



How do search engines work?

KS3 Ages 11-14 🕒 4 min read

Type something into Google and within a fraction of a second you get a ranked list of the most relevant pages out of the billions that exist. How on earth does that work?

Step 1: Crawling

Google uses programmes called **crawlers** or **spiders** — automated bots that browse the web constantly, following links from page to page. Every time a crawler visits a page, it reads all the text and follows every link it finds, building a map of what exists on the web. Google's crawlers process billions of pages every day.

Step 2: Indexing

The information crawlers find gets stored in a massive database called the **index**. Think of it as a library catalogue — every word on every page gets logged, along with where it appears and how often. Google's index is estimated to contain hundreds of billions of pages and is stored across huge numbers of servers around the world.

Before Google, imagine trying to find a specific topic in the world's biggest library, but none of the books are organised and there's no catalogue. Google's crawlers read every book, and the index is the enormous card catalogue that results: "these books mention 'black holes' on pages 4, 17, and 203." When you search, it checks the catalogue, not the books.

Step 3: Ranking

The hard part. When you search, thousands of pages match your query. Google's algorithm decides which to show you first, using hundreds of factors. The original big idea was **PageRank** — the more other reputable pages link to a page, the more trustworthy that page probably is. But modern ranking involves over 200 factors including how fast the page loads, whether it works on mobile, how recently it was updated, and whether the content genuinely answers the question.

How does it do this in 0.4 seconds?

It doesn't search the web in real time. When you hit search, Google queries its pre-built index — not the actual web. The index is stored on thousands of servers, parts of which are kept in memory (fast RAM) for instant access. It's already read everything; it's just looking up your query in the catalogue. The speed is engineering, not magic — though the scale of the engineering is genuinely mind-boggling.