Checklist della sitemap

I controlli per scrivere, inviare e verificare una sitemap XML, in ordine di lavoro. Spunta una voce solo quando l'hai verificata sul file vero.

Fase 1: decidere e creare

- [ ] Ho deciso se il sito ha bisogno di una sitemap: grande, nuovo con pochi link, o con molti video, immagini o notizie.
- [ ] Ho scelto il formato: XML per quasi tutti i casi, testo solo per siti piccoli.
- [ ] Ho verificato se il CMS genera già una sitemap e a quale indirizzo.
- [ ] Ho tenuto una sola sitemap attiva, senza sovrapposizioni tra CMS e plugin.

Fase 2: il contenuto del file

- [ ] Il file si apre con risposta 200, è in UTF-8 e non ha testo prima di <?xml.
- [ ] Gli URL sono assoluti, completi, con https e l'host della versione principale.
- [ ] I caratteri speciali sono con l'escape XML (& diventa &).
- [ ] Ci sono solo URL canonici, indicizzabili e con risposta 200.
- [ ] Non ci sono redirect, errori 4xx o 5xx, pagine in noindex o bloccate dal robots.txt.
- [ ] Non ci sono URL con parametri di tracciamento, ordinamento o sessione.
- [ ] Non ci sono pagine di servizio: carrello, account, ricerca interna, ringraziamento.
- [ ] lastmod cambia solo quando cambia davvero la pagina; altrimenti l'ho tolto.
- [ ] Ho tolto priority e changefreq, che Google ignora.

Fase 3: limiti e struttura

- [ ] Ogni file ha al massimo 50.000 URL e 50 MB non compressi.
- [ ] Se serve, ho una sitemap indice, con le sitemap nella stessa cartella o in una più specifica.
- [ ] Ho diviso le sitemap per tipo di contenuto, per leggere i problemi nel rapporto Indicizzazione delle pagine.
- [ ] Per immagini, video, notizie o lingue uso solo i tag che la documentazione di Google prevede oggi.

Fase 4: invio

- [ ] La riga Sitemap: con l'URL completo è nel robots.txt.
- [ ] Ho inviato la sitemap dal rapporto Sitemap di Search Console, da proprietario.

- [ ] Non uso il ping: Google ne ha annunciato la fine e la documentazione di invio non lo menziona.
- [ ] Ho controllato con Controllo URL che Googlebot recuperi il file.

Fase 5: controllo

- [ ] Il rapporto Sitemap mostra lo stato Riuscita, senza errori.
- [ ] Ho filtrato il rapporto Indicizzazione delle pagine per sitemap e so quante pagine sono indicizzate.
- [ ] Ho scansionato gli URL della sitemap con un crawler e ogni URL dà 200 ed è canonico.
- [ ] Rifaccio il controllo dopo ogni migrazione, cambio di plugin o ristrutturazione.
- [ ] Se ho una sitemap HTML, i link sono veri, annidati e con gli stessi nomi del menu.