**Section 15: Web Scaping with Python**

**17.03.**

**116. Introduction to Web Scraping**

Web scraping is a general term for automating the gathering of data from a website

Rules of scraping:

Get permission before scraping

If there are many scraping attempts, the IP address may get blocked

Some sites automatically block scraping software

Limitations of scraping:

For each site that is unique a different script is needed

Updates inside a site that are frequent may break the script

While scraping a site, the interaction is manly with HTML and CSS, just perhaps with JS

With HTML and CSS main info can be in a page

Web scaping in Python is done with BeautifulSoup and request libraries: requests, lxml, bs4

**117. Setting Up Web Scraping Libraries**

Before web scraping the elements of interest and the source code should be inspected

For this whole page source code can be viewed and particular elements inspected