



SEO Fundamentals

A Complete 10-Chapter Web Development Course

Topics covered:

How search engines crawl, index, and rank · on-page SEO · content strategy & E-E-A-T
Structured data & rich snippets · technical SEO & sitemaps · performance as a ranking factor
URL structure & internal linking · backlinks & authority · measuring real results

Closing chapter: a real audit of osztromok.com, not a hypothetical page

Exercises: 30 hands-on scenarios with worked solutions

Format: A4 · Dark-theme code examples

Philip Osztromok · Generated with Claude

Table of Contents

- 01 What SEO Actually Is & How Search Engines Work
- 02 On-Page SEO: Titles, Meta Descriptions & Headings
- 03 Content Strategy, Search Intent & E-E-A-T
- 04 Structured Data & Rich Snippets
- 05 Technical SEO: Crawlability, Indexability & Sitemaps
- 06 Site Performance as a Ranking Factor
- 07 URL Structure, Internal Linking & Site Architecture
- 08 Off-Page SEO: Backlinks & Authority
- 09 Measuring & Monitoring SEO
- 10 Capstone: A Real Audit of osztromok.com

CHAPTER 1 OF 10

What SEO Actually Is & How Search Engines Work

SEO Fundamentals

Chapter 1 · What SEO Actually Is & How Search Engines Work

SEO gets talked about as a bag of tricks — magic words to sprinkle on a page. It isn't. It's the practical consequence of three separate systems, each with its own real constraints: a page has to be **crawled**, then **indexed**, then **ranked** — and nearly everything this course covers maps onto making one of those three stages go well.

The Three-Stage Pipeline

STAGE	WHAT ACTUALLY HAPPENS	WHAT CAN GO WRONG
Crawling	A bot (Googlebot) discovers a page by following links or reading a sitemap	A page nobody links to, and that isn't in any sitemap, may never be found at all
Indexing	A crawled page's content is processed and stored in the search engine's own massive index	Crawling doesn't guarantee indexing — thin or duplicate content can be crawled and still left out
Ranking	Among every indexed page, an algorithm orders results for a specific query	Being indexed doesn't guarantee visibility — hundreds of signals decide the order

These are genuinely separate hurdles. A page can fail at any one of them independently of the others — beautifully written, keyword-perfect content that nothing links to never gets past stage one, no matter how good stage three would have gone.

SEO vs. Paid Search

The results in a search page split into two genuinely different systems: **organic** results, produced by the crawl/index/rank pipeline above, and **paid** results (search engine marketing, SEM), produced by an entirely separate real-time auction where advertisers bid for placement. This course is about the organic side only — SEM is a real, substantial topic of its own, deliberately out of scope here rather than covered thinly.

An Honest Caveat, Stated Early

NO ONE OUTSIDE GOOGLE HAS THE EXACT ALGORITHM

Ranking uses hundreds of signals, weighted in ways Google treats as a trade secret and updates continuously. What's actually known publicly comes from Google's own official statements, patent filings, and controlled experiments run by the SEO industry over many years — genuinely useful, well-supported knowledge, but not a leaked source file. Any course, including this one, that claims certainty about the exact ranking formula is overselling what's actually knowable.

This Course's Own Roadmap

WHERE EACH PIPELINE STAGE GETS ADDRESSED

- **Crawling** — [seo1-5](#) (robots.txt, sitemaps) and [seo1-7](#) (URL structure, internal linking — how pages actually get discovered)
- **Indexing** — [seo1-5](#) again (canonical tags, noindex) and [seo1-4](#) (structured data helping a search engine understand what was indexed)
- **Ranking** — [seo1-2](#) (titles/headings), [seo1-3](#) (content/E-E-A-T), [seo1-6](#) (performance), [seo1-8](#) (backlinks) — the largest share of chapters, since ranking draws on the most signals

[seo1-9](#) covers measuring whether any of this actually worked, and [seo1-10](#)'s own capstone traces this exact pipeline for a real site.

Hands-On Exercises

EXERCISE 1

Explain why a page can be crawled but still not indexed, and give a plausible real reason a search engine might make that decision.

 [View solution](#)

EXERCISE 2

Explain the structural difference between organic and paid search results — not just "one costs money" but what actually produces each set of results.

 [View solution](#)

EXERCISE 3

Explain why "no one outside Google has the exact algorithm" doesn't mean SEO knowledge is guesswork — what kinds of evidence does the field actually rely on instead?

 [View solution](#)

CHAPTER 1 QUICK REFERENCE

- **Crawl → Index → Rank** — three genuinely separate stages; a page can fail at any one independently of the others
- **SEO ≠ SEM** — organic ranking vs. a real-time paid auction; this course covers organic only
- No exact ranking algorithm is public — known SEO practice rests on official statements, patents, and controlled experiments, not a leaked formula
- This course's own roadmap: crawling (seo1-5/seo1-7), indexing (seo1-4/seo1-5), ranking (seo1-2/3/6/8), measurement (seo1-9), and a real capstone audit (seo1-10)
- **Next chapter:** On-Page SEO: Titles, Meta Descriptions & Headings

CHAPTER 2 OF 10

On-Page SEO: Titles, Meta Descriptions & Headings

SEO Fundamentals

Chapter 2 · On-Page SEO: Titles, Meta Descriptions & Headings

`htm12-5`'s own "Meta Tags" chapter named SEO basics as one of its topics without ever teaching how to actually write a good title or description. This chapter delivers on that preview in full — the highest-leverage on-page work a page can have, and genuinely quick to get right once the reasoning behind it is clear.

