

**candour**

# Low-value pages and crawling

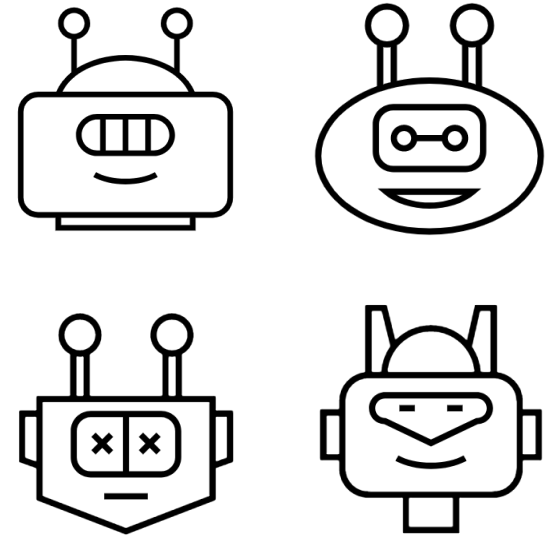
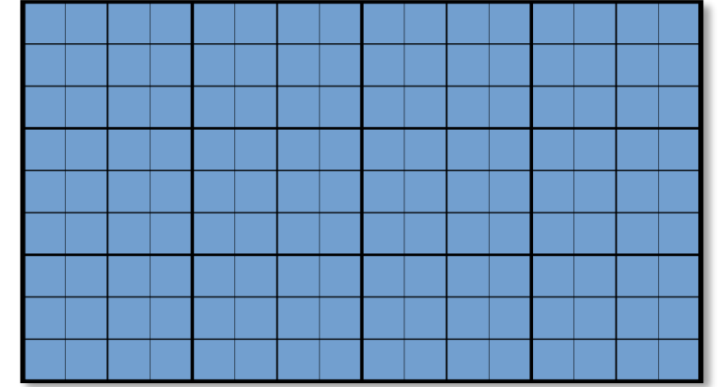


**Trainer:**

Mark Williams-Cook  
Candour

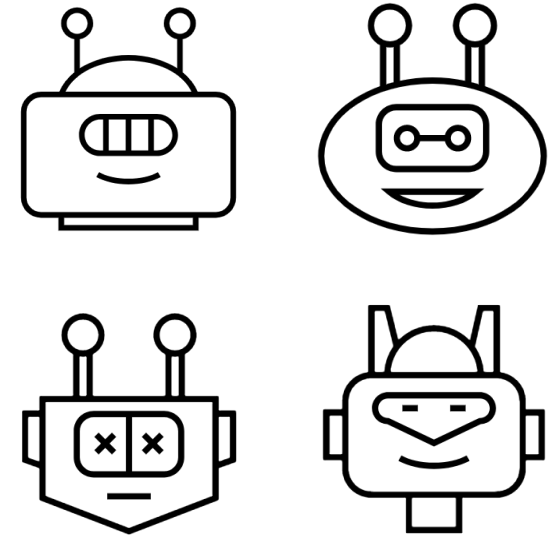
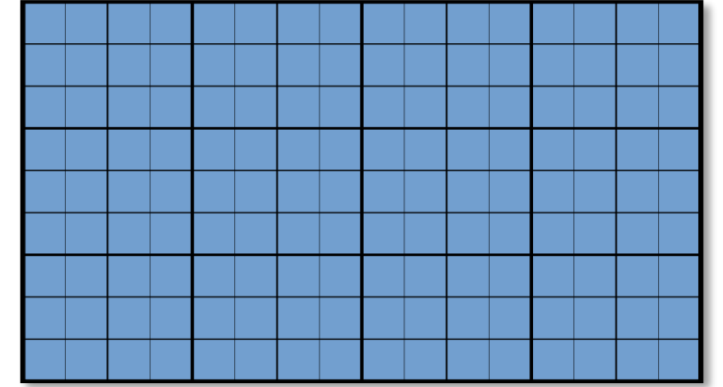


