

User Guide

Part 2: The data and the filters

Version 1.0 · October 2026
Designed and built by Raj Meena

Part 2 explains the data in a report: what every tab and section shows, what each column means, and what every filter chip means and what to do about it. Part 1 covers the features, the buttons and the rules.

COVERS

Every tab and section of a report

ALSO COVERS

Every column and every filter chip

LEVELS

Error, Warning, Notice, Info

NUMBERS SHOWN

The app's default limits

PART 1

Features, buttons, rules

WHERE TO OPEN

Settings, User Guide

PICTURES

Taken from a sample website

COST TO READ

Nothing. It is a PDF

The numbers and sample sites in the pictures are examples only.

Contents

1. How to read a report	3
2. The Overview and All Issues	8
3. The Crawl tab	11
4. The On-Page tab	41
5. The Links & Media tab	65
6. The Markup tab	91
7. The AI Search tab	110
8. The Crawl Log tab	121
9. The Compare Crawl tab	128
10. Words used in the data	131

1. How to read a report

The few ideas behind every number.

A Silkpath report is a set of tables. Each table is a section (for example Page Titles), and the sections are grouped into tabs (for example On-Page). Above every table is a row of filter chips. This chapter explains the ideas that apply everywhere. The chapters after it go through the tabs and sections one by one.

Rows, URLs and issues

- **A row is one thing that was found**

In most sections a row is a web address (a URL). In All Links a row is one link, and in Images a row is one image file.

- **A URL counts once**

The three boxes on the Overview (Errors, Warnings, Notices) count URLs, not issues. A page that fails five checks is still one URL.

- **A filter chip is one check**

Each chip shows the rows that fail that check, with the number of rows. One chip is active at a time. The word All shows every row again.

- **The search box narrows the chip**

Type in the filter box to keep only the rows whose address (or text) contains what you typed. It works together with the chip you picked.

- **Sorting**

Press a column heading to sort by it. Press again to reverse. Numbers sort from largest first.

- **Paging and exports**

Tables show 100 rows a page. A section export saves the rows as you have filtered them, with no page limit.

Levels

Level	Chip dot	Meaning	Counts in the score
Error	Red	Broken or harmful. Fix these first.	A URL with an error costs one full point
Warning	Yellow	A real weakness that is worth fixing.	A URL with only warnings costs one third of a point
Notice	Blue	A small improvement, or a fact worth knowing.	Nothing
Info	No dot	Not a problem. A way to slice the list, such as only images or only external links.	Nothing

The health score is 100 minus the points as a share of all URLs crawled, kept between 0 and 100. The AI Search checks (AEO and GEO) appear in the report but never change the score. 80 and above is good, 60 to 79 needs work, below 60 is poor.

Reading a table

Every section is built the same way, from the top down. The pictures of each section below show these parts.

Part	What it is
Heading and note	The name of the section and, for some, a line saying what it checks.
Filter box	A box that says Filter by URL. Type to keep only the rows that contain what you typed. Next to it, a count such as 286 of 286 URLs says how many rows are shown out of all the rows.
Filter chips	One chip for each check, plus All. Chips are explained for every section below.
The first column	The thing each row is about, usually the URL. The URL is a link to the page itself.
The Issues column	The second column in any section that has checks. It lists, as small tags, every check that row fails. A tag has the same name as its filter chip, coloured by level: red for Error, yellow for Warning, blue for Notice. A dash means the row passes every check there. Sorting by Issues puts the rows with most problems first.
The other columns	The facts about each row. Each column is explained for every section below.

How a cell looks	What it means
Yes or No	Whether something is true for that row, such as whether a tag is present.
A small red tag such as missing	The thing was looked for and not found.
A dash	Nothing to show for this cell.
A number in colour	The number is outside the recommended range, for example a title with too many characters, or an inlink count of zero.
Right-aligned numbers	Counts and sizes. Press the column heading to sort from largest to smallest.

Numbers in chip names

Some chip names contain a number, such as Title Over 60 Characters or URL Over 115 Characters. These are the app's default limits and are the same for every report. The explanations in this part use the same numbers.

Sample pictures

The pictures in this part come from a sample website, so the numbers on them are examples. Your own report will show different counts. A chip that shows 0 in the picture may show a number in yours, and the other way round.

2. The Overview and All Issues

The two pages that summarise everything.

WHERE TO FIND IT A report, Overview tab (the first tab). All Issues is the second item in the Overview sidebar

The Overview shows the health score, the number of URLs with errors, warnings and notices, a comparison with the previous crawl of the same website when there is one, a chart of the score over your last 30 crawls, and totals for the crawl. Press an Errors, Warnings or Notices box to open All Issues filtered to that level.

All Issues lists every check that found something, across all sections, with the worst first. The chips at the top are All, Errors, Warnings and Notices. Pressing an issue opens its own section with the right filter chip already chosen, so you see exactly the affected URLs.

All Issues column	What it shows
Severity	Error, Warning or Notice.
Issue	The name of the check, the same as the filter chip in its section. Press it to go there.
Section	The section the check belongs to.
Affected	How many rows have the problem.
Share	That number as a share of all rows in the section.

Part of the Overview	What it shows
Site Health	The score from 0 to 100 with a bar.
Errors, Warnings, Notices	How many URLs have each level of problem.
Compare with the previous crawl	Health then and now, and the See What Changed button. Only shown when the same website was crawled before.
Site Health Over Time	A small chart of the score across your last 30 crawls of this website.

Part of the Overview	What it shows
Crawl, Status Codes, Site Checks	Totals, the split of response codes, and the checks that apply to the whole site: robots.txt, sitemap, HTTPS, preferred host, schema, llms.txt, JavaScript content and Safe Browsing.
Top Issues, Crawl Depth, External Domains	The eight biggest problems, how many clicks deep the pages sit, and which other websites are linked most.

3. The Crawl tab

How the crawl went: every address, the server's replies, security, what can be indexed, canonical tags and sitemaps.

WHERE TO FIND IT A report, Crawl tab (press the tab to see its main page)

No issues	20 with errors	1 with errors
Indexability 310 URLs 5 with warnings	Canonicals 286 URLs 29 with warnings	XML Sitemaps 310 URLs 9 with errors
PageSpeed - Not run yet		

The main page of the Crawl tab: one card for each section, with its count and its problems.

All URLs

WHERE TO FIND IT A report, Crawl tab, All URLs in the left list

All URLs

Every URL found during the crawl, in crawl order.

Q Filter by URL

310 of 310 URLs

All 310

HTML 306

PDF 1

Other 2

Unknown 1

● Uppercase in URL 1

● Non-ASCII Characters in URL 1

Show 13 empty filters

The top of All URLs: its heading, the filter box and the filter chips.

Every address the crawler found: pages, images, scripts, styles, PDFs and more, in the order it found them. Start here to see how big the site is and to spot odd addresses.

The tiles above the table show the URLs crawled, the HTML pages, the share that can be indexed, the average depth, the average and the median response time, whether http redirects to https, and whether the www and non-www forms redirect to one preferred form.

URL	ISSUES	STATUS	CONTENT TYPE	INDEXABLE	ORDER	DEPTH	INLINKS
/	-	200	text/html	Indexable	1	0	284
/about/	-	200	text/html	Indexable	2	1	284
/faq/	-	200	text/html	Indexable	3	1	284
/old-about/	-	301	-	Non-indexable	4	1	285

The table in All URLs: the first rows. Each row is one URL.

Column	What it shows
URL	The address.
Issues	The checks this row fails, shown as small tags coloured by level. Each tag has the same name as a filter chip. A dash means the row passes every check in this section.

Column	What it shows
Status	The server's reply code. 200 is fine, 301 means moved, 404 means not found.
Content type	The kind of file the server said it sent, such as text/html or image/png.
Indexable	Whether search engines may list the address. Indexable or Non-indexable.
Order	The position in which the crawler found it.
Depth	Clicks from the homepage. 0 is the homepage. Sitemap means it was found only in the XML sitemap and no link leads to it.
Inlinks	How many pages link to it.
Outlinks	How many links it contains.
Time (ms)	How long the server took to answer, in milliseconds. A page that answered far slower than the rest is asked once more, and the faster time is kept, because the first requests of a crawl are often slow while connections open and the site wakes up.
URL length	The number of characters in the address.
Last modified	The date in the Last-Modified header, when the server sends one.
Compression	How the server compressed the page for the trip (gzip or br). None means it sent the page uncompressed.

Column	What it shows
Transferred	The size the server announced for the trip, when it announces one.
Folder depth	How many folders deep the address is. /about-us/ is 1 and /blog/2024/post/ is 3. Not the same as Depth, which counts clicks.

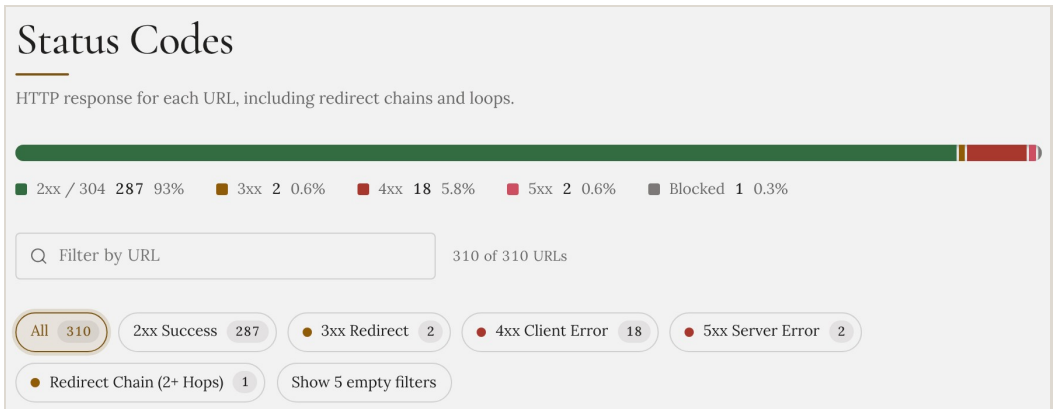
Filter chip	Level	What it means, and what to do
HTML, JavaScript, CSS, Images, Media, Fonts, XML, PDF, Other, Unknown	Info	Show only that kind of file. A chip is left out when the crawl found none. Unknown means the server did not say what it sent.
HTTP Address (Not HTTPS)	Warning	An address found in the crawl starts with http:// and not https://. Links to it should be updated to the secure address, and the old address should redirect with a 301.
URL Over 115 Characters	Notice	Very long addresses are hard to read and share and can be cut off in results. Shorten them where you can.
URL Has Parameters	Notice	The address has a ? part, such as ?sort=price. Parameters can create many copies of the same page. Make sure important pages also have a clean address.

Filter chip	Level	What it means, and what to do
Uppercase in URL	Notice	Capital letters in the address. Servers may treat /About and /about as different pages, which creates duplicates. Prefer lowercase.
Underscores in URL	Notice	Google reads a hyphen as a word break but not an underscore. Prefer hyphens.
Trailing-Slash Duplicate	Warning	The page works with and without a final slash (/shop and /shop/). Two addresses for one page split its strength. Redirect one to the other.
Non-ASCII Characters in URL	Notice	Letters such as é or other scripts in the address. They work, but become %-codes when shared. Plain letters are safer.
Space in URL (%20)	Notice	A space in the address. Use a hyphen instead.
Double Slash in Path	Notice	Two slashes in a row inside the address (/blog//post). Usually a mistake that makes a duplicate address.
Repeated Folder in Path	Notice	The same folder name twice in a row (/blog/blog/post). Often a link written wrongly.