The Title Tag — Doing Two Jobs at Once

```
<title>Hiragana Stroke Order Guide – Philip's Learning Blog</title>
```

The `<title>` is both a real ranking signal (`seo1-1`'s own ranking stage reads it directly) and the clickable blue link shown in a results page. A title stuffed with repeated keywords can work against both jobs at once — it reads as manipulative to a ranking algorithm built to detect exactly that pattern, and it looks unappealing enough to a real person that it lowers the click-through rate even for a page that did rank well.

WEAK TITLE	WHY IT'S WEAK
"Hiragana Hiragana Learn Hiragana Guide Hiragana"	Repetition reads as manipulation, not relevance
"Home"	Says nothing about the page's own actual content
A 200-character title	Search results truncate around ~60 characters — the rest is wasted

Meta Descriptions — Not a Ranking Signal, but a Real Lever

```
<meta name="description" content="A clear, illustrated guide to hiragana stroke order, with example words for each character.">
```

AN HONEST DISTINCTION WORTH GETTING RIGHT

Google has stated directly that the meta description text itself doesn't factor into ranking. What it does is shape the snippet shown under a result — and a well-written one measurably improves click-through rate for a given rank,

which is a real, indirect benefit worth the two minutes it takes to write. Skip it and the search engine will auto-generate a snippet from the page's own text, which is often an awkward, out-of-context fragment.

Headings — Structure a Crawler Can Actually Use

[web-accessibility1-2](#) already covered heading hierarchy in depth for screen-reader users navigating by landmark. The exact same structure a screen reader relies on to understand a page's own outline is what a crawler uses to understand which content is the main subject versus a subsection — one clean `<h1>`, logically nested `<h2>`/`<h3>` underneath, no skipped levels.

THE SAME STRUCTURE, TWO DIFFERENT BENEFICIARIES

[web-accessibility1-2](#) taught this pattern for a screen reader's own navigation model. Nothing about the markup changes here — the accessibility win and the SEO win come from the identical heading structure, built once.

Hands-On Exercises

EXERCISE 1

Explain why a keyword-stuffed title tag can hurt a page in two separate ways — not just one — using this chapter's own reasoning.

 [View solution](#)

EXERCISE 2

Explain why a meta description can be worth writing carefully even though it isn't a ranking signal, and what happens if a page skips it entirely.

 [View solution](#)

EXERCISE 3

Explain why the exact same heading structure serves both an accessibility purpose ([web-accessibility1-2](#)) and an SEO purpose, rather than needing two separate markup approaches.

 [View solution](#)

CHAPTER 2 QUICK REFERENCE

- Delivers in full on html2-5's own preview — titles and descriptions taught as a real skill, not a mention
- `<title>` is both a real ranking signal and the clickable result text — keyword-stuffing hurts both at once
- Meta descriptions don't affect ranking directly, but shape click-through rate — worth writing, honestly framed
- Heading hierarchy is the exact same structure web-accessibility1-2 taught — one build, two beneficiaries
- **Next chapter:** Content Strategy, Search Intent & E-E-A-T

CHAPTER 3 OF 10

Content Strategy, Search Intent & E-E-A-T

SEO Fundamentals

Chapter 3 · Content Strategy, Search Intent & E-E-A-T

`seo1-1`'s own ranking stage draws on hundreds of signals — but the content itself, what a page actually says, is what most of them are ultimately trying to evaluate. This chapter covers writing content a search engine (and a real reader) has genuine reason to trust.

Search Intent — What Is the Searcher Actually Trying to Do?

INTENT TYPE	EXAMPLE QUERY	WHAT THE PAGE NEEDS TO DELIVER
Informational	"how does hiragana work"	An explanation — not a product, not a login form
Navigational	"philip's learning blog hiragana"	The searcher already knows the destination; get out of the way
Transactional	"buy hiragana flashcards"	A clear path to actually completing the action

A page that ranks for a query but doesn't match what the searcher was actually trying to do produces a bad outcome for everyone — the searcher bounces immediately, and that bounce is itself a real signal ranking systems can pick up on over time. Matching intent isn't a nice-to-have layered on top of good content; it's what "relevant" actually means.

Correcting an Outdated Myth

KEYWORD DENSITY IS NOT A REAL RANKING LEVER

Older SEO advice pushed hitting a specific keyword percentage — repeating a target phrase some fixed number of times per hundred words. Modern ranking systems evaluate meaning and topical coverage, not a literal count.