# What are low-value pages?



# What are low-value pages?

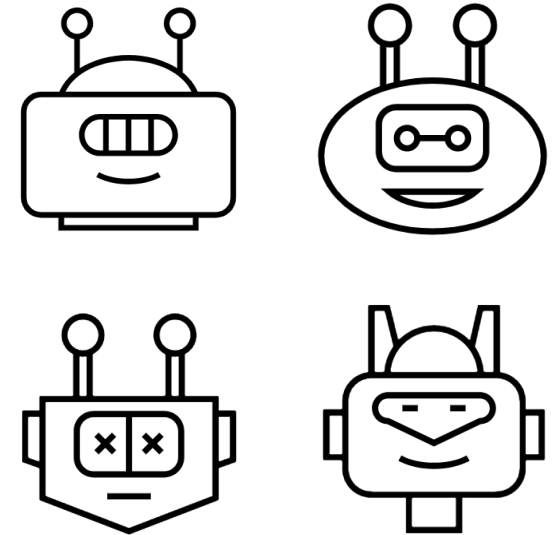
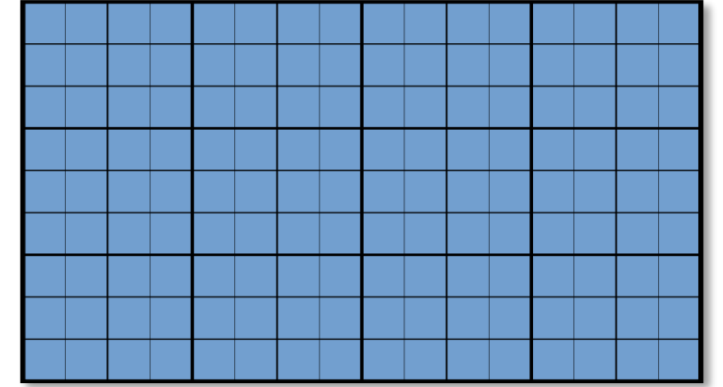
When we talk about “pages” we are actually talking about URLs.



# What are low-value pages?

When we talk about “pages” we are actually talking about URLs.

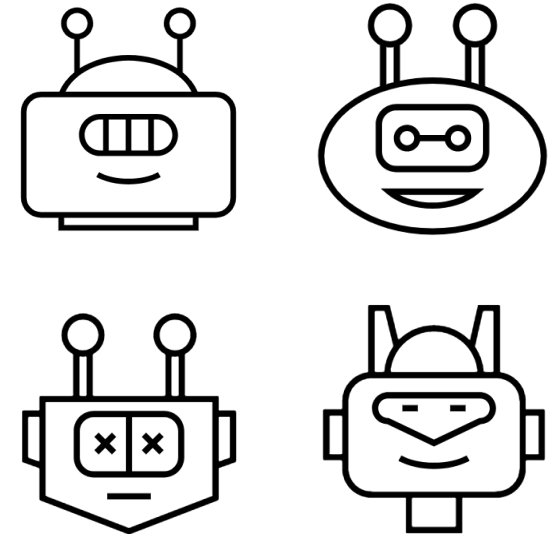
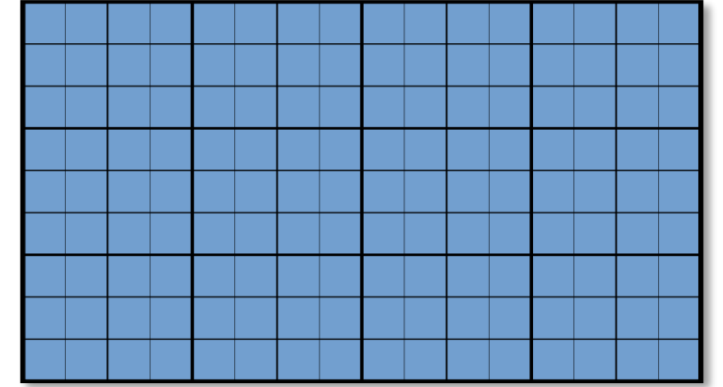
- Not worried if it appears in search
- Could be a duplicate/near duplicate page
- Might be generated because of technical debt



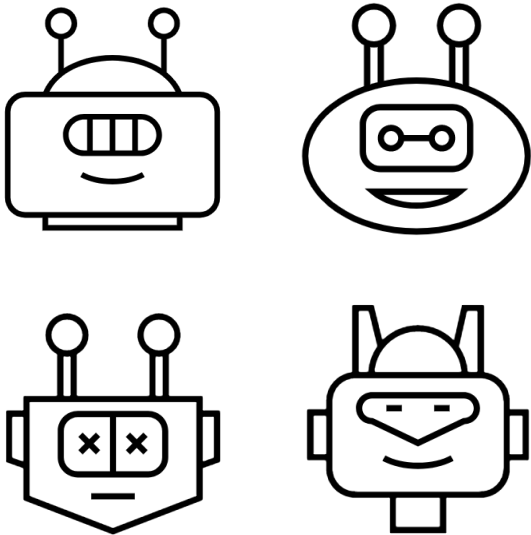
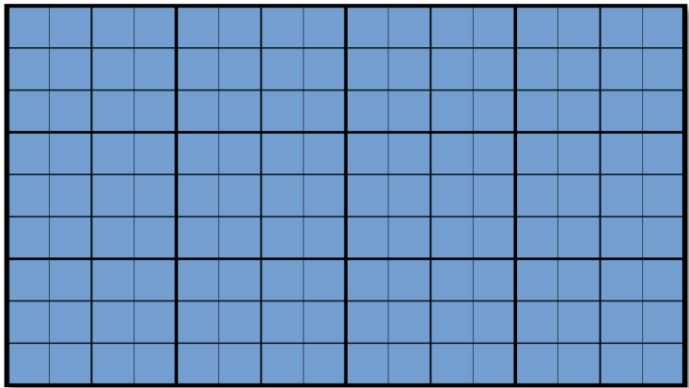
# Low-value pages and crawling

There are situations where a large website may be generating lots of “low-value” pages that are potentially causing crawling to not be as effective as it could be.

