
















# Strategic Searching Online: How Search Engines Work: Web Crawlers and Web Indexing

<b>National Curriculum</b> Use search technologies effectively, appreciate how results are selected and ranked, and be discerning in evaluating digital content.		<b>Lesson Duration</b> It is estimated that this lesson will take approximately 60 minutes.
<b>Aim</b> To understand how search engines work.		
<b>Success Criteria</b> I can explain what a web crawler is. I can explain how web indexing works. I can explain how search engines work.	<b>Key Vocabulary:</b> Web crawler, spider, web indexing, index, search engine index, hyperlink.	
<b>Resources</b> <a href="#">Lesson Pack</a> Whiteboard and pens Pens or pencils	<b>Preparation</b> Differentiated <a href="#">Classroom Crawlers Activity Sheet</a> - one per child	

**Prior Learning:** Throughout this [Lesson Pack](#) children have used search engines and therefore should have a good understanding of what they are and what they are used for. It is important that children understand the difference between a web page and a website.

## Learning Sequence

	<b>Remember It:</b> Using the <a href="#">Lesson Presentation</a> , complete the quiz questions. The quiz asks children to decide which keywords would return the most efficient search results. This can be completed as a whole class, pair work or individually, asking the children to write their answers down on a whiteboard.	
	<b>Search Engines:</b> Using the <a href="#">Lesson Presentation</a> , introduce the term index and explain how this relates to search engines. <i>Can the children explain why they think having an index could be useful?</i> Use the <a href="#">Lesson Presentation</a> to introduce the term web crawlers and explain how search engines use them to create their search engine index. <i>Can children explain to a partner, in their own words, what a web crawler is?</i>	
	<b>Web Crawlers:</b> Using the <a href="#">Lesson Presentation</a> , explain what web crawlers are in more depth. Explain to the children what information web crawlers store, including the location of the web page. The <a href="#">Lesson Presentation</a> also explains some of the limitations of a web crawler programme.	
	<b>Classroom Crawlers:</b> Using the differentiated <a href="#">Classroom Crawlers Activity Sheet</a> , children must find the various classroom equipment listed. They must note down all the locations they find the equipment and the total number of each object found. This activity can be done in groups, pairs or independently. It is recommended that the safety of children is taken into account. An editable/blank version of this resource is available to be adapted as required.	
	Children need to index four pieces of classroom equipment, noting the location/s and total found.	
	Children need to index six pieces of classroom equipment, noting the location/s and total found. The items listed are likely to have multiple locations within the classroom.	
	<b>Evaluate:</b> Using the <a href="#">Lesson Presentation</a> , ask the children to discuss the questions provided with someone they didn't work with. Children should discuss the indexing process and relate this to search engines.	

**Explore it**

**Index it:** Children could be given the blank differentiated [Classroom Crawlers Activity Sheet](#) and index another area in the school.

**Use it:** Children could experience using an index. Ask the children to pick up a non-fiction book with an index page. In pairs the children could take turns to use the index page to find a specific word in a book. This could be done as part of a game.

