légifrance references across many pages;
- BOSS present on relevant topics;
- 102/199 pages with at least one link to a public reference source according to the analysis;
- strong operational legitimacy in calculation, declaration and payment.

## Field performance

**Public CrUX passes Core Web Vitals** for the homepage URL and origin, on mobile and desktop. Optimization must preserve this result while reducing the cost observed in the lab.

## Mobile rendering

No root horizontal overflow was observed on the homepage, DPAE, general reduction and tools/documentation at 390 × 844. Visible editorial heading sequences show no skipped levels across the 12 templates.

PRIORITIZATION

# Decision priorities

| Priority | Decision                         | Why now                                                             | Completion evidence                                                  |
|----------|----------------------------------|---------------------------------------------------------------------|----------------------------------------------------------------------|
| P0       | Normalize the 3 URLs with spaces | Confirmed issue, low volume, clear fix.                             | Single-hop 301/308, new canonical, clean sitemap and links.          |
| P0       | Connect first-party measurement  | Without GSC/GA4, impossible to quantify indexing, demand and value. | 16-month baseline, documented scopes, reproducible exports.          |
| P1       | Lighten template and assets      | ≈7 MB in the lab, median HTML 351 KB, converging images/CSS/JS.     | Median of 3 runs, reduced weight, CrUX preserved.                    |
| P1       | Repair H1 and hreflang           | Systemic and mechanically testable problems.                        | 1 editorial H1; 16 complete and reciprocal FR/EN pairs.              |
| P1       | Structure entity and provenance  | Rich YMYL content, but weak machine signals.                        | Visible, accurate, dated and validated JSON-LD on 5 pilot templates. |
| P1       | Pilot 3 task journeys            | Demand spans several administrations and tools.                     | Tested journeys, "who does what?", sources and next steps.           |
| P2       | Metadata, mobile, security       | Quality, sharing, compliance and robustness gains.                  | Visual, social, header and CMP tests without regression.             |

Do not do

Do not deploy a megamenu or cache redesign without a test environment, backup, accessibility check and rollback.

Do not promise

Neither rich results, ranking, AI citations nor indexing time. Criteria concern what is controllable and verifiable.

P0 / P1

# Technical foundations — P0/P1 issues

## Three URLs with literal spaces

P0 · CONFIRMED

The French sitemap exposes three paths containing a space. Their encoded variants return 200 and the canonical retains that form.

**Fix:** clean target → 301/308 → canonical, internal links and sitemap aligned in the same batch.

## 302 chains at the four entry points

P1 · CONFIRMED

HTTP/apex/www go through one to three temporary redirects before /accueil.html.

**Fix:** after validating the durable target, one direct permanent redirect per variant.

## Incomplete hreflang on 16 FR/EN pairs

P1 · CONFIRMED

32 entries declare only the other language: no self-reference in the sitemap. Nine French pages in the crawl show the same HTML profile.

**Fix:** self-reference + reciprocity + 200 + canonical per version; x-default only if a real neutral page exists.

## Headers and CORS to harden

P2 · STRONG

HSTS and nosniff are present. CSP, anti-framing, Referrer-Policy and Permissions-Policy are not observed; Access-Control-Allow-Origin: \* is sent with credentials.

**Fix:** third-party inventory, CSP Report-Only, per-resource policy and integration tests. This is not a pentest.

## TLS

Valid chain, TLS 1.2/1.3 accepted and TLS 1.0/1.1 rejected. The observed certificate expires on September 30, 2026: confirm automated renewal in operational monitoring.

ARCHITECTURE

# Architecture and semantics

649–676

rendered links per page

628 to 653 are internal, mainly because of the megamenu. The complete graph becomes almost flat and weakly discriminating.

≈351 KB

median decompressed HTML

The template accounts for a significant share of each response. This weight is not the compressed transfer.