Content written naturally, actually covering the subject a reader searched for, outperforms content padded to hit an arbitrary repetition target — and reads better too, which itself feeds back into engagement signals.

E-E-A-T — Experience, Expertise, Authoritativeness, Trustworthiness

A framework Google's own quality-rating guidelines use to describe what separates genuinely reliable content from content that merely looks reliable:

- **Experience** — has the author actually done the thing being described, not just researched it secondhand?
- **Expertise** — does the content demonstrate real, specific knowledge of the subject?
- **Authoritativeness** — is this source recognized, by others, as a legitimate voice on this topic?
- **Trustworthiness** — is the site and the content itself honest, accurate, and transparent?

AN HONEST NOTE ON WHAT E-E-A-T IS AND ISN'T

E-E-A-T isn't a literal, measurable score a page receives — it's a quality framework human raters use, and one Google has stated broadly informs ranking systems trained partly on those same quality judgments. Treat it as a genuine design principle for how to build trustworthy content, not a checklist item with a numeric target.

Hands-On Exercises

EXERCISE 1

Explain why a page ranking for a query but mismatching the searcher's actual intent produces a bad outcome, and why that outcome itself becomes a signal ranking systems can use.

 [View solution](#)

EXERCISE 2

Explain specifically why keyword density targets are outdated advice, contrasting what older systems could evaluate against what modern ranking systems actually evaluate.

 [View solution](#)

EXERCISE 3

Explain why this chapter describes E-E-A-T as "not a literal, measurable score" — what would be lost by treating it as a numeric checklist instead of a design principle?

 [View solution](#)

CHAPTER 3 QUICK REFERENCE

- **Search intent** — informational/navigational/transactional; matching it is what "relevant" actually means
- **Keyword density is dead** — modern ranking evaluates meaning and topical coverage, not repetition counts
- **E-E-A-T** — Experience, Expertise, Authoritativeness, Trustworthiness; a quality framework, not a literal score
- Content written to genuinely serve a reader outperforms content engineered to satisfy an outdated formula

- **Next chapter:** Structured Data & Rich Snippets

CHAPTER 4 OF 10

Structured Data & Rich Snippets

SEO Fundamentals

Chapter 4 · Structured Data & Rich Snippets

[html_lesson_09_2](#) already taught the markup mechanics of microdata and Schema.org. What it didn't cover is why a search engine cares — this chapter delivers that half in full: structured data as a direct line to a richer, more clickable search result.

What Structured Data Actually Does

A page's own visible text tells a human reader what it's about. Structured data tells a machine the exact same thing, unambiguously — this is a recipe, this field is the cook time, this field is the rating. Without it, a search engine has to infer meaning from plain prose; with it, the meaning is stated directly, in a format built specifically to be machine-read.

JSON-LD — The Practical Default

```
<script type="application/ld+json">
{
  "@context": "https://schema.org",
  "@type": "Article",
  "headline": "Hiragana Stroke Order Guide",
  "author": { "@type": "Person", "name": "Philip Osztromok" },
  "datePublished": "2026-01-15"
}
</script>
```

[html_lesson_09_2](#) covered inline microdata ([itemscope](#) / [itemprop](#) attributes woven directly into visible HTML). JSON-LD does the identical job — describing the same [schema.org](#) vocabulary — as one self-contained script block, entirely separate from the visible markup. This is now the format Google's own documentation recommends by default, specifically because it doesn't require touching the page's own display HTML at all.

The Actual Payoff — Rich Snippets

WHAT STRUCTURED DATA CAN UNLOCK

A recipe page with correctly marked-up structured data can show its own star rating and cook time directly in the search result, before a single click. A how-to guide can show numbered steps. An event page can show its own date and location. None of this is guaranteed — eligibility and actual display are decided by the search engine — but the data has to be present and correctly marked up before any of it becomes possible at all.

STRUCTURED DATA IS NOT A RANKING FACTOR BY ITSELF

Adding `schema.org` markup does not, on its own, move a page up in rankings. What it does is unlock the possibility of a richer search-result display, which — exactly like `seo1-2`'s own meta description — can meaningfully improve click-through rate at whatever rank the page already earns. Treat it as a click-through lever, not a ranking shortcut.

A Real Gotcha: Marked-Up Data Must Match Visible Content

Structured data claiming a 5-star rating on a page that shows no visible rating anywhere is a real, documented policy violation, not a clever trick — search engines actively check for this mismatch and can remove rich-result eligibility (or apply a manual penalty) for structured data that misrepresents what's actually on the page.

Hands-On Exercises

EXERCISE 1

Explain the practical difference between the inline microdata `html_lesson_09_2` taught and this chapter's own JSON-LD approach, and why JSON-LD has become the recommended default.

 [View solution](#)

EXERCISE 2

Explain why structured data is described as a click-through lever rather than a ranking factor, drawing the parallel to `seo1-2`'s own meta description material.

 [View solution](#)

EXERCISE 3

Explain why marking up a 5-star rating that isn't actually visible on the page is a real policy violation rather than a clever shortcut, and what the practical consequence can be.

