

SAMPLE REPORT · SEO / GEO / AEO

example-shop.com

4tech — web development agency

TECHNICAL SEO + AI SEARCH AUDIT · 232 pages reviewed · sample

SEO

4/10

Needs improvement

GEO

3/10

Critical

AEO

2/10

Critical

This is a sample report.

The company **Example Company d.o.o.** and the domain **example-shop.com** do not exist. Every URL, measurement, number and quoted string in this document is made up. We never publish our clients' reports, not in part and not anonymized.

Everything else is real: **the methodology, the scope, the document structure, the design and the level of detail are identical** to what a client receives.

This is deliberately the fullest example we could put together. The test domain carries every error we run into most often on Slovenian websites at once, from a wrong `lang` attribute and duplicate titles to an indexed staging copy and empty prices in the schema. No real website has all of them at the same time; yours will have some. The sample is built this way so you see *the full range* of what the audit covers, not just what one client happened to have.

The report runs to **50 pages**. It is delivered in English or Slovenian, as PDF and DOCX, in three to five business days.

Executive summary

example-shop.com is a website that was **built competently and then left alone for five years**. The foundations are right: valid HTTPS, a working XML sitemap with all 218 URLs, clean Slovenian URLs without parameters, WooCommerce already emitting `Product` schema on every product (most Slovenian stores have none at all), and a blog with 24 genuinely expert posts signed by their authors.

Almost no signals are attached to those foundations. **156 of 218 pages have no meta description. 31 products share a single title**, Example Company. **411 of 693 images have no alt text**, and 209 are still named `IMG_20230514_112233.jpg`. The `<html>` element on all 218 pages declares `lang="en-US"`, the WordPress default that nobody ever changed. So the site tells every search engine and every AI assistant that it is American and in English.

Below that sits a layer of structural errors Lighthouse does not report, because each individual tag looks correct to it. **Not one of the 218 pages has exactly one `<h1>` element**. The footer wraps the logo in an `<h1>`, so a product category has up to thirteen. **There are three XML sitemaps, and they contradict each other**. `robots.txt` **blocks all 693 images** from Google Images and denies Googlebot the files it needs to render the pages. **Nine pages load resources over `http://`**, so the browser blocks the gallery on them. And there is no hreflang anywhere, even though the same store exists in three language versions.

Four findings stand out, because a standard audit does not find them.

First, the previous company still lives on the site. The `Organization` block in the structured data on all 218 pages states `"legalName": "Old Company d.o.o."` ["Old Company Ltd."] and a phone number that isn't yours. An old XML sitemap, `/sitemap-staro.xml`, still returns HTTP 200 and offers Google 41 URLs, all of which return 404. Article 4 of the terms and conditions names the previous company as the contracting party. To a search engine, this is someone else's website using your name.

Second, 63 products have an empty price in the schema. For out-of-stock products WooCommerce emits `"price": ""`. An empty string is not a price, so Google discards the whole offer on that product, not just the stock information. The schema is present, which looks like a strength, but for these 63 products it is invalid.

Third, `robots.txt` blocks AI crawlers. It contains five blocks for AI agents. Four of them govern model training and cost you no visibility. The fifth, `0AI-SearchBot`, governs ChatGPT search and removes you from it. Someone once added all five at the same time, most likely without knowing they were two different decisions. The fix is to remove **one line**, not the file.

Fourth, a staging copy of the site is publicly reachable. `test.example-shop.com` runs a complete copy of all 218 pages, with no password, no `noindex`, and older prices. Google has two identical websites and has to choose which one to rank, and it may not choose your production site. On top of that, non-existent URLs return status 200 instead of 404, category filters generate about 2,400 URLs and WordPress adds another 693 attachment pages. In total that is **about 3,600 reachable URLs for 232 real pages**.

None of the 45 items in the priority matrix requires a website rebuild. Thirteen are critical, and nine of those are a change to one setting, one line of code or one file. The gap between what this site is today and what it could rank for is unusually cheap to close. That is the reason this report exists.

Dimension	Score	Status	Key finding
SEO	4/10	Needs improvement	Clean URLs and canonicals hold up. On top of them: 218 pages with more than one <code><h1></code> , two conflicting sitemaps, a <code>robots.txt</code> that blocks images and rendering resources, mixed content, and <code>lang="en-US"</code> .
GEO	3/10	Critical	The schema declares the previous company, and <code>robots.txt</code> shuts AI crawlers out. The entity is both wrong and unreachable.
AEO	2/10	Critical	The blog answers real questions, but none is framed as an answer and no heading is followed by a direct answer.
Total	9/30	—	A well-built store with no signals attached. Settings, copy and one file. Not a rebuild.

Five numbers that sum up the report

Number	What it means
156 / 218	Pages without a meta description. Google writes its own. For products, it builds one from the price line and the Add to cart button.
4,180 ms	Field LCP for real visitors (CrUX, 75th percentile). Google's threshold for "good" is 2,500 ms. The lab data shows 3,640 ms, so real users are worse off.
2 / 8	How many of eight real buyer questions an AI engine answered by citing this site. The other six cited competitors.
232 : 3,600	How many pages the site has , and how many URLs a search engine can reach : filters, attachment pages, archives and a publicly reachable staging copy. For every real page there are about fourteen that aren't.
9 / 13	How many of the thirteen critical items are a change to one setting, one line of code or one file. No new content, no template rework. Two of them fix all 218 pages at once.

Method and scope

All 218 URLs in the XML sitemap were downloaded as raw HTML and parsed programmatically. Meta tags, structured data, heading structure, images, links and response headers were read from the source code. **Nothing is judged by eye, and no number in this report is an estimate.** Every fetch is logged with its HTTP status, time, size and final URL, so redirects and 404 errors show up where they actually are.

Step	What was done	Why
Discovery	<code>robots.txt</code> , <code>sitemap_index.xml</code> and all five child sitemaps. We checked 14 URLs that usually exist on a site of this kind but aren't linked.	To establish the real shape of the site before fetching it.
Fetch	218 pages, one file per URL, with a real browser user agent and rate limiting. We separately checked 47 internal links that led outside the XML sitemap.	A complete audit means every page, not a sample.
JavaScript rendering	40 pages reloaded in real Chrome and compared with the raw HTML.	To separate what is in the code from what JavaScript adds. Search engines don't all see both the same way.
Technical layer	Duplicate titles, placeholder text in production, broken and redirected internal links, unminified resources, image file names, readability, security headers, leftovers of the previous brand.	Findings a developer can fix, not general advice.
Speed	CrUX field data (75th percentile, 28 days) and a real Chrome trace under mobile emulation (412×915, slow 4G, 4× CPU slowdown). Lighthouse for mobile and desktop.	Where both exist, field data wins. The speed section explains why.
AI visibility	Eight questions a buyer would actually ask, put to a search-grounded language model on September 9, 2026, with the cited sources recorded.	A GEO score from the schema alone is inference. This is a measurement.

What the Lighthouse score doesn't prove

Lighthouse gives the homepage **SEO 83**. That score checks whether tags are present and correctly formed. It does not check whether their content is correct, sufficient or true. It can't see that 31 products share one title, that the declared entity is the wrong company, that 63 products carry an empty price, or that `robots.txt` shuts out AI crawlers. All of these are valid tags with the wrong content.

Read the 83 as confirmation that the technical base works. It is not a verdict on rankings and not a substitute for a content review.

How to implement the fixes

This section matters more than any single finding. It decides whether the report produces results or a list of half-finished tasks.

1. **Make a full backup of the files and the database** before you touch anything. Not just the host's automatic backup. A backup you know the location of and know how to restore.
2. **Work on a password-protected staging copy** (for example `test.example-shop.com` , with the `X-Robots-Tag: noindex` header). Never directly on production.
3. **Ship all the fixes at once, not spread over weeks.** This is the most overlooked instruction in this document. When changes go out one a week, Google re-evaluates a site that is still half-finished each time, and you have no point at which you can measure what worked. One release, one date, one before-and-after comparison.
4. **Record baseline values before release:** Search Console (impressions, clicks, average position), PageSpeed Insights for three representative pages, and an export of the current titles and meta descriptions. Without a baseline you can't prove the effect.
5. **After release:** submit the XML sitemap in Search Console, request reindexing of the homepage and the five most important pages, validate the schema with the Rich Results Test and rerun PageSpeed Insights on the same three pages.
6. **Measure after 30 days, not after three.** Recrawling and re-evaluating 218 pages takes weeks. Anything earlier shows noise.

If only part of this list gets done, make it steps 1 and 3.

Pages reviewed

All 218 URLs in the XML sitemap, grouped by template type. The groups are condensed. The full list with HTTP response, size and final URL for every page is attached to the report as `fetchLog.txt`.

Page type	No.	HTTP	Notes
/ — homepage	1	200	Hero slider with 6 full-resolution images. The only page with a complete meta description.
Static pages	7	200	/o-nas/ , /kontakt/ , /storitve/ , /dostava/ , /nacini-placila/ , /splosni-pogoji/ , /zasebnost/ . Two have no <code><h1></code> .
Product categories	12	200	None has intro text. Nine have no meta description.
Products	174	200	31 share the same title. 63 have an empty price in the schema. 48 have a meta description that is only the price.
Blog	24	200	The best content on the site. Signed, dated, expert, and unmarked.
English pages /en/	14	200	Not in any XML sitemap. The Kärcher catalog in English. No hreflang, no link from the Slovenian section.
Total reviewed	232	200	218 from the sitemap + 14 found while crawling. None returns an error, but all 218 sitemap entries are written without a trailing slash and redirect.

What else was checked

URL	HTTP	Finding
/sitemap-staro.xml	200	The previous company's old XML sitemap. Lists 41 URLs; all 41 return 404. Details in the “Migration leftovers” section.
/robots.txt	200	Exists, and blocks images, rendering resources and five AI agents. Broken down in the SEO section.
/wp-sitemap.xml	200	A second, competing XML sitemap. WordPress core serves it alongside the plugin's sitemap. It lists 41 tag archives and 2 author archives, all <code>noindex</code> .
test.example-shop.com	200	A publicly reachable staging copy of the site. No password, no <code>noindex</code> , older prices. All 218 pages are duplicated.
/karkoli-kar-ne-obstaja/	200	Soft 404. Non-existent URLs return status 200, not 404, so Google indexes them.
/izdelki/*/slika-1/	200	693 attachment pages, one for every uploaded image. No text, duplicate title.
/wp-content/uploads/2023/	200	Directory listing is on. The server prints the list of files.

/llms.txt	404	Not present. A proposed standard that no major engine has yet confirmed as an input. Listed for completeness, not as a recommendation.
/pogosta-vprasanja/ , /faq/	404	Checked because it is the fastest route to visibility in AI answers. It doesn't exist.
/cenik/ , /reference/ , /galerija/ , /zaposlitev/	404	Checked because they are standard for this line of business. None exists.
/vrtnarstvo/ , /trgovina/ , /shop/	301	Redirect correctly to /izdelki/ . No chains.
http://example-shop.com , www.	301	Both redirect in one hop to https://example-shop.com/ . Correct.

We separately checked 47 internal links that pointed outside the XML sitemap: **23 return 404** and **17 point to a URL that redirects**. All are listed in Appendix A3, together with the page each link comes from.

SEO — 4/10

Working technical foundations with almost no content signals attached.