Filter chip	Level	What it means, and what to do
Looks Like a Copy or Draft Page (Copy, Draft, Staging, Temp in URL)	Warning	The address contains copy, draft, staging, temp, dummy or untitled. A test page may be live by mistake.
Deep Page (6 or More Clicks from the Homepage)	Warning	It takes six clicks or more to reach this page. Deep pages are found and ranked less. Link to it from nearer pages.
Paginated Page Is Noindex	Warning	A page of a numbered series (such as /page/3) tells search engines not to list it, so the items on it may never be found.
Pagination Gap (Previous Page Was Not Found)	Warning	Page 4 exists but page 3 was not found. A broken series stops crawlers from following it.
Pagination Link Not in an Anchor Tag	Warning	The page names its next or previous page in a rel="next" or rel="prev" tag but has no normal link to it. Crawlers follow links, so the series may not be found. Add real Next and Previous links.

Status Codes

WHERE TO FIND IT **A report, Crawl tab, Status Codes in the left list**



The top of Status Codes: its heading, the filter box and the filter chips.

A status code is the server's reply to a request. 2xx means all is well, 3xx means the address has moved, 4xx means the request cannot be met (page not found, forbidden), and 5xx means the website's server failed. A bar chart at the top shows how many URLs got each kind of reply, with No response for addresses that never answered and Blocked for addresses robots.txt kept the crawler out of.

URL	ISSUES	STATUS	REDIRECTS TO	HOPS	INLINKS	TIME (MS)
/	-	200	-	0	284	1
/about/	-	200	-	0	284	2
/faq/	-	200	-	0	284	1
/old-about/	3xx Redirect Redirect Chain (2+ Hops)	301	/older-about/	2	285	2

The table in Status Codes: the first rows. Each row is one URL.

Column	What it shows
URL	The address.
Issues	The checks this row fails, shown as small tags coloured by level. Each tag has the same name as a filter chip. A dash means the row passes every check in this section.
Status	The reply code.

Column	What it shows
Redirects to	Where a redirect sends visitors.
Hops	How many redirect steps it takes to reach the final page.
Inlinks	How many pages link to it.
Time (ms)	How long the server took to answer.

Filter chip	Level	What it means, and what to do
2xx Success	Info	The server answered fine, usually with 200.
304 Not Modified	Info	The server said the page has not changed since the last visit. Normal.
3xx Redirect	Warning	The address sends visitors somewhere else. A few are normal. Internal links should point straight at the final address.
4xx Client Error	Error	Page not found (404), forbidden (403) and similar. Fix the link or restore the page, or redirect to a good page.
5xx Server Error	Error	The website's server failed (500, 502, 503 and so on). Check the server. If it keeps happening, search engines drop the page.

Filter chip	Level	What it means, and what to do
No Response (Failed, Timed Out or Refused)	Error	The website did not answer at all: the connection failed, timed out or was refused. Check that the address works, then rescan. If many addresses fail, lower Parallel Requests in Crawl Options.
Redirect Chain (2+ Hops)	Warning	A redirects to B which redirects to C. Each step slows visitors and loses strength. Make A go straight to C.
Redirect Loop	Error	Redirects that lead back to where they started, so the page never loads. Fix the redirect rule.
Rate Limited (429): the Site Blocked the Crawler	Notice	The website told the crawler to slow down, or blocked it. Lower Parallel Requests or add a Delay Between Requests in Crawl Options, then rescan.
Slow Response (Over 1 Second)	Warning	The server took more than a second to answer. Slow pages lose visitors and rank lower.
Page Could Not Be Read Completely	Warning	The page answered, but its HTML could not be read: it is nested thousands of levels deep, or the parser stopped part way. Silkpath skips such a page so the rest of the crawl can finish; the Crawl Log gives the reason. Check the page in a browser and simplify its markup.

Security

WHERE TO FIND IT **A report, Crawl tab, Security in the left list**

The screenshot shows the 'Security' report interface. At the top, the heading 'Security' is followed by a subtext: 'HTTPS, mixed content, unsafe forms and links, and the security headers each page is sent with. Crawls made before this check existed need a rescan to fill it in.' Below this is a search bar with the placeholder 'Filter by URL' and a count '286 of 286 URLs'. A row of filter chips is displayed, including 'All 286' (selected), 'Form or Password Field Sent Over HTTP 1', and 'Missing HSTS Header 286'. Subsequent rows contain chips for 'Missing Content-Security-Policy 286', 'Missing X-Content-Type-Options: Nosniff 286', 'Missing X-Frame-Options or CSP Frame-Ancestors 286', and 'Missing Referrer-Policy 286'. The final row shows 'Links Open a New Tab Without rel="noopener" 1' and a 'Show 4 empty filters' button.

Security

HTTPS, mixed content, unsafe forms and links, and the security headers each page is sent with. Crawls made before this check existed need a rescan to fill it in.

Q Filter by URL 286 of 286 URLs

All 286

Form or Password Field Sent Over HTTP 1

Missing HSTS Header 286

Missing Content-Security-Policy 286

Missing X-Content-Type-Options: Nosniff 286

Missing X-Frame-Options or CSP Frame-Ancestors 286

Missing Referrer-Policy 286

Links Open a New Tab Without rel="noopener" 1

Show 4 empty filters

The top of Security: its heading, the filter box and the filter chips.

Checks that protect visitors. HTTPS is the secure form of a web address. Mixed content is a secure page that still loads something over an insecure address. Headers are instructions the server sends along with each page. This section reads them for every page. Crawls made before this check existed need a rescan to fill it in.

URL	ISSUES	PROTOCOL	HSTS	CSP	NOSNIFF	FRAMING RULE
/	Missing HSTS Header	HTTPS	No	No	No	No
	Missing Content-Security-Policy					
	Missing X-Content-Type-Options: Nosniff					
	Missing X-Frame-Options or CSP Frame-Ancestors					
	Missing Referrer-Policy					
/about/	Missing HSTS Header	HTTPS	No	No	No	No
	Missing Content-Security-Policy					
	Missing X-Content-Type-Options: Nosniff					
	Missing X-Frame-Options or CSP Frame-Ancestors					
	Missing Referrer-Policy					
/faq/	Missing HSTS Header	HTTPS	No	No	No	No

The table in Security: the first rows. Each row is one URL.

Column	What it shows
URL	The address.
Issues	The checks this row fails, shown as small tags coloured by level. Each tag has the same name as a filter chip. A dash means the row passes every check in this section.
Protocol	HTTPS (secure) or HTTP (not secure).
HSTS	Whether the page sends a Strict-Transport-Security header.
CSP	Whether it sends a Content-Security-Policy header.
nosniff	Whether it sends X-Content-Type-Options: nosniff.
Framing rule	Whether it limits being shown inside another site's frame.
Referrer-Policy	Whether it sends a Referrer-Policy header.
Mixed content	How many insecure items a secure page loads.
New-tab links	Links that open a new tab without rel="noopener".
// links	Links that start with // instead of https://.

Filter chip	Level	What it means, and what to do
Page Served Over HTTP Instead of Redirecting to HTTPS	Warning	The http:// address answers with the page itself instead of sending visitors to the https:// address. Browsers warn visitors, the page exists twice, and every file it loads (styles, scripts, pictures) is loaded over HTTP too. Redirect the address to HTTPS with a 301.
Mixed Content: HTTP Script, Stylesheet or Frame on an HTTPS Page	Error	A secure page loads code over an insecure link. Browsers block it, so the page may break. Change the link to HTTPS.
Form or Password Field Sent Over HTTP	Error	A form that sends data over an insecure connection can be read by others. On an http:// page every form counts, not only a password box. On an https:// page it counts when the form posts to an http:// address. Serve the page and the form address over HTTPS.
Mixed Content: HTTP Image or Media on an HTTPS Page	Warning	A secure page shows an image or media from an insecure address. Browsers may warn or hide it. Use an HTTPS link.

Filter chip	Level	What it means, and what to do
Missing HSTS Header	Notice	HSTS tells browsers to always use HTTPS for this site. Adding the header makes the secure connection automatic.
Missing Content-Security-Policy	Notice	A content security policy lists where a page may load scripts and styles from. It limits the damage from injected code.
Missing X-Content-Type-Options: Nosniff	Notice	Stops browsers guessing a file's type, which blocks some attacks. Add the header with the value nosniff.
Missing X-Frame-Options or CSP Frame-Ancestors	Notice	Stops other websites from showing your page inside a frame, a trick used for click-hijacking.
Missing Referrer-Policy	Notice	Controls how much of the address is passed to other sites when a visitor clicks a link. A privacy setting.

Filter chip	Level	What it means, and what to do
Links Open a New Tab Without rel="noopener"	Notice	A link that opens a new tab without noopener lets the new page reach back into the old one. Add rel="noopener".
Protocol-Relative Links (//)	Notice	Links that begin with // copy the page's own connection type. It is an old style. Write the full https:// address.

Indexability

WHERE TO FIND IT A report, Crawl tab, Indexability in the left list

Indexability

Whether each URL can be indexed, and why not.

Q Filter by URL

310 of 310 URLs

All 310

Non-Indexable 32

Noindex 4

Blocked by robots.txt 1

Noarchive 1

Show 7 empty filters

The top of Indexability: its heading, the filter box and the filter chips.

Whether search engines are allowed to list each address, and if not, why. A page can be kept out of results by a noindex instruction in its code, by the same instruction in the server's reply (the X-Robots-Tag header), by robots.txt, by a redirect, or by an error. The tiles at the top show how many URLs are indexable, how many are not, whether a robots.txt file was found and how many Disallow rules it holds.

URL	ISSUES	STATUS	INDEXABLE	REASON	META ROBOTS	X-ROBOTS
/	-	200	Indexable	-	index, follow	-
/about/	-	200	Indexable	-	index, follow	-
/faq/	-	200	Indexable	-	index, follow	-
/old-about/	-	301	Non-indexable	Redirected	-	-

The table in Indexability: the first rows. Each row is one URL.

Column	What it shows
URL	The address.
Issues	The checks this row fails, shown as small tags coloured by level. Each tag has the same name as a filter chip. A dash means the row passes every check in this section.
Status	The reply code.

Column	What it shows
Indexable	Indexable or Non-indexable.
Reason	Why it is not indexable, such as noindex, redirect, error, or canonical elsewhere.
Meta robots	The robots instruction found in the page's code.
X-Robots-Tag	The robots instruction sent in the server's reply.

Filter chip	Level	What it means, and what to do
Non-Indexable	Info	Every address search engines will not list. The Reason column says why.
Noindex	Warning	The page has a robots noindex instruction, so search engines will not list it. Fine for private pages. A mistake on pages you want found.
X-Robots-Tag Noindex	Warning	The same instruction, sent in the server's reply instead of the page. Easy to miss because it is not in the page's code.
Blocked by robots.txt	Warning	The site's robots.txt asks crawlers to stay out. Search engines may still list the address but cannot read the page.
Nofollow	Notice	The page tells crawlers not to follow its links.

Filter chip	Level	What it means, and what to do
Nosnippet	Notice	Search engines may not show text snippets from the page in results.
Noarchive	Notice	Search engines may not show a saved copy of the page.
"None" Directive	Notice	A short way of writing noindex and nofollow together.
Noimageindex	Notice	Images on the page should not be listed in image search.
Max-Image-Preview Limited (None or Standard)	Notice	The page limits how large its images may appear in results. Large previews usually get more clicks.
Page Expiry Date Set (unavailable_after)	Notice	The page asks to be dropped from results after a date. Check that the date is intended.

Canonicals

WHERE TO FIND IT **A report, Crawl tab, Canonicals in the left list**

Canonicals

The rel="canonical" each page declares. A missing canonical is only reported on indexable pages:
a noindex page does not need one.