 [View solution](#)

CHAPTER 4 QUICK REFERENCE

- Delivers in full on `html_lesson_09_2`'s own microdata chapter — the "why a search engine cares" half that chapter didn't cover
- JSON-LD describes the same `schema.org` vocabulary as inline microdata, as a separate script block — now the recommended default
- Rich snippets (star ratings, cook times, event dates) require correct structured data, but aren't guaranteed by it
- Structured data is a click-through lever, not a ranking factor — the same honest framing as `seo1-2`'s own meta description
- Marked-up data must match visible content — mismatches are a real, enforced policy violation
- **Next chapter:** Technical SEO: Crawlability, Indexability & Sitemaps

CHAPTER 5 OF 10

Technical SEO: Crawlability, Indexability & Sitemaps

SEO Fundamentals

Chapter 5 · Technical SEO: Crawlability, Indexability & Sitemaps

This is the course's own central technical chapter — where `seo1-1`'s own crawling and indexing stages get their real, concrete tools.

robots.txt — The Other Side of dsproj1-1's Own Warn-Box

```
# /robots.txt
User-agent: *
Disallow: /admin/
Disallow: /tmp/
Allow: /

Sitemap: https://example.com/sitemap.xml
```

THE MIRROR IMAGE OF A CHAPTER THIS SITE ALREADY HAS

`dsproj1-1` taught what a *scraper* must respect when it encounters a site's own `robots.txt` — checking it before crawling, honoring its disallow rules. This chapter teaches the other half: what a *site owner* actually writes there. Same file, same protocol, opposite side of the exact same interaction.

`robots.txt` is a request, not a lock — well-behaved crawlers (including Googlebot) honor it, but nothing technically prevents a page from being fetched by something that ignores it. It manages crawl behavior for cooperative bots; it is not a security or access-control mechanism.

Sitemaps — Making Pages Findable Without Relying on Links Alone

```
<?xml version="1.0" encoding="UTF-8"?>
<urlset xmlns="http://www.sitemaps.org/schemas/sitemap/0.9">
  <url>
    <loc>https://example.com/japanese/hiragana</loc>
    <lastmod>2026-07-17</lastmod>
```

```
</url>  
</urlset>
```

WHY A SITEMAP MATTERS EVEN ON A WELL-LINKED SITE

`seo1-1`'s own crawling stage relies primarily on following links. A sitemap is the explicit backup — a direct, machine-readable inventory of every URL that exists, submitted straight to a search engine, regardless of how well internal linking happens to be structured. Pages that are genuinely orphaned from a site's own navigation are exactly the case a sitemap is built to rescue.

Canonical Tags — Telling a Search Engine Which Version Is Real

```
<link rel="canonical" href="https://example.com/japanese/hiragana">
```

The same content is often reachable at more than one URL — with and without a trailing slash, with a tracking parameter, over HTTP versus HTTPS. Without a canonical tag, a search engine has to guess which version is the "real" one to index and rank, and duplicate versions can dilute ranking signals that should have been concentrated on a single URL. The canonical tag states the answer directly.

noindex — Removing a Page From the Index on Purpose

```
<meta name="robots" content="noindex">
```

TOOL	WHAT IT ACTUALLY DOES
<code>robots.txt Disallow</code>	Asks crawlers not to fetch the page at all
<code>noindex meta tag</code>	Allows crawling, but asks the page not to be added to the index

A genuinely common mistake: blocking a page in `robots.txt` and also adding a `noindex` tag. Since `robots.txt` already prevents the page from being crawled, the crawler never sees the `noindex` instruction at all — the two tools solve different problems and using them together on the same page is usually redundant at best, contradictory at worst.

Mobile-First Indexing

Search engines now primarily use a site's *mobile* version to crawl and index content, not the desktop version — meaning a page whose mobile layout hides or omits content the desktop version shows is effectively invisible for that missing content, indexing-wise. This is fundamentally a crawling/indexing concern, not a new

design topic — the actual responsive-design techniques that solve it are `css_beginner_11`'s own material in full; nothing here needs re-teaching, only applying with this specific consequence in mind.

Hands-On Exercises

EXERCISE 1

Explain why robots.txt is described as "a request, not a lock," and what this implies about using it as a security measure.

 [View solution](#)

EXERCISE 2

Explain the specific difference between robots.txt Disallow and a noindex meta tag, and why combining both on the same page is usually a mistake rather than "extra safety."

 [View solution](#)

EXERCISE 3

Explain why a sitemap still matters on a site with good internal linking, and describe the specific kind of page a sitemap is built to rescue.