Technical on-page signals

Signal	Finding	Status
Title tags	31 products share a single title , Example Company, because the product template falls back to the site name when the field is empty. Another 44 titles are longer than 65 characters and will be truncated, mostly because of the suffix Example Company — spletna trgovina [" Example Company — online store"], which uses up 44 characters on every title.	Duplicated
Meta description	Present on 62 of 218 pages. Of those 62, 48 are only the price , for example 29,90 €, because the product template takes the first line of the listing as the description. Categories: 9 of 12 have none. Blog: all 24 posts have one, and all 24 are well written.	156 missing
H1 heading	Not one of the 218 pages has exactly one <h1> element. The footer template wraps the logo in an <h1>, so every page has at least one before the content even starts. The homepage has four, a product page two, and a category up to thirteen , because the product grid uses an <h1> for every product name. Details below.	Critical
Heading order	96 pages skip a level, most often <h2> to <h4>, because the WPBakery product block uses <h4> for product names. Lighthouse flags this as an error.	Needs attention
lang attribute	All 218 pages declare <html lang="en-US">, the default of a fresh WordPress install that nobody ever changed. The content is entirely Slovenian. The fix is one setting under <i>Settings</i> → <i>General</i> → <i>Site Language</i> .	Wrong
Canonical tags	Present and self-referencing on all 218 pages. On paginated category pages (?page=2) they correctly point to themselves, not to page one.	Good
hreflang	None anywhere. Under /en/ there are 14 English pages with the Kärcher catalog, and the sister domain example-shop.hr carries the same catalog in Croatian. None of the three versions declares the others and none carries x-default, so Google treats them as three competing copies of the same thing, not as three language versions.	Missing
Robots meta	No accidental noindex anywhere. Tag archives are correctly excluded. The author archive is noindex, which isn't ideal for a blog with signed posts. See the GEO section.	Good
Mobile viewport	Correct tag on every page. No mobile usability errors apart from undersized tap targets in the category filter.	Good

HTTPS and mixed content	The certificate is valid, and the redirect from the unencrypted URL and from <code>www.</code> is correct, in one hop. But nine pages load resources over the unencrypted protocol : six images from the previous company's server and one gallery script. The browser blocks active mixed content, so the gallery doesn't work on those pages at all.	Mixed content
URL structure	Clean, Slovenian, readable: <code>/izde/ki/vrtne-skarje-gardena-b-s-xl/</code> . No parameters, no session IDs, no duplicate paths. This is above average.	Good
Duplicate analytics	Google Analytics 4 loads twice , with two different measurement IDs (G-4KX2L9QRTP through a plugin and G-8ND7VC3MEZ hard-coded in the theme's <code>header.php</code>). Every pageview and every event is counted twice, so every conversion number to date is wrong.	Critical

Crawling and indexing

Signal	Finding	Status
XML sitemap	Three XML sitemaps at once, and they contradict each other. <code>/sitemap_index.xml</code> (plugin) lists 218 URLs, every one without a trailing slash , so all 218 entries are 301 redirects. <code>/wp-sitemap.xml</code> (WordPress core) also returns 200 and lists a different set, including 41 tag archives and two author archives that are <code>noindex</code> . Details below.	Critical
Abandoned sitemap	<code>/sitemap-staro.xml</code> returns HTTP 200 and lists 41 URLs of the previous company. All 41 return 404. It isn't in <code>robots.txt</code> , but the file exists and Google has known it before.	Critical
robots.txt	The file exists but blocks four things that shouldn't be blocked : the <code>/wp-content/uploads/</code> folder (all 693 images are unreachable for Google Images), the <code>/wp-content/plugins/</code> and <code>/wp-includes/</code> folders (Googlebot can't render the pages), and ChatGPT search through OAI-SearchBot. On top of that, the <code>Sitemap:</code> directive points to <code>http://</code> and to the dead sitemap. Details below.	Critical
Broken internal links	23 internal links return 404 . Eighteen point to products that were deleted instead of being taken off sale; five point to old blog posts. All are in Appendix A3, with the page they come from.	23
Redirected internal links	17 internal links point to a URL that redirects . Each one costs the crawler an extra hop and dilutes the link equity passed on. The links sit in the footer template and in three blog posts, so the fix is in one place.	17
Orphan pages	Nine pages have no inbound internal link : seven products and two categories (<code>/kategorija/zimska-zascita/</code> , <code>/kategorija/semena-zelenjave/</code>). They can only be reached through the XML sitemap and the site search.	9
Low-value pages	The plugin's sitemap contains nothing superfluous: no tag archives, no search results, no <code>?add-to-cart=</code> URLs. The excess is everywhere else. See the next group.	Sitemap only
Click depth	Every product is reachable within three clicks of the homepage. The deepest page sits at level four.	Good

Duplication and crawl budget

Signal	Finding	Status
Indexed staging copy	<code>test.example-shop.com</code> is publicly reachable, returns HTTP 200, has no password protection and no <code>noindex</code> . It contains a complete copy of all 218 pages, including older prices. So Google has two identical websites and has to choose which one to rank.	Critical
Filter URLs	The category filter generates URLs like <code>?filter_znamka=gardena&filter_cena=20-50&orderby=price</code> . There are about 2,400 combinations. All return 200, all are indexable, and none has <code>noindex</code> or a canonical to the base category.	2,400
Attachment pages	WordPress creates a separate page for every uploaded image. 693 attachment pages are reachable at <code>/izdelki/ime-izdelka/slika-1/</code> . Each holds one image and no text, and carries a title that duplicates the product.	693
Archives	41 tag archives (most with one or two posts), 24 monthly archives and 2 author archives are indexable. All are thin, duplicate pages. Tags are correctly left out of the plugin's sitemap, but the core sitemap lists them.	67
Search results	Internal search at <code>?s=</code> returns 200 and isn't blocked. Google has already picked up URLs of this kind. They come from links that people share.	Open
Soft 404	A non-existent URL returns HTTP 200 , not 404. <code>/karkoli-ne-obstaja/</code> serves a nicely designed “Stran ni najdena” [“Page not found”] page with status 200, so Google treats it as valid content and indexes it. This is why pages that don't exist show up in Search Console.	Critical
Upper and lower case	The server serves <code>/Izdelki/</code> and <code>/izdelki/</code> with the same content and status 200. The canonical softens this, but there is no redirect to lowercase.	Partial
PDF files	Eleven PDF files in <code>/wp-content/uploads/</code> are indexable, among them <code>cenik-2022.pdf</code> , a two-year-old price list that can outrank your current pricing page.	11
Total reachable	The site has 232 real pages. A search engine can reach about 3,600 . The difference is crawl budget spent on content nobody wants, and the reason new pages get indexed slowly.	~3,600

Content quality

Signal	Finding	Status
Category text	None of the 12 product categories has intro text. The product grid starts right under the heading. A category is a page that can rank for terms like <i>vrtna škarija</i> [“garden shears”], and right now it says nothing about itself.	Missing
Product descriptions	78 words per product on average. 52 products have fewer than 30 words , and 14 have no description at all, just a name, a price and an image. For 39 products the description is copied word for word from the manufacturer's catalog and also appears on competitors' sites.	Too thin

Blog	The best content on the site. 24 posts, 1,140 words on average, signed by name, dated, with concrete advice for the Slovenian climate and planting seasons. The latest is 11 months old.	Good
Image alt text	411 of 693 images have no alt attribute , including all six hero slider images and 174 main product images. 96 images have alt="", which is correct only for decorative images; here product images are among them.	411
Image file names	209 images carry the name the camera or phone gave them: IMG_20230514_112233.jpg, DSC_0431.jpg, WhatsApp Image 2024-03-02 at 11.14.07.jpeg. The last one contains spaces and periods, which is bad for a URL. The other 484 are descriptive.	209
Untranslated interface strings	WooCommerce's default English strings are visible to Slovenian visitors on all 174 products and 12 categories: Add to cart, Showing 1-12 of 174 results, Related products, You may also like..., Sale!. The theme's translation file is installed; the WooCommerce translations are not.	Visible error
Placeholder text in production	Two hits. /dostava/ contains 84 words of Lorem ipsum under its second heading. On three pages the footer template prints the phone number +386 40 123 456, which is a placeholder, not your number.	Critical
Brand name spelling	Kärcher is spelled three ways: Karcher (61×), Kärcher (34×) and KARCHER (12×). Gardena appears as Gardenna on seven products. Entity matching is literal. A buyer searching for <i>Kärcher visokotlačni čistilec</i> ["Kärcher pressure washer"] isn't matched as well by the spelling Karcher.	Inconsistent
Readability	Measured with the Gunning Fog index. The blog reads comfortably (average 14.2). Static pages are harder (/splosni-pogoji/ reaches 31.4, which is expected for legal text). No commercial page is a problem.	Good

Signals missing from the code

Signal	Finding	Status
Opening hours as an image	The opening hours for both locations are published as a JPEG image (urnik-2024.jpg). A visitor can read them. A search engine, a screen reader and an AI assistant cannot . This is both the most common and the easiest error to fix for local businesses: text that exists only as pixels.	Unreadable
BreadcrumbList schema	Breadcrumbs are visible on 205 pages and correctly built, but without BreadcrumbList markup . So Google can't show them in results in place of the bare URL.	Missing
og:url and og:locale	og:url on all 218 pages points to the homepage . Every shared product link therefore resolves to the homepage on Facebook. og:locale isn't set.	Wrong
Twitter Card	No twitter:card tags. Links shared on X and in some messaging apps render without an image.	Missing
Phone isn't a link	The phone number appears in seven places as plain text, without . On mobile it can't be called with one tap. For a local business that is a direct loss of calls.	Missing

Site icon	Only a 16×16 <code>favicon.ico</code> from the theme's default install is present. There is no <code>apple-touch-icon</code> and no 512×512 icon, so phone home screens and browser tabs show a generic icon.	Incomplete
Thank-you page	<code>/hvala-za-narocilo/</code> is indexable and in the XML sitemap. Besides appearing in results, it lets visitors land on it straight from search, which falsely triggers the conversion goal and further corrupts analytics numbers that are already doubled.	Indexable
Contact form	No consent checkbox for data processing and no link to the privacy policy next to the button. After submission only a message on the same page appears, with no separate page or event, so form submissions are not measured .	Missing
Exposed server settings	Directory listing is on: <code>/wp-content/uploads/2023/</code> prints the list of files. <code>xmlrpc.php</code> responds. Three plugins are behind the current version, one of them by three major releases.	Hygiene
JavaScript-dependent content	Comparing 40 pages in raw HTML and after rendering: the product grid on 12 categories and the “related products” block don't exist in the raw HTML . Googlebot usually renders them, but <code>robots.txt</code> doesn't let it (see below), so in the current state it doesn't see them.	Invisible

Findings in detail

Structural errors the tools don't report

Five errors that Lighthouse and similar tools miss, because each individual tag looks correct to them. The structure is wrong, not the syntax. These are the reason the SEO score is 4 and not 7 despite solid foundations.

1. No page has exactly one `<h1>` element

The footer template wraps the logo in an `<h1>` element:

```
<footer class="site-footer">
  <h1 class="footer-logo"><a href="/"></a></h1>
```

This causes two separate problems. First, **each of the 218 pages gets at least one extra `<h1>` from the footer**. Second, and worse, **that `<h1>` is empty**. It contains only an image with no `alt` attribute, so a search engine reads it as `<h1></h1>`. On the four pages with no heading in the content, it is the **only `<h1>`** on the page. That means `/dostava/` and `/nacini-placila/` tell Google nothing about themselves.

Page type	No.	H1 count	Source
Homepage	1	4	Slider heading, welcome block, newsletter block, footer.
Product category	12	3–13	Category heading, the name of every product in the grid (the WPBakery block uses <code><h1></code>), footer.
Product page	174	2	Product name, footer.
Blog post	24	2	Post title, footer.

Static pages 7 1–2 Four have only the empty `<h1>` from the footer.

Pages with exactly one `<h1>` — 0 Out of 218.

There are two fixes, each one line. In the footer, replace `<h1>` with `<div>`. A logo isn't a heading. In the product grid, replace `<h1>` with `<h3>`. That fixes 218 pages at once and also clears the 96 skipped-heading-level warnings from the accessibility section.

2. Three XML sitemaps that contradict each other

The search engine gets three lists of what exists on the site, and none matches the other two.

