

AI Index Checklist

30-min diagnostika: kako pripraviti spletno stran, da jo AI lahko prebere, razume in jo pri pravih vprašanjih tudi citira.

Kaj je cilj? V 30 minutah najdeš 3 stvari: (1) ali te boti sploh lahko berejo (crawl), (2) ali so ključne strani pravilno indeksirane (index), (3) ali so strani dovolj jasne in strojno razumljive, da jih AI lahko citira.

Kaj potrebuješ:

- Dostop do robots.txt (tvoja-domena.si/robots.txt)
- Dostop do sitemap (tvoja-domena.si/sitemap.xml ali sitemap_index.xml)
- Osnovni dostop do analitike (GA4/Matomo) in po možnosti Search Console / Bing Webmaster Tools
- 2-3 najpomembnejše URL-je (top kategorija, top produkt, top vodič)

Kako uporabljati checklist:

- 1) Začni pri Modulu A in pojdi po vrsti.
- 2) Pri vsaki točki označi [] kot OK ali problem.
- 3) Če je problem, takoj zabeleži akcijo v zadnjem stolpcu.
- 4) Na koncu uporabi "Prioriteta popravil" in izberi 3 največje blokade.

Opomba: Ta checklist ne daje garancij za omembe. Daje pa sistem, da postaneš "viden in razumljiv vir" v ekosistemu iskalnikov in AI iskanja.

Časovni razpored (30 min):

A. Nastavitev (2 min) | B. Crawl dostop (6 min) | C. Index čistost (8 min) | D. Podatki (schema/feed) (7 min) | E. Citabilnost vsebine (5 min) | F. Merjenje (2 min)

Modul A: Priprava in fokus (2 min)

	Kontrolna točka	Kje preveriš	Če ne drži, naredi to
[]	Imaš izbranih 3-5 "money" strani (kategorija, produkt, vodič/primerjava).	Zapiši URL-je na vrh te strani ali v opombe.	Izberi strani z največ prometa/prodaje ali z največjim potencialom (top izdelki, top kategorije).
[]	Za vsako stran imaš jasno "vprašanje z namenom" (kaj uporabnik vpraša).	Poglej Search Console poizvedbe ali uporabi interno logiko (npr. "najboljši X za Y", "X vs Y").	Naredi seznam 10 vprašanj; vsako poveži z eno ciljno stranjo.

Hitri rezultat:

Če ne veš, katera stran je "cilj" za določeno vprašanje, bo optimizacija razpršena. Najprej določi "vprašanje -> ciljna stran".

Modul B: Crawl dostop (6 min)

	Kontrolna točka	Kje preveriš	Če ne drži, naredi to
[]	robots.txt obstaja in se naloži (HTTP 200).	Odpri tvoja-domena.si/robots.txt.	Če robots.txt manjka ali je 4xx/5xx, ga ustvari/popravi. Dodaj tudi Sitemap: vrstico.
[]	Nimaš globalne blokade, ki ustavi vse bote (npr. User-agent: * + Disallow: /).	Preveri robots.txt za "Disallow: /" v skupini User-agent: *.	Odstrani globalno blokado ali naredi izjemo za bote, ki jih želiš.
[]	Ne blokiraš ključnih sekcij: produkti, kategorije, vodiči.	Preveri Disallow pravila (npr. /product/, /category/, /blog/).	Odblokiraj ključne poti; blokiraj le košarico/checkout, interne iskalnike in parametre.
[]	WAF/CDN ne blokira crawlerjev (ni 403/429 za legitimne bote).	V strežniških logih ali WAF dashboardu preveri zadnje obiske botov.	Dodaj izjeme/rate limit prilagoditve za legitimne crawlerje; odpravi 403/429.
[]	Ključne strani se naložijo brez praznega HTML (kritični tekst ni samo v JS).	Odpri izvorno kodo strani in preveri, ali je glavni tekst v HTML.	Dodaj SSR ali izpis ključnih delov v HTML; ne skrivaj vsebine za klientskim renderjem.

Rdeče zastavice:

- 403/429 za crawlerje (WAF)
- Globalni Disallow: /
- Blokada /product/ ali /category/
- Strani z veliko vsebine, ki je vidna samo po JS

Modul C: Index čistost in tehnični signali (8 min)

	Kontrolna točka	Kje preveriš	Če ne drži, naredi to
[]	Sitemap obstaja in je v robots.txt pravilno naveden.	Odpri sitemap.xml ali sitemap_index.xml in preveri, da je dostopen.	Dodaj/posodobi Sitemap: URL v robots.txt. Preveri, da so v njem ključni URL-ji.
[]	Ključne strani nimajo noindex.	Preveri meta robots (view-source) ali SEO plugin nastavitve.	Odstrani noindex z osnovnih strani (kategorije, produkti, vodiči).
[]	Canonical na ključnih straneh kaže na "glavni" URL (brez parametrov).	Poglej <link rel="canonical"> v izvorni kodi.	Popravi canonical logiko (posebej pri filtrih/sortiranju in variacijah).
[]	Parametri (filter/sort/utm) niso indeksirani in ne ustvarjajo duplikatov.	Preveri, ali imaš veliko URL-jev z ?filter=, ?orderby= ipd. v indexu.	Blokiraj nepotrebne parametre v robots, uvedi canonical na osnovni URL, po potrebi noindex param. strani.
[]	Kategorije imajo dovolj interne povezljivosti (meni, breadcrumbs, povezave iz vodičev).	Klikaj po strani: ali do kategorij prideš v 2-3 klike?	Dodaj interne linke: vodič -> kategorija -> top produkti; izboljšaj navigacijo.