 [View solution](#)

CHAPTER 5 QUICK REFERENCE

- **robots.txt** — the site-owner side of dsproj1-1's own scraper-side warn-box; a request cooperative crawlers honor, not an access-control mechanism
- **Sitemaps** — an explicit URL inventory, the backup for pages links alone might not surface
- **Canonical tags** — resolve duplicate-URL ambiguity so ranking signals concentrate on one real version
- **noindex** — allows crawling, blocks indexing; combining with robots.txt Disallow is usually redundant or contradictory
- **Mobile-first indexing** — a crawling/indexing consequence, solved with `css_beginner_11`'s own already-complete responsive-design techniques
- **Next chapter:** Site Performance as a Ranking Factor

CHAPTER 6 OF 10

Site Performance as a Ranking Factor

SEO Fundamentals

Chapter 6 · Site Performance as a Ranking Factor

`web-vitals1` already taught LCP, INP, and CLS in full technical depth, framed entirely as a user-experience concern. This chapter doesn't re-teach the mechanics — it delivers the other half: these three metrics are also real, Google-confirmed ranking signals, not just a UX nicety with no bearing on visibility.

The Core Web Vitals, Reused Directly

METRIC	WEB-VITALS1'S OWN COVERAGE	THE RANKING ANGLE THIS CHAPTER ADDS
LCP	Largest Contentful Paint — full technical treatment (web-vitals1-2)	Part of the confirmed Page Experience ranking signal set
INP	Interaction to Next Paint — full technical treatment (web-vitals1-4)	Same — replaced FID in that same official signal set
CLS	Cumulative Layout Shift — full technical treatment (web-vitals1-3)	Same

Nothing about how these are measured or improved changes here — `web-vitals1-2` through `web-vitals1-4` remain the authoritative source for the technical work itself. What this chapter adds is the missing context for *why* that work matters beyond visitor satisfaction.

An Honest Note on Weight

A REAL SIGNAL — BUT NOT THE DOMINANT ONE

Google has been explicit that Core Web Vitals are one signal among many, and a comparatively minor one next to genuine content relevance. A page with excellent Core Web Vitals scores and thin, unhelpful content will not outrank a page with strong content and mediocre performance. Performance is a real, worthwhile lever — it is not a substitute for everything `seo1-3` already covered about content quality.

Where Performance and Crawling Actually Intersect

A SECOND, LESS-DISCUSSED REASON SPEED MATTERS

A search engine allocates a finite amount of time and resource to crawling any given site — its own "crawl budget." A slow-responding server means fewer pages get crawled within that budget on any given pass, which connects performance directly back to `seo1-1`'s own crawling stage, not just to ranking once a page is already indexed.

Hands-On Exercises

EXERCISE 1

Explain what this chapter adds on top of web-vitals1's own coverage of LCP/INP/CLS, and what it deliberately does not re-teach.

[View solution](#)**EXERCISE 2**

Explain why a page with excellent Core Web Vitals scores but thin content would not outrank a page with strong content and mediocre performance — what does this reveal about how ranking signals are weighted relative to each other?

[View solution](#)**EXERCISE 3**

Explain the crawl-budget connection between site performance and `seo1-1`'s own crawling stage — how is this a genuinely separate mechanism from the ranking-signal effect?

[View solution](#)**CHAPTER 6 QUICK REFERENCE**

- A direct, undiluted extension of web-vitals1's own LCP/INP/CLS material — no re-teaching, only the missing ranking context
- Core Web Vitals are a real, confirmed ranking signal — but a comparatively minor one next to content relevance (`seo1-3`)
- A second, separate mechanism: slow response times reduce how much of a site gets crawled within its own crawl budget
- Performance is a genuine lever, not a substitute for the content-quality work this course already covered

- **Next chapter:** URL Structure, Internal Linking & Site Architecture

CHAPTER 7 OF 10

URL Structure, Internal Linking & Site Architecture

SEO Fundamentals

Chapter 7 · URL Structure, Internal Linking & Site Architecture

Everything so far has treated a page mostly in isolation. This chapter zooms out to how pages relate to each other — a real, distinct layer of SEO that no earlier chapter has touched.

Clean URLs

WEAK	STRONG	WHY
<code>/page?id=4471&cat=3</code>	<code>/japanese/hiragana</code>	Readable to both a person and a crawler; no dependence on an internal ID scheme
<code>/Japanese_Hiragana_Guide!!.html</code>	<code>/japanese/hiragana</code>	No unnecessary punctuation, casing, or file extensions

A URL is itself a small, persistent piece of context — shown in search results, shared directly, referenced in links. It doesn't need keywords stuffed into it (`seo1-3` 's own caution against stuffing applies here too), just a clean, stable, human-readable path.

Internal Linking & Link Equity

Every internal link is a small vote of relevance from one page to another — search engines use this internal linking pattern to judge which pages a site itself treats as important. A page linked from dozens of other pages across a site sends a different signal than a page linked from nowhere at all.

A PAGE WITH NO INTERNAL LINKS IS NEARLY INVISIBLE

A page that's never linked to from anywhere else on the same site gets no internal link equity at all, regardless of how good its own content is — and, per `seo1-5` 's own sitemap material, is exactly the kind of page that depends entirely on a sitemap (or an external backlink) to be discovered in the first place. Strong internal linking and a complete sitemap are solving genuinely different parts of the same underlying problem.

