

Searching with Sherlock Holmes

Powered by @Brilliant App

**Lecture prepared and delivered by
Dr Syed Khaldoon Khurshid**

- A ***search engine*** is a tool that helps you find the documents you're looking for among a potentially huge selection of documents. You use a search engine by giving it a ***query***.
- You are most likely to encounter search engines in three places. Most obviously, search engines like **Google, Bing, and Baidu** allow you to search across the entire web: the documents are web pages. Additionally, most computer operating systems have a **search feature** that allows you to inspect all the files on your hard drive.
- Finally, there are many **special-purpose search engines** which are used to offer information about the books in your local library, the donations given to political candidates in your country, or even the stars in the sky.
- The goal of this module is to learn how to use these search engines more effectively — and along the way you'll put the concepts in place to understand how they do their job so well.

- It's difficult to learn about search engines the way you learn about Newton's laws of motion or the rules of logic. A falling object will accelerate towards the earth at the same rate in the year 1930 or the year 2040.
- **Search engines, on the other hand, change very quickly.** The search results you get for searching for "**Sherlock and Irene**" on the web today may not be the same search results you get next month or next year.
- Which of the following might be possible reasons for this variation?

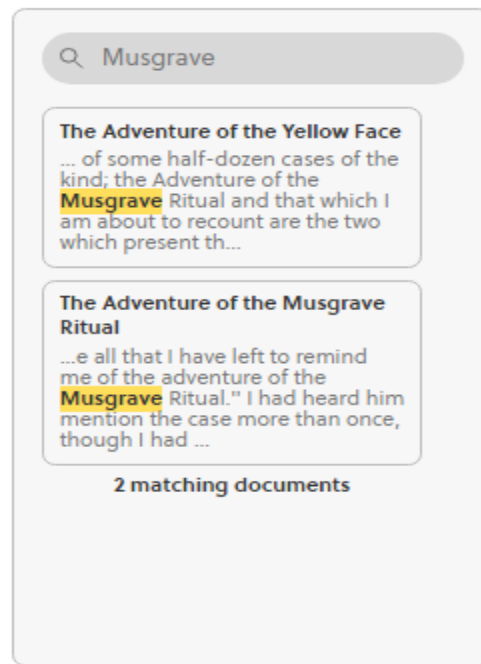
- A new movie with that title is released.
- A popular webpage that shows up on that query is taken down due to a copyright violation.
- The search engine developers change their algorithm for which results to display at the top.
- All of the above.

Correct answer: **All of the above**

Explanation

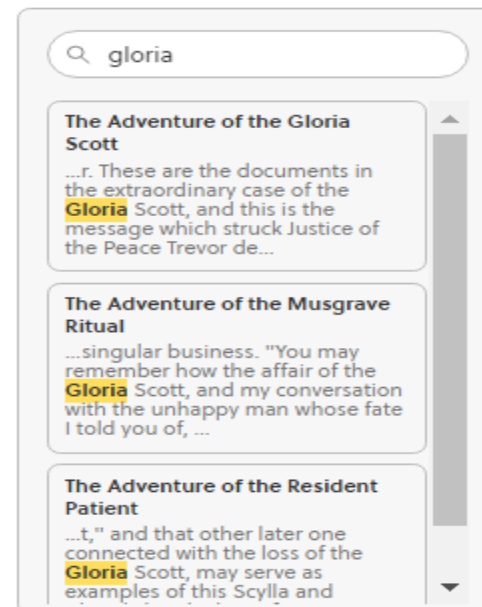
- Search results to a particular query can change both because the underlying document collection has changed or because the algorithm for finding and displaying them has changed.

- Because it's impossible to write a lasting module that asks you to use Google, in this module you'll use a search engine based on **4 novels and 47 short stories** by **Sir Arthur Conan Doyle**, featuring the great detective **Sherlock Holmes**.
- Most of the Sherlock Holmes stories are told from the perspective of Holmes' friend **Dr. John Watson**. The adventures had colorful names, like *The Adventure of the Blue Carbuncle*. The narrator, John Watson, frequently refers to previous adventures as he writes, sometimes referring to them as "affairs" or "problems."

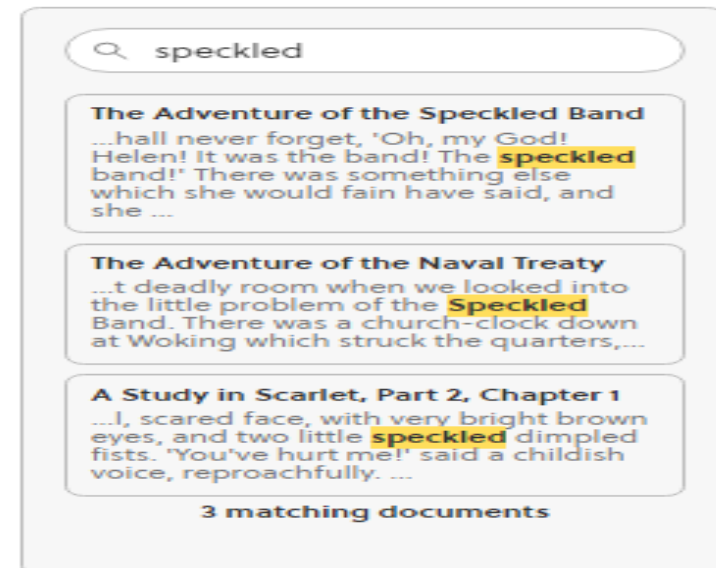


- Here, the search engine reveals that John Watson talks about *The Adventure of the Musgrave Ritual* in one other story, *The Adventure of the Yellow Face*. The search engine returned all the stories that contain the word "**Musgrave**" anywhere, as well as a **snippet** of text containing the first time the word "**Musgrave**" appeared in each story.