Sitemap	HTTP	What it lists	Problem
/sitemap_index.xml	200	218 URLs in five child sitemaps	All 218 entries lack a trailing slash, while the site canonicalizes with the slash. So every entry is a 301 redirect.
/wp-sitemap.xml	200	A parallel WordPress core set: 174 products, 24 posts, 41 tag archives and 2 author archives	43 entries point to pages that are <code>noindex</code> . The search engine is told “index this” and “don't index this” at the same time.
/sitemap-staro.xml	200	41 URLs of the previous company	All 41 return 404. See the “Migration leftovers” section.
(none)	—	14 English pages under /en/	Not in any sitemap, and no link from the Slovenian section points to them.

The slash error is the quietest and the most expensive. The sitemap says:

```
<loc>https://example-shop.com/izdelki/vrtne-skarje-gardena-b-s-xl</loc>    ← no trailing slash
https://example-shop.com/izdelki/vrtne-skarje-gardena-b-s-xl/              ← canonical, 301
```

So for each of the 218 URLs the crawler spends **two requests instead of one**: the first returns a redirect, only the second returns content. That halves the crawl budget, and the `lastmod` dates refer to a URL that isn't canonical.

The fix: turn on trailing slashes in the sitemap plugin's settings, disable the core sitemap `/wp-sitemap.xml` (disable it, don't block it, or it stays reachable), set `/sitemap-staro.xml` to `410 Gone`, and either add the 14 English pages to the sitemap with proper hreflang or remove them.

3. robots.txt blocks images and the files Google needs to render pages

The whole file, annotated:

```

User-agent: *
Disallow: /wp-admin/
Allow: /wp-admin/admin-ajax.php
Disallow: /wp-content/plugins/      ← Googlebot can't read plugin CSS and JS
Disallow: /wp-includes/             ← same; this holds jQuery, which the product grid depends on
Disallow: /wp-content/uploads/      ← all 693 images unreachable for Google Images
Crawl-delay: 20

User-agent: GPTBot
Disallow: /
User-agent: CCBot
Disallow: /
User-agent: Google-Extended
Disallow: /
User-agent: anthropic-ai
Disallow: /

Sitemap: http://example-shop.com/sitemap_index.xml    ← http, not https
Sitemap: http://example-shop.com/sitemap-staro.xml    ← the previous company's dead sitemap

```

Line	What it actually does	Priority
Disallow: /wp-content/uploads/	All 693 images are invisible to Google Images. For a garden center, where buyers search by a picture of a plant or tool, that is an entire channel lost. The block also covers product images in the schema: <code>image</code> in <code>Product</code> points to a URL the crawler isn't allowed to fetch.	Critical
Disallow: /wp-content/plugins/ Disallow: /wp-includes/	Googlebot renders a page by loading its CSS and JavaScript. These two lines forbid that, so it sees an unstyled page without the product grid . Google explicitly advises against this. It is also why the product grid doesn't appear in the rendering comparison of 40 pages.	Critical
Sitemap: http://...	Both directives use <code>http://</code> , while the site runs on <code>https://</code> . The crawler gets a redirect, which works but is unnecessary. And the second line actively advertises the previous company's dead sitemap.	High
Crawl-delay: 20	Google ignores the directive; Bing respects it. 20 seconds between requests means at most 4,320 pages a day. For 232 pages that isn't a bottleneck today, but the directive serves no purpose and should go.	Low
Five AI agents	Only one of the five costs you visibility: <code>0AI-SearchBot</code> . The other four govern model training. Broken down in the GEO section. The decision is yours, but make it line by line, not for all five at once.	Critical

The correct file contents for this site are in Appendix C.

4. Mixed content: nine pages, one broken feature

The site runs on HTTPS, but nine pages load resources over unencrypted `http://`. All of them come from the previous company's server.

Resource	Type	Pages	What the browser does
<code>http://stara-domena.si/wp-content/uploads/2019/staro-galerija-0{1..6}.jpg</code>	Images	6	Passive mixed content. Chrome tries to upgrade them to https. When that fails, it blocks them and the slot stays empty.
<code>http://code.jquery.com/jquery-migrate-1.4.1.min.js</code>	Script	9	Active mixed content. The browser blocks it unconditionally. The gallery on these pages doesn't open; clicking an image does nothing.

The affected pages are `/o-nas/`, three product pages and five blog posts. Two points here deserve attention beyond security itself.

First, a site feature is broken and nobody notices, because the browser console only logs a quiet warning. A visitor clicks an image and nothing happens.

Second, six images are served from a domain that isn't yours. If the previous company drops the domain or the server goes down, those six images disappear from the site, with no warning and no way to recover them. Move the images into your own media library and fix the links. That fixes the mixed content and removes the dependency on someone else's server in one step.

5. No hreflang, despite three language versions

The same catalog exists in three forms, and none of them knows about the others:

Version	Size	Lang	Status
<code>example-shop.com</code>	218 pages	en-US	Slovenian content declaring itself English. No hreflang.
<code>example-shop.com/en/</code>	14 pages	en-US	English content, so here <code>Lang</code> happens to be right. In no sitemap, no inbound links.
<code>example-shop.hr</code>	191 pages	en-US	Croatian content, same catalog, same images. No hreflang in either direction.

Without `hreflang` tags, Google doesn't read the three versions as language counterparts but as **three competing copies of the same catalog**. As a result they take rankings from each other: a Croatian buyer may get the Slovenian version and vice versa, and the search engine has to guess which one to show.

The fix is a set of three reciprocal tags on each matching page, plus `x-default`. The key word is **reciprocal**: if the Slovenian page declares the Croatian one, the Croatian page must declare the Slovenian one, or Google ignores the whole set. A template is in Appendix C.

6. The staging copy of the site is publicly indexable

`test.example-shop.com` runs a complete copy of the site. It returns HTTP 200, has no password protection, no `X-Robots-Tag` header and no `noindex` tag:

```
$ curl -sI https://test.example-shop.com/izdelki/gardena-b-s-xl/
HTTP/2 200
content-type: text/html; charset=UTF-8
      ← no x-robots-tag, no login required
```

So Google sees **two complete copies of the same catalog** and has to choose which one to rank. It may not choose your production site. The staging copy is often older, faster or better linked from external sites, because links to it remain from the development period. The staging copy also carries **older prices**, which a visitor can land on from search results.

This is one of the most common errors we find, and one of the quickest to fix. The fix is password protection at the server level (`.htpasswd`). **Not** `noindex` , and **not** `robots.txt` . A `robots.txt` block stops the crawler from reading a page, not from indexing it: the URL can still stay in results, just without a description. A password is the only reliable fix.

7. Non-existent URLs return HTTP 200

The site has a nicely designed “Stran ni najdena” [“Page not found”] page, but it is served with the wrong status:

```
$ curl -sI https://example-shop.com/karkoli-kar-ne-obstaja/
HTTP/2 200          ← should be 404
```

A visitor sees no difference. For a search engine the difference is total: status 200 means “this content exists and is valid”, so Google indexes the page. Because all such pages contain the same text, you get any number of duplicate pages with identical content. Every mistyped link creates one, as does every old bookmark and every one of the 23 broken internal links in Appendix A3.

This is also why URLs show up in Search Console that nobody can trace. The fix is a setting in the `404.php` template so it sends the correct status code before output.

8. Filters and attachment pages create ~3,400 excess URLs

The site has 232 pages that someone created on purpose. A search engine can reach about 3,600.

Source	URLs	Why they exist	Fix
Category filters	~2,400	Every combination of brand, price range and sort order is its own URL that returns 200.	Canonical to the base category + <code>noindex</code> on filtered variants.
Attachment pages	693	WordPress creates a page for every uploaded image. Each holds one image and no text.	Redirect attachments to the parent page. A setting in the SEO plugin.
Staging copy	218	See point 6.	Password protection.
Tag, month and author archives	67	On by default in WordPress; most tags have one post.	<code>noindex</code> for tag and month archives.
Search results and PDF files	11+	Internal search isn't blocked; the PDFs include a two-year-old price list.	Disallow: <code>/*?s=</code> and X-Robots-Tag: <code>noindex</code> for outdated PDFs.
Excess	~3,390	For every real page there are about fourteen that aren't.	

The practical consequence is indexing speed. Google's crawl budget for a site of this size and authority is limited. Once it is spent on filter variants and attachment pages, a new product or blog post takes days to get indexed instead of hours. None of the items above requires development. They are all settings.

9. The opening hours are an image

The opening hours for both locations are published on `/kontakt/` as an image:

```

```

A visitor can read it perfectly well. For everyone else it doesn't exist. A search engine can't read the opening hours, a screen reader can't tell a blind visitor when you're open, and an AI assistant can't answer *kateri vrtni center ima odprto ob sobotah* [“which garden center is open on Saturdays”]. That happened in the AI visibility measurement. And if the hours change, someone has to make a new image.

The image also has no alt attribute, so even that fallback is missing. The fix is to write the hours as text in a table and mark them up with openingHoursSpecification (Appendix C2). The content already exists and is accurate. It's just in the wrong format.

Title tags: 31 duplicated and 44 too long

The product template builds the title from the pattern {ime izdelka} | Example Company – spletna trgovina [“{product name} | Example Company — online store”]. The suffix uses 44 characters of the roughly 60 Google displays, so for longer product names all that's left of the title is the site name, which is the same on every title.

Worse, 31 products have an empty title field, so the template falls back to the site name alone. In Google's results these 31 products are indistinguishable from one another:

```
<title>Example Company</title>          × 31 products
<title>Vrtno škarje Gardena B/S XL | Example Company – spletna trgovina</title>    (79 characters)
<title>Kärcher K5 Premium Smart Control visokotlačni čistilec | Vrtni cent...</title> (94 characters)
```

The fix has two parts, and both are settings, not development. Shorten the suffix to | Zeleni Val (12 characters instead of 44), and change the title pattern so that it uses the product name and category instead of the site name. That fixes all 31 empty ones automatically.

Page	Now	Proposed
/	Example Company – spletna trgovina	Vrtni center Domžale in Celje – vrtnarstvo Example Company (53) [“Garden center in Domžale and Celje. Example Company gardening”]
/kategorija/vrtno-orodje/	Vrtno orodje Example Company – spletna trgovina	Vrtno orodje – škarje, lopate, grablje Example Company (52) [“Garden tools: shears, spades, rakes”]
/izdelki/gardena-b-s-xl/	Example Company	Gardena B/S XL vrtno škarje za veje do 42 mm Example Company (58) [“Gardena B/S XL garden shears for branches up to 42 mm”]

Meta descriptions: 156 missing and 48 that are only the price

When there's no meta description, Google writes its own from whatever text it finds on the page. On a product page that means the price line, the button and the stock notice. The result in search looks like this:

```
29,90 € Add to cart Related products You may also like... Sale!
```

For 48 products it's even more clear-cut: **the meta description exists and is only the price**, because the template takes the first line of the product listing as the description.

Page group	Has one	Missing	Note
Homepage	1	0	Well written, 148 characters. The only model example on the site.
Static pages (7)	3	4	Missing on /storitve/, the most commercially important static page.
Categories (12)	3	9	A category can rank for a broad term. Nine of them say nothing about themselves.
Products (174)	48	126	All 48 existing ones are only the price. In practice, 0 products have a useful description.
Blog (24)	24	0	All present, all well written, all under 150 characters.
Total	79	139	Of the 79 present, 48 are useless. 31 are useful.

The full list of 139 URLs without a useful description is in the attached file `manjkajoci-opisi.csv`, because it's too long for a table in the report. The first 20 are in Appendix A1.

Language declaration: `lang="en-US"` on a Slovenian website

The `<html>` element on all 218 pages reads:

```
<html lang="en-US" prefix="og: https://ogp.me/ns#">
```

This is the default value of a fresh WordPress install. The site's content is entirely Slovenian, but the code tells every search engine and every AI assistant that the content is American English. It is the cheapest high-value fix in the whole report: one setting under *Settings* → *General* that changes all 218 pages at once.

