

The Ten-Minute Door Check

Four quick checks that show whether search engines and AI assistants can reach your website and read it.

Before you write new content, check that the door is open. This sheet will not catch everything, but it spots the problems that come up most often, and most owners have never looked.

Best completed on a computer. Two of these checks involve looking at your website's code, which is difficult on a phone.

HOW TO USE THIS SHEET

Checks 1, 2, and 3 you can do yourself in about ten minutes. Check 4 is one message you send, so that answer comes back later. Tick a result for each, and screenshot what you find today before anything is changed.

Website: _____ Date: _____

1 Read your robots.txt

DO THIS

In your browser address bar, type your website address followed by `/robots.txt`. Example:
`yourbusiness.com/robots.txt`

WHAT THIS FILE IS

It tells crawlers which parts of your site they may visit. Rules come in groups, each starting with a crawler name, and apply only to that crawler. **User-agent: *** means all crawlers.

YOU ARE FINE IF

- You get a "not found" page. No restrictions were found in robots.txt. This is common. Other settings can still block access, which is check 4.
- The only **Disallow** lines point at back office pages such as `/admin` or `/checkout`.

FLAG IT IF

You see **Disallow: /** (a slash on its own) under **User-agent: *** or under any of these names:

- **Googlebot, Bingbot.** Regular search engines.
- **OAI-SearchBot, Claude-SearchBot, PerplexityBot.** AI search. OpenAI names OAI-SearchBot as the setting that controls whether you can appear in ChatGPT's search answers.
- **ChatGPT-User, Claude-User, Perplexity-User.** These fetch a page when a person asks an assistant about you. Leave them open. robots.txt does not reliably control them, so check 4 matters more here.

A SEPARATE DECISION

GPTBot and ClaudeBot control AI training. Google-Extended covers training and some use of your content in Gemini's answers. Blocking these does not remove you from Google Search.

RESULT ☐ Clear ☐ Needs fixing ☐ Not sure

2 Read your own page source

DO THIS

Open your homepage. Right click on the page and choose **View Page Source**. Search that code twice, with Ctrl-F on Windows or Cmd-F on a Mac:

- A full sentence you can see on your homepage.
- The word **noindex**.

YOU ARE FINE IF

- Your sentence is in the code, and **noindex** is not there.

FLAG IT IF

- **Your sentence is missing.** Your text may be built by JavaScript after the page loads, and not every crawler runs JavaScript.
- **noindex appears** in a line starting with **meta name="robots"**. That asks search engines to leave the page out.

BEFORE YOU REMOVE A NOINDEX

Some pages are kept out of search on purpose: thank you pages, checkout pages, staging copies, client portals. Only remove noindex from pages you want found. On a homepage or main service page it is almost always there by accident.

Note that noindex only keeps a page out of search results. It does not make the page private, and anyone with the link can still open it.

RESULT

☐

Clear

☐

Needs fixing

☐

Not sure

3 Check your structured data

DO THIS

Paste your homepage address into the Schema.org validator: validator.schema.org

WHAT STRUCTURED DATA IS

A hidden summary of your business written for machines rather than visitors: name, phone, address or service area, hours.

YOU ARE FINE IF

- Something is found, and those facts match what a visitor sees.

FLAG IT IF

- Any fact is out of date or contradicts the page. For example, one phone number in the summary and a different one on your contact page.

IF NOTHING IS FOUND

Nothing is broken. It means you do not have a summary written for machines yet, and it is the easiest item here to add. Use this validator rather than Google's Rich Results Test, which only checks the types Google can display as rich results and can report nothing on a site that does have structured data.

RESULT

☐

Clear

☐

Needs fixing

☐

Not sure

4 Send your web person this message

DO THIS

Copy the message below and send it. Your web person can confirm these three points and fix them if needed.

COPY AND SEND

Three quick questions about our website:

1. Is anything in our security, firewall, hosting, or bot protection settings blocking AI crawlers or AI assistants?
2. Can search crawlers read our page text without running JavaScript?
3. Is there an X-Robots-Tag noindex header on any page we want appearing in search?

Short answers are fine. Thank you.

WHAT A GOOD ANSWER LOOKS LIKE

Question	Good answer
1. Security settings blocking AI crawlers	No
2. Text readable without JavaScript	Yes
3. noindex header on a page you want found	No

YOU ARE FINE IF

- All three answers match the table above, and they can tell you where they looked.

FLAG IT IF

- Any answer is the opposite of the table, or comes back “not sure.” Not sure means it still needs checking.

WHY THIS CANNOT BE SKIPPED

Security and hosting services often block AI crawlers by default, without anyone choosing it. A site can pass check 1 and still be blocked here.

RESULT ☐ Clear ☐ Needs fixing ☐ Not sure

If anything came back “needs fixing”

Work through them in this order. The first two are usually quick. Making your text readable without JavaScript can be a larger piece of work, so ask what is involved before agreeing to it.

1. Remove *noindex* from any page you want people to find. Leave it on pages intentionally excluded from search.
2. Unblock the crawlers, in robots.txt and in your security settings.
3. Make your visible text readable without JavaScript.
4. Correct your structured data, or add it if there was none.

The door being open is step one, not the whole job. Once AI assistants can read your site, the next question is whether they can describe you well enough to name you. That is the NAMED Method, and the whole of Become the Answer: becometheanswerbook.com.