Q Filter by URL

286 of 286 URLs

All 286

● Missing Canonical 29

Canonicalised (Points Elsewhere) 4

Show 6 empty filters

The top of Canonicals: its heading, the filter box and the filter chips.

A canonical tag names the main version of a page when several addresses show the same or very similar content. It tells search engines which one to list. A page normally names itself.

URL	ISSUES	CANONICAL	TYPE	INDEXABLE
/	-	/	Self-referencing	Indexable
/about/	Missing Canonical	missing	missing	Indexable
/faq/	Missing Canonical	missing	missing	Indexable
/espresso-beans/	Missing Canonical	missing	missing	Indexable

The table in Canonicals: the first rows. Each row is one URL.

Column	What it shows
URL	The address.
Issues	The checks this row fails, shown as small tags coloured by level. Each tag has the same name as a filter chip. A dash means the row passes every check in this section.
Canonical	The address the page names as the main version.
Type	Self-referencing (names itself), Points elsewhere, or Missing.
Indexable	Whether the page itself can be listed.

Filter chip	Level	What it means, and what to do
Missing Canonical	Warning	No canonical tag on an indexable page. Add one that points at the page's own preferred address. A page with noindex does not need one, so it is not reported.
Canonicalised (Points Elsewhere)	Info	The canonical names a different address as the main version. Normal for duplicates. This page is expected to drop out of results.
Canonical Target Returns 4xx/5xx	Error	The canonical points at a page that is missing or broken. Fix the address.
Canonical Points to a Noindex Page	Error	The canonical names a page as the main version, but that page tells search engines not to list it. The two signals clash and the canonical is likely to be ignored. Point it at an indexable page, or remove noindex from the target.
Canonical Target Blocked	Warning	The page it points to is blocked by robots.txt, so search engines cannot check it.
Canonical Chain	Warning	The canonical points at a page whose own canonical or redirect points somewhere else. Point straight at the final page.

Filter chip	Level	What it means, and what to do
Multiple Canonicals	Error	Two or more canonical tags on one page. Search engines may ignore all of them. Keep one.
Noindex with Canonical Elsewhere	Warning	The page says do not list me and also says the main version is elsewhere. The two signals clash. Choose one.

XML Sitemaps

WHERE TO FIND IT **A report, Crawl tab, XML Sitemaps in the left list**

XML Sitemaps

Sitemap URLs compared against what the crawl found.

- robots.txt

Found

2 disallow rules · sitemap listed

- sitemap.xml

Found

287 URLs listed

IN SITMAP, NOT LINKED INTERNALLY (1)

/orphan-landing/

310 of 310 URLs

All 310

Indexable, Not in Sitemap 4

Non-Indexable in Sitemap 4

Redirect in Sitemap 1

Error in Sitemap 4

Canonicalised URL in Sitemap 4

The top of XML Sitemaps: its heading, the filter box and the filter chips.

A sitemap is a file that lists the pages you want search engines to find. This section compares the sitemap with what the crawl found. A panel at the top shows the sitemap and robots.txt checks, and lists pages that are in the sitemap but not linked from anywhere inside the site.

URL	ISSUES	STATUS	INDEXABLE	IN SITEMAP
/	-	200	Indexable	Yes
/about/	Indexable, Not in Sitemap	200	Indexable	No
/faq/	Indexable, Not in Sitemap	200	Indexable	No
/old-about/	Redirect in Sitemap	301	Non-indexable	Yes

The table in XML Sitemaps: the first rows. Each row is one URL.

Column	What it shows
URL	The address.
Issues	The checks this row fails, shown as small tags coloured by level. Each tag has the same name as a filter chip. A dash means the row passes every check in this section.
Status	The reply code.

Column	What it shows
Indexable	Whether the page can be listed.
In sitemap	Yes if the sitemap lists it.

Filter chip	Level	What it means, and what to do
Indexable, Not in Sitemap	Warning	A page that can be listed but is missing from the sitemap. Add it.
Non-Indexable in Sitemap	Error	The sitemap lists a page that search engines are told not to list. Remove it from the sitemap, or allow the page.
Redirect in Sitemap	Error	The sitemap lists an address that redirects. List the final address instead.
Error in Sitemap	Error	The sitemap lists a broken address (4xx or 5xx). Remove or fix it.
Canonicalised URL in Sitemap	Warning	The sitemap lists a page whose canonical points elsewhere. List the main version instead.

PageSpeed

Google's speed test for your top pages. Needs the owner's Google key.

WHERE TO FIND IT **A report, Crawl tab, PageSpeed in the left list**

PageSpeed

Google PageSpeed Insights for your top 10 pages on mobile, with real-visitor Core Web Vitals from the Chrome UX Report where Google has enough data. Runs only when you ask, then the result is saved with this crawl.

Run PageSpeed Check (10 pages)

Takes about a minute. Each server has a daily limit, so a check may stop early.

PageSpeed before it is run: the button that starts the check.

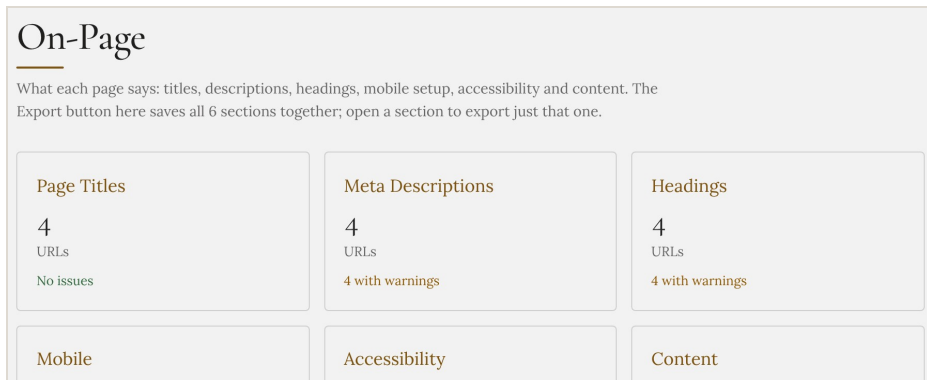
PageSpeed runs Google PageSpeed Insights on your most-linked pages as a phone visitor. It only runs when you press Run PageSpeed Check, and the result is saved with the crawl. It has no filter chips.

Column	What it shows
URL	The page that was tested.
Score	The phone speed score from 0 to 100. 90 and above is good, 50 to 89 needs work, below 50 is poor. It comes from one simulated visit, so it moves a little each run.
LCP	Largest Contentful Paint: how many seconds until the main content appears. Under 2.5 is good.
CLS	Cumulative Layout Shift: how much the page jumps around while loading. Under 0.1 is good.
TBT	Total Blocking Time: how long the page was too busy to respond, in milliseconds.
Real LCP, Real INP, Real CLS	The same measures from real Chrome visitors over the last 28 days, when Google has enough data. INP is Interaction to Next Paint: how quickly the page reacts to a tap or click. Under 200 milliseconds is good.
Core Web Vitals	Pass or fail on the real-visitor numbers for LCP, INP and CLS together.

4. The On-Page tab

What each page says: titles, descriptions, headings, phone set-up and the words.

WHERE TO FIND IT A report, On-Page tab (press the tab to see its main page)



The main page of the On-Page tab: one card for each section.

Page Titles

WHERE TO FIND IT **A report, On-Page tab, Page Titles in the left list**

Page Titles

Pixel width is estimated from character widths. Google doesn't publish its exact cut-off. These checks look at indexable pages only. Pages with noindex, a redirect or a canonical elsewhere do not show in search results, so they are left out.

Q Filter by URL

286 of 286 URLs

All 286

Missing Title 5

Duplicate Title 6

Title Over 60 Characters 7

Title Under 30 Characters 11

Title Over 561px (Estimated) 7

Title Under 200px (Estimated) 2

Title Same as H1 1

Show 4 empty filters

The top of Page Titles: its heading, the filter box and the filter chips.

The title is the headline in search results and on the browser tab. Each page should have one clear title that is different from every other page's. Pixel width is an estimate from the width of each letter, because Google cuts titles by width, not by letters, and does not publish its exact limit. These checks look at indexable pages only. Pages with noindex, a redirect or a canonical elsewhere do not show in search results, so they are left out. Such a page carries a small "not indexable" tag next to its address.

URL	ISSUES	TITLE	CHARS	PIXELS
/	Title Under 30 Characters	Harbour Roasters Home	23	228
/about/	Title Under 30 Characters Title Under 200px (Estimated)	About us	8	80
/faq/	Title Under 30 Characters	FAQ Harbour Roasters	22	216
/espresso-beans/	-	espresso-beans – page 1 Harbour Roasters	42	395

The table in Page Titles: the first rows. Each row is one URL.

Column	What it shows
URL	The address. A page that is not indexable has a small "not indexable" tag after it, and the checks in this section skip it.
Issues	The checks this row fails, shown as small tags coloured by level. Each tag has the same name as a filter chip. A dash means the row passes every check in this section.
Title	The title text.
Chars	The number of characters. Turns coloured outside 30 to 60.
Pixels	The estimated width in pixels. Turns coloured over 561.

Filter chip	Level	What it means, and what to do
Missing Title	Error	No title tag. Search engines will invent one. Add a clear title.
Duplicate Title	Warning	Another indexable page has exactly the same title. Give each page its own. The pages of a numbered series (/page/2/, /page/3/ and so on) are not compared with each other, because they share a title by design.
Title Over 60 Characters	Warning	Long titles get cut off in results. Shorten to the point.
Title Under 30 Characters	Notice	A very short title wastes space that could describe the page.

Filter chip	Level	What it means, and what to do
Title Over 561px (Estimated)	Warning	Estimated to be wider than Google shows. Wide letters fill the space faster than narrow ones.
Title Under 200px (Estimated)	Notice	Estimated to be very narrow. The same idea as a short title.
Title Same as H1	Notice	Not wrong, but a missed chance to use different words for the result headline and the page heading.
Multiple Title Tags	Warning	Two or more title tags in one page. Keep one.
Line Break Inside the Title Tag	Error	A line break inside the title tag, usually from a template mistake. It can break how the title is read.
Title Contains Leaked "Meta Description:" Text	Error	The words Meta description appear inside the title, a sign that a tag was left open and swallowed the next one. Fix the template.
Title and H1 Share No Words	Notice	The title and the main heading talk about different things. Check the page is about what the title promises.

Meta Descriptions

WHERE TO FIND IT A report, On-Page tab, Meta Descriptions in the left list

Meta Descriptions

These checks look at indexable pages only. Pages with noindex, a redirect or a canonical elsewhere do not show in search results, so they are left out.

Q Filter by URL

286 of 286 URLs

All 286

● Missing Description 39

● Description Under 70 Characters 25

● Description Under 400px (Estimated) 25

Show 5 empty filters

The top of Meta Descriptions: its heading, the filter box and the filter chips.

The meta description is the grey summary under the title in search results. It is not a ranking factor on its own, but a good one makes people more likely to click. These checks look at indexable pages only. Pages with noindex, a redirect or a canonical elsewhere do not show in search results, so they are left out. Such a page carries a small "not indexable" tag next to its address.

URL	ISSUES	META DESCRIPTION	CHARS	PIXELS
/	-	Small-batch coffee roasted every Tuesday in Bristol. Beans, brewing...	125	771
/about/	Missing Description	missing	missing	-
/faq/	-	Frequently asked questions about the test shop, shipping, returns a...	102	637
/espresso-beans/	Description Under 70 Characters Description Under 400px (Estimated)	All products in espresso-beans.	31	194

The table in Meta Descriptions: the first rows. Each row is one URL.