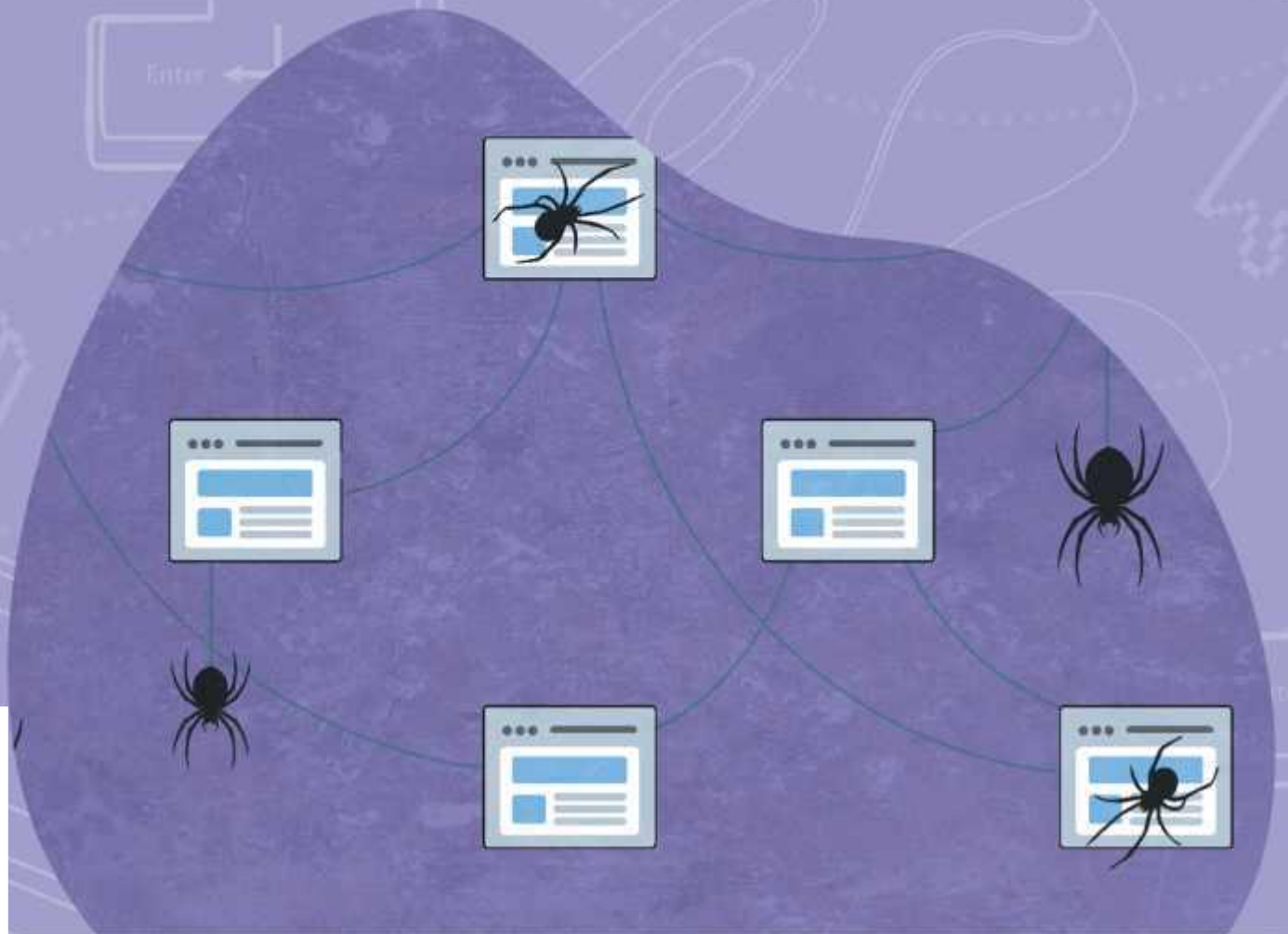
**Assessment Notes:**



# Computing

## Strategic Searching Online

# How Search Engines Work: Web Crawlers and Web Indexing



# Question Marks

You will spot question marks at certain points in this **Lesson Presentation**.

**Clicking the question marks will bring up key questions.**



The assessment questions that appear will enable you to check your understanding against the lesson aim and success criteria.

## **Aim**

- To understand how search engines work.

## **Success Criteria**

- I can explain what a web crawler is.
- I can explain how web indexing works.
- I can explain how search engines work.

# Remember It

Thinking about all you have learnt about search engines, can you complete the following quiz?

Research questions will be displayed at the top of the screen. You must decide, from the multiple choice answers, which keywords would return the most efficient search results.



# Remember It

What is the fastest animal on land?

a) fastest animal

**Not quite. Have another try.**

b) the fastest animal on land

**Not quite. Have another try.**

c) fastest animal land

**Well done! That's correct!**

When using a search engine, your search terms need to be specific to your desired search results. Remember, you only need to include keywords in your search terms.



# Remember It

How many flightless species of birds are there?

a) bird AND flightless

**Well done! That's correct!**

When using the Boolean operator AND, all searches that include both keywords are included in the search results.

b) bird OR flightless

**Not quite. Have another try.**

c) bird -flightless

**Not quite. Have another try.**

# Remember It

What are the names of some UK cities, not including London?

a) UK cities

**Not quite. Have another try.**

b) uk cities -london

**Well done! That's correct!**

When using the Boolean operator NOT (-) searches that include the keywords UK and not London will be included in the search results. Remember, search engines do not recognise capital letters for proper nouns.

c) UK cities but not London

**Not quite. Have another try.**

# Search Engines

The Internet is made up of nearly 2 billion websites. Thanks to search engines, these websites can be easily accessed through our inputted search terms. This is done through a program called 'web crawler' or 'spider'.



But how does this work?

When we input our search terms, a search engine will scan its **index** of web pages to find results that relate to your search terms.

