

# JOB OPENING

DATA  
SCIENCE



## Python Web Scraper Specialist (Full Time / Part Time / Casual Hours)

### Job Description:

We are looking for a skilled Python Web Scraper Specialist to join our team. The ideal candidate will have experience in developing and maintaining web scraping scripts and tools. You will be responsible for extracting and processing data from various websites, ensuring accuracy, efficiency, and compliance with legal and ethical standards.

### Key Responsibilities:

- **Web Scraping Development:**
  - Design, develop, and maintain Python scripts to scrape data from various websites.
  - Utilize libraries such as BeautifulSoup, Scrapy, Selenium, and Requests to extract and process data.
  - Handle dynamic content using tools like Selenium for websites that require interaction or JavaScript rendering.
- **Data Processing & Management:**
  - Clean, format, and store scraped data into databases or other data storage solutions.
  - Ensure data integrity and accuracy during extraction and storage processes.
  - Automate data extraction tasks to run on a regular schedule.
- **Website Analysis:**
  - Analyze website structures to identify the best approach for data extraction.
  - Monitor website changes and update scraping scripts accordingly.
  - Ensure compliance with websites' robots.txt files and respect their terms of service.
- **Optimization & Performance:**
  - Optimize scraping scripts for speed and efficiency.
- **Documentation & Reporting:**
  - Document the scraping process, including code, methodologies, and any special configurations.
  - Generate reports or dashboards to display the scraped data and insights.

### Required Skills:

- **Proficiency in Python:**
  - Strong understanding of Python programming, especially in the context of web scraping.
  - Experience with libraries like BeautifulSoup, Scrapy, Selenium, Requests, and lxml.
- **Web Technologies:**
  - Knowledge of HTML, CSS, and JavaScript to analyze web pages and determine scraping strategies.
  - Familiarity with JSON, XML, and other data formats often used in web APIs.
- **Data Handling:**
  - Experience with data manipulation libraries such as Pandas.
  - Ability to work with databases (e.g., SQL, NoSQL) and data storage formats (e.g., CSV, JSON).

- **Problem Solving:**
  - Strong analytical skills to troubleshoot and solve issues related to web scraping.
- **Knowledge of Web Scraping Ethics & Legalties:**
  - Understanding of the ethical considerations and legal restrictions surrounding web scraping.
  - Ability to design scrapers that respect robots.txt and comply with terms of service.
- **Automation & Scheduling:**
  - Experience with scheduling tools like Cron or task queues like Celery to automate scraping tasks.
  - Knowledge of cloud services (e.g., AWS, GCP) for deploying and running scrapers.

## Preferred Qualifications:

- **Experience with Scrapy Framework:**
  - Advanced use of Scrapy for large-scale web scraping projects.
  - Familiarity with Scrapy pipelines, middlewares, and settings optimization.
- **Continuous Integration/Deployment (CI/CD):**
  - Familiarity with CI/CD pipelines for automating the deployment and updating of scrapers.
- **Personal Attributes:**
  - **Detail-Oriented:** Strong attention to detail, ensuring accuracy in data extraction and processing.
  - **Self-Motivated:** Ability to work independently and manage multiple scraping projects simultaneously.
  - **Collaborative:** Willingness to work closely with data analysts, developers, and other team members.

Interested applicants should send their CV and an application letter to [humanresources@ayadata.ai](mailto:humanresources@ayadata.ai) or upload their CV to our website ([www.ayadata.ai](http://www.ayadata.ai)). The subject of your email application should be **Python Web Scraper Specialist**.

**Terms and Conditions apply.**

**Head Office:**  
25 Southampton buildings,  
London, WC1A 2AL  
+44-33-377-21194

**Ghana Office:**  
87 Spintex Road  
Accra, Ghana  
(+233) 558 123029  
(+233) 302 975370

info@ayadata.ai  
LinkedIn: Aya-Data  
www.ayadata.ai