There are three consequences, all silent. Google matches Slovenian queries to this content less well. Screen readers pronounce Slovenian text with English phonetics. An AI assistant looking for a Slovenian provider writes the site off on this signal before it reads any text.

Duplicate analytics: every event is counted twice

Every page loads **two different Google Analytics 4 containers**:

```

<!-- analytics plugin -->
<script async src="https://www.googletagmanager.com/gtag/js?id=G-4KX2L9QRTP"></script>

<!-- hard-coded in the child theme's header.php -->
<script async src="https://www.googletagmanager.com/gtag/js?id=G-8ND7VC3MEZ"></script>

```

These aren't two versions of the same container. They are two separate properties. Every pageview, every click and every completed purchase is recorded twice, in two reports that don't match each other. Every decision made on these numbers in recent years was made on wrong data. And every measurement of the effect of this report's

fixes would be just as wrong unless this is fixed **before** implementation starts. That's why this item ranks high in the priority matrix even though it doesn't win a single ranking by itself.

Pick one property, remove the other, and check in Google Analytics whether both have been collecting data. If they have, the history is usable from only one of them.

Images: 411 without alt text, 209 with camera file names

411 of 693 images have no alt attribute. They include all six hero slider images and all 174 main product images. Those are exactly the images that could bring traffic from image search, where purchase intent in gardening is high.

```



```

These are two separate problems. The first is the missing alt text. The second is that **209 files carry the name the camera or phone gave them.** The file name is part of the image URL and one of the signals a search engine uses to work out what the image shows. The name WhatsApp Image 2024-03-02 at 11.14.07.jpeg also contains spaces and periods, which get encoded in the URL.

Renaming 209 files means fixing the links in the content, so it is ranked medium priority and applies mainly to new images from now on. Writing 411 alt texts is ranked high, because it takes nothing but time and also clears the largest accessibility item.

Visible errors in production

Two hits, both on a page the buyer reads right before purchasing.

/dostava/ : 84 words of placeholder text

Pogoji dostave

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat.

The delivery page is one of the three pages buyers read most often in an online store. Right now, under the heading “Pogoji dostave” [“Delivery terms”], it shows Latin placeholder text.

Placeholder phone number in the footer

On three pages, /dostava/, /nacini-placila/ and /zasebnost/, the footer template prints +386 40 123 456. This is a sample number from the theme template, not yours. The other 215 pages have the correct one. A buyer who calls from these pages doesn't reach you.

GEO — 3/10

The schema is present and technically valid. It declares the wrong company, and AI crawlers are locked out.

Signal	Finding	Status
AI crawler access	robots.txt blocks five AI agents. Four of them govern model training and cost you no visibility. The fifth, OAI-SearchBot, governs ChatGPT search , and that one line removes you from it. The difference is broken down below.	1 of 5 hurts
Organization schema	Present on all 218 pages and technically valid, yet it declares the wrong company : "legalName": "Old Company d.o.o.", "telephone": "+386 1 542 33 10". The name field was correctly updated to Example Company. Nothing else was.	Wrong
LocalBusiness schema	None. The company has two physical locations (Domžale, Celje) with published opening hours, and neither is declared as a location a machine can read. No GardenStore, no Store, no openingHoursSpecification, no geo.	Missing
Product schema validity	Product is present on all 174 products, which is better than average. On 63 out-of-stock products it reads "price": "" . An empty string is not a price, so Google discards the whole offer. Beyond that: brand is missing on 88 products, neither gtin nor mpn appears on any product, and shippingDetails and hasMerchantReturnPolicy appear nowhere.	63 invalid
NAP consistency	Three versions for two locations. The footer gives Kamniška cesta 41, Domžale with no postal code; /kontakt/ gives Kamniška cesta 41a, 1230 Domžale; the schema gives the previous company's phone number. The Celje postal code is not published anywhere.	Critical
Opening hours	Published as text on /kontakt/ for both locations, with summer and winter schedules. The content is complete and accurate. It is not marked up anywhere, so machines can't read it.	Unmarked
Blog markup	24 posts are marked up as plain WebPage. No Article, no author, no datePublished, no publisher, even though this is signed, dated expert content. The author archive is also noindex.	Unmarked
E-E-A-T signals	The /o-nas/ page exists and states the founding year, 1998, and 14 employees. Nobody is named , there are no team photos and no qualifications. The blog is signed with a name, but that name has no profile anywhere.	Partial
Verifiable endorsements	The company is an authorized Gardena and Kärcher dealer and service center . That is a verifiable third-party endorsement, and it is mentioned once, in one sentence on /o-nas/. It is not in the schema, not on the homepage, not on any product page.	Underused
Customer reviews	There are no reviews, ratings or testimonials anywhere on the site. The product template has no review interface. None of the 174 products has aggregateRating or review.	None

Open Graph	og:title, og:description and og:type are present. og:image is the same file on all 218 pages: a 300×90 logo, below Facebook's minimum recommendation, which renders as a thin strip.	Partial
Entity links (sameAs)	The footer links to Facebook and Instagram, and both profiles are active. The schema has no sameAs, so the search engine is never told explicitly that these are the same entity.	Unmarked
llms.txt	Not present. A proposed standard that no major engine has yet confirmed it uses as input. Experimental; listed for completeness, not as a recommendation.	Absent

Findings in detail

robots.txt shuts out AI crawlers

The `/robots.txt` file is otherwise well built. It points to the XML sitemap and does not block anything that should be indexed. But it also contains this:

```
User-agent: OAI-SearchBot           ← ChatGPT search
Disallow: /

User-agent: GPTBot                 ← model training
Disallow: /

User-agent: CCBot
Disallow: /

User-agent: Google-Extended
Disallow: /

User-agent: anthropic-ai
Disallow: /
```

This is not a misconfiguration. It is a **decision** someone made at some point, most likely when there was a lot of talk about AI companies taking content without permission. The decision was understandable. What matters is that **these lines do not block the same thing**, and that distinction decides what this file actually costs you.

Crawler	What it governs	What blocking it costs you
OAI-SearchBot	Whether the site can appear and be cited in ChatGPT search .	Visibility. This is the only one of the five lines that removes you from AI answers directly.
GPTBot	Whether the content may be used for model training .	No visibility. Blocking it is a legitimate decision about your own content.
Google-Extended	Use of content in Google's generative products (Gemini, Vertex).	No visibility in Search. It does not remove you from Google's AI Overviews. Those are served from regular Googlebot crawling.
CCBot	Crawling for Common Crawl, which third parties use.	Indirect and hard to measure.
anthropic-ai	An older training agent.	No visibility.

The practical conclusion differs from what a quick read of the file would suggest. **If you want to be visible in AI answers, one line has to go: 0AI-SearchBot**. The other four can stay. They keep your content out of model training and cost you no visibility. This middle path is the one most articles on the topic miss, because they treat all AI crawlers as a single switch.

The decision is yours, not ours. It is worth making line by line, though, not for all five at once.

The schema declares the previous company

The `Organization` block is present on all 218 pages. It is technically valid and passes Google's Rich Results Test. Here is what it says:

```
{
  "@context": "https://schema.org",
  "@type": "Organization",
  "name": "Example Company",
  "legalName": "Old Company d.o.o.",           ← previous company
  "telephone": "+386 1 542 33 10",             ← not your number
  "email": "info@stara-domena.si",             ← not your email
  "url": "https://example-shop.com"
}
```

Only the `name` field was corrected. Everything else survived from the site this one was built from. To Google's Knowledge Graph this is **one entity with two names and contact details that match nothing on the site**. That is exactly the state that stops a company from settling into a recognized entity.

It is also why the NAP fix is rated critical. Correcting the address on the `/kontakt/` page is not enough. As long as the schema on 218 pages says otherwise, the search engine has two sources, and the machine-readable one is wrong.

63 products with an empty price

WooCommerce outputs `Product` schema on every product. That is an advantage, and it deserves saying. For products currently out of stock, however, it outputs this:

```
"offers": {
  "@type": "Offer",
  "price": "",                               ← empty string
  "priceCurrency": "EUR",
  "availability": "https://schema.org/OutOfStock"
}
```

What happens next is the key point. An empty string is **not** a price of zero and **not** a missing price. It is an invalid value. When a required field is invalid, Google discards **the whole offer**, not just the missing value. In Google's eyes these 63 products have no price and no stock information, even though the schema looks complete.

The fix is a template rule: when there is no price, don't output the `price` field at all. Schema that omits an optional field is valid. Schema with an empty required field is not.

Check with the [Rich Results Test](#) and in Search Console under *Enhancements* → *Products*, where these 63 should be listed as errors.

Two locations machines can't see

The company runs two garden centers with published opening hours. The site has **no** `LocalBusiness` schema at all. For a business whose customers search by location (*vrtni center Domžale* [“garden center Domžale”], *vrtnarija blizu mene* [“garden center near me”]), this is the largest single gap in the report after the closed `robots.txt`.

The published addresses also contradict each other:

Where	What it says	Problem
Footer (218 pages)	Kamniška cesta 41, Domžale	No house letter and no postal code.
/kontakt/	Kamniška cesta 41a, 1230 Domžale	Complete, and the only complete version.
Schema (218 pages)	No address; the phone number belongs to the previous company.	The machine-readable source is wrong.
Celje location	Mariborska cesta 88, Celje	No postal code anywhere on the site.

The full fix: settle on one version of each address, publish it identically everywhere, and output separate `GardenStore` schema for each location. Ready-made code is in Appendix C.

The blog is the best content and isn't marked up as content

24 posts, averaging 1,140 words, with specific planting dates for the Slovenian climate, signed and dated. This is exactly what AI engines cite as a source, and every one of them is marked up as a plain `WebPage`.

No `Article`, no `author`, no `datePublished`, no `publisher`. The author archive is also `noindex`, so the byline has no page anywhere for an authorship signal to attach to. Three years of expert writing build no authority, because nothing says whose it is.

The fix is markup in the post template plus one author profile page. It doesn't take a single new word of content.

What AI engines actually say

A GEO score built from schema alone is inference. This is measurement: eight questions a customer would actually ask, put to a search-grounded language model on September 9, 2026, with the cited sources recorded.

Question asked	Cited?	Cited instead of you
Kje v Domžalah lahko kupim vrtno orodje Gardena? [Where in Domžale can I buy Gardena garden tools?]	Yes, in the answer	bigbang.si, mercator-tehnika.si, obi.si, hoferpro.si
Kateri vrtni center je najboljši v okolici Ljubljane? [Which garden center near Ljubljana is best?]	Sources only	kalia.si, semenarna.si, obi.si, gardencenter-brdo.si. The answer recommends three others by name
Kdaj je najboljši čas za sajenje česna v Sloveniji? [When is the best time to plant garlic in Slovenia?]	No	kmetijskizavod-lj.si, semenarna.si, mojvrt.si, delo.si. You have a blog post on this

Kako pogosto zalivati paradižnik v rastlinjaku? [How often should I water tomatoes in a greenhouse?]	No	mojvrt.si, agroruse.si, vrtnarstvo-info.si. You have a blog post on this
Kje kupiti visokotlačni čistilec Kärcher s servisom v Sloveniji? [Where can I buy a Kärcher pressure washer with service in Slovenia?]	No	karcher.si, mimovrste.com, bigbang.si, merkur.si
Katera semena zelenjave so primerna za začetnike? [Which vegetable seeds suit beginners?]	No	semenarna.si, mojvrt.si, kalia.si
Kateri vrtni center v Celju ima odprto ob sobotah? [Which garden center in Celje is open on Saturdays?]	No	No sources cited. The model replied that it didn't have the information
Kdo v Sloveniji servisira vrtno opremo Gardena? [Who services Gardena garden equipment in Slovenia?]	Yes, named third	gardena.si, servis-vrt.si, example-shop.com, kalia.si

Two out of eight. On the Gardena service question the site was named as the third provider. That proves the model still finds the authorization, even though it is mentioned in a single sentence on `/o-nas/`. It is the strongest signal the company has, and right now it sits in one sentence on one page.