Breadcrumbs

A visible trail — [Home > Japanese > Hiragana](#) — that does double duty: it helps a visitor understand where they are and navigate up a level, and it gives a crawler an explicit, low-effort signal of the page's own place in the site's overall hierarchy. Breadcrumbs are also a common target for the structured-data markup [seo1-4](#) already covered, unlocking a breadcrumb trail shown directly in the search result itself.

Site Architecture & Content Silos

A "silo" groups closely related content under a shared structural path — every hiragana-related page living under [/japanese/hiragana/](#), every kanji-related page under [/japanese/kanji/](#) — rather than scattering topically related pages across unrelated, disconnected paths. This isn't just tidiness: a search engine reading a site's own URL and linking structure can use consistent silos as a signal of genuine topical depth on a subject, one instance of [seo1-3](#)'s own E-E-A-T material made structural rather than purely textual.

Hands-On Exercises

EXERCISE 1

Explain what makes a URL "clean" beyond just being readable, and why keyword-stuffing a URL is subject to the same caution [seo1-3](#) raised for content.

 [View solution](#)

EXERCISE 2

Explain why strong internal linking and a complete sitemap are described as solving "genuinely different parts of the same underlying problem" rather than being redundant with each other.

 [View solution](#)

EXERCISE 3

Explain how content silos connect back to [seo1-3](#)'s own E-E-A-T material — specifically, how can a purely structural choice (URL/linking organization) serve as a signal usually associated with content quality?

 [View solution](#)

CHAPTER 7 QUICK REFERENCE

- **Clean URLs** — readable, stable, no unnecessary parameters/punctuation; the same anti-stuffing caution as [seo1-3](#)

- **Internal linking** — each link is a small relevance signal; a page with none is nearly invisible even with great content
- **Breadcrumbs** — serve both navigation and crawling, and are a real structured-data (seo1-4) target
- **Content silos** — consistent topical grouping in URLs/links as a structural E-E-A-T signal (seo1-3)
- **Next chapter:** Off-Page SEO: Backlinks & Authority

CHAPTER 8 OF 10

Off-Page SEO: Backlinks & Authority

SEO Fundamentals

Chapter 8 · Off-Page SEO: Backlinks & Authority

`seo1-7` covered internal links — votes a site casts for itself. This chapter covers the external kind: links from other sites, and why they carry real weight precisely because they're harder to fake than anything on a page itself.

Why a Backlink Is a Genuinely Different Kind of Signal

Every earlier chapter covered signals a site controls directly — its own title tags, its own content, its own structure. A backlink is different: it's another, independent site choosing to link to this one. That independence is exactly what makes it valuable as a trust signal — it's much harder to fabricate at scale than anything living entirely within one site's own control.

Not All Links Carry Equal Weight

FACTOR	WHY IT MATTERS
Linking site's own authority	A link from a well-established, trusted site carries more weight than one from an obscure or low-quality one
Topical relevance	A link from a genuinely related site is a stronger signal than one from something unrelated
<code>rel="nofollow"</code> / <code>rel="sponsored"</code>	Explicitly tells a search engine not to pass ranking credit through this specific link

```
<a href="https://example.com" rel="sponsored">a paid placement</a>
```

`rel="sponsored"` and `rel="nofollow"` exist specifically so a site can link out honestly — to a paid partner, to user-submitted content it can't fully vouch for — without that link being read as an unqualified endorsement.

Real Link-Building Value

Genuine backlinks accumulate as a natural consequence of content worth referencing — original research, a uniquely useful resource, something no one else has built quite the same way. Earning links this way takes real time and real quality, which is exactly why it's a meaningful signal rather than a hollow one.

The Other Side: Real Manipulation Risk

LINK SCHEMES ARE A REAL, ACTIVELY ENFORCED VIOLATION

Buying links in bulk, participating in link-exchange networks built purely to inflate link counts, or generating links through automated/low-quality directories are all patterns search engines actively detect and penalize — sometimes with a manual action that suppresses a site's own rankings broadly, not just the specific manipulated signal. Unlike the honest trade-offs covered in earlier chapters, this isn't a "works less well than hoped" risk — it's a genuine, documented way to actively damage a site that otherwise did everything right.

AN HONEST BOTTOM LINE

A handful of genuinely relevant, earned backlinks from real, related sites is worth more than hundreds of low-quality or purchased ones — and carries none of the penalty risk. This chapter isn't a "get more links" pitch; it's a "understand what makes a link valuable, and why the shortcuts are a real trap" one.

Hands-On Exercises

EXERCISE 1

Explain why a backlink is described as a genuinely different kind of signal than anything covered in earlier chapters, specifically in terms of who controls it.

[!\[\]\(51e24a718c479f61046a6569471331fa_img.jpg\) View solution](#)

EXERCISE 2

Explain what `rel="nofollow"`/`rel="sponsored"` actually communicate to a search engine, and why a site would want to use them rather than just not linking out at all.

