

Middleware

- Cookies and sessions
- HTTP compression
- HTTP authentication
- HTTP cache
- User-agent spoofing
- robots.txt handling
- Crawl depth control

Alternatives of Scrapy