The pattern behind the six misses is the real finding here. Look at the third and fourth questions: *when to plant garlic* and *how often to water tomatoes*. **Your blog answers both**, and answers them better than the sources the model cited: the posts `/blog/sadilni-koledar-jesen/` and `/blog/zalivanje-v-rastlinjaku/`. The model didn't cite them for three reasons, all covered in this report:

- **robots.txt locks out** the crawlers this model uses for search.
- **The posts aren't marked up as Article**, so the engine reads them as ordinary pages rather than dated, authored content.
- **Neither uses the question as its heading** or answers it in the first 60 words. That is the subject of the next section.

None of this is a content problem. The content exists and it is good. It is unreachable and unmarked.

What this measurement is and isn't. One model, one day, eight questions, each asked once. AI answers vary between engines, sessions and weeks, so treat this as a baseline, not a ranking. It is repeatable, though. The same eight questions can be asked again 30 days after the fixes go in, and that is what we recommend as the measure of impact for GEO, where Search Console can't help.

AEO — 2/10

Nothing on the site can become a featured snippet or a voice answer, even though the content for it is already written.

Signal	Finding	Status
Question-style headings	None of the 1,847 headings on the site is phrased as a question. Blog headings are descriptive (<i>Sadilni koledar za jesen</i> ["Autumn planting calendar"], <i>Zalivanje v rastlinjaku</i> ["Watering in the greenhouse"]), while customers type <i>kdaj saditi česen</i> ["when to plant garlic"] and <i>kako pogosto zalivati paradižnik</i> ["how often to water tomatoes"] into search engines and AI assistants.	None
Direct answers	No heading is followed by a concise 40–60 word answer. Blog posts open with a three- to five-paragraph introduction, and the answer often turns up only halfway through. The search engine has nothing to summarize.	Missing
Question-and-answer content	There is no FAQ page. <i>/pogosta-vprasanja/</i> and <i>/faq/</i> both return 404. One page with ten real customer questions, each question a heading with the answer directly below it, is the strongest move available in this dimension. The content does the work, not the markup.	Missing
Pricing and delivery information	Product prices are published and clear. Delivery cost and delivery time are not. <i>/dostava/</i> holds placeholder text. "How much is delivery" is among the most common pre-purchase questions, and the site doesn't answer it.	Missing
Lists and tables	Three blog posts contain well-built tables (planting dates by vegetable), which is the right format for a table snippet. They aren't marked up and have no heading that anticipates a question.	Partial
Local answers	The question <i>which garden center is open on Saturdays</i> goes unanswered, even though opening hours are published. Without <code>LocalBusiness</code> and <code>openingHoursSpecification</code> schema, machines can't read them.	Missing
Fact density	Blog posts contain specific, quotable numbers: planting depths in centimeters, temperature thresholds, plant spacing. The raw material for citations is there. Only the format is missing.	Good

Findings in detail

The content answers the right questions in the wrong format

This is not a site without content. It is a site with **good content shaped like an article rather than an answer**. The difference lies in the first 60 words under the heading, and it decides whether a page can appear as a featured snippet.

Post	Current heading	A heading that can become an answer
/blog/sadilni-koledar-jesen/	Sadilni koledar za jesen [Autumn planting calendar]	Kdaj saditi česen v Sloveniji? [When to plant garlic in Slovenia?]
/blog/zalivanje-v-rastlinjaku/	Zalivanje v rastlinjaku [Watering in the greenhouse]	Kako pogosto zalivati paradižnik v rastlinjaku? [How often to water tomatoes in a greenhouse?]
/blog/zimska-zascita-rastlin/	Zimska zaščita rastlin [Winter plant protection]	Kdaj in kako zaščititi rastline pred zmrzaljo? [When and how to protect plants from frost?]
/blog/izbira-zemlje/	Izbira zemlje [Choosing soil]	Katera zemlja je primerna za zelenjavni vrt? [Which soil suits a vegetable garden?]

Changing the heading alone is not enough. Each of these headings needs to be followed by **a 40–60 word answer that responds to the question immediately**, and only then the explanation. A template is in Appendix B.

Seven posts are also built as step-by-step guides: transplanting, pruning, preparing a bed. For these, put the effort into the format, not the markup: an ordered list `<ol>`, a question-style heading, and `Article` markup with a named author and date. Google's special display for `HowTo` no longer exists, so that markup on its own gains nothing.

The FAQ page that doesn't exist

The strongest move available in this dimension is one page with eight to twelve real customer questions, each question a heading with the answer directly below it. The questions aren't hypothetical. They are the ones your sales staff at both locations hear every day:

- How much does delivery cost and how fast do you deliver? (*Unanswered today. The delivery page holds placeholder text.*)
- Is the product in stock in Domžale or in Celje?
- Do you service Gardena and Kärcher equipment even if it wasn't bought from you? (*This is your strongest differentiator.*)
- What is the warranty and how do I make a claim?
- Can you advise on choosing plants for a specific garden position?
- Can I order online and pick up at the garden center?
- When are you open on Saturdays and public holidays?
- Do you prepare and deliver bulk substrate?

Each answer 40–60 words, specific, with a number wherever there is one. This single page addresses three findings in this report at once: the missing `FAQPage` schema, the missing delivery information, and the unanswered questions from the AI visibility measurement.

Speed: field data first, then lab data

Where both exist, field data wins. A lab measurement is one run on one device. Field data (CrUX) is 28 days of your real visitors on their own devices and networks.

Field data (CrUX, 75th percentile, 28 days)

Metric	Field	Lab	"Good" threshold	Rating
Largest Contentful Paint	4,180 ms	3,640 ms	< 2,500 ms	Poor. Real users are 540 ms worse off than the lab.
Cumulative Layout Shift	0.21	0.04	< 0.10	Poor, and the lab doesn't show it. See below.
Interaction to Next Paint	312 ms	—	< 200 ms	Needs improvement. The category filter reloads the entire product grid on every click.
Time to First Byte	1,240 ms	890 ms	< 800 ms	Over the threshold. A share of server cache misses.

Why lab and field differ so much on CLS. The lab run measures 0.04, which is excellent. Real users see 0.21, which is poor. The cause is the cookie consent banner. It appears only after the page has rendered and pushes the content down. A lab run has no previous session, and the banner behaves differently than it does for a returning visitor. **Had we reported only the lab number, this would have been one of two poor metrics that went unnoticed.**

What exactly takes 4.2 seconds

From a real Chrome trace of the homepage under mobile emulation:

Phase	Time	Share	What happens
Time to First Byte	890 ms	24%	The server builds the page. The cache misses on roughly every third visit.
Resource load delay	770 ms	21%	The hero image doesn't even start downloading until 14 render-blocking files have loaded.
Resource load time	1,610 ms	44%	The hero image is a 2400×1600 JPEG weighing 812 KB , displayed in a 412 px wide frame.
Render delay	370 ms	11%	The main thread is busy parsing the slider's JavaScript.
LCP total	3,640 ms	100%	Field: 4,180 ms.

The hero slider: 4.7 MB before the first word

The homepage loads a slider with **six full-resolution images** before anything renders:

File	Resolution	Size	Note
slider-pomlad-2024.jpg	2400×1600	812 KB	The LCP element. Displayed in a 412 px frame.
slider-akcija-gardena.jpg	2400×1600	906 KB	Not visible without interaction, yet it loads immediately.
slider-semena.jpg	2400×1600	744 KB	Same.
slider-substrat.jpg	2400×1600	688 KB	Same.
slider-karcher.jpg	2400×1600	921 KB	Same.
slider-jesen.jpg	2400×1600	679 KB	Same.
Total	—	4.75 MB	One is visible. Six are loaded.

Three fixes, in order of impact. **First**, load only the first image immediately and the other five with `loading="lazy"`. That alone takes roughly 3.9 MB off the critical path. **Second**, serve WebP or AVIF at display size, with `srcset` for three screen widths. Product images are already served as WebP; the slider is not. **Third**, give the first image `fetchpriority="high"` and a `<link rel="preload">` so it starts downloading before the render-blocking files.

Estimated combined effect: LCP under 2 seconds without a single change to how the page looks.

Render-blocking files: fourteen

Type	File	Note
CSS	/wp-content/themes/woodmart/css/parts/woo-single.min.css	Also loads on the homepage, where there is no product.
CSS	/wp-content/plugins/js_composer/assets/css/js_composer.min.css	WPBakery, 412 KB unminified, compressed to 61 KB.
CSS	/wp-content/plugins/woocommerce/assets/css/woocommerce-layout.css	Unminified version.
CSS	/wp-content/plugins/contact-form-7/includes/css/styles.css	The form is only on /kontakt/ ; the stylesheet loads everywhere.
CSS	fonts.googleapis.com/css2?family=Poppins:wght@300..700	Five weights loaded, two used.
CSS	+ 3 more plugin files	Slider, gallery, cookie banner.
JS	/wp-includes/js/jquery/jquery.min.js	No <code>defer</code> . Required by the theme and three plugins.

JS	/wp-content/plugins/js_composer/assets/js/dist/js_composer_front.min.js	No defer .
JS	/wp-content/themes/woodmart/js/scripts/.../swiper.js	The slider script. No defer .
JS	+ 3 more scripts	Cookies, analytics ×2 (see the SEO section).

Two things deserve attention beyond the count itself. **Page-specific stylesheets load everywhere.** The form stylesheet and the single-product stylesheet both load on the homepage, where neither is used. Conditional loading is a caching plugin setting, not development work. **The Poppins font loads in five weights** while only two are used. Limiting it to 400 and 700 and adding `font-display: swap` saves about 180 KB.

What speed already gets right

Signal	Finding
Compression	gzip is enabled on all text resources: 738 KB transfers as 94 KB.
Protocol	HTTP/2 with HTTP/3 advertised. No duplicate requests.
Product images	174 main product images are already served as WebP with <code>srcset</code> . Only the slider isn't.
Browser cache	Static resources carry <code>cache-control: max-age=31536000, immutable</code> with a version string. Correct.
No redirect chains	No request on the critical path is redirected.

Accessibility: Lighthouse 68

Accessibility is not a separate topic from SEO. Six of the ten items below are also SEO findings, because they involve the same missing descriptions.

Issue	Scope	Detail
Missing <code>alt</code> text	411 images	The largest single item. On a product, a screen reader reads out the file name.
Wrong <code>lang</code> attribute	218 pages	<code>en-US</code> on Slovenian text. Screen readers pronounce Slovenian with English phonetics.
Insufficient color contrast	7 elements	The gray price on a light background reaches 3.1:1 (4.5:1 required). The same applies to the text under products in the grid.
Buttons without an accessible name	4 types	The cart, search, wishlist and banner-close buttons have only an icon and no <code>aria-label</code> .
Touch targets too small	12 elements	The category filter checkboxes are 28×28 px; the recommended minimum is 44×44 px.
Headings out of sequence	96 pages	An <code><h2></code> followed by an <code><h4></code> : product names in the WPBakery grid.
Two <code><h1></code> elements	Homepage	One in the slider, one in the welcome block.
Form without labels	/kontakt/	Three fields have only a <code>placeholder</code> , which disappears as you type.
Iframe without a title	2 maps	The embedded Google Maps for both locations have no <code>title</code> attribute.
Focus not visible	Entire site	The theme removes the focus outline and doesn't replace it. Keyboard use is practically impossible.

Lighthouse scores for the homepage: **Performance 41 · Accessibility 68 · Best Practices 79 · SEO 83** (mobile).
Desktop: Performance 72.