Hitri test "ali je stran v igri":

Če je stran crawlable, brez noindex in z normalnim canonicalom, ima osnovno možnost, da se pojavi kot vir. Če pade na enem od teh treh, jo najprej popravi, preden pišeš novo vsebino.

Modul D: Strojno razumljivi podatki (7 min)

	Kontrolna točka	Kje preveriš	Če ne drži, naredi to
[]	Produktne strani imajo Product + Offer structured data (cena, valuta, zaloga).	Preveri z Rich Results Test ali Schema validatorjem.	Dodaj/uredi schema (JSON-LD). Poskrbi, da se ujema z vidno ceno/zalogo.
[]	Atributi (brand, SKU, GTIN/MPN) so tam, kjer je smiselno.	Preveri structured data polja.	Dodaj brand + identifikatorje; izboljša ujemanje produktov v iskalnikih.
[]	Merchant feed je posodobljen (če si e-commerce) in nima kritičnih napak.	Google Merchant Center / Microsoft Merchant Center status.	Popravi disapprovals: manjkajoči atributi, neujemanje cene/zaloge, slabe slike.
[]	Strukturirani podatki se ujemajo z vidno vsebino (brez "skritih" cen/ocen).	Primerjaj schema vrednosti s tem, kar vidi uporabnik.	Poenoti izvor podatkov (en vir za ceno/zalogo) in odstrani nasprotja.

Zakaj je to pomembno:

AI in iskalniki hitreje zaupajo, ko imajo nedvoumne podatke: kaj je izdelek, koliko stane, ali je na zalogi, komu je namenjen. To zmanjša ugibanje.

Modul E: Citabilnost in uporabnost vsebine (5 min)

	Kontrolna točka	Kje preveriš	Če ne drži, naredi to
[]	Na vrhu strani je "answer-first" blok (2-3 stavki direktnega odgovora).	Odpri stran in preveri prvih 10 sekund branja.	Dodaj kratek odgovor: komu, kaj izbrati, zakaj. Potem šele razlaga.
[]	Stran ima kriterije izbire ali primerjavo (ne samo opis).	Poglej, ali uporabnik dobi pomoč pri odločitvi.	Dodaj 3-6 bulletov kriterijev + mini primerjavo ("najboljše za...").
[]	FAQ odgovori na realna vprašanja (kompatibilnost, uporaba, dostava, garancija).	Preveri support email/chat: kaj ljudje sprašujejo.	Dodaj 5-10 FAQ in jih poveži s produkti/kategorijami.
[]	Stran ima jasne specifikacije in dokazila zaupanja.	Poglej: ocene, kontakt, dostava/vračila, podjetje.	Dodaj trust blok: dostava, vračila, garancija, kontakt, reference/ocene.

Mini pravilo:

AI pogosto citira najbolj direkten in strukturiran odgovor. Če mora "iskati" po strani, si manj verjetna izbira.

Modul F: Merjenje in ponovljivost (2 min)

	Kontrolna točka	Kje preveriš	Če ne drži, naredi to
[]	V analitiki lahko ločiš AI referral promet (kjer je mogoče).	GA4: Acquisition -> Traffic acquisition; filtriraj po source/medium.	Dodaj report za AI referral (npr. chatgpt.com, perplexity, bing/copilot).
[]	Imaš tedensko rutino (10 min) za pregled: top strani, index status, konverzije.	Naredi en list: URL, index OK, schema OK, promet, konverzija.	Vsak teden popravi 3 stvari: 1 tehnično, 1 podatkovno, 1 vsebinsko.

Prioriteta popravil (izberi 3):

- 1) Blokade crawlinga (robots/WAF/403)
- 2) noindex ali slab canonical na ključnih straneh
- 3) Parametri/duplikati, ki razpršijo signal
- 4) Manjka Product+Offer schema / feed napake
- 5) Manjka answer-first + kriteriji + FAQ
- 6) Manjka trust (dostava/vračila/kontakt/ocene)

Reference (za nadaljnje branje):

- OpenAI: Overview of OpenAI crawlers (OAI-SearchBot, GPTBot): <https://platform.openai.com/docs/bots>
- OpenAI: Publishers and developers FAQ (robots.txt, GPTBot): <https://help.openai.com/en/articles/12627856-publishers-and-developers-faq>
- Google Search Central: AI features and your website: <https://developers.google.com/search/docs/appearance/ai-features>
- IETF RFC 9309: Robots Exclusion Protocol: <https://datatracker.ietf.org/doc/html/rfc9309>
- Perplexity: Crawlers and robots.txt: <https://docs.perplexity.ai/guides/bots>
- Perplexity: How Perplexity follows robots.txt: <https://www.perplexity.ai/help-center/en/articles/10354969-how-does-perplexity-follow-robots-txt>
- Anthropic: Crawling & opt-out (Claude-SearchBot): <https://privacy.claude.com/en/articles/8896518-does-anthropic-crawl-data-from-the-web-and-how-can-site-owners-block-the-crawler>
- IndexNow: Protocol documentation: <https://www.indexnow.org/documentation>