## The visible H1 is good; the DOM is not

| Layer                      | Observed H1s | Assessment                                         |
|----------------------------|--------------|----------------------------------------------------|
| Crawl source HTML          | 9 per page   | 8 menu audiences + 1 content                       |
| Rendered DOM, 12 templates | 14 per page  | 1 visible content + 1 visible consent + 12 hidden  |
| Visible editorial content  | 1 per page   | Readable hierarchy, no skip detected in the sample |

## Architecture decision

- 1. retain a single editorial H1;
- 2. replace navigation/modal H1s without changing their appearance;
- 3. measure menu groups that are truly useful to each audience;
- 4. strengthen contextual links in main before any overall reduction;
- 5. test crawlability, keyboard, screen reader and mobile together.

Beware of false depth

The crawl subgraph makes 199 URLs reachable in one click via the menu. This does not prove that users discover them easily or that no page on the domain is orphaned.

FIELD ≠ LAB

# Performance: favorable field data, heavy laboratory results

## Public Core Web Vitals — 28-day period

| Scope            | CWV  | LCP   | INP    | CLS  | TTFB  |
|------------------|------|-------|--------|------|-------|
| URL · mobile     | PASS | 2.0 s | 119 ms | 0    | 0.5 s |
| URL · desktop    | PASS | 1.4 s | 59 ms  | 0.02 | 0.3 s |
| Origin · mobile  | PASS | 1.3 s | 161 ms | 0    | 0.4 s |
| Origin · desktop | PASS | 1.1 s | 80 ms  | 0    | 0.3 s |

## Lighthouse — single run on the homepage

| Device           | Perf. | LCP    | TBT      | CLS   | Transfer  |
|------------------|-------|--------|----------|-------|-----------|
| Mobile · Slow 4G | 53    | 35.1 s | 170 ms   | 0     | 7 186 KiB |
| Desktop          | 42    | 3.3 s  | 1 100 ms | 0.027 | 7 385 KiB |

**≈2.4 MB**

images

estimated savings

**≈1.35 MB**

unused CSS

Lighthouse estimate

**≈0.86 MB**

unused JS

Lighthouse estimate

**≈2.1 MB**

cache

estimated savings

### Correct interpretation

Field data answers “what do users experience over 28 days?”. Laboratory data helps diagnose a controlled run. The extreme mobile run must be repeated three times; it does not replace CrUX.

YMYL

# Entities, structured data, and provenance

11/12

templates without JSON-LD

Homepage, hubs, regulatory articles, services, and institutional pages tested.

1

minimal NewsArticle

Headline and image, without dates, author, publisher, or mainEntityOfPage.

102/199

public sources

At least one reference link identified by the analysis.

## Recommended data architecture

| Template           | Candidate types                           | Non-negotiable condition                                                       |
|--------------------|-------------------------------------------|--------------------------------------------------------------------------------|
| Site / homepage    | Organization, WebSite                     | Legal name, logo, and valid identifiers.                                       |
| Internal pages     | BreadcrumbList, WebPage                   | Trail and URLs consistent with visible content.                                |
| Regulatory article | Article                                   | Real date, author/publisher, and sources.                                      |
| News               | NewsArticle                               | datePublished, dateModified, publisher, image, and URL.                        |
| Visible FAQ        | FAQPage if useful beyond Google's promise | Questions/answers actually visible; Google's FAQ rich result has been removed. |

## Editorial provenance

166/199 pages have a date label somewhere in the content, but no <time> tag or author meta is observed. On hubs, the date may belong to a card: distinguish the page date from the linked resource date.

Objective

Make the chain “who publishes, who validates, when the content was reviewed, based on which sources” readable to the user and consistent in the DOM and JSON-LD.

EDITORIAL QUALITY

# Content and metadata

## Substantial base

Median main content is close to 985 words; 94/199 pages exceed 1 000 words according to the source review. No exact main-content duplicate is detected.

## Hubs to judge by intent

Ten pages fall below 250 words in our extraction. This is not evidence of “thin content”: their ability to guide users must be tested.