Security headers

Header	Value	Rating
<code>strict-transport-security</code>	missing	Recommended. Without it, the first visit over <code>http://</code> is vulnerable, even though it redirects.
<code>x-content-type-options</code>	missing	Add <code>nosniff</code> . One line.
<code>x-frame-options</code>	missing	Add <code>SAMEORIGIN</code> . Stops the store from being embedded in someone else's frame.

referrer-policy	strict-origin-when-cross-origin	Present and correct.
content-security-policy	missing	Worth setting up for a store with a checkout, but it needs preparation. Low priority in this set.
WordPress version	exposed	<meta name="generator" content="WordPress 6.6.2"> and ?ver=6.6.2 on resources. Remove the generator tag.

Migration leftovers

This site was built by adapting another one. We ask about this in every audit, because it is the thing a rebuilt site most often gets wrong, and because no tool looks for it.

We found **six** leftovers in five different places. Each is small on its own. Together they mean the site tells search engines the wrong legal entity.

#	Where	What was left behind	Scope
1	Organization schema	"legalName": "Old Company d.o.o.", "telephone": "+386 1 542 33 10", "email": "info@stara-domena.si"	218 pages
2	/sitemap-staro.xml	An old XML sitemap that returns HTTP 200 and lists 41 URLs. All 41 return 404.	1 file
3	/splosni-pogoji/	Article 4 names the previous company as the contracting party, with its tax number. This is binding legal text.	1 page
4	Delivery text on /nacini-placila/	»... na naslov cvetličarne« ["... to the florist's address"], a leftover from the previous business in visible text.	1 paragraph
5	Image file names	31 images in /wp-content/uploads/2019/ are named staro-*.jpg and predate the site.	31 files
6	Mixed content	Six images are still served from the external server http://stara-domena.si . If that domain goes down, the images disappear. See the section on mixed content.	9 pages

Why this isn't cosmetic. Google builds an **entity graph**. It tries to work out which legal entity stands behind a domain. Right now it gets two answers: the visible text says Example Company, the machine-readable data says the previous company, with its phone number and email. When the sources disagree, the search engine doesn't pick one. It trusts neither enough to settle the entity. This is one reason the company doesn't show up in local results the way its size and age say it should.

Item 3 is also not an SEO finding but a **legal** one. The terms and conditions are text the customer accepts at checkout, and they currently name as the contracting party a company that plays no part in the sale. We found this while reviewing content and mention it because it matters, not because it belongs in an audit.

The fix for item 2, because it is the least obvious: don't delete the file. Set it to return **HTTP 410 Gone** and remove the old XML sitemap in Search Console. A 410 tells the search engine the content was removed deliberately and permanently, which is a faster signal than a 404.

What already works

An audit that reports only problems is not an accurate description of this site. It isn't useful either, because it doesn't say what must not be broken. Every item comes with evidence.

Strength	Evidence
The XML sitemap exists and covers the whole catalog	218 URLs in five child sitemaps, no content page left out. The trailing-slash errors and the second sitemap are real problems, but the set is complete, which is half the work already done.
Product schema exists at all	Product on all 174 products. The 63 empty prices are a real error, but the foundation is in place. Most Slovenian stores don't have it and have to build it from scratch.
Clean Slovenian URLs	/izdelki/vrtne-skarje-gardena-b-s-xl/. No parameters, no session IDs, no duplicate paths, no English slugs. Better organized than most.
Canonical tags everywhere and correct	Self-referencing on all 218 pages. Paginated pages correctly point to themselves and not to page one, which is a common mistake.
The main sitemap contains no junk	The plugin's sitemap lists neither search results nor ?add-to-cart= URLs. The redundant archives come only from the second, core sitemap, which needs to be switched off, not fixed.
The blog is genuinely good	24 posts, averaging 1,140 words, specific planting dates for the Slovenian climate, signed and dated. This is content most stores pay for and don't get.
Product images are modern	174 main images are served as WebP with srcset at three widths. The slider fix repeats something the theme already knows how to do.
A verifiable endorsement	Authorized Gardena and Kärcher dealer and service center. An AI engine has already found the company through this signal. See the eighth question in the visibility measurement.
Opening hours are complete and accurate	For both locations, with summer and winter schedules. Only the markup is missing. The content is already written.
Real identifiers are published	Tax number and company registration number on /kontakt/, valid HTTPS, correct single-hop redirect for http and www.
Shallow structure	Every product is three clicks from the homepage. For a catalog of 174 products, that is well designed.

This has a practical consequence for the order of work: **don't rebuild this site**. The foundations are right. This report describes fixes to the existing structure, not a reason for a new one.

Proposed new URLs

Not “add more content”. Specific pages, each with a stated purpose and a reason drawn from what the site already says about itself and from the questions the AI engine couldn't answer. The slugs are for the Slovenian site, so they stay in Slovenian.

Group one: pages the site needs now

Proposed URL	Search intent it answers	Source of the proposal
/pogosta-vprasanja/	[FAQ] Eight to twelve real customer questions with FAQPage schema. The fastest route to visibility in AI answers.	AEO
/dostava/ (rewrite)	[Delivery] “How much is delivery” and “when do you deliver”. The page exists and holds placeholder text.	Visible error
/servis-vrtnne-opreme/	[Garden equipment service] Your strongest differentiator , authorized Gardena and Kärcher service, currently mentioned in one sentence on /o-nas/ .	AI measurement
/vrtni-center-domzale/	[Garden center Domžale] Location searches, with the full address, opening hours, map and GardenStore schema.	GEO
/vrtni-center-celje/	[Garden center Celje] The same for the second location. Celje has no page of its own on the site today.	GEO
/o-nas/ekipa/	[About us / team] Named staff with qualifications, including the blog author, who has no profile today.	E-E-A-T

Group two: content already written that needs reshaping

These aren't new pages. They are existing blog posts that get a question-style heading, an answer in the first 60 words, and schema. The work is editing, not writing.

Existing post	The question it takes on	Format
/blog/sadilni-koledar-jesen/	Kdaj saditi česen v Sloveniji? [When to plant garlic in Slovenia?]	Question + answer
/blog/zalivanje-v-rastlinjaku/	Kako pogosto zalivati paradižnik v rastlinjaku? [How often to water tomatoes in a greenhouse?]	Question + answer
/blog/priprava-gredice/	Kako pripraviti zelenjavno gredico spomladi? [How to prepare a vegetable bed in spring?]	Article
/blog/obrezovanje-sadnega-drevja/	Kdaj in kako obrezati sadno drevje? [When and how to prune fruit trees?]	Article

</blog/zimska-zascita-rastlin/>

Kako zaščititi rastline pred zmrzaljo? [How to protect plants from frost?]

Article

Group three: categories that need copy

All 12 categories are bare product grids today. A category is a page that can rank for a broad commercial term (*vrtna škarija* [garden shears], *semena zelenjave* [vegetable seeds], *substrati* [potting substrates]) and right now it says nothing about itself. Add 250–400 words of intro copy above the grid, explaining what a buyer is choosing between in that category and how to decide.

This is a starting point drawn from what the site already says about itself, not from search volume data. The final set and the priority order within it should be confirmed against real volumes in Ahrefs or Semrush, which this audit doesn't have access to.

Google Business Profile

Outside the site, but for a company with two physical locations it is very likely worth more than anything on the site itself. We couldn't verify it from the outside. Make it **the first thing you check before work begins**, because it ties directly into the address and schema recommendations.

Check	Why it connects to this report
Are there two separate, verified profiles , one for Domžale and one for Celje?	Two locations need two profiles. One profile for two locations is a common mistake that weakens both.
Does the address on the profile match exactly the one that will be published on the site?	Google compares name, address and phone across sources. A mismatch is worse than an omission, and right now you have three different versions.
Is the phone number on the profile the same as in the schema?	The schema currently lists the previous company's number. Fix 1 and the profile fix have to go together.
Is the primary category <i>Vrtni center</i> [Garden center], with <i>Servis vrtna opreme</i> [Garden equipment service] among the additional ones?	Service is your differentiator, and today it is declared nowhere: not on the site, not on the profile.
Are opening hours and holiday hours filled in?	The question “which garden center is open on Saturdays” went unanswered in the visibility measurement.
How many reviews does the profile have, and have they been answered?	The site has no reviews at all. If the profile has some, they are worth showing on the site too, with markup.
Does the profile link to the site, and does the link point to the right location page?	The proposed pages /vrtni-center-domzale/ and /vrtni-center-celje/ are exactly where these links should go.

Publishing one consistent address on the site and adding `GardenStore` schema that matches the profile exactly is what connects the two. Doing one without the other wastes most of the benefit.

Priority matrix

Ranked by impact. Everything rated Critical is a change to a setting, to copy, or to a single file. None of it requires a site rebuild. Order matters: some items are prerequisites for others, and each row says so where it applies.

#	Action	Priority	Expected impact
1	Remove the three incorrect <code>Disallow</code> lines from <code>robots.txt</code> : <code>uploads</code> , <code>plugins</code> , <code>wp-includes</code>	Critical	Opens 693 images to Google Images and lets Googlebot render the pages. Three lines.
2	Set <code>lang="sl-SI"</code> in <i>Settings</i> → <i>General</i>	Critical	Fixes all 218 pages at once. One setting.
3	Fix the <code>Organization</code> schema. <code>legalName</code> , <code>telephone</code> and <code>email</code> belong to the previous company	Critical	Stops declaring the wrong legal entity on 218 pages.
4	Replace the <code><h1></code> in the footer with a <code><div></code> , and the <code><h1></code> in the product grid with an <code><h3></code>	Critical	For the first time in the site's history, every page has exactly one <code><h1></code> . Also clears 96 accessibility warnings.
5	Fix the empty price in the schema. When there is no price, don't output the <code>price</code> field	Critical	63 products stop being rejected. The fix is one template rule, not 63 manual edits.
6	Fix the duplicated analytics: two GA4 containers	Critical	Do this before anything else , or you can't measure what the other fixes achieve.
7	Fix mixed content: move 6 images off the external server, load jQuery Migrate over https or remove it	Critical	Repairs the broken gallery on 9 pages and removes the dependency on someone else's server.
8	Rewrite <code>/dostava/</code> . It contains 84 words of Lorem ipsum	Critical	Removes the most visible error on the site, on the page a customer reads right before buying.
9	Replace the placeholder phone number <code>+386 40 123 456</code> in the footer	Critical	Right now a customer on three pages calls the wrong number.
10	Fix Article 4 of the terms and conditions. It names the previous company	Critical	A legal item, not an SEO one. The contracting party has to be correct.
11	Remove <code>OAI-SearchBot</code> from <code>robots.txt</code> ; decide on the other four separately	Critical	One line. While it stays, no schema fix will do anything for ChatGPT search. The other four govern model training and can stay.
12	Put <code>test.example-shop.com</code> behind a server-level password	Critical	Removes a second full copy of the site from search. A password, not <code>noindex</code> and not <code>robots.txt</code> .
13	Fix the soft 404. URLs that don't exist should return a 404 status	Critical	Stops the indexing of an unlimited number of nonexistent pages. One change in <code>404.php</code> .