An index is used to categorise items. They can be found in lots of different places, such as at the back of a non-fiction book and in a library. An index can help us find specific items that we are searching for by providing their location.

X

Why do you think having an index can be useful?



# Search Engines

## How does indexing work on the Internet?

A search engine makes its own index through a program called spider or web crawler.

A spider or web crawler is programmed to crawl through the World Wide Web. They visit web pages through hyperlinks and store information about each page they visit.

When the spider or web crawler stores information about a web page, it is creating an index of these web pages. This information is then all put together and stored. The web crawler program takes a copy of the web pages they visit and its location on the World Wide Web. These are then stored on the search engine's servers and becomes the search engine's index.

When a web page appears in your search results, this means a spider or web crawler has visited it and is recommending it based on the information it has collected from the web page. Another part of the web crawler or 'spider' program is to decide whether or not the web pages are good enough to deliver search results.

X

What is a web crawler? In your own words, explain it to a partner.



# Web Crawlers

The spiders or web crawlers look at a web page more than once. This means that if a web page gets an update, the search engine index has the newest copy stored.

The web crawlers look at all of the information on a website, such as:

keywords

images

website name

hyperlinks

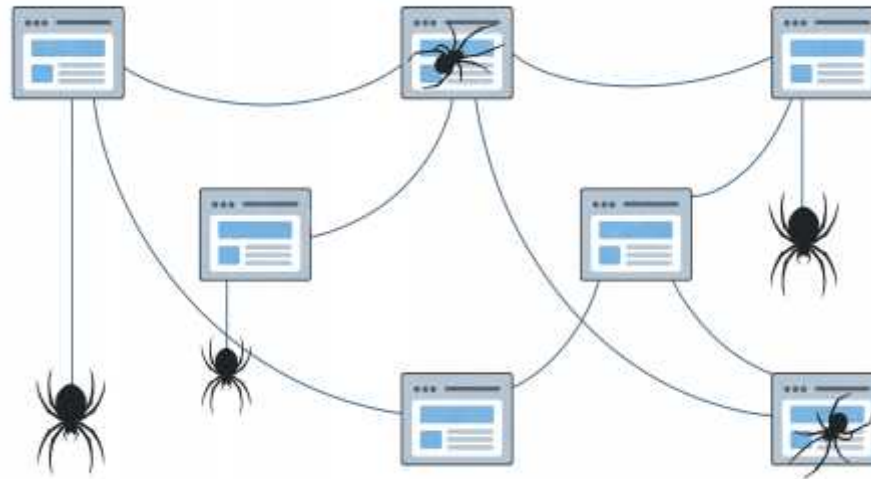
headings and  
sub-headings

web page titles

# Web Crawlers

Each search engine has its own web crawler program. This means that each search engine has a different search index.

When the web crawler program crawls through web pages, they are looking for hyperlinks on each web page. The hyperlinks are stored on a robots.txt file. This file is created by the web page developer and cannot be seen when viewing the website.



# Web Crawlers

Sometimes a website may not show up anywhere in search results, this could be because:

- the web page is new and hasn't been crawled yet.
- the web page has no hyperlinks to it from other web pages.
- the web page is too difficult to navigate around for the web crawlers.
- the website contains some code called 'crawler directives' that is blocking it from the search engine.
- the web crawlers have detected some scam software.



# Classroom Crawlers

You are now going to crawl the classroom, just like a web crawler. You have been given a list of equipment you would find in the classroom. On the sheet, note down the locations and the total number found for each item.

**Top Tip:** Sometimes equipment is not always put back in the correct place. Make sure you look really carefully around the classroom.



The worksheet is titled "Classroom Crawlers" and includes a URL "http://www.regentstudies.com". It features a table with three columns: "Object", "Location/s", and "Total Number Found". The table has four rows, with the first row containing the text "pencil", the second "yellow rubbering pencil", the third "marker", and the fourth "marker cap". At the bottom of the worksheet, there is an illustration of a person holding a globe.

