

web scraping :

Web Scraping with Python, 2nd Edition [Book]

<https://www.oreilly.com/library/view/web-scraping-with/9781491985564/>
<https://scrapeops.io/python-scrapy-playbook/scrapy-beginners-guide/>
<https://scrapeops.io/python-scrapy-playbook/scrapy-selenium/>
<https://scrapeops.io/python-scrapy-playbook/scrapy-managing-user-agents/>
<https://scrapeops.io/python-scrapy-playbook/scrapy-rotating-proxy-guide/>
<https://www.zenrows.com/blog/scrapy-proxy#prerequisites>
<https://www.zyte.com/blog/scrapy-proxy/>
<https://scrapeops.io/python-scrapy-playbook/extensions/scrapy-scrapyd-guide/>
<https://scrapeops.io/python-scrapy-playbook/best-scrapyd-dashboards-ui/>

kafka :

Kafka: The Definitive Guide [Book]

<https://www.oreilly.com/library/view/kafka-the-definitive/9781491936153/>
<https://www.youtube.com/watch?v=1JAljxpm-Q&t=2968s>
<https://www.youtube.com/watch?v=2YJEWPD8yk&t=3422s>
<https://kafka.apache.org/quickstart>
<https://developer.confluent.io/get-started/python/>
<https://pypi.org/project/os-scrapy-kafka-pipeline/>

spark :

Stream Processing with Apache Spark [Book]

<https://www.oreilly.com/library/view/stream-processing-with/9781491944233/>
<https://liliasfaxi.github.io/Atelier-Spark/>
<https://neptune.ai/blog/saving-trained-model-in-python>

elasticsearch + kibana :

<https://gitlab.com/xavki/presentations-elk>
<https://www.elastic.co/guide/en/elasticsearch/reference/current/deb.html>
<https://www.elastic.co/guide/en/kibana/current/deb.html>