AIM

Program to implement a program to scrap the web page of any popular website – suggested python package is scrappy (ensure ethical conduct).

Programming code:

```
class BlogSpider(scrapy.Spider):

    name = 'blogspider'

    start_urls = ['https://www.zyte.com/blog/']


    def parse(self, response):

        for title in response.css('.oxy-post-title'):

            yield {'title': title.css('::text').get()}


        for next_page in response.css('a.next'):

            yield response.follow(next_page, self.parse)
```

OUTPUT

```
[
{&quot;title&quot;: &quot;Zyte named as one of Deloitte Technology Fast 50&quot;},
{&quot;title&quot;: &quot;Zyte named as one of Deloitte Technology Fast 50&quot;},
{&quot;title&quot;: &quot;How to extract data from an HTML table&quot;},
{&quot;title&quot;: &quot;What is a proxy server and how do they work?&quot;},
{&quot;title&quot;: &quot;Extract Summit 2021: Highlights and key
takeaways&quot;},
{&quot;title&quot;: &quot;How does a headless browser help with web scraping and
data extraction?&quot;},
{&quot;title&quot;: &quot;Proxies versus VPNs: What\u2019s the difference, and
which one is right for my
use case?&quot;},
{&quot;title&quot;: &quot;Manage bans and get your data with Zyte Data API Smart
Browser&quot;},
{&quot;title&quot;: &quot;How to reduce noise in the data through data
```

parsing"},

{"title": "What is web data harvesting?"},
{"title": "In pursuit of perfection: measuring web product data quality"},
{"title": "Zyte named as one of Deloitte Technology Fast 50"},
{"title": "Web Data Extraction Summit 2021"},

  {"title": "Residential Proxies: How are they different to data center proxies & how to
manage them"},
{"title": "Zyte Developers Community newsletter issue #10"},
{"title": "What is data mining? How is it different from web scraping?"},
{"title": "Zyte Developers Community newsletter issue #9"},
{"title": "How Scrapy makes web crawling easy"},
]