Column	What it shows
URL	The address. A page that is not indexable has a small "not indexable" tag after it, and the checks in this section skip it.

Column	What it shows
Issues	The checks this row fails, shown as small tags coloured by level. Each tag has the same name as a filter chip. A dash means the row passes every check in this section.
Meta description	The description text.
Chars	The number of characters. Turns coloured outside 70 to 155.
Pixels	The estimated width. Turns coloured over 985.

Filter chip	Level	What it means, and what to do
Missing Description	Warning	No description. Search engines will pick text from the page. Write one that sells the page.
Duplicate Description	Warning	Another indexable page has exactly the same description. Write a different one for each page. The pages of a numbered series are not compared with each other.
Description Over 155 Characters	Warning	Likely to be cut off in results. Put the important words first.
Description Under 70 Characters	Notice	Short descriptions leave useful space unused.

Filter chip	Level	What it means, and what to do
Description Under 400px (Estimated)	Notice	Estimated to be narrower than a typical result line, so part of the space is unused. Screaming Frog has the same check.
Description Over 985px (Estimated)	Warning	Estimated to be wider than Google shows, so it will be cut.
Multiple Descriptions	Warning	Two or more description tags. Search engines may choose the wrong one. Keep one.
Meta Keywords Tag Present (Ignored by Google)	Notice	An old tag Google ignores. It does no harm but can be removed, and it shows rivals your keywords.

Headings

WHERE TO FIND IT **A report, On-Page tab, Headings in the left list**

Headings

These checks look at indexable pages only. Pages with noindex, a redirect or a canonical elsewhere do not show in search results, so they are left out.

Q Filter by URL

286 of 286 URLs

All 286

● Missing H1 8

● Missing H2 28

Show 9 empty filters

The top of Headings: its heading, the filter box and the filter chips.

Headings are the outline of a page. H1 is the main title, H2s are its sections, H3s are parts of a section. People, screen readers and search engines all read the outline. These checks look at indexable pages only.

Pages with noindex, a redirect or a canonical elsewhere do not show in search results, so they are left out. Such a page carries a small "not indexable" tag next to its address.

URL	ISSUES	H1	H1 CHARS	H1 COUNT	H2 COUNT
/	Missing H2	Harbour Roasters	16	1	
/about/	-	About	5	1	
/faq/	-	FAQ	3	1	
/espresso-beans/	Missing H2	espresso-beans	14	1	

The table in Headings: the first rows. Each row is one URL.

Column	What it shows
URL	The address. A page that is not indexable has a small "not indexable" tag after it, and the checks in this section skip it.
Issues	The checks this row fails, shown as small tags coloured by level. Each tag has the same name as a filter chip. A dash means the row passes every check in this section.
H1	The first H1 on the page.
H1 chars	Its length in characters.

Column	What it shows		
H1 count	How many H1s the page has. One is right.		
H2 count	How many H2s the page has.		
Filter chip	Level	What it means, and what to do	
Missing H1	Error	The page has no main heading. Add one that says what the page is about.	
Multiple H1	Warning	More than one H1. Use one, and make the others H2.	
Duplicate H1	Warning	Another indexable page has the same H1, ignoring capital letters. Give each page its own. Unlike titles and descriptions, the pages of a numbered series are compared too, the way Screaming Frog does it.	
H1 Over 70 Characters	Notice	A very long main heading. Shorten it.	
Missing H2	Warning	No section headings. Long pages are easier to read and scan with them.	
Duplicate H2 on Page	Notice	The same H2 text twice on one page. Make each section heading different.	
H2 Over 70 Characters	Notice	A very long section heading.	

Filter chip	Level	What it means, and what to do
H1 Not the First Heading	Notice	Another heading, such as an H2 in a sidebar or menu, comes before the H1 in the page code. Put the H1 first so the outline starts with the page's main heading.
Multiple H2	Notice	More than one H2 on the page. This is normal for long pages and not a fault. It is listed because Screaming Frog lists it; use it to see which pages are split into sections.
Heading Levels Skipped (H1 to H3)	Notice	The outline jumps a level, such as H1 straight to H3, on a page that has H2s. Keep the steps in order so the outline makes sense. A page with no H2 at all is reported under Missing H2 instead, so it is not counted twice.
H2 Shared with Other Pages (First Two H2s)	Notice	One of the page's first two section headings has the same text as a section heading on another indexable page (such as Welcome), ignoring capital letters. Make section headings specific.

Mobile

WHERE TO FIND IT **A report, On-Page tab, Mobile in the left list**

Mobile

Checks the viewport tag in the page source. Tap-target size and text size need a real browser and are not checked.

Q Filter by URL

286 of 286 URLs

All 286

● Missing Viewport Meta Tag 1

● Homepage Missing Favicon 1

● Homepage Missing Web App Manifest 1

Show 2 empty filters

The top of Mobile: its heading, the filter box and the filter chips.

Checks that a page is set up for phones. It reads the viewport tag in the page's code. It does not measure tap-target size or text size, which need a real browser.

URL	ISSUES	VIEWPORT TAG	RESPONSIVE WIDTH	ZOOM ALLOWED
/	Homepage Missing Favicon Homepage Missing Web App Manifest	width=device-width,initial-scale=1	Yes	Yes
/about/	-	width=device-width,initial-scale=1	Yes	Yes
/faq/	-	width=device-width,initial-scale=1	Yes	Yes
/espresso-beans/	-	width=device-width,initial-scale=1	Yes	Yes

The table in Mobile: the first rows. Each row is one URL.

Column	What it shows
URL	The address.
Issues	The checks this row fails, shown as small tags coloured by level. Each tag has the same name as a filter chip. A dash means the row passes every check in this section.

Column	What it shows
Viewport tag	The viewport instruction found in the page.
Responsive width	Yes if the page asks for the width of the device.
Zoom allowed	Yes if visitors can pinch to zoom.

Filter chip	Level	What it means, and what to do
Missing Viewport Meta Tag	Warning	Phones show the page as a tiny desktop page. Add a viewport tag.
Viewport Without width=device-width	Warning	The tag is there but does not ask for the device width, so the page may not fit a phone.
Zooming Disabled (user-scalable=no or maximum-scale=1)	Warning	Visitors cannot pinch to zoom, which is hard for people with poor eyesight. Allow zooming.
Homepage Missing Favicon	Warning	The small icon shown on browser tabs and in some results. Add one.
Homepage Missing Web App Manifest	Notice	A small file that lets the site be installed on a phone's home screen.

Accessibility

WHERE TO FIND IT **A report, On-Page tab, Accessibility in the left list**

Accessibility

Quick accessibility checks read from the page source: the page language, form fields with no label, buttons and links with nothing in them, and IDs used twice. Colour contrast, keyboard use and anything a script builds after the page loads need a real browser and are not checked. Crawls made before these checks existed need a rescan to fill them in.

Q Filter by URL

286 of 286 URLs

All 286

Show 6 empty filters

The top of Accessibility: its heading, the filter box and the filter chips.

Quick accessibility checks that are read from each page's source code: the language of the page, form fields with no label, buttons and links with nothing in them, and IDs used twice. Colour contrast, keyboard use and anything a script builds after the page loads need a real browser and are not checked. Crawls made before these checks existed need a rescan to fill them in.

URL	ISSUES	PAGE LANGUAGE	UNLABELED FIELDS	EMPTY BUTTONS	EMPTY LINKS	DUPI
/	-	en	0	0	0	
/about/	Missing Lang Attribute on the Page Form Fields Without a Label Buttons with No Text Links with No Accessible Name Duplicate IDs	missing	1	1	1	
/shop/	Form Fields Without a Label Buttons with No Text Links with No Accessible Name	en	1	1	1	
/contact/	-	en	0	0	0	

The table in Accessibility: the first rows. Each row is one URL.

Column	What it shows
URL	The address.
Issues	The checks this row fails, shown as small tags coloured by level. Each tag has the same name as a filter chip. A dash means the row passes every check in this section.
Page language	The language the page declares in its opening tag, such as en or en-GB.
Unlabeled fields	How many form fields have no label, aria-label, title or placeholder to name them.
Empty buttons	How many buttons have no text, image description or aria-label.
Empty links	How many links have no text, image description or aria-label.
Duplicate IDs	How many ID names are used more than once on the page.

Filter chip	Level	What it means, and what to do
Missing Lang Attribute on the Page	Warning	The page does not say what language it is in. Screen readers and search engines use this. Add lang="en" (or the right code) to the opening html tag.
Invalid Lang Value	Notice	The lang value is not a valid language code such as en or en-GB. Fix the code.

Filter chip	Level	What it means, and what to do
Form Fields Without a Label	Notice	A form field has no label that says what to type. People who use screen readers cannot tell what the field is for. Add a label, or an aria-label.
Buttons with No Text	Notice	A button has nothing in it that says what it does, such as an icon with no description. Add text or an aria-label.
Links with No Accessible Name	Notice	A link has no text, image description or aria-label, so a screen reader can only say “link”. Describe where it goes.
Duplicate IDs	Notice	The same ID is used for more than one element. Labels and scripts that point at an ID can pick the wrong one. Make each ID unique.

Content

WHERE TO FIND IT **A report, On-Page tab, Content in the left list**

Content

These checks look at indexable pages only. Pages with noindex, a redirect or a canonical elsewhere do not show in search results, so they are left out.

Q Filter by URL

286 of 286 URLs

All 286

● Thin Content (Under 200 Words) 58

● Near Duplicate (90%+ Similar) 201

● Low Text-to-HTML Ratio (Under 10%) 3

Show 2 empty filters

The top of Content: its heading, the filter box and the filter chips.

The words on each page: how many, how readable, and whether they repeat another page. The tiles show total words, average words per page, total HTML size and the average text-to-HTML ratio. These checks look at indexable pages only. Pages with noindex, a redirect or a canonical elsewhere do not show in search results, so they are left out. Such a page carries a small "not indexable" tag next to its address.

URL	ISSUES	WORDS	HTML SIZE	TEXT RATIO	READING EASE	S
/	Thin Content (Under 200 Words)	35	2.3 KB	18.8%	-	
/about/	Thin Content (Under 200 Words)	28	687 B	22.1%	-	
/faq/	Thin Content (Under 200 Words)	28	1.1 KB	16.2%	-	
/espresso-beans/	Thin Content (Under 200 Words)	17	1.2 KB	12.7%	-	

The table in Content: the first rows. Each row is one URL.

Column	What it shows
URL	The address. A page that is not indexable has a small "not indexable" tag after it, and the checks in this section skip it.
Issues	The checks this row fails, shown as small tags coloured by level. Each tag has the same name as a filter chip. A dash means the row passes every check in this section.

Column	What it shows
Words	The number of words of visible text, not counting the navigation menu and the footer. This is how Screaming Frog counts, so the two tools agree.
HTML size	The size of the page's code.
Text ratio	The share of the page's code that is visible text.
Reading ease	The Flesch score. 0 to 100, higher is easier. Under 40 is difficult.
Similarity	How much of the text repeats another page. 100% is an exact copy.