14	Clean up the XML sitemaps: trailing slashes, disable <code>/wp-sitemap.xml</code> , 410 for <code>/sitemap-staro.xml</code> , <code>https</code> in the directive	High	Halves wasted crawl budget and removes conflicting signals.
15	Add <code>GardenStore</code> schema for both locations, with NAP, opening hours and geo coordinates	High	The biggest local search gain available. The opening-hours text already exists.
16	Make both location addresses identical everywhere: footer, contact page, schema	High	Prerequisite for item 13 and for matching Google Business Profile.
17	Fix the title template: shorten the suffix, use the product name instead of the site name	High	Fixes 31 duplicate and 44 truncated titles in one go.
18	Write meta descriptions for the 139 pages that have no usable one	High	Directly changes what Google shows in results. Start with the 12 categories and 7 static pages.
19	Optimize the hero slider: lazy-load five images, WebP, <code>fetchpriority</code>	High	Takes 3.9 MB off the critical path. LCP under 2 s with no visual change.
20	Fix the layout shift from the cookie banner by reserving its space	High	Field CLS from 0.21 to below 0.10. This is the measurement Google ranks on.
21	Install the Slovenian WooCommerce translations. Add to cart appears on 174 products	High	Removes English strings from every commercial page.
22	Write 411 image alt texts	High	Image search and accessibility at once. After item 1, the images can actually be reached.
23	Fix 23 broken and 17 redirected internal links	High	The list, with source pages, is in Appendix A3.
24	Deal with filter URLs: canonical to the base category and <code>noindex</code> on filtered views	High	Removes about 2,400 URLs from the crawl budget.
25	Redirect 693 attachment pages to their parent page	High	One setting in the SEO plugin. Removes 693 empty pages at once.
26	Publish the opening hours as text. Today they are a JPEG	High	Search engines, screen readers and AI assistants can read them for the first time. Prerequisite for the location schema item.
27	Fix <code>og:url</code> . On all 218 pages it points to the homepage	High	Shared product links stop resolving to the homepage.
28	Turn phone numbers into <code>tel:</code> links (7 places)	High	One-tap calling on mobile. For a local business, every missed tap is a lost call.
29	Set <code>noindex</code> on <code>/hvala-za-narocilo/</code> and remove it from the sitemap	High	Search visits stop firing the conversion goal by mistake.
30	Mark up the 24 blog posts with <code>Article</code> and publish an author page	Medium	Three years of expert writing starts building authority. No new words needed.

31	Publish <code>/pogosta-vprasanja/</code> with 8–12 questions and <code>FAQPage</code> schema	Medium	The fastest route into AI answers and voice search.
32	Set up <code>hreflang</code> across the three language versions, reciprocal, with <code>x-default</code>	Medium	The three copies stop competing with each other.
33	Write 250–400 words of intro copy for all 12 categories	Medium	A category can start ranking for its broad commercial term.
34	Restructure 9 blog posts into question-and-answer format and mark them up with <code>Article</code> (author, date)	Medium	Editorial work on existing content. The titles are listed in the AEO section.
35	Publish <code>/servis-vrtne-opreme/</code> : authorized Gardena and Kärcher service	Medium	Your strongest signal is currently one sentence. An AI engine has already found you through it.
36	Cut the number of render-blocking files: conditional loading, two font weights, <code>defer</code>	Medium	Removes about 180 KB and 770 ms of resource delay.
37	Add per-page <code>og:image</code> and <code>sameAs</code> pointing to the two existing profiles	Medium	Shared links stop rendering as a narrow strip with a logo.
38	Add <code>BreadcrumbList</code> schema. Breadcrumbs are visible on 205 pages but not marked up	Medium	Google can show them in results instead of a bare URL.
39	Set <code>noindex</code> on 41 tag archives and 24 monthly archives	Medium	Removes 65 thin, duplicate pages. One plugin setting.
40	Add consent and submission tracking to the contact form	Medium	Form submissions are currently not tracked anywhere, and there is no consent.
41	Disable directory browsing and <code>xmlrpc.php</code> , update three plugins	Low	Server hygiene. One of the plugins is three major versions behind.
42	Complete the site icons: <code>apple-touch-icon</code> and 512×512	Low	Phone home screens and browser tabs currently show a generic icon.
43	Add security headers: <code>HSTS</code> , <code>nosniff</code> , <code>SAMEORIGIN</code>	Low	Three lines in the server configuration.
44	Fix contrast, accessible button names, visible focus and touch target size	Low	Lighthouse accessibility from 68 toward 95+.
45	Rename 209 images that still carry camera file names. Mainly a rule for new images from now on	Low	Requires fixing links in the content. Low return for the effort.

Where to start. Items 1 to 13 are critical, and it makes sense to do them in a single pass. None of them needs new content or template rework: nine are a change to one setting, one line or one file. If only this block gets done, every error that makes the site tell search engines wrong things about itself is gone. Everything after that is improvement, not repair.

Appendix A: work list by URL

For the developer. Each item with its URL, current value and what it needs to become. Where a list runs past 40 rows, the full list comes as a CSV file and the table shows the first 20.

A1: Pages without a usable meta description (139)

First 20 shown; the full list is in `manjkajoci-opisi.csv`.

URL	Current	Characters
/storitve/	empty	0
/dostava/	empty	0
/nacini-placila/	empty	0
/o-nas/	empty	0
/kategorija/vrtno-orodje/	empty	0
/kategorija/semena-zelenjave/	empty	0
/kategorija/substrati-in-gnojila/	empty	0
/kategorija/zimska-zascita/	empty	0
/kategorija/lonci-in-korita/	empty	0
/kategorija/namakanje/	empty	0
/kategorija/visokotlacni-cistilci/	empty	0
/kategorija/sadike-sadnega-drevja/	empty	0
/kategorija/travne-mesanice/	empty	0
/izdelki/gardena-b-s-xl/	34,90 €	7
/izdelki/karcher-k5-premium/	289,00 €	8
/izdelki/substrat-univerzalni-70l/	12,50 €	7
/izdelki/semi-cesna-jesenski/	4,20 €	6
/izdelki/namakalni-sistem-osnovni/	empty	0
/izdelki/lonci-terakota-40cm/	empty	0

/izdelki/travna-mesamica-sencna-5kg/

empty

0

+ 119 more pages in manjkajoci-opisi.csv

A2: Products with a duplicate title (31)

All of them have `<title>Example Company</title>`, because the title field is empty and the template falls back to the site name. Item 15 of the priority matrix fixes all of them, not 31 manual edits. Full list in `podvojeni-titli.csv`.

URL	Product name on the page
/izdelki/gardena-b-s-xl/	Gardena B/S XL vrtné škarije [pruning shears]
/izdelki/gardena-comfort-500/	Gardena Comfort 500 škropilnica [sprayer]
/izdelki/karcher-k5-premium/	Kärcher K5 Premium Smart Control
/izdelki/karcher-sc3-parni-cistilec/	Kärcher SC 3 parni čistilec [steam cleaner]
/izdelki/substrat-univerzalni-70l/	Univerzalni substrat 70 l [all-purpose potting soil]
/izdelki/lonec-terakota-40cm/	Lonec terakota 40 cm [terracotta pot]
/izdelki/namakalni-sistem-osnovni/	Osnovni namakalni sistem [basic irrigation system]
+ 24 more products in podvojeni-titli.csv	

A3: Broken and redirected internal links (40)

Each link is listed with the page it comes from. Without that, nobody can fix them.

Link points to	HTTP	Linked from
/izdelki/gardena-set-2022/	404	/blog/darilne-ideje-za-vrtnarje/
/izdelki/karcher-k4-compact/	404	/kategorija/visokotlacni-cistilci/ , /blog/ciscenje-terase/
/izdelki/semi-paradiznik-cherry-2023/	404	/blog/zalivanje-v-rastlinjaku/
/blog/akcija-pomlad-2023/	404	Footer, on all 218 pages
/cenik-storitev/	404	/storitve/
/trgovina/	301	Main menu, on all 218 pages
/kategorija/orodje/	301	Footer, /o-nas/

`/blog/sadilni-koledar/`

301

Three product pages

+ 18 more 404 errors and 14 redirects in `povezave.csv`

Two items fix most of it. The links to `/blog/akcija-pomlad-2023/` and `/trgovina/` sit in the footer and the main menu, so they appear on all 218 pages. Two fixes in one place clear 436 of the 476 bad links on the site.

A4: Products with an empty price in the schema (63)

All of them are currently out of stock. The fix is a template rule (item 5), not 63 edits. Full list in `prazne-cene.csv`. After the fix, check Search Console under *Enhancements* → *Products*, where these 63 should be listed as errors.

A5: Orphan pages with no inbound link (9)

URL	Suggested inbound link
<code>/kategorija/zimska-zascita/</code>	Main menu (seasonal) + <code>/blog/zimska-zascita-rastlin/</code>
<code>/kategorija/semena-zelenjave/</code>	Main menu + <code>/blog/sadilni-koledar-jesen/</code>
7 product pages	A “related products” block on the matching category; full list in <code>osirotele.csv</code>

A6: Pages with mixed content (9)

`/o-nas/`, three product pages and five blog posts. All of them load `http://code.jquery.com/jquery-migrate-1.4.1.min.js`; six also load images from `http://stara-domena.si`. Full list in `mesana-vsebina.csv`.

Appendix B: content specs

For whoever does the writing. Specs, not theory.

Element	Spec
Title	Unique for every URL, 50–60 characters including the suffix. Main keyword near the start. Suffix no longer than Example Company . Example: Gardena B/S XL vrtne škarje za veje do 42 mm Example Company [Gardena B/S XL pruning shears for branches up to 42 mm]
Meta description	130–155 characters including spaces. Contains the keyword and ends with a call to action. Never just the price. Slovenian inflection is not a problem; Google handles Slovenian morphology.
H1	Exactly one per page , in the content and not in the footer. It says what the page is, contains the keyword and matches what the URL says.
Category copy	250–400 words above the product grid. What a customer is choosing between in this category, how to decide, which brands you carry. Not a description of each product.
Product description	At least 120 words, of which at least 40 are your own, not copied from the manufacturer's catalog. Specifications as a list, because a machine can quote a list verbatim.
Answer to a question	The heading is the question as a customer would ask it. Directly under it, 40–60 words that answer the question fully and on their own. The explanation comes after. This is the format that can become a featured snippet.
Image file name	Lowercase, hyphens, descriptive, contains the keyword: gardena-b-s-xl-vrtne-skarje.webp . Never spaces or dots in the name.
Alt text	Describes what the image shows, as a sentence fragment. Not a keyword list. Empty alt="" only for purely decorative images.
Image weight	WebP or AVIF, at display size, with srcset for three widths. loading="lazy" below the fold, fetchpriority="high" for the LCP element.

Appendix C: copy-paste code

Prepared for your site, with made-up values in the places you need to replace. In a real report those places hold your actual data.

C1: A correct robots.txt

```
User-agent: *  
Disallow: /wp-admin/  
Allow: /wp-admin/admin-ajax.php  
Disallow: /*?s=  
Disallow: /*?add-to-cart=  
# cart, checkout and my-account are deliberately NOT listed here. See the note below  
  
Sitemap: https://example-shop.com/sitemap_index.xml
```

This removes all three blocks that prevented rendering and image indexing, both Sitemap: directives on http:// , the dead sitemap and Crawl-delay . Decide on the AI crawler blocks line by line. The version above contains none of them.

Why cart, checkout and account are not on this list. Most advice gets this wrong: a Disallow in robots.txt **does not remove** a URL from the index. It stops the crawler from reading the page, so Google never sees the noindex tag on it and can still show the URL, just without a description. For cart, checkout and my account, the correct fix is the **opposite**: leave them open to crawling and set noindex on them. WooCommerce does this by default. Check that nobody has turned it off.

