Scrapy

	Fill in	Note
Name of the software	Scrapy	
Developer	Scrapinghub, Ltd.	
Version	Scrapy 1.8.0	
Year of publication	First release: 2008 Newest version: 29.10.2019	
Description of the software	Scrapy is a free and open-source web-crawling framework written in Python. Originally designed for web scraping, it can also be used to extract data using APIs or as a general-purpose web crawler. Scrapy project architecture is built around "spiders", which are self-contained crawlers that are given a set of instructions. Following the spirit of others don't repeat yourself frameworks, it makes it easier to build and scale large crawling projects by allowing developers to reuse their code. Scrapy also provides a web-crawling shell, which can be used by developers to test their assumptions on a site's behavior. It does not support continuous crawls and it can be difficult to export big amounts of data.	
Link to the software	https://scrapy.org/	GitHub: https://github.com/scrapy/scrapy
Is it open-source	YES	

software?		
Is it free software?	YES	
License	[BSD License](#)	
Does it cost to use?	NO	
Does it collect any personal data?	It sends and receives cookies like a normal browser would do, but no form of log in or personal information is required.	
Does it respect the right to forget?		Software should have a function to delete all data. The deletion should be permanent and not reversible.
Application logs	Scrapy uses Python's logging system.	https://docs.scrapy.org/en/latest/topics/logging.html
Does it contain an online user guide?	https://docs.scrapy.org/en/latest/	The amount of documentation is one of the strengths of scrapy.
Hardware configuration	No information on minimal hardware configuration, but it is possible to Benchmark on scrapy. See note for more information.	https://docs.scrapy.org/en/latest/topics/benchmarking.html
Software configuration	Runs on Linux, Windows, Mac OS X, BSD.	
Text data visualisation	NO	You need to use other scripts in Python to visualize.
Does it analyse pictures/photos?	NO	
Multiple languages	YES	Read more: https://www.w3.org/Protocols/rfc2616/rfc2616-sec14.html#sec14.4
Third-party software products		

integration		
Cross-platform support for desktop, tablet and mobile users	NO	
Scalability	Scalable	
Automatically creates database	Yes. Can store data in different formats	http://doc.scrapy.org/en/latest/topics/feed-exports.html
Uses AI	NO	