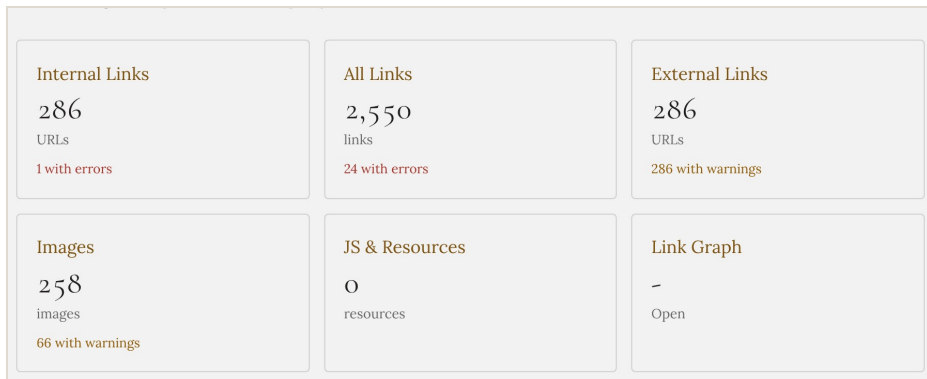
Filter chip	Level	What it means, and what to do
Thin Content (Under 200 Words)	Warning	Very little text. Pages with almost nothing to read rarely rank. Add useful content, or merge the page into another. The 200 can be changed in Settings, Check Thresholds.
Exact Duplicate	Error	The text is identical to another page's. Keep one, then redirect the others or point their canonical at it.
Near Duplicate (90%+ Similar)	Warning	The text is almost the same as another page's. Make each page say something of its own.

Filter chip	Level	What it means, and what to do
Low Text-to-HTML Ratio (Under 10%)	Notice	The page is mostly code with little visible text.
Hard to Read (Flesch Score Under 50)	Notice	Long sentences and long words. Shorter sentences and simpler words help most readers. Each heading, list item and short line counts as its own sentence, the way Screaming Frog counts, so pages built from short lines score as easy to read.

5. The Links & Media tab

How pages link to each other, to other sites, and the pictures and files they use.

WHERE TO FIND IT A report, Links & Media tab (press the tab to see its main page)



The main page of the Links & Media tab: one card for each section.

Internal Links

WHERE TO FIND IT A report, Links & Media tab, Internal Links in the left list

Internal Links

Q Filter by URL

286 of 286 URLs

All 286

● Orphan Page (In Sitemap, No Inlinks) 1

● Only One Internal Inlink 6

Show 3 empty filters

The top of Internal Links: its heading, the filter box and the filter chips.

Links between the pages of your own site, counted for each page. Links are how visitors and crawlers move around, so a page with few links in is hard to find.

URL	ISSUES	INLINKS	NOFOLLOW IN	OUTLINKS	UNIQUE ANCHORS	DEPTH
/	-	284	0	33	33	0
/about/	-	284	0	4	4	1
/faq/	-	284	0	4	4	1
/espresso-beans/	-	11	0	14	14	1

The table in Internal Links: the first rows. Each row is one URL.

Column	What it shows
URL	The address.
Issues	The checks this row fails, shown as small tags coloured by level. Each tag has the same name as a filter chip. A dash means the row passes every check in this section.
Inlinks	How many pages link to it. Zero is shown in red.
Nofollow in	How many of those links carry nofollow, which tells search engines not to pass strength.
Outlinks	How many internal links it holds.
Unique anchors	How many different link texts point at it.
Depth	Clicks from the homepage.

Filter chip	Level	What it means, and what to do
Orphan Page (In Sitemap, No Inlinks)	Error	The sitemap lists the page but no page links to it. Visitors cannot reach it by clicking. Link to it from a related page.
Only One Internal Inlink	Warning	Only one page links here. If that link goes, the page is lost. Add more links from related pages.
Over 150 Internal Outlinks	Notice	Too many links on one page dilute each one and are hard to use.
Nofollow Internal Inlinks	Notice	Some links to this page carry nofollow, so they pass no strength. Remove nofollow from your own links unless there is a reason.
Links to the Non-Preferred Host (Www or Non-Www)	Warning	The page links to the other form of your address (www or non-www), which then redirects. Link to the preferred form.

All Links

WHERE TO FIND IT **A report, Links & Media tab, All Links in the left list**

All Links

Every link found in the crawled HTML, with the page it is on. Internal links show the status the crawler got for the destination; external links were each requested once.

2,550 of 2,550 links

All 2,550

Broken Internal Links (4xx, 5xx, Failed) 24

Broken External Links (404, 410, 5xx, Failed) 594

Links to Redirects 285

Non-Descriptive Anchor Text (Internal Links) 1

To URLs Blocked by robots.txt 1

Internal 1,956

External 594

Show 3 empty filters

The top of All Links: its heading, the filter box and the filter chips.

Every link found in the crawled pages, one row per link, with the page it sits on. Internal links show the status the crawler got for the destination. External links show a status only when Check External Links was on.

SOURCE PAGE	ISSUES	DESTINATION	STATUS	ANCHOR TEXT
/	-	/	200	Home
/	-	/about/	200	About
/	-	/faq/	200	FAQ
/	Links to Redirects	/old-about/	301	Old about

The table in All Links: the first rows. Each row is one link.

Column	What it shows
Source page	The page the link is on.
Issues	The checks this row fails, shown as small tags coloured by level. Each tag has the same name as a filter chip. A dash means the row passes every check in this section.
Destination	Where the link goes.
Status	The destination's reply code.

Column	What it shows
Anchor text	The clickable words.
Type	Internal (same site) or External.
Follow	Follow, or Nofollow.

Filter chip	Level	What it means, and what to do
Broken Internal Links (4xx, 5xx, Failed)	Error	A link on your site that leads to a missing or broken page. Fix the link or restore the page.
Broken External Links (404, 410, 5xx, Failed)	Warning	A link to another website that is gone (404 or 410), erroring (5xx) or does not answer. A failure is tried a second time before it is reported. Update the link or remove it.
External Links That Block the Crawler (403, 429, 999)	Notice	The other website answered 403, 429 or 999, which usually means it turns crawlers away. The page may be fine for visitors, so these are not counted as broken. Open a few in a browser to be sure.
Links to Redirects	Warning	The link goes to an address that redirects. Point it at the final address.

Filter chip	Level	What it means, and what to do
Non-Descriptive Anchor Text (Internal Links)	Notice	Link text such as click here or read more says nothing about where it goes. Use words that describe the page.
Empty Anchor Text	Notice	A link with no words, such as an image link without alt text. Add text or alt text.
Nofollow Internal Links	Notice	An internal link marked nofollow. Usually not needed.
To URLs Blocked by robots.txt	Info	Links to addresses that robots.txt keeps crawlers out of.
Internal	Info	Only links within your own site.
External	Info	Only links to other websites.

External Links

WHERE TO FIND IT A report, Links & Media tab, External Links in the left list

External Links

Each external link was requested once to check its status.

Q Filter by URL

286 of 286 URLs

All 286

● Links to a Broken External URL 286

Show 4 empty filters

The top of External Links: its heading, the filter box and the filter chips.

Links from your pages to other websites, counted for each page. Turn on Check External Links in Crawl Options to find the broken ones.

URL	ISSUES	EXTERNAL	NOFOLLOW	BROKEN	BLOCKED	REDIRECTED	DOMAIN
/	Links to a Broken External URL	2	0	2	0	0	twitter.com link.invalid
/about/	Links to a Broken External URL	2	0	2	0	0	twitter.com link.invalid
/faq/	Links to a Broken External URL	2	0	2	0	0	twitter.com link.invalid
/espresso-beans/	Links to a Broken External URL	2	0	2	0	0	twitter.com link.invalid

The table in External Links: the first rows. Each row is one URL.

Column	What it shows
URL	The page on your site.
Issues	The checks this row fails, shown as small tags coloured by level. Each tag has the same name as a filter chip. A dash means the row passes every check in this section.
External	How many links to other websites it holds.
Nofollow	How many are nofollow.

Column	What it shows
Broken	How many lead to a missing or broken page.
Blocked	How many lead to a website that answered 403, 429 or 999 and may be turning crawlers away.
Redirected	How many lead to an address that redirects.
Domains	How many different websites it links to.

Filter chip	Level	What it means, and what to do
Links to a Broken External URL	Warning	The page links to a broken address on another website. Update or remove the link.
Links to External Sites That Block the Crawler (403, 429, 999)	Notice	The page links to a website that answered 403, 429 or 999, which usually means it turns crawlers away. These links are probably fine for visitors, so they are not counted as broken.
High Number of External Links (Over 10)	Warning	The page has more than ten followed links to other websites, counting a link again each time it appears (in the header and the footer, for example). Check that they are all useful and trusted.

Filter chip	Level	What it means, and what to do
Redirected External Links	Notice	The page links to an address that now redirects. Update the link to the final address.
No External Links	Info	Pages that link to no other website. Not a problem, although linking to good sources can help readers.

Images

WHERE TO FIND IT **A report, Links & Media tab, Images in the left list**

Images

Each image is listed once, with the pages it appears on.

258 of 258 images

All 258

● Missing Alt Attribute 66

● Missing Width and Height Attributes 1

● Over 100 KB 249

● Not WebP, AVIF or SVG 257

Show 9 empty filters

The top of Images: its heading, the filter box and the filter chips.

Every image file found, one row per image, with the number of pages that use it. Sizes need Fetch Image Sizes on in Crawl Options.

IMAGE	ISSUES	FOUND ON	PAGES	FORMAT	SIZE	DIMENSIONS	ALT TEXT
/img/logo.png	Not WebP, AVIF or SVG	/ +284 more	285	PNG	114 KB	100×30	Harbour
/img/icon.svg	Missing Alt Text Missing Width and Height Attributes	/ +284 more	285	SVG	114 KB	-	missing
/img/p24.jpg	Missing Alt Text Not WebP, AVIF or SVG	/shop/coffee-24/	1	JPG	105 KB	400×300	missing
/img/p48.jpg	Missing Alt Text Not WebP, AVIF or SVG	/shop/coffee-48/	1	JPG	105 KB	400×300	missing

The table in Images: the first rows. Each row is one image.

Column	What it shows
Image	The image address.
Issues	The checks this row fails, shown as small tags coloured by level. Each tag has the same name as a filter chip. A dash means the row passes every check in this section.

Column	What it shows	
Found on	A page that uses it.	
Pages	How many pages use it.	
Format	The file type, such as jpg, png or webp.	
Size	The file size.	
Dimensions	The width and height.	
Alt text	The text that describes the image.	

Filter chip	Level	What it means, and what to do
Missing Alt Attribute	Warning	The image has no alt attribute at all, so screen readers may read out the file name. Add one to every image: describe it if it carries meaning, or leave it empty (alt="") if it is only decoration.
Empty Alt Text	Notice	The alt attribute is there but empty. That is right for pictures that are only decoration, and wrong for pictures that carry meaning. Check each one.
Generic or Duplicate Alt Text	Notice	Alt text such as image, photo or logo, or the same alt text on several different images. Describe what each picture shows.

Filter chip	Level	What it means, and what to do
Missing Width and Height Attributes	Notice	Without sizes the page jumps as images load. Add width and height.
Over 100 KB	Warning	A heavy image slows the page. Compress it or use a smaller one. The 100 can be changed in Settings, Check Thresholds.
Broken Image	Error	The image address returns an error, or answers with a web page instead of a picture (a missing image sent to the home page), so visitors see a gap. Fix the address or restore the file. The image is followed through any redirect first, and a broken image is not also judged for alt text, size or format.
Image Address Redirects	Notice	The address in the page sends the browser somewhere else before the picture loads. Point the page at the final address.
Blocked by robots.txt	Warning	robots.txt keeps crawlers away from the image, so it cannot appear in image search.
Not WebP, AVIF or SVG	Notice	Old formats (jpg, png, gif). Newer formats are smaller for the same quality.
HTTP Address (Not HTTPS)	Warning	The image is loaded from an http:// address. Browsers may block or warn about it. Use the https:// address.
Uppercase in URL	Notice	Capital letters in the image address. Servers may treat the upper and lower case forms as different files. Prefer lowercase.