Object	Location/s	Total Number Found
pencil		
yellow rubbering pencil		
marker		
marker cap		



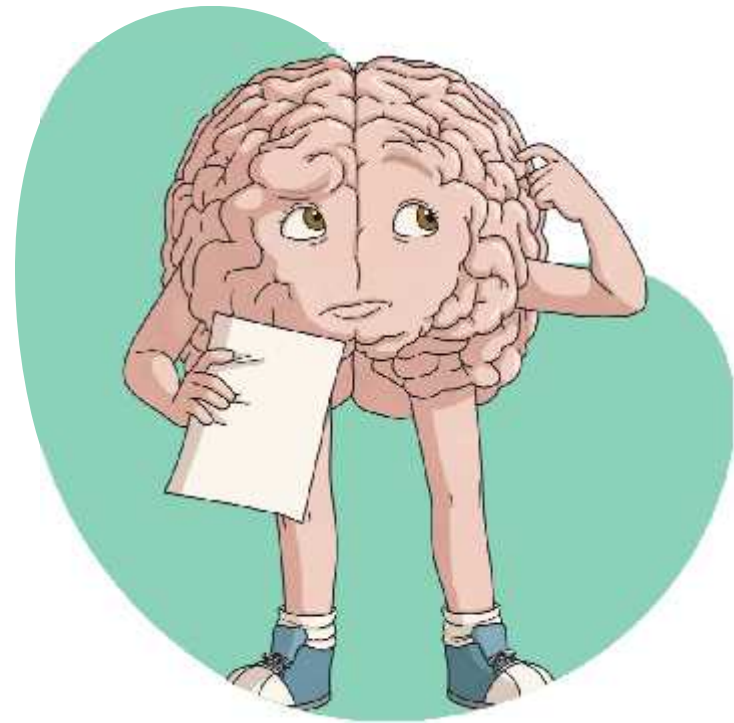
# Evaluate

What problems did you encounter?

Do you think web crawlers would also face a similar problem?

Did you create a method for 'crawling' the classroom?

Check with someone you didn't work with. Did you both find the equipment in the same places?