Sometimes the robots.txt is the best solution.



# Example 1: MailOnline

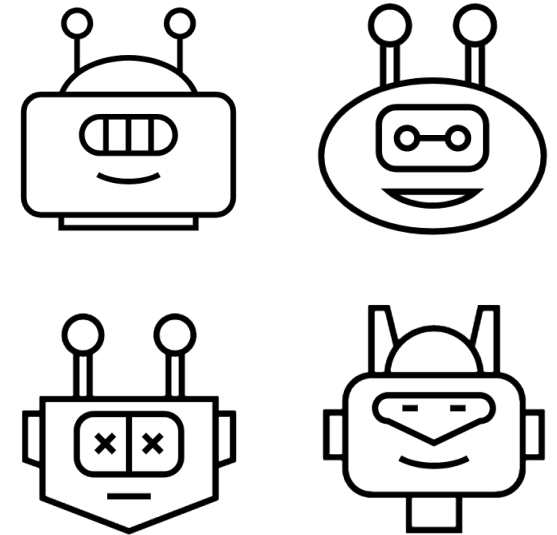
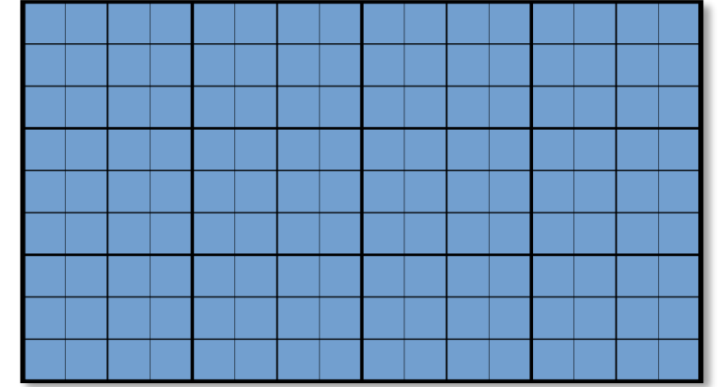


Low-value pages and crawling

# Technical debt

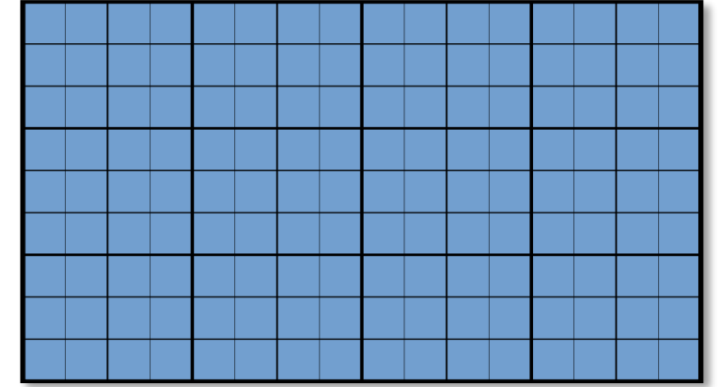
This is an example of **technical debt**.

There are other solutions to this problem that negate the need to use robots.txt and may result in better SEO performance.



# Key concept

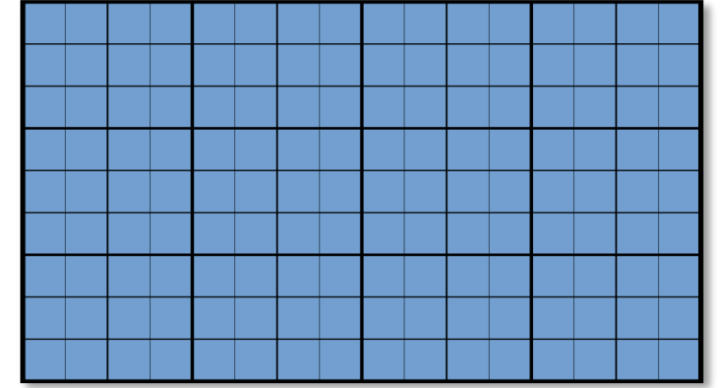
This kind of robots.txt management is only necessary on large sites where these kind of links scale with the website.





# Key concept

Log file analysis can be used to confirm if you are having issues with crawling and crawl budget.



# Challenge



Try visiting some of your favourite large websites and inspecting their robots.txt – it can be a really interesting “behind the scenes” look at their SEO!