Filter chip	Level	What it means, and what to do
Underscores in URL	Notice	An underscore in the image address. Google reads a hyphen as a word break but not an underscore.
Non-ASCII Characters in URL	Notice	Letters such as é or other scripts in the image address. They work, but become %-codes when shared.

JS & Resources

WHERE TO FIND IT A report, Links & Media tab, JS & Resources in the left list

JS & Resources

CSS, JavaScript and font files linked in the HTML. Files that scripts load later are not listed.

The top of JS & Resources: its heading, the filter box and the filter chips.

The style, script and font files linked in each page. Files that scripts load later are not listed.

Column	What it shows
URL	The file address.
Issues	The checks this row fails, shown as small tags coloured by level. Each tag has the same name as a filter chip. A dash means the row passes every check in this section.
Type	CSS, JS or Font.
Size	The file size.

Column	What it shows
Status	The reply code.
Party	First-party (your site) or Third-party (another website).

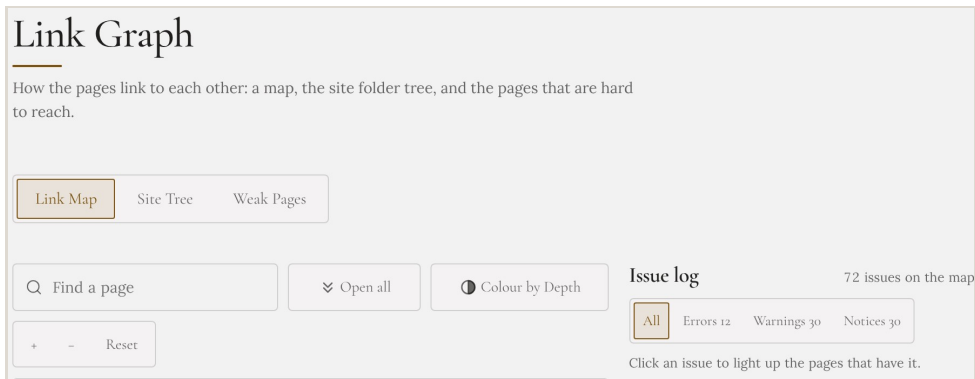
Filter chip	Level	What it means, and what to do
Resource Returns Error	Error	A style or script file is missing or broken, so the page may look or work wrong. Fix the address or restore the file.
Blocked by robots.txt	Warning	Search engines cannot read the file, so they may not see the page as visitors do. Allow it in robots.txt.
Third-Party	Info	Files that come from other websites, such as analytics, fonts and ads. They can slow the page and are outside your control.
JavaScript Over 250 KB	Notice	A heavy script slows the page. Split it, remove what is not used, or load it later.
HTTP Address (Not HTTPS)	Warning	The file is loaded from an http:// address. Browsers may block or warn about it. Use the https:// address.
Uppercase in URL	Notice	Capital letters in the file address. Servers may treat the upper and lower case forms as different files. Prefer lowercase.

Filter chip	Level	What it means, and what to do
Underscores in URL	Notice	An underscore in the file address. Google reads a hyphen as a word break but not an underscore.
Non-ASCII Characters in URL	Notice	Letters such as é or other scripts in the file address. They work, but become %-codes when shared.

Link Graph

A map of how the pages link to each other, and a log of what is wrong with them.

WHERE TO FIND IT A report, Links & Media tab, Link Graph in the left list



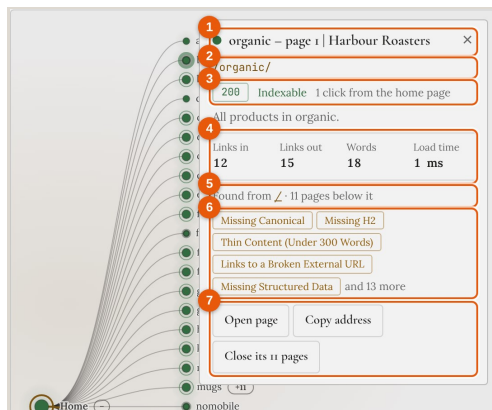
The Link Graph with its three views and the tools above the map.

Link Graph has no filter chips. It has three views, and the Link Map has its own tools. Part 1 shows every button; this part explains what the picture and the numbers mean.

View	What it shows
Link Map	A mind map of the site. Home is on the left. Each row is a page, and the pages under it open to the right when you click it. Lines join a page to the page it was found from. It draws up to 2,000 pages, the ones nearest the home page first.
Site Tree	The pages grouped by folder, so you can see the shape of the site.
Weak Pages	Orphan pages, pages with only one link in, pages far from the homepage, and the most-linked pages.

In the Link Map	What it means
Dot colour, Colour by Status	Green is a page that loaded well (2xx). Amber is a redirect (3xx). Red is an error, a server failure or a page that did not answer.
Dot colour, Colour by Depth	Depth is the number of clicks from the home page. Gold is Home, then blue, green, purple, orange, and grey for five or more clicks.
Hollow dot, and the Not linked row	A page that is in the XML sitemap but that no page links to (an orphan). Search engines may find it only through the sitemap.
A ring around a dot	The page has pages under it. Home has a gold ring that breathes slowly.

In the Link Map	What it means
The number on a closed row	How many pages are hidden under it. Red means some of them are broken, amber means some are redirected, grey means all are fine.
Gold dashed line	The click path from Home to the page you picked or the page you are pointing at.
Small lights moving along the lines	Only decoration, to show how the pages connect. They are switched off when your computer asks for less motion.
Dimmed rows	Pages without the issue or search you lit. The pages that match glow with a pulsing ring.



The summary of a page, opened by clicking its row.

In the summary	What it means
Title and address	The page title, then its address. The address opens the real page in a new tab.
Status line	The response code, then Indexable or the reason it is not (for example Noindex or Redirected), then how many clicks it is from the home page.
Description	The meta description, as search engines read it, or No meta description.

In the summary	What it means
Links in, Links out	How many links on other pages point to this page, and how many links this page has to pages of the same site.
Words, Load time	The number of words in the main text, and how many milliseconds the page took to load.
Found from	The page it hangs under on the map. If several pages link to it, the one in the same folder wins. Click it to jump there.
Issue tags	The worst checks the page fails, red for errors, amber for warnings, blue for notices. Click one to light up every page with the same issue.

In the Issue log	What it means
A card	One check that fails on at least one page. It shows the check, the tab and section it belongs to, how many pages have it, and its level.
The triangle and the number	The level (Error, Warning or Notice) and the number of pages. Errors are listed first.
All, Errors, Warnings, Notices	Show every card or only one level.

In the Issue log	What it means
Clicking a card	Lights up the pages with that issue on the map and opens the first one. The strip above the map says which issue is lit and has arrows to go from page to page.
Issues without a page	Checks about files rather than pages, such as an oversized script, are not in the log. They are still listed in the section they belong to.

6. The Markup tab

Hidden labels in a page's code: what the page is, how it looks when shared, and which language it is in.

WHERE TO FIND IT A report, Markup tab. Its three sections are in the left list

Markup

Structured data, social tags and language versions. The Export button here saves all 3 sections together; open a section to export just that one.

Structured Data

286

URLs

43 with errors

Open Graph & Twitter

286

URLs

286 with warnings

Hreflang

286

URLs

30 with errors

The Markup tab: one card for each of its three sections.

Markup is information written into a page's code for machines to read. Visitors never see it on the page, but search engines, social networks and AI tools use it. A page can be perfect to look at and still have no markup at all. The tab number is the count of URLs with a problem in any of its three sections.

Section	What it is	Who reads it	What it changes
Structured Data	Labels that say what the page is: an article, a product, a business, an FAQ and so on.	Search engines and AI tools	Whether the page can appear with extras in results, such as star ratings, prices and questions, and how well it is understood.
Open Graph & Twitter	Tags that choose the title, description and picture of the preview when the page is shared.	Facebook, LinkedIn, WhatsApp, X and similar	How the page looks when someone shares it, and how many people click.
Hreflang	Tags that say which language and country each version of a page is for.	Search engines	Which version of the page each visitor is shown. Only needed on sites with more than one language.

Structured Data

Labels that tell machines what a page is.

WHERE TO FIND IT **A report, Markup tab, Structured Data in the left list**

Structured Data

Checked against schema.org rules. Google's Rich Results Test has no public API, so rich-result eligibility isn't confirmed.

286 of 286 URLs

All 286

● Missing Structured Data 198

● Validation Errors 43

● Product Page Missing Product Schema 170

● No BreadcrumbList on Pages 2+ Clicks Deep 256

Show 4 empty filters

The top of Structured Data: its heading, the filter box and the filter chips.

Structured data is a set of labels in a page's code that say what the page is: an article, a product, a business, a recipe, an FAQ and so on. It uses a shared vocabulary called schema.org. The usual way to write it is JSON-LD, a block of data placed in the page. Search engines read it to understand the page and, for some types, to show richer results such as star ratings, prices and questions.

A row of chips above the table lists each schema type found, such as Organization, Article, Product or BreadcrumbList, with how many pages use it. The check is made against schema.org rules. Google's Rich Results Test has no public interface, so the app cannot confirm that a page will get a rich result.

WHAT STRUCTURED DATA LOOKS LIKE (JSON-LD)

```
<script type="application/ld+json">
{ "@context": "https://schema.org",
  "@type": "Organization",
  "name": "Harbour Roasters",
  "url": "https://harbourroasters.example/",
  "logo": "https://harbourroasters.example/logo.png" }
</script>
```

That block says: this is an Organization called Harbour Roasters, with this address and this logo. Visitors do not see it. @type is the schema type. The other words are the details that type can have.

What comes inside structured data

Each page can carry one or more types. These are the ones the report looks for and the ones you will see in the Types column and in the row of chips above the table.

Schema type	What it says	Where it belongs
Organization, LocalBusiness, Corporation	Who runs the site: name, logo, address, phone and links to its social profiles.	The homepage
WebSite	The site itself, and its own search box if it has one.	The homepage
Person	A person, for a personal website or an author page.	The homepage or an author page
Article, BlogPosting, NewsArticle	An article: headline, author, published and modified dates, picture.	Blog and news pages
Product, Offer, Review, AggregateRating	A product: its name, price and availability, and its reviews and star rating.	Product pages
BreadcrumbList	The trail that shows where the page sits in the site, such as Home, Shop, Shoes.	Pages two or more clicks from the homepage

Schema type	What it says	Where it belongs
FAQPage	A list of questions with their answers.	Pages that really answer questions
HowTo	Step-by-step instructions.	Pages that teach a task
Other types	Event, Recipe, VideoObject, ImageObject and many more are valid. The report lists any type it finds, and checks the ones above.	Where they fit

The messages in the Validation column

The Validation column is blank when the data passes. Otherwise it lists what is wrong. These are the checks, and what to do.

Message	What is wrong	What to do
Invalid JSON-LD (does not parse)	The block has a typing mistake, such as a missing comma or quote, so no machine can read it.	Paste it into a JSON checker, fix the mistake, and publish again.

Message	What is wrong	What to do
Product: missing "offers", "review" or "aggregateRating"	A Product with none of the details Google needs to show a price or stars.	Add the price (offers), or reviews, or the rating.
Article: missing "headline"	An article with no title in its data.	Add the headline.
Article: missing "datePublished"	An article with no publish date in its data.	Add the date.
FAQPage: question without "acceptedAnswer" text	A question in the FAQ has no answer text.	Give each question an answer, or remove the question.
BreadcrumbList: empty "itemListElement"	A breadcrumb trail with nothing in it.	Add the steps of the trail.
Organization: "logo" is not an absolute URL	The logo address is written without the full https:// start.	Write the full address of the logo.