## Signals to fix

| Signal               | Observation                                                                                   | Decision                                                                      |
|----------------------|-----------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------|
| Titles               | 3 strictly duplicated groups; 31 below 30 characters, 42 above 65 in our inventory.           | Fix confirmed duplicates; use GSC before a massive rewrite.                   |
| Descriptions         | 199/199 present and unique; 41 above 170 characters.                                          | Prioritize intent and the actual snippet, not a rigid threshold.              |
| Open Graph / Twitter | 0/199 in source HTML despite share buttons.                                                   | Add cards per template with authorized images.                                |
| News                 | “...don’t miss!1” in the list; description almost identical to the title for Prime carburant. | Immediate editorial fix and template check.                                   |
| Freshness            | 532/839 lastmod values are over one year old; 340 over two years.                             | Review according to legal volatility, without declaring these pages obsolete. |

## YMYL principle

An old rule may remain accurate. The priority is not to “refresh the date”, but to prove a subject-matter review proportionate to the risk and only change lastmod when a significant change occurs.

## PUBLIC SEARCH

# Public demand and visibility competition

The sample of 40 queries measures neither volume nor rank. It mainly reveals how needs are distributed between authorities, tools, comparison services, and friction-related content.

| User moment               | Visible sources                                | Urssaf's natural role                                      |
|---------------------------|------------------------------------------------|------------------------------------------------------------|
| Create / modify / close   | INPI, Guichet unique, Bercy, Service Public    | Social consequences and post-formality steps.              |
| Calculate / declare / pay | Urssaf, Service Public, private actors         | Main operational source and official calculators.          |
| Hiring / DPAE / DSN       | Urssaf, Ameli, France Travail, Net-entreprises | Cost → DPAE → payroll → DSN journey.                       |
| Health / taxation         | Ameli, impots.gouv.fr                          | Clearly distinguish social, health, and tax matters.       |
| International             | CLEISS, Ameli, impôts, Urssaf                  | Urssaf step after determining the competent country.       |
| Comparison / edge cases   | Firms, comparison services, YouTube, Reddit    | Assumptions, dated examples, limits, and error correction. |

## Three editorial pilots

### PILOT 1

#### Create, then declare

Creation → affiliation → first declaration → calculation → payment.

### PILOT 2

#### First hire

Cost → DPAE → payroll → DSN, with referrals to the right services.

### PILOT 3

#### International

Competent country → body → Urssaf process → supporting documents.

### Reusable block

**Who does what?** Name the authority, the action it owns, the handoff moment, and the official link. Urssaf becomes more useful by coordinating referrals, not by duplicating every administration.

Context: 35 strictly unbranded queries, 1 semi-branded, and 4 site: checks. Integrated web search surface; geolocation and personalization cannot be certified.

AI VISIBILITY

# AI visibility: actual status

**V = unmeasured**

No panel of responses in fresh, homogeneous sessions was run. Without repetitions by model/surface/locale, no mention, citation, or link frequency is published.

## Public technical accessibility

| Platform   | Search / response                     | Training                                     | Safe interpretation                                                     |
|------------|---------------------------------------|----------------------------------------------|-------------------------------------------------------------------------|
| Google     | Googlebot + GSC Generative AI control | Google-Extended for certain uses, not Search | Indexing and snippet eligibility; no special AI file required.          |
| OpenAI     | OAI-SearchBot                         | GPTBot                                       | ChatGPT-User is user-triggered, not the Search crawler.                 |
| Anthropic  | Claude-SearchBot / Claude-User        | ClaudeBot                                    | Distinct functions; Anthropic states that it honors robots.txt.         |
| Perplexity | PerplexityBot / Perplexity-User       | Not PerplexityBot                            | User fetch generally ignores robots.txt according to the documentation. |

## Urssaf findings

- robots.txt specifies no block specific to these bots;
- permission makes access possible, without guaranteeing crawling, indexing, citation, or traffic;
- llms.txt returns 404: this is not a Google defect and the format remains experimental;
- the GSC “Generative AI” control and report are unmeasured due to lack of access;
- the best current levers remain crawlability, useful content, sources, provenance, and structure.

**Future panel**

At least three repetitions per homogeneous context, capture of prompt/response/sources, distinction between mention–citation–link–attribution and publication of intervals, never a rate mixing all models.

