

Lesson 4: Searching the Web

Computing

Developing for the Web

Allen Heard

Materials from the Teach Computing Curriculum created by the National Centre for Computing Education



Task 1 - Crawling the web - part 1

Here is some HTML code. Can you determine the keywords, content and last updated from the information given?

```
<head>
<title>Scratch (programming language) - Wikipedia</title>
<meta name="description" content="Information about Scratch, a visual programming language">
</head>
<body>
<h1>Scratch (programming language)</h1>
<p>Scratch is a block-based visual programming language and website targeted primarily at children.<p> <em>Last update: 20/04/20</em>
</body>
```

Credit: Wikipedia



Task 1 - Crawling the web - part 1

Results stored in search engine index

Read the content of the web page as if you were a crawler. Fill in the table below to show what information you believe is most important when cataloguing this page:

Five keywords for page in order of relevance	Type of content found (images, text, etc)	Date of last update
1.		
2.		
3.		
4.		
5.		



Task 1 - Crawling the web - part 2

```
<head>
```

```
<title>Chapter 4: Plate Tectonics</title>
```

```
<meta name="description" content="How plate tectonics shaped the world from its beginnings as Pangea">
```

```
</head>
```

```
<body>
```

```
<h1>Early Evidence for Continental Drift</h1>
```

```

```

```
<p>The first piece of evidence is that the shape of the coastlines of some continents fit together like pieces of a jigsaw puzzle.<p> <em>Last update: May 1, 2019</em>
```

```
</body>
```

Credit: SLCC Pressbooks



Task 1 - Crawling the web - part 2

Results stored in search engine index

Read the content of the web page as if you were a crawler. Fill in the table below to show what information you believe is most important when cataloguing this page:

Five keywords for page in order of relevance	Type of content found (images, text, etc)	Date of last update
1.		
2.		
3.		
4.		
5.		



Task 2 - Build a high-quality web page

Considering all that you have learnt, you now need to create a web page that would rank highly at the top of a list of search results.

Your page needs to summarise the how search engines work, including:

- How crawlers work
- How web pages are indexed
- How web pages could be ranked and why this is necessary



What makes a quality web page?

Clear heading

Using the heading tags `<h1>`, `<h2>`, etc.

Include meta-tags

In the `<head>` of the page

Include keywords

Regularly and near the top of the page

Images

Suitable images that illustrate the content



What makes a quality web page?

Junk

Avoid using unnecessary information

Colour

Use complementary colours

Readability

Key information obvious



Task 3 - What makes a quality web page

Tick the descriptions that are key to building a quality web page.

Clear headings using the heading tags	
Suitable meta-tags	
Plenty of unrelated keywords to drive traffic to your page	
Important keywords near the top of each section of the page	
Links to popular websites unrelated to yours	
Key information obvious on the page (e.g. uses bold, italics, etc.)	
Suitable images	
As much colour as possible to make it stand out	
No unnecessary information	
A complementary colour that isn't too strong	