URL	ISSUES	FORMAT	TYPES	VALIDATION
/	-	JSON-LD	Organization	Valid
/about/	Missing Structured Data	missing	-	-
/faq/	-	JSON-LD	FAQPage	Valid
/espresso-beans/	Missing Structured Data	missing	-	-

The table in Structured Data: the first rows. Each row is one URL.

Column	What it shows
URL	The address.
Issues	The checks this row fails, shown as small tags coloured by level. Each tag has the same name as a filter chip. A dash means the row passes every check in this section.
Format	JSON-LD (recommended), Microdata or RDFa. The last two are older ways of writing the same labels.
Types	The schema types found on the page.
Validation	Problems found in the data, or OK. For example: JSON-LD that does not parse; a Product without offers, review or rating; an Article without a headline or publish date; a FAQ question with no answer; an empty breadcrumb list; an Organization logo that is not a full address.

Filter chip	Level	What it means, and what to do
Missing Structured Data	Warning	The page has no labels at all. Add the types that fit it.
Validation Errors	Error	The labels have mistakes that stop them being used. The Validation column says what is wrong.
Homepage Missing Organization or WebSite Schema	Warning	The homepage should say who the business is. Add Organization, WebSite, LocalBusiness, Person or Corporation.
Article Page Missing Article Schema	Warning	A page that looks like a blog or news article (its address holds /blog/, /news/ or similar, or it says it is an article) has no Article or BlogPosting data.
Product Page Missing Product Schema	Warning	A page that looks like a product page (its address holds /product/ or /shop/, or it says it is a product) has no Product data.
No BreadcrumbList on Pages 2+ Clicks Deep	Notice	Breadcrumb data shows where the page sits in the site (Home, then Shop, then Shoes). Useful for pages two or more clicks deep.

Filter chip	Level	What it means, and what to do
Phone Link but No LocalBusiness or Organization Schema	Notice	The page has a telephone link but no business data. Add LocalBusiness or Organization.
Microdata or RDFa (Not JSON-LD)	Notice	Older ways of writing schema. JSON-LD is easier to keep correct and is the format Google recommends.

Open Graph & Twitter

The preview a page gets when it is shared.

WHERE TO FIND IT A report, Markup tab, Open Graph & Twitter in the left list

The screenshot shows a web interface for filtering Open Graph and Twitter data. At the top, the heading "Open Graph & Twitter" is displayed with a small orange underline. Below the heading is a search bar with a magnifying glass icon and the placeholder text "Filter by URL". To the right of the search bar, the text "286 of 286 URLs" is shown. Below the search bar, there are several filter chips. The first chip is "All" with a count of "286" and is highlighted with an orange border. The other chips are "Missing og:title" (30), "Missing og:description" (286), "Missing og:image" (286), "Missing og:url" (286), "Missing og:type" (286), and "Missing Twitter/X Card" (286). A "Show 2 empty filters" button is also present.

Open Graph & Twitter

Q Filter by URL 286 of 286 URLs

All 286 Missing og:title 30 Missing og:description 286 Missing og:image 286 Missing og:url 286 Missing og:type 286 Missing Twitter/X Card 286 Show 2 empty filters

The top of Open Graph & Twitter: its heading, the filter box and the filter chips.

Open Graph tags and Twitter (X) card tags control how a page looks when someone shares it on Facebook, LinkedIn, WhatsApp, X and similar places: the title, the short description and the picture. Without them, the other site guesses, and the preview is often poor.

WHAT THESE TAGS LOOK LIKE

```
<meta property="og:title" content="Fresh coffee, roasted weekly" />
<meta property="og:description" content="Small-batch coffee delivered to your door." />
<meta property="og:image" content="https://harbourroasters.example/share.jpg" />
<meta property="og:url" content="https://harbourroasters.example/" />
<meta property="og:type" content="website" />
<meta name="twitter:card" content="summary_large_image" />
```

When someone pastes the address into a message, the other site reads these tags and draws a card: a picture, a bold title and a line of text. Open Graph (og:) is read by almost everyone. Twitter, now X, reads its own twitter: tags and falls back to the Open Graph ones. In the table, Yes means the tag is there and No means it is missing. For og:type and the Twitter/X card, the column shows the value found (such as website or summary_large_image), or a small red missing tag.

URL	ISSUES	OG:TITLE	OG:DESCRIPTION	OG:IMAGE	OG:URL	OG:TYPE	TWITTER
/	Missing og:title Missing og:description Missing og:image Missing og:url Missing og:type Missing Twitter/X Card	No	No	No	No	missing	missin
/about/	Missing og:title Missing og:description Missing og:image Missing og:url Missing og:type	No	No	No	No	missing	missin

The table in Open Graph & Twitter: the first rows. Each row is one URL.

Column	What it shows
URL	The address.
Issues	The checks this row fails, shown as small tags coloured by level. Each tag has the same name as a filter chip. A dash means the row passes every check in this section.
og:title	The title shown in the preview.

Column	What it shows
og:description	The short text in the preview.
og:image	The picture in the preview.
og:url	The official address of the page.
og:type	What kind of thing the page is, such as website, article or product.
Twitter/X card	The card type for X, for example summary_large_image.

Filter chip	Level	What it means, and what to do
Missing og:title	Warning	Without it the shared preview may show a poor title.
Missing og:description	Warning	Without it the shared preview may show no text, or the wrong text.
Missing og:image	Warning	The preview has no picture. Shared posts with a picture get far more attention.
Missing og:url	Notice	The official address for the page. Add it so shares are counted together.
Missing og:type	Notice	Says what kind of thing the page is (website, article, product).
Missing Twitter/X Card	Notice	X uses Open Graph tags where it can. Add a twitter:card tag to choose a large preview.

Filter chip	Level	What it means, and what to do
og:url Differs from the Canonical URL	Warning	The two disagree about the main address of the page. Make them the same.
Twitter/X Card Without twitter:image	Notice	A card with no picture, and no og:image to fall back on.

Hreflang

Which language or country a page is for.

WHERE TO FIND IT **A report, Markup tab, Hreflang in the left list**

Hreflang

286 of 286 URLs

All 286

Missing Return Tag 30

Invalid Language or Country Code 0

Hreflang URL Returns Error 0

Missing X-Default 30

Conflicting Hreflang 0

No Hreflang on Page 256

The top of Hreflang: its heading, the filter box and the filter chips.

Hreflang tags tell search engines which language or country version of a page to show to which visitor, for example en-gb for English in the United Kingdom, fr-ca for French in Canada, or de for German. They only matter for sites that have several language versions. Every version must list all the others, including itself, and each must point back.

WHAT THESE TAGS LOOK LIKE

```
<link rel="alternate" hreflang="en-gb" href="https://example.com/en-gb/" />  
<link rel="alternate" hreflang="fr-ca" href="https://example.com/fr-ca/" />  
<link rel="alternate" hreflang="x-default" href="https://example.com/" />
```

Code	Means
en	English, for everyone
en-gb	English for the United Kingdom
fr-ca	French for Canada
de	German
x-default	The version to show to anyone who matches none of the others

A language code is two letters (en, fr, de). A country can follow after a hyphen (en-gb, fr-ca). Every language version of a page should list all the versions, itself included, and every version must point back at the others. That two-way link is what Return tags OK checks.

URL	ISSUES	LANGUAGES	X-DEFAULT	RETURN TAGS OK
/	No Hreflang on Page	-	No	No
/about/	No Hreflang on Page	-	No	No
/faq/	No Hreflang on Page	-	No	No
/espresso-beans/	No Hreflang on Page	-	No	No

The table in Hreflang: the first rows. Each row is one URL.

Column	What it shows
URL	The address.
Issues	The checks this row fails, shown as small tags coloured by level. Each tag has the same name as a filter chip. A dash means the row passes every check in this section.
Languages	The language and country codes the page lists.

Column	What it shows	
x-default	Whether it names a default version for everyone else.	
Return tags OK	Whether every page it lists points back at it.	

Filter chip	Level	What it means, and what to do
Missing Return Tag	Error	Page A says B is its French version, but B does not point back to A. Search engines then ignore the pair. Every version must list every other.
Invalid Language or Country Code	Error	A code that is not a real language or country, such as en-UK. The right one is en-GB.
Hreflang URL Returns Error	Error	A listed language version is missing or broken.
Missing X-Default	Warning	x-default names the version to show to visitors who match none of the others. Add one.
Conflicting Hreflang	Warning	The same language code points at two different addresses. Keep one.
No Hreflang on Page	Notice	The page lists no other languages. Normal for a site with one language.

7. The AI Search tab

How ready the site is for answer boxes, voice assistants and AI tools. These checks never change the health score.

AI Search

How ready the site is for answer engines and AI tools. These checks do not change the health score. The Export button here saves all 3 sections together; open a section to export just that one.

AEO

286

URLs

256 with warnings

GEO

286

URLs

No issues

AI Visibility

-

Not run yet

The AI Search tab: its three sections.

AEO

Answer engine optimization.

WHERE TO FIND IT **A report, AI Search tab, AEO in the left list**

AEO

Answer engine optimization: how easily answer boxes and voice assistants can lift a clear answer from each page.

Q Filter by URL

286 of 286 URLs

All 286

● Missing JSON-LD 198

● No Direct Answer in First Paragraph 255

● No Author Byline 286

● No FAQ or HowTo Schema 285

● No Question-Style Headings 284

The top of AEO: its heading, the filter box and the filter chips.

AEO is about answer boxes and voice assistants: can they lift a clear answer straight from the page? Pages that open with a short, direct answer, name their author, and use question headings or FAQ data are easier to quote.

URL	ISSUES	JSON-LD	DIRECT ANSWER	AUTHOR	FAQ SCHEMA	QUESTION HEADINGS
/	No Author Byline No FAQ or HowTo Schema No Question-Style Headings	Yes	Yes	No	No	No
/about/	Missing JSON-LD No Author Byline No FAQ or HowTo Schema	No	Yes	No	No	Yes
/faq/	No Author Byline	Yes	Yes	No	Yes	Yes
/espresso-beans/	Missing JSON-LD No Direct Answer in First Paragraph No Author Byline	No	No	No	No	No

The table in AEO: the first rows. Each row is one URL.

Column	What it shows
URL	The address.
Issues	The checks this row fails, shown as small tags coloured by level. Each tag has the same name as a filter chip. A dash means the row passes every check in this section.
JSON-LD	Whether the page has JSON-LD structured data.
Direct answer	Whether the first real paragraph is a short answer.
Author	Whether an author is named.
FAQ schema	Whether the page has FAQ or HowTo data.
Question headings	Whether any heading is a question.

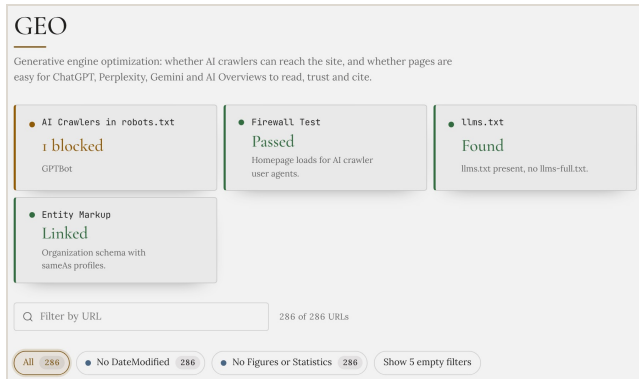
Filter chip	Level	What it means, and what to do
Missing JSON-LD	Warning	No JSON-LD structured data on the page. Microdata and RDFa do not count here. Add JSON-LD.
No Direct Answer in First Paragraph	Warning	The first real paragraph is not an answer of 15 to 80 words. Open the page with a short, direct answer to its main question.
No Author Byline	Notice	No author found (a byline, an author tag, or an author in the schema). Named authors build trust.