MEASUREMENT HONESTY

# Blind spots and confirmation protocol

| Unverified                   | Why it matters                                                                | How to confirm                                                                |
|------------------------------|-------------------------------------------------------------------------------|-------------------------------------------------------------------------------|
| Google indexing and coverage | A 200 + canonical does not prove indexing.                                    | GSC Pages, sitemaps, sampled URL inspections, and dated exports.              |
| Queries, positions, CTR      | Impossible to prioritize titles, cannibalization, and opportunities.          | 16-month GSC export by query/page/device/country, homogeneous comparisons.    |
| Traffic and conversion       | No quantifiable user or business impact.                                      | GA4 or compliant measurement: journeys, events, forms, and exits to services. |
| Logs and bot crawling        | robots.txt does not prove actual crawling.                                    | Anonymized logs: UA + verified IP, status, URL, frequency, bytes.             |
| CWV by template              | Public CrUX covers homepage/origin, not every template.                       | GSC CWV, CrUX, and RUM; laboratory repeated on five templates.                |
| Exhaustive backlinks         | External authority and lost links unknown.                                    | Export from an authorized tool + GSC Links, deduplication and quality review. |
| AI visibility                | No reliable mention/citation frequency.                                       | Fixed GEO panel, fresh sessions, ≥3 repetitions, and strict segmentation.     |
| GSC Generative AI control    | It can control inclusion and its report has a limited rollout.                | Check the GSC property with an authorized owner; otherwise remain unmeasured. |
| Uncrawled families           | Employee, artist-author, seafarer, and third-party declarant not represented. | Targeted additional crawl by sitemap and template matrix.                     |

Rule

Lack of access is neither a site defect nor a zero score. Each blind spot remains isolated from published scores and has a confirmation method.

0-90 DAYS

## 0-90 day roadmap

0-7

### Secure and establish the baseline

**P0** Normalize the three URLs. Validate the canonical target of redirects. Obtain read access. Repeat Lighthouse. Freeze screenshots and criteria before changes.

8-30

### Fix the testable template

**P1** Menu/consent H1, hreflang, responsive images, hashed cache, first CSS/JS gains. Pilot JSON-LD on five templates with rollback.

31-60

### Strengthen provenance and journeys

**P1** Machine-readable dates, responsibilities, sources; three task journeys; "who does what?" blocks. Review the most critical content, not just a date change.

61-90

### Measure, learn, expand

**P1** Compare GSC/analytics with baseline, expand winning templates, launch the clean GEO panel, review megamenu and security based on evidence.

## Critical dependencies

| Action           | Depends on                         | Stop condition                                  |
|------------------|------------------------------------|-------------------------------------------------|
| URL / redirects  | Validation of the durable target   | A target does not return 200 or changes intent. |
| JSON-LD          | Legal name, publisher, logo, dates | Data not validated or invisible on the page.    |
| Cache / assets   | Reliable versioning                | Unhashed resource or functional regression.     |
| Content journeys | Subject-matter/legal validation    | Ambiguous institutional referral.               |
| Measurement      | Owner authorization                | Access not granted: action remains blocked.     |

## SUCCESS CRITERIA

# Measurement and success criteria

Each action has technical completion evidence before any interpretation of results. SEO/business impact can only be concluded after a sufficient window and a homogeneous comparison.

| Layer         | Indicator                 | Baseline                             | Controllable criterion                                              |
|---------------|---------------------------|--------------------------------------|---------------------------------------------------------------------|
| URL           | Spaces / chains           | 3 URLs; max 3 hops 302               | 0 spaces, 1 hop 301/308, target 200.                                |
| Template      | H1 / HTML / links         | 14 H1 in DOM; ≈351 KB; 649–676 links | 1 editorial H1; documented reduction without loss of crawlability.  |
| Performance   | Weight, lab LCP/TBT, CWV  | ≈7 MB; CrUX passed                   | Median of 3 runs improved; CrUX not regressed at 28 days.           |
| Entity        | Valid JSON-LD             | 11/12 without; 1 minimal             | 5 exact, visible pilot templates, without blocking errors.          |
| International | FR/EN clusters            | 16 pairs without self                | 32 pages self + reciprocal + 200 + canonical.                       |
| Content       | Pilot journeys            | 0 proprietary baseline               | 3 validated journeys; then impressions, clicks, exits, conversions. |
| GEO           | Mention / citation / link | Unmeasured                           | ≥3 repetitions per context, published denominator and interval.     |

## Observation windows

### Immediate

status, canonical, DOM, JSON-LD, hreflang, security, rendering.

