

第三课 网络数据的获取与表示

第九节 Scrapy使用步骤 (3)

主 讲: Robin

Scrapy使用步骤

- 安装: `conda install -c conda-forge scrapy`
- 1. 创建工程
- 2. 定义Item, 构造爬取的对象 (可选)
- 3. 编写Spider, 爬虫主体
- 4. 编写配置和Pipeline, 用于处理爬取的结果 (可选)
- 5. 执行爬虫

Scrapy使用步骤

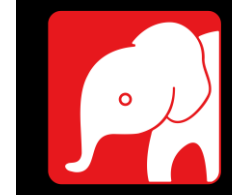
2. 定义Item

- scrapy.Field()

```
class CityItem(scrapy.Item):  
    city_name = scrapy.Field() # 城市名称  
    city_link = scrapy.Field() # 城市链接
```

3. 编写Spider

- 使用BeautifulSoup解析网页内容
- 调用自定义的Item



只为遇见明天更优秀的你！