SEARCH ENGINES AND WEB SEARCH



1. What is the primary function of a web crawler or spider?
A. To find new content on the internet
B. To index the words and content of websites
C. To visit the links and URLs found on websites
D. All of the above

2. Which of the following is NOT a common term used to describe a web crawler?
A. Robot
B. Search bot
C. Indexer
D. Scraper

3. What is the first thing a web crawler does when it visits a website?
A. Loads the site's content into a database
B. Looks for a robots.txt file
C. Indexes the words and content of the site
D. Visits the links and URLs found on the site

4. Which statement is true about the web crawling process?
A. The crawler only visits a website once
B. The crawler checks again to verify if a page is offline
C. The crawler ignores the instructions in the robots.txt file
D. The crawler only indexes the words on a website

5. What is the typical crawl rate of a major search engine like Google?
A. 5 requests per second
B. 10,000 pages per day
C. 30 trillion pages per month
D. All of the above

6. What is the purpose of the robots.txt file on a website?
A. To instruct the crawler on which parts of the site to index
B. To verify that a page is offline
C. To control the crawl rate of the spider
D. To load the site's content into a database

7. How frequently does Google crawl the web?
A. 5 requests per second
B. 10,000 pages per day
C. 30 trillion pages per month
D. All of the above

LIVEWORKSHEETS