[!\[\]\(5fbc60decb3ef3e18ac6af3c11358089_img.jpg\) View solution](#)

EXERCISE 3

Explain why this chapter treats link schemes as a fundamentally different kind of risk than the honest trade-offs covered in earlier chapters (like seo1-6's performance-vs-content weighting).

 [View solution](#)

CHAPTER 8 QUICK REFERENCE

- A backlink is valuable specifically because it's independent — much harder to fabricate than anything a site controls directly
- Not all links are equal: linking-site authority, topical relevance, and nofollow/sponsored attributes all shape a link's own real weight
- Genuine link-building comes from content worth referencing — a real, time-consuming, quality-driven process
- Link schemes are an actively detected, actively penalized violation — a real trap, not just a weaker version of a good strategy
- **Next chapter:** Measuring & Monitoring SEO

CHAPTER 9 OF 10

Measuring & Monitoring SEO

SEO Fundamentals

Chapter 9 · Measuring & Monitoring SEO

Every technique in this course so far has been a change to make. This chapter is about finding out whether any of it actually worked — and being honest about how long that answer takes to arrive.

Google Search Console — The Primary Tool

REPORT	WHAT IT ACTUALLY SHOWS	WHICH EARLIER CHAPTER IT VERIFIES
Coverage	Which pages are indexed, and why others were excluded	seo1-5's own crawling/indexing material, made visible
Performance	Real queries a site appears for, average position, click-through rate	seo1-2's own title/description work, measured directly
Core Web Vitals	Real-user LCP/INP/CLS data, aggregated across actual visitors	seo1-6's own performance material, using real field data
Links	A sample of backlinks Google has discovered pointing to the site	seo1-8's own off-page material, partially observable

Search Console is free, and it's the only tool giving a direct view into how Google itself sees a specific site — not a third-party estimate, the actual data.

The Coverage Report — Closing the Loop on seo1-5

seo1-5 covered robots.txt, sitemaps, canonical tags, and noindex as tools that shape what gets indexed. The Coverage report is where the actual, real-world result of those choices becomes visible — which pages Google has indexed, which it found but excluded, and, critically, the stated reason for each exclusion.

An Honest Framing, Stated Directly

SEO IS SLOW

A meaningful ranking change from real on-page or content work typically takes weeks to months to show up, not days — search engines re-crawl and re-evaluate pages on their own schedule, and ranking systems themselves incorporate new signals gradually rather than instantly. Any claim of overnight SEO results should be treated with real skepticism. This isn't a flaw in the techniques this course has covered; it's an honest, structural property of how the whole pipeline actually operates.

WHY THIS MAKES MEASUREMENT DISCIPLINE MATTER MORE, NOT LESS

Because results arrive slowly, it's tempting to change several things at once and lose track of what actually caused an improvement. Tracking one change at a time, and giving each one enough time to genuinely show up in Search Console's own data before judging it, is the only reliable way to learn which techniques from this course actually mattered for a specific site.

Hands-On Exercises

EXERCISE 1

Explain how the Coverage report closes the loop on seo1-5's own robots.txt/sitemap/noindex material — what does it show that those tools alone couldn't confirm on their own?

 [View solution](#)

EXERCISE 2

Explain why "SEO is slow" is described as a structural property of the pipeline rather than a limitation of the specific techniques this course teaches.

 [View solution](#)

EXERCISE 3

Explain why slow feedback makes disciplined, one-change-at-a-time measurement more important, not less — what specifically goes wrong if several changes are made at once?

 [View solution](#)

CHAPTER 9 QUICK REFERENCE

- Google Search Console is the primary tool — a direct view of how Google itself sees a specific site, not a third-party estimate

- Coverage report verifies seo1-5's own crawling/indexing choices; Performance report verifies seo1-2's own title/description work; Core Web Vitals verifies seo1-6 with real field data
- **SEO is slow** — weeks to months for real results, a structural property of the pipeline, not a technique flaw
- Slow feedback makes one-change-at-a-time discipline more important, not less — otherwise cause and effect become impossible to separate
- **Next chapter:** Capstone — A Real Audit of osztromok.com

CHAPTER 10 OF 10

Capstone: A Real Audit of osztromok.com

SEO Fundamentals

Chapter 10 · Capstone: A Real Audit of osztromok.com

Every prior chapter used illustrative examples. This one doesn't — it's a genuine audit of Philip's own live site (`debserver\website\`), based on a real source-level review conducted 2026-07-17. Every finding below is real; every fix is a genuine recommendation, not a hypothetical exercise.

Finding 1 — Crawling Is Effectively Broken for Most of the Site (seo1-5)

THE SINGLE BIGGEST FINDING

Roughly 300 static Japanese kanji/hiragana/katakana pages exist on disk but are linked from *nowhere* in the live site's own navigation — and their own internal "back to grid" links point to a `kanji-tiles` / `hiragana-tiles` page that doesn't exist anywhere in the codebase. Per `seo1-1`'s own pipeline and `seo1-7`'s own internal-linking material, these pages are structurally orphaned: no crawl path finds them, and the one link each page does have is broken.

