# The five-minute skill check

Before you install an agent skill from someone you don’t know

Agentic. Schule

1 **Read the `SKILL.md`**

The file your agent will read as instructions, every line. Not the README of the marketplace.

2 **Search for addresses**

Does every domain really belong to the vendor? Is the spelling exactly right? Every URL is a place where something different may sit tomorrow.

3 **Look into the scripts**

Not only the Markdown files. And everything the skill wants to install: “prerequisites”, “setup”, “install it globally”.

4 **Origin over popularity**

Who submitted it? Stars belong to the repository, not to a single skill in it.

## Three rules that matter more than the check

- **A skill that fetches instructions from the internet cannot be reviewed once and for all.** What gets executed is whatever sits behind the link right now.
- **Treat untrusted skills as inspiration, never as a dependency.** Let your agent rewrite them in its own words, then read the result.
- **Keep the safety nets on.** Know which mode your agent starts in (Claude Code: auto mode). Sandbox on, no Turbo, `--dangerously-skip-permissions`, bypass permissions or `--yolo` with a new skill, and read every permission prompt. In Claude Code, check `allowed-tools` in the frontmatter.

From the talk “Dangerous AI Skills” by Johannes Hoppe · Slides, prompt and sources: talks.agentic.schule/dangerous-ai-skills/materials/