- The Gloria Scott was a convict ship bound for Australia. It was blown up following a mutiny in 1885. Years later, this event would be relevant to Sherlock Holmes' first case.
- The search engine will return all the stories that contain the word "gloria" anywhere, as well as a snippet of text that includes the first time the word "gloria" appears in the story.



- Many search engines, including our Sherlock Searcher, show you some of the text surrounding the place where the first occurrence of the search term was found, which can be useful.
- Search for "speckled" in the Sherlock Search window below. The search engine will return all the stories that contain the word "speckled" anywhere, as well as a snippet of text that includes the first time the word "speckled" appears in the story.
- Based on those results, **how many times did Holmes mention the adventure (or problem) of the Speckled Band in other stories besides The Adventure of the Speckled Band?**

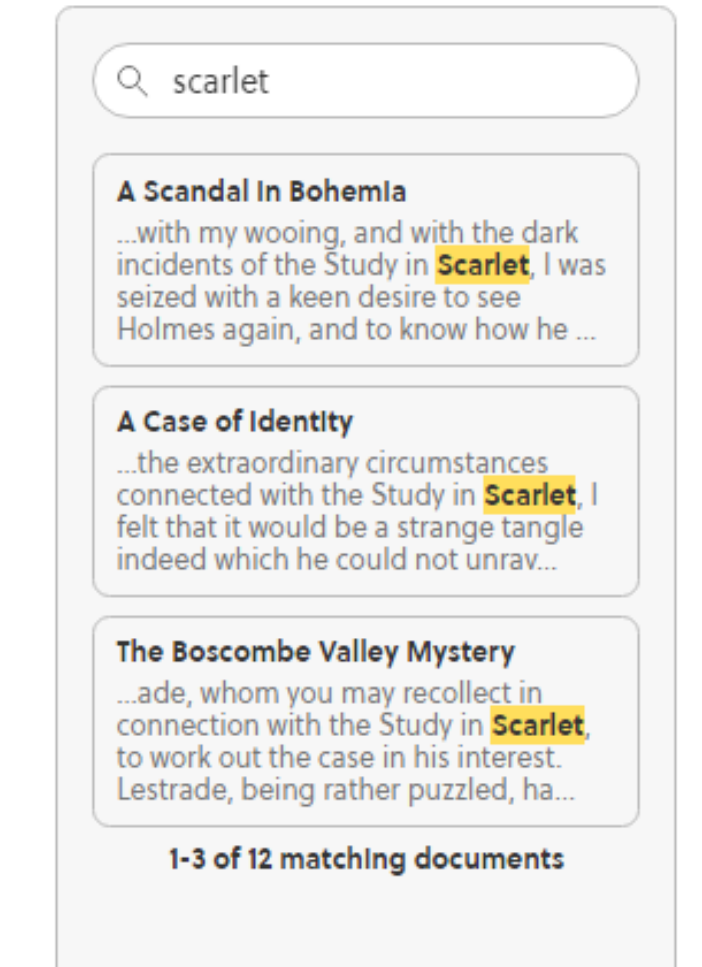


Correct answer: **One or two other stories**

Explanation

- Searching for "**speckled**" gives three results. One of them is *The Adventure of the Speckled Band* itself, that doesn't count.
- From the context around *The Adventure of the Naval Treaty*, it is referring to the other story as the problem of the Speckled Band. So there's at least one story.
- The last result is a chapter in the book *A Study in Scarlet*. This only uses the word "speckled" as an adjective. **It might or might not be the case that the chapter talks about *The Adventure of the Speckled Band* elsewhere, but there's no way to know that from this single search.** Therefore, we know either one or two stories refer back to *The Adventure of the Speckled Band*.

- Search for "**scarlet**". Based only on the results you get, how many other stories refer to *A Study in Scarlet*?



A screenshot of a search interface. At the top is a search bar with a magnifying glass icon and the text 'scarlet'. Below the search bar are three result cards. Each card has a title in bold, followed by a snippet of text where the word 'Scarlet' is highlighted in yellow. The first card is titled 'A Scandal In Bohemia', the second 'A Case of Identity', and the third 'The Boscombe Valley Mystery'. At the bottom of the interface, it says '1-3 of 12 matching documents'.

Q scarlet

A Scandal In Bohemia
...with my wooing, and with the dark incidents of the Study in **Scarlet**, I was seized with a keen desire to see Holmes again, and to know how he ...

A Case of Identity
...the extraordinary circumstances connected with the Study in **Scarlet**, I felt that it would be a strange tangle indeed which he could not unrav...

The Boscombe Valley Mystery
...ade, whom you may recollect in connection with the Study in **Scarlet**, to work out the case in his interest. Lestrade, being rather puzzled, ha...

1-3 of 12 matching documents

Correct answer: **Three or more**

Explanation

- The Sherlock Searcher only returns three results. None of them are *A Study in Scarlet*, but all of them refer to that book. You can also see that there were 12 documents — short stories or book chapters — where the word "scarlet" appeared, but **these may or may not have been references to the book.**
- (If you were curious, six short stories and one book aside from *A Study in Scarlet* refer to that title.)
- The Sherlock Searcher limits the number of results to help you refine your searching skills. These explorations only include **102 documents — 47 short stories and 55 book chapters.** Usually, you will use search engines to search thousands or millions of documents, and you'll need to use search to remove all but the tiniest fraction of these documents.

- Now that you know the basics of the Sherlock search engine, it's time to enter the story of Sherlock yourself!
- In the next module, you'll be a research assistant to the great-granddaughter of John Watson. **You'll develop your search engine skills by answering questions with the search engine and your own powers of deduction.**