Filter chip	Level	What it means, and what to do
No FAQ or HowTo Schema	Notice	The page has no FAQ or HowTo data. Add it where the page really answers questions or gives steps.
No Question-Style Headings	Notice	No H2, H3 or H4 is written as a question (ending in a question mark, or starting with what, how, why and so on).

GEO

Generative engine optimization.

WHERE TO FIND IT **A report, AI Search tab, GEO in the left list**



The top of GEO: its heading, the filter box and the filter chips.

GEO is about AI tools such as ChatGPT, Perplexity, Gemini and AI Overviews: can their crawlers reach the site, and are the pages easy to read, trust and cite? The top of the section also shows which AI crawlers robots.txt allows and whether an llms.txt file exists.

URL	ISSUES	IN RAW HTML	UPDATED	CITATIONS	STATISTICS	LISTS / TABLES
/	No DateModified No Figures or Statistics	Yes	missing	Yes	No	Yes
/about/	No DateModified No Figures or Statistics	Yes	missing	Yes	No	No
/faq/	No DateModified No Figures or Statistics	Yes	missing	Yes	No	No
/espresso-beans/	No DateModified No Figures or Statistics	Yes	missing	Yes	No	Yes

The table in GEO: the first rows. Each row is one URL.

Column	What it shows		
URL	The address.		
Issues	The checks this row fails, shown as small tags coloured by level. Each tag has the same name as a filter chip. A dash means the row passes every check in this section.		
In raw HTML	Whether the main text is in the page's code, not built later by scripts.		
Updated	How long ago the page says it was last changed.		
Citations	Whether it links to sources.		
Statistics	Whether it contains figures.		
Lists / tables	Whether it has lists or tables.		
Author	Whether an author is named.		
Filter chip	Level	What it means, and what to do	
Main Content Needs JavaScript	Error	The page's code holds almost no text and looks like an app shell. AI crawlers mostly read the raw code and would see an empty page. Serve the text in the page itself.	
Snippets Restricted (Nosnippet / Max-Snippet)	Warning	The page limits how much text may be quoted, so AI tools may quote less or nothing.	

Filter chip	Level	What it means, and what to do
Not Updated in 12+ Months	Warning	The page says it was last changed over a year ago. AI tools prefer fresh pages.
No DateModified	Notice	No last-changed date found. Add one, in the structured data.
No Outbound Citations to Sources	Notice	The page links to no outside sources. Social sites do not count. Linking to good sources helps people and AI tools trust it.
No Figures or Statistics	Notice	No percentages, large counts or similar figures. Facts and numbers are easier to quote.
Long Sections Without Subheadings	Notice	More than 900 words with fewer than one H2 or H3 for every 450 words. Break the text up with subheadings.

AI Visibility

Asks an AI whether it knows and cites the site. Needs the owner's Gemini key.

WHERE TO FIND IT **A report, AI Search tab, AI Visibility in the left list**

Questions to Ask, One per Line (up to 10)

What is Harbour Roasters?

Is Harbour Roasters trustworthy? What do people say about it?

Where can I find the best cold-brew?

Where can I find the best single-origin?

Where can I find the best blends?

Where can I find the best organic?

Where can I find the best fair-trade?

Where can I find the best capsules?

Where can I find the best hot-chocolate?

Where can I find the best grinders?

Run AI Visibility Check

Takes about 10 seconds per question. Gemini's free plan has a daily limit, so a check may stop early.

AI Visibility before it is run: the questions and the button that starts the check.

AI Visibility asks Google's Gemini up to 10 questions about the website, using Google Search for its answers, and shows whether the site is named and which sources are cited. It only runs when you press Run AI Visibility Check, the result is saved with the crawl, and it has no filter chips. Answers from AI change from day to day, so treat the result as a sample.

Part	What it shows
Questions to Ask	A box with one question per line. The first lines are made from the site's own name. You can edit them. 10 is the most.
Cited	How many answers named or linked to your website.
Answers	Each question, the answer, and the sources used. Your own website is highlighted among the sources.

8. The Crawl Log tab

Your own rules, and the record of every request.

Crawl Log

Your own extraction and search rules, and the list of requests the crawler made. The Export button here saves all 3 sections together; open a section to export just that one.

Custom Extraction

0
URLs

Custom Search

286
URLs
No issues

Crawl Log

310
requests listed

The Crawl Log tab: its three sections.

Custom Extraction

Pulls your own values out of every page.

WHERE TO FIND IT A report, Crawl Log tab, Custom Extraction in the left list. The rules are set in Crawl Options before the crawl

Custom Extraction

Rules:

Custom Extraction when no rules were written: the heading and the Rules line, which would list your rules.

Custom Extraction shows the result of the extraction rules you wrote in Crawl Options before the crawl, for example the author or the price on each page. There is one row for each HTML page and one column for each rule. The line under the heading lists your rules. Without rules this section is empty.

Filter chip	Level	What it means
No Match: (the name of a rule)	Info	One chip for each rule you wrote. It shows the pages where that rule found nothing. For example, No Match: Author shows the pages with no author.

Custom Search

Counts the pages that contain, or lack, some text.

WHERE TO FIND IT A report, Crawl Log tab, Custom Search in the left list. The rules are set in Crawl Options before the crawl

Custom Search

Searches the page source. Rules: Analytics tag (missing, text: googletagmanager) · Free delivery (contains, text: ship free) · Price in text (contains, regex: \d+(\.\d\d)? pounds)

Custom Search with its rules: the Rules line lists them, and each rule gets a column and a filter chip.

Custom Search searches the source code of every HTML page for the text or pattern in each rule you wrote in Crawl Options. There is one row for each page and one column for each rule, showing how many times it was found. The line under the heading lists your rules. Without rules this section is empty.

Filter chip	Level	What it means
Contains: (the name of a rule)	Info	One chip for each rule of the kind contains. It shows the pages where the text was found.
Missing: (the name of a rule)	Info	One chip for each rule of the kind missing. It shows the pages where the text was not found, for example the pages with no Google Analytics code.

Crawl Log

The record of the crawl itself.

WHERE TO FIND IT **A report, Crawl Log tab, Crawl Log in the left list**

```
#10 200 /decaf/ 1.2 KB 1ms
#11 200 /cold-brew/ 1.2 KB 3ms
#12 200 /single-origin/ 1.2 KB 7ms
#13 200 /blends/ 1.2 KB 12ms
#14 200 /organic/ 1.2 KB 1ms
#15 200 /fair-trade/ 1.2 KB 8ms
#16 200 /capsules/ 1.2 KB 12ms
#17 200 /tea/ 1.2 KB 1ms
#18 200 /chai/ 1.2 KB 7ms
#19 200 /hot-chocolate/ 1.2 KB 11ms
#20 200 /grinders/ 1.2 KB 2ms
#21 200 /kettles/ 1.2 KB 5ms
#22 200 /filters/ 1.2 KB 13ms
#23 200 /scales/ 1.2 KB 2ms
#24 200 /mugs/ 1.2 KB 6ms
#25 200 /travel-cups/ 1.2 KB 13ms
```

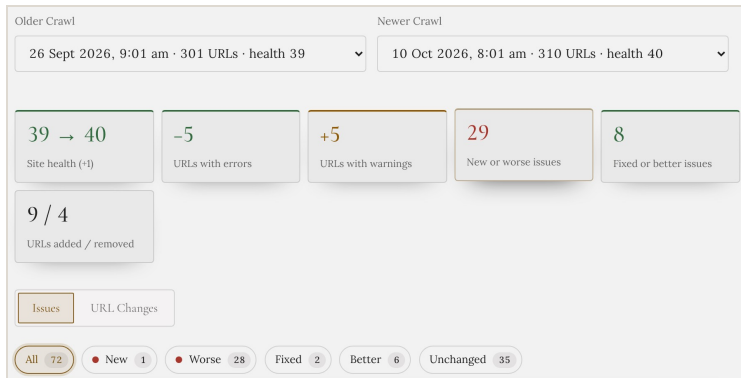
The Crawl Log.

The Crawl Log is a numbered text log of the crawl, with colours: green is fine, yellow is a redirect or a warning, red is an error. It shows the start settings, whether robots.txt and a sitemap were found, one line for each of the first 400 addresses (reply code, address, size, time and where a redirect leads), how many requests were retried and recovered, which kinds of failure happened, and a last line that says whether the crawl finished or was stopped, and how long it took. Use it to understand why a crawl was slow or incomplete. It has no filter chips.

9. The Compare Crawl tab

Two crawls of the same website, side by side.

WHERE TO FIND IT A report, Compare Crawl tab. Or press See What Changed on the Overview



Compare Crawl: the two crawls, the tiles, and the filter chips.

Pick an older and a newer crawl of the same website. The tiles show the change in site health, URLs with errors, URLs with warnings and the number of new, worse, fixed and better issues. Below, two views are shown with their own filter chips: Issues and URL Changes.

Issues

Each check, compared between the two crawls.

Filter chip	Level	What it means
New	Error or Warning	The check found nothing in the older crawl and finds something now. Look at these first.
Worse	Error or Warning	The check found something before, and finds more now.
Fixed	Good news	The check found something before and finds nothing now.
Better	Good news	The check still finds something, but fewer URLs than before.
Unchanged	Info	The check finds the same number as before.

URL Changes

Which addresses were added, removed or changed.

Filter chip	What it means
Added	The address is in the newer crawl and was not in the older one.
Removed	The address was in the older crawl and is gone from the newer one. Check that it is meant to be gone, and redirect it if people may still arrive there.
Status Changed	The server now replies with a different code, for example from 200 to 404.
Title Changed	The page title is different.
Description Changed	The meta description is different.
H1 Changed	The main heading is different.

10. Words used in the data

Plain meanings for the terms in the sections above.

Word	Meaning
Anchor text	The clickable words of a link.
Canonical	The tag that names the main version of a page.
Click depth	How many clicks from the homepage it takes to reach a page.
Core Web Vitals	Google's three measures of page experience: LCP (how fast the main content appears), INP (how fast the page reacts) and CLS (how much it jumps).
Flesch score	A reading-ease score from 0 to 100. Higher is easier. Under 50 is difficult (Screaming Frog calls 30 to 50 Hard and 50 to 60 Fairly Hard).
Header	An instruction the server sends along with a page, before the page itself.
Hop	One step in a chain of redirects.
Hreflang	A tag that says which language or country a page is for.
Inlink and outlink	A link coming into a page, and a link going out of it.
JSON-LD	The usual way to write structured data: a block of data placed in the page.

Word	Meaning
Mixed content	A secure page that loads something over an insecure address.
Noindex and nofollow	Instructions that say do not list this page, and do not follow its links.
Open Graph	Tags that control the preview when a page is shared.
Orphan page	A page that no other page links to.
Redirect	An address that sends visitors to another address.
robots.txt	A file that tells crawlers which parts of a site to stay out of.
Schema.org	The shared vocabulary used for structured data.
Sitemap	A file that lists the pages you want search engines to find.
Status code	The server's three-digit reply: 2xx fine, 3xx moved, 4xx request failed, 5xx server failed.
Viewport	The tag that tells a phone how wide to draw the page.

Need to change this guide?

Part 2 is built from tools/guide-data.js and Part 1 from tools/guide-content.js, with the command: node tools/build-guide.js. The filter names are checked against the app by node tools/check-guide-filters.js.