C2: Schema for both locations

One block per location page. The @id has to be stable and unique, because entity linking points to it.

```
{
  "@context": "https://schema.org",
  "@type": "GardenStore",
  "@id": "https://example-shop.com/vrtni-center-domzale/#store",
  "name": "Example Company – Domžale",
  "parentOrganization": { "@id": "https://example-shop.com/#organization" },
  "url": "https://example-shop.com/vrtni-center-domzale/",
  "telephone": "+386 1 721 45 60",
  "email": "info@example-shop.com",
  "address": {
    "@type": "PostalAddress",
    "streetAddress": "Kamniška cesta 41a",
    "postalCode": "1230",
    "addressLocality": "Domžale",
    "addressCountry": "SI"
  },
  "geo": { "@type": "GeoCoordinates", "latitude": 46.1389, "longitude": 14.5947 },
  "openingHoursSpecification": [
    { "@type": "OpeningHoursSpecification",
      "dayOfWeek": ["Monday", "Tuesday", "Wednesday", "Thursday", "Friday"],
      "opens": "08:00", "closes": "19:00" },
    { "@type": "OpeningHoursSpecification",
      "dayOfWeek": "Saturday", "opens": "08:00", "closes": "13:00" }
  ],
  "sameAs": [
    "https://www.facebook.com/exampleshop",
    "https://www.instagram.com/example_shop"
  ]
}
```

The corrected `Organization` block that the one above refers to. It replaces the block that currently names the previous company:

```
{
  "@context": "https://schema.org",
  "@type": "Organization",
  "@id": "https://example-shop.com/#organization",
  "name": "Example Company",
  "legalName": "Example Company d.o.o.",
  "vatID": "SI12345678",
  "foundingDate": "1998",
  "url": "https://example-shop.com",
  "logo": "https://example-shop.com/wp-content/uploads/logo.png",
  "telephone": "+386 1 721 45 60",
  "email": "info@example-shop.com",
  "sameAs": ["https://www.facebook.com/exampleshop",
    "https://www.instagram.com/example_shop"]
}
```

C3: Reciprocal hreflang

On **every** matching page of all three versions, including the self-reference. If one is missing, Google ignores the whole set.

```
<link rel="alternate" hreflang="sl-SI" href="https://example-shop.com/izdelki/gardena-b-s-xl/">
<link rel="alternate" hreflang="en" href="https://example-shop.com/en/products/gardena-b-s-xl/">
<link rel="alternate" hreflang="hr-HR" href="https://example-shop.hr/proizvodi/gardena-b-s-xl/">
<link rel="alternate" hreflang="x-default" href="https://example-shop.com/izdelki/gardena-b-s-xl/">
```

C4: FAQPage schema for the new page

```
{
  "@context": "https://schema.org",
  "@type": "FAQPage",
  "mainEntity": [{
    "@type": "Question",
    "name": "Ali servisirate opremo Gardena in Kärcher, kupljeno drugje?",
    "acceptedAnswer": {
      "@type": "Answer",
      "text": "Da. Kot pooblaščen servis za Gardena in Kärcher sprejmemo opremo
              ne glede na to, kje je bila kupljena. Oceno popravila pripravimo
              v dveh delovnih dneh, garancijska popravila so brezplačna."
    }
  }]
}
```

[The question reads “Do you service Gardena and Kärcher equipment bought elsewhere?” The answer says yes, as an authorized service center, with a repair estimate in two business days and free warranty repairs.] The rule is simple: **the text in acceptedAnswer must match** the visible text on the page word for word. If they differ, Google ignores the markup, and repeat cases get penalized.

C5: Fixing the empty price in the template

```
// wrong: current state
"price": ""

// correct: when there is no price, omit the field
<?php if ( $product->get_price() !== '' ) : ?>
  "price": "<?php echo esc_js( $product->get_price() ); ?>",
<?php endif; ?>
"priceCurrency": "EUR",
"availability": "https://schema.org/<?php echo $product->is_in_stock()
                ? 'InStock' : 'OutOfStock'; ?>"
```

What we can't measure from outside

Stated explicitly so nobody reads this report as the complete picture. The items below need access to the company's accounts, and we have not guessed at them.

Signal	Where to get it
Backlink profile and domain authority	Ahrefs, Semrush or Moz. It isn't visible in the site's code.
Actual impressions, clicks and queries	Google Search Console. It also has the <i>Products</i> report, which after fix 5 will show whether the 63 errors are gone.
Conversion data	Google Analytics, but every number to date is doubled (see the SEO section). The second container has to go before anything is measured.
State of the Google Business Profile listings	Check this first; it ties directly into recommendations 13 and 14.
Search volumes for the proposed new URLs	Ahrefs or Semrush. Our proposal comes from the site's content and from the AI visibility test, not from volume data.
What happens in checkout	The payment flow needs a real session with a product in the cart. We didn't trigger it; the audit places no orders.
Stock data per location	The back-office system. "Do you have this in Domžale?" is one of the most common questions, and right now nothing on the site answers it.

Glossary

Every technical term used in this report, in plain English. The names match what you will see in the tools and what developers call these things.

Term	What it means
SEO	Search Engine Optimization. Everything that decides where a page ranks in Google's regular list of results.
GEO	Generative Engine Optimization. How readable and citable a site is for AI search: ChatGPT, Perplexity, Google AI Overviews, Gemini. These engines compose an answer and cite sources, so they reward clarity, verifiable facts and a declared identity.
AEO	Answer Engine Optimization. Whether a page can serve as a direct answer: a featured snippet, a “People also ask” box, or the spoken reply of a voice assistant.
Crawl / crawler	A search engine fetching pages automatically. Googlebot is a crawler. A page the crawler can't reach effectively doesn't exist in search.
Crawl budget	How many requests a search engine is willing to spend on your site in a given period. Redirects, 404 errors and duplicate paths use it up and return nothing.
Index	The search engine's stored copy of a page. Crawled doesn't mean indexed, and indexed doesn't mean ranking.
Title tag	The <code><title></code> element. The blue clickable line in search results and the text in the browser tab. About 50–60 characters get displayed.
Meta description	The gray summary under the title in results. It doesn't affect ranking directly, but it decides how many people click. If it's missing, Google writes its own from whatever text it finds.
H1	The main heading of a page, marked up with <code><h1></code> . One per page , saying what the page is about. H2 and H3 are subheadings under it. More than one H1 leaves the search engine unsure which one is the main one.
lang attribute	<code><html lang="sl-SI"></code> declares the language a page is written in. Search engines and screen readers rely on it.
hreflang	Tags that tell a search engine which pages are language versions of the same content. They must be reciprocal : if A declares B, B must declare A, or the set is ignored.
Canonical	A tag that sets the definitive URL of a page, so the search engine knows which version to rank when the same content is reachable at several addresses.
robots.txt	A text file in the site root that tells crawlers what they may not fetch and points them to the XML sitemap. It does not hide content . It only forbids fetching.
sitemap.xml	A machine-readable list of every page. Without it, search engines have to find pages by following links. Two different sitemaps at once are worse than one.

410 Gone	A server response meaning “this content was removed on purpose, permanently.” For a search engine it is a clearer and faster signal than a 404.
Orphan page	A page that exists and loads, but no other page links to it. If it isn't in the XML sitemap either, it is practically invisible.
Mixed content	An HTTPS page that loads resources over unencrypted <code>http://</code> . Browsers usually block images silently; scripts are always blocked, so any feature that depends on them stops working.
Alt text	A written description of an image in the <code>alt</code> attribute. Screen readers use it, it appears when the image fails to load, and it is the only way a search engine knows what the image shows.
Schema / JSON-LD	Structured data: a block of machine-readable facts embedded in the page. Who the company is, where it is, what it sells. <code>JSON-LD</code> is the format Google recommends.
LocalBusiness / GardenStore	Schema types for a business with a physical location. <code>GardenStore</code> is the specific type for a garden center and carries the address, opening hours and services.
Product / Offer	Product and offer schema. <code>Offer</code> carries the price, currency and availability. An empty price is not a price: when a value is invalid, Google discards the whole offer.
FAQPage	Markup around a list of questions and answers. Google limits the rich result to health and government sources, so on a commercial site the gain comes from the content (the question as a heading, the answer right below it), not from the markup itself.
Open Graph	Tags (<code>og:title</code> , <code>og:image</code>) that control how a link looks when someone shares it on Facebook, WhatsApp or Viber. Without them, a shared link is a bare URL.
sameAs	A schema property that links the company to its profiles elsewhere and confirms to search engines that they are the same entity.
NAP	Name, Address, Phone. Three details that must match exactly everywhere, on the site and off it. Inconsistent NAP is the most common cause of weak local rankings.
E-E-A-T	Experience, Expertise, Authoritativeness, Trustworthiness. Google's framework for judging whether a source deserves trust: named people, real qualifications, verifiable contact details.
Featured snippet	A boxed answer Google sometimes shows above the regular results, taken from a page that answered the question concisely.
Core Web Vitals	Google's three measured indicators of how a page performs in use: LCP, CLS and responsiveness to interaction (INP).
LCP	Largest Contentful Paint. The time until the largest visible element is rendered. Under 2,500 ms is good.
CLS	Cumulative Layout Shift. How much the layout moves while loading. Under 0.1 is good, 0 is perfect.
INP	Interaction to Next Paint. How long it takes from a click to a visible response. Under 200 ms is good.
TTFB	Time to First Byte. How long the server takes to start responding. Under 800 ms is good. It measures the server, not the design.

Render-blocking	A file the browser has to download and process before it can draw anything. Each one delays the moment a visitor sees the page.
CrUX / field data	Chrome User Experience Report. Google's collection of measurements from real visitors, 28 days, 75th percentile. When it disagrees with lab data, field data wins: it is what Google ranks on.
Lab data	One run on one device under controlled conditions (Lighthouse). Useful for finding a cause, not for judging where you stand.
HSTS	A security header that tells browsers to always use HTTPS for this domain and never plain HTTP.
Gunning Fog	A readability index that estimates how many years of schooling it takes to understand a text on first reading. Because of its morphology, Slovenian scores two to four points higher than English, so we compare values across pages rather than against the English scale.
Soft 404	A page that tells the visitor “not found” while the server returns it with status 200 OK . The search engine treats it as valid content and indexes it instead of dropping it.
Attachment page	A page WordPress creates automatically for every uploaded image. It holds one image and nothing else, and its title duplicates the page where the image is actually used.
Filter URL	A URL with parameters, created when someone applies a filter in the store (<code>?filter_znamka=...</code>). To a search engine every combination is a separate page, even though it is the same set of products in a different order.
Staging copy	A parallel install of the site used for development (often on <code>test.</code> or <code>staging.</code>). Without password protection, search engines index it as a separate site that competes with production.
BreadcrumbList	Schema for breadcrumbs, the path a visitor took to reach a page. It lets Google show a readable path in results instead of a bare URL.
Text in images	Content that exists only as pixels: opening hours, a price list or an address saved as a JPEG. A visitor can read it. A search engine, a screen reader or an AI assistant can't.
Lighthouse	Google's audit tool built into Chrome. It checks whether good practices are in place, not whether the content is correct. A site can score 100 and still be wrong.

What you receive

To make the sample complete: this is the full scope of delivery, with no surcharges and no add-on packages. The report is delivered in English or Slovenian, as PDF and DOCX, in three to five business days.

Document	Scope	Format
Full audit , in English	This document: every page, every signal, the priority matrix, the appendices	PDF + DOCX
Full audit , in Slovenian	Available if you want both languages. Same findings, same numbers, same order. Not a machine translation but a second version of the same report.	PDF + DOCX

Attached CSV files	Lists too long for a table: missing descriptions, duplicate titles, broken links, empty prices, orphan pages, mixed content.	CSV
Raw crawl data	fetchLog.txt with the response for every page, a structured data dump, the Lighthouse reports.	on request

The document is yours. If you decide to make the fixes yourself or with your current developer, that's fine. The report is written so you can, and the priority matrix is deliberately laid out as a task list ranked by impact. Nothing is held back to push you into ordering the implementation.

What would we find on your site?

Every site has its own set. Some have three critical items, others fifteen. A site with nothing to fix on day one practically doesn't exist.

It starts with a free, quick look at your domain: we tell you how many pages the site has, what your structured data says about you, and which three things we would fix first. If that convinces you, the full audit follows. If not, you still walk away with three concrete findings.

Book a 20-minute call.

info@4tech-agency.com · +386 41 221 265 · 4tech-agency.com