### 7–28 days

logs, crawling, indexing, errors, laboratory performance, and early field data.

### 28–90 days

CrUX, queries, CTR, journeys, and conversions with seasonal context.

### Cautious attribution

Annotate each deployment, keep a comparison group when possible, and check regulatory updates, seasonality, and search engine changes before attributing a gain.

APPENDIX V3

# Evidence register — summary

| Source                 | Scope                                 | Status          | Confidence          |
|------------------------|---------------------------------------|-----------------|---------------------|
| RosoAuditBot crawl     | 200 responses / 199 unique URLs       | Observed        | Confirmed           |
| XML sitemaps           | 839 URLs, lastmod, hreflang, spaces   | Observed        | Confirmed           |
| Rendered browser       | 12 desktop templates, 4 mobile        | Observed        | Strong              |
| Headers / TLS          | homepage, redirects, sampled assets   | Observed        | Strong to confirmed |
| Public PageSpeed       | 28-day CrUX + 1 lab run/device        | Public proxy    | Strong / moderate   |
| Public search          | 40 queries, 9 families                | Volatile proxy  | Moderate            |
| Platform docs          | Google, OpenAI, Anthropic, Perplexity | Primary sources | Strong              |
| GSC / analytics / logs | no access                             | Unmeasured      | Confirmed absence   |
| GEO panel              | no homogeneous run                    | Unmeasured      | Confirmed absence   |

## Traceability delivered with the project

- evidence.jsonl register : 201 initial collection evidence items + 33 analysis evidence items;
- facts.json : 10 facts separated from interpretations;
- findings.json : 23 findings with severity, confidence, basis, and limits;
- actions.json : 16 actions with procedure, risk, backup, rollback, and criteria;
- score\_v3.json : deterministic F/V/O/E/M calculation without a global average;
- crawl and sitemap CSV inventories for filtering and resumption.

### Integrity check

Raw captures carry a SHA-256 and QA recalculates scores from the registers. Missing private data is never simulated.

Full technical identifiers remain in the machine register; the report body retains only readable evidence and its denominators.

## PRIMARY SOURCES

# Sources and glossary

## Official sources

- [Google — AI features and SEO fundamentals](#)
- [Google — crawlers and Google-Extended](#)
- [Search Console — Generative AI control](#)
- [Search Console — Generative AI report](#)
- [OpenAI — OAI-SearchBot, GPTBot, ChatGPT-User](#)
- [Anthropic — Claude bots](#)
- [Perplexity — crawlers](#)
- [Google — FAQ rich result removal](#)
- [web.dev — Core Web Vitals thresholds](#)
- [web.dev — field and laboratory](#)
- [llms.txt — experimental proposal v2](#)

## Glossary

**Canonical** — declared preferred URL for a page.

**CrUX** — Chrome field data aggregated over a rolling period.

**CWV** — LCP, INP, and CLS evaluated at the 75th percentile.

**GEO** — optimization and measurement of visibility in generative answers.

**Hreflang** — relationship between language or regional versions.

**JSON-LD** — structured data syntax embedded in the page.

**Lighthouse** — controlled laboratory test, distinct from field data.

**Proxy** — useful but indirect or limited indicator.

**YMYL** — content that may affect finances, rights, health, or safety.

## Conclusion

The priority is to make existing value lighter, more structured, and measurable. The site does not need an AI gimmick; it needs clean foundations, explicit provenance, and a durable evidence system.

### Version

SEO/GEO Audit V3.1.0 · data cutoff: 21 August 2026 · representative external public scope.