Fix: either build the missing tile-grid pages and link them from the main site nav, or add these ~300 URLs directly to a sitemap (`seo1-5`) as the explicit backup `seo1-7`'s own exercise already named for exactly this scenario.

Finding 2 — No robots.txt, No Sitemap (seo1-5)

Neither file exists anywhere on the site. Per `seo1-5`, a sitemap is the direct fix for Finding 1; a `robots.txt` (even a minimal one excluding `/admin/`) is the site-owner-side counterpart to `dsproj1-1`'s own scraper-side material.

Finding 3 — No Canonical Tags, No HTTPS-Enforcing Redirect (seo1-5)

`config.php` still hardcodes `SITE_BASE_URL = 'http://localhost'`, and the one existing `$canonicalBase` variable in the anime sub-app is defined but never rendered into an actual `<link rel="canonical">` tag anywhere in the codebase.

Finding 4 — No Meta Descriptions Anywhere on the Main Blog (seo1-2)

Neither `template.php` nor any of the static kanji/kana pages define `<meta name="description">`. The separate Anime Vault sub-app *does* have dynamic meta descriptions — direct, real evidence this is a gap in execution, not a gap in knowledge.

Finding 5 — `<html lang="en">` Hardcoded Everywhere, Including on Japanese-Language Pages

Every page, including ones whose entire content is Japanese vocabulary, declares `lang="en"`. For a multi-language learning site, this is a real, fixable mismatch between what a page says and what it claims to be.

Findings Summary Table

FINDING	SEVERITY	CHAPTER
~300 orphaned kanji/kana pages, broken internal links	High	seo1-1, seo1-5, seo1-7
No robots.txt / no sitemap	High	seo1-5
No canonical tags / no HTTPS redirect	Medium	seo1-5
No meta descriptions on the main blog	Medium	seo1-2
Hardcoded lang="en" on non-English pages	Medium	seo1-2
No structured data anywhere	Medium	seo1-4
No caching/compression headers on the main site's own .htaccess	Medium	seo1-6

A REAL STRENGTH WORTH CREDITING

Heading structure is consistently well-formed across every sampled page (one clean `<h1>`, logical nesting), the DB-driven blog uses genuinely clean URLs with no query strings, and internal navigation for the main blog (nav/sidebar/breadcrumbs/prev-next) is solid — `seo1-2`'s and `seo1-7`'s own material is already largely satisfied there.

Chapter Attribution Table

CHAPTER	WHAT IT DIAGNOSED IN THIS AUDIT
seo1-1	The crawl/index/rank framing used to explain why orphaned pages are invisible

CHAPTER	WHAT IT DIAGNOSED IN THIS AUDIT
seo1-2	Missing meta descriptions, the lang="en" mismatch
seo1-4	Absent structured data across the whole codebase
seo1-5	Missing robots.txt/sitemap/canonical tags, the orphaned-page fix
seo1-6	Missing caching/compression on the main site's own .htaccess
seo1-7	The broken internal-linking pattern behind the orphaned pages
seo1-9	Search Console's Coverage report as the next real step to confirm these findings live

Hands-On Exercises

EXERCISE 1

Of the two fixes proposed for Finding 1 (build the missing tile-grid pages, or add the URLs to a sitemap), explain what each one actually solves and what it doesn't — are they mutually exclusive, or complementary?

 [View solution](#)

EXERCISE 2

Explain why the Anime Vault sub-app having meta descriptions while the main blog doesn't is described as "evidence this is a gap in execution, not a gap in knowledge" — why does that distinction matter for how the fix gets prioritized?

 [View solution](#)

EXERCISE 3

Using seo1-9's own material, describe the specific, real next step that would confirm whether fixing Finding 1 actually worked, and how long it would be reasonable to wait before checking.

 [View solution](#)

Scope Note — What This Capstone Deliberately Doesn't Cover

HONESTLY OUT OF SCOPE

- No paid-search/SEM audit — [seo1-1](#) scoped this course to organic search only.
- No local-SEO/Google Business Profile review — the site has no physical-location component to evaluate.

- No international/multi-language SEO strategy (hreflang, regional targeting) — a genuinely separate topic from the lang-attribute fix named in Finding 5.
- No live Search Console data — this audit is based on source-code review; a real Coverage/Performance report requires the site actually being verified in Search Console first, which is itself the recommended next step.

CHAPTER 10 QUICK REFERENCE — COURSE SUMMARY

- A genuine audit of osztromok.com, not a hypothetical exercise — real findings from a real 2026-07-17 source review
- **Biggest finding:** ~300 orphaned kanji/kana pages with broken internal links, invisible to crawling
- Also found: no robots.txt/sitemap/canonical tags, no meta descriptions on the main blog, hardcoded lang="en" on Japanese pages, no structured data
- Real strengths confirmed: clean URLs, solid heading structure, good main-blog internal navigation
- This completes the SEO Fundamentals course (10 chapters) — closing with a real case study rather than a hypothetical one