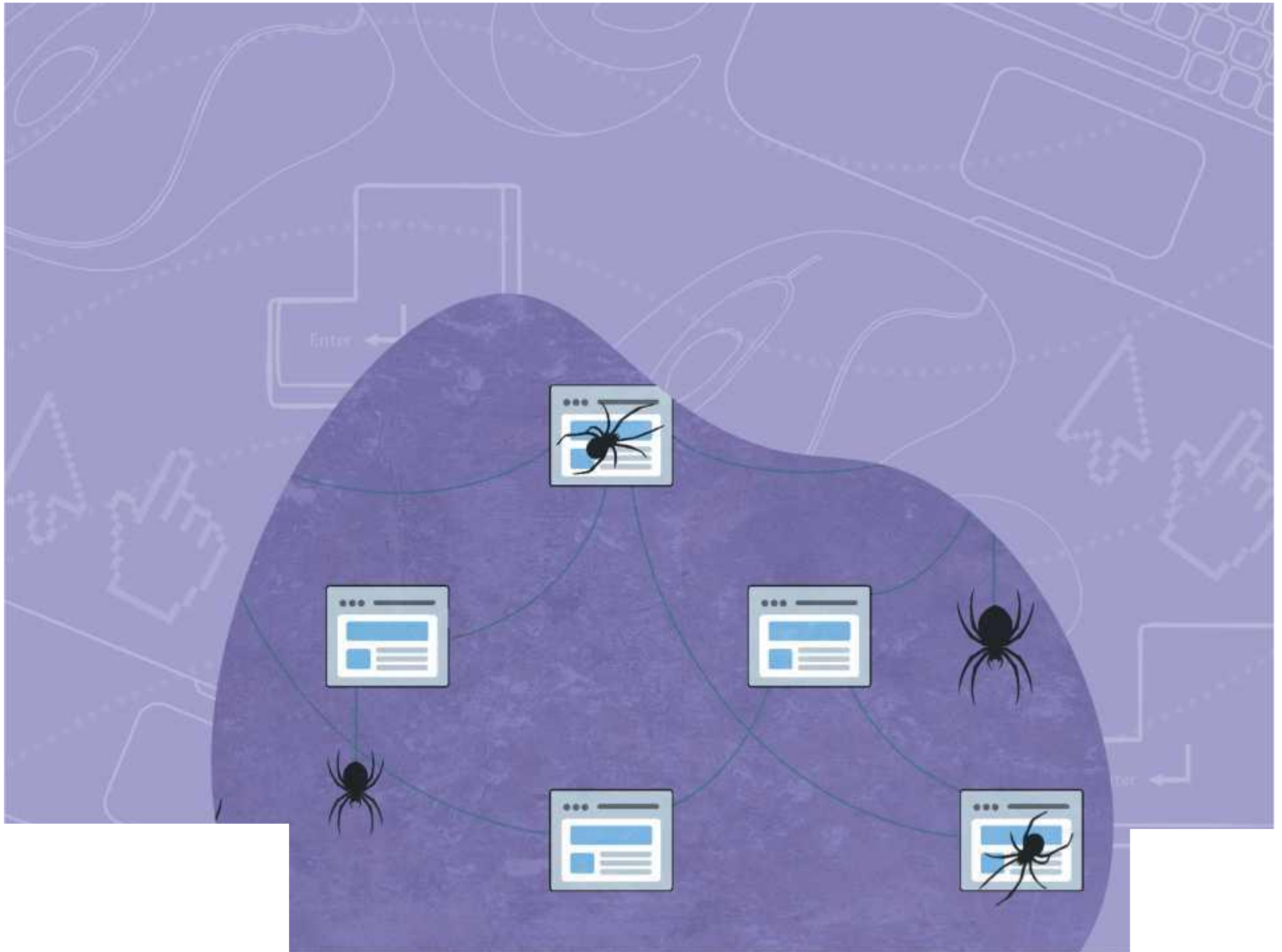


## **Aim**

- To understand how search engines work.

## **Success Criteria**

- I can explain what a web crawler is.
- I can explain how web indexing works.
- I can explain how search engines work.



# Classroom Crawlers

To understand how search engines work.



Object	Location/s	Total Number Found
pencil		
yellow colouring pencil		
mirror		
teacher pen		



# Classroom Crawlers

To understand how search engines work.



Object	Location/s	Total Number Found
pencil		
pen		
ruler		
glue stick		
whiteboard		



# Classroom Crawlers

To understand how search engines work.



Object	Location/s	Total Number Found
pencil		
pen		
ruler		
glue stick		
whiteboard		
paper		



# Classroom Crawlers

To understand how search engines work.



Object	Location/s	Total Number Found

# Classroom Crawlers

To understand how search engines work.



Object	Location/s	Total Number Found
pencil		
yellow colouring pencil		
mirror		
teacher pen		





# Classroom Crawlers

To understand how search engines work.



Object	Location/s	Total Number Found
pencil		
pen		
ruler		
glue stick		
whiteboard		



# Classroom Crawlers

To understand how search engines work.



Object	Location/s	Total Number Found
pencil		
pen		
ruler		
glue stick		
whiteboard		
paper		



# Classroom Crawlers

To understand how search engines work.



Object	Location/s	Total Number Found
pencil		
yellow colouring pencil		
mirror		
teacher pen		



# Classroom Crawlers

To understand how search engines work.



Object	Location/s	Total Number Found
pencil		
pen		
ruler		
glue stick		
whiteboard		



# Classroom Crawlers

To understand how search engines work.



Object	Location/s	Total Number Found
pencil		
pen		
ruler		
glue stick		
whiteboard		
paper		



Strategic Searching Online | How Search Engines Work: Web Crawlers and Web Indexing

To understand how search engines work.		
I can explain what a web crawler is.		
I can explain how web indexing works.		
I can explain how search engines work.		

Strategic Searching Online | How Search Engines Work: Web Crawlers and Web Indexing

To understand how search engines work.		
I can explain what a web crawler is.		
I can explain how web indexing works.		
I can explain how search engines work.		

Strategic Searching Online | How Search Engines Work: Web Crawlers and Web Indexing

To understand how search engines work.		
I can explain what a web crawler is.		
I can explain how web indexing works.		
I can explain how search engines work.		

Strategic Searching Online | How Search Engines Work: Web Crawlers and Web Indexing

To understand how search engines work.		
I can explain what a web crawler is.		
I can explain how web indexing works.		
I can explain how search engines work.		

Strategic Searching Online | How Search Engines Work: Web Crawlers and Web Indexing

To understand how search engines work.		
I can explain what a web crawler is.		
I can explain how web indexing works.		
I can explain how search engines work.		

Strategic Searching Online | How Search Engines Work: Web Crawlers and Web Indexing

To understand how search engines work.		
I can explain what a web crawler is.		
I can explain how web indexing works.		
I can explain how search engines work.		

Strategic Searching Online | How Search Engines Work: Web Crawlers and Web Indexing

To understand how search engines work.		
I can explain what a web crawler is.		
I can explain how web indexing works.		
I can explain how search engines work.		

Strategic Searching Online | How Search Engines Work: Web Crawlers and Web Indexing

To understand how search engines work.		
I can explain what a web crawler is.		
I can explain how web indexing works.		
I can explain how search engines work.		