

Scrapy **|version|** documentation

Scrapy is a fast high-level [web crawling](#) and [web scraping](#) framework, used to crawl websites and extract structured data from their pages. It can be used for a wide range of purposes, from data mining to monitoring and automated testing.

Getting help

Having trouble? We'd like to help!

- Try the `:doc:`FAQ <faq>`` -- it's got answers to some common questions.

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs] index.rst, line 22); [backlink](#)
Unknown interpreted text role "doc".

- Looking for specific information? Try the `ref`genindex`` or `ref`modindex``.

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs] index.rst, line 23); [backlink](#)
Unknown interpreted text role "ref".

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs] index.rst, line 23); [backlink](#)
Unknown interpreted text role "ref".

- Ask or search questions in [StackOverflow](#) using the scrapy tag.
- Ask or search questions in the [Scrapy subreddit](#).
- Search for questions on the archives of the [scrapy-users mailing list](#).
- Ask a question in the [#scrapy IRC channel](#),
- Report bugs with Scrapy in our [issue tracker](#).
- Join the Discord community [Scrapy Discord](#).

First steps

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs] index.rst, line 42)

Unknown directive type "toctree".

```
.. toctree::
   :caption: First steps
   :hidden:

   intro/overview
   intro/install
   intro/tutorial
   intro/examples
```

`:doc:`intro/overview``

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs] index.rst, line 52); [backlink](#)
Unknown interpreted text role "doc".

Understand what Scrapy is and how it can help you.

`:doc:`intro/install``

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-

resources\scrapy-master\docs\[scrapy-master] [docs] index.rst, line 55); [backlink](#)

Unknown interpreted text role "doc".

Get Scrapy installed on your computer.

`:doc:`intro/tutorial``

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs] index.rst, line 58); [backlink](#)

Unknown interpreted text role "doc".

Write your first Scrapy project.

`:doc:`intro/examples``

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs] index.rst, line 61); [backlink](#)

Unknown interpreted text role "doc".

Learn more by playing with a pre-made Scrapy project.

Basic concepts

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs] index.rst, line 68)

Unknown directive type "toctree".

```
.. toctree::
   :caption: Basic concepts
   :hidden:

   topics/commands
   topics/spiders
   topics/selectors
   topics/items
   topics/loaders
   topics/shell
   topics/item-pipeline
   topics/feed-exports
   topics/request-response
   topics/link-extractors
   topics/settings
   topics/exceptions
```

`:doc:`topics/commands``

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs] index.rst, line 86); [backlink](#)

Unknown interpreted text role "doc".

Learn about the command-line tool used to manage your Scrapy project.

`:doc:`topics/spiders``

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs] index.rst, line 89); [backlink](#)

Unknown interpreted text role "doc".

Write the rules to crawl your websites.

`:doc:`topics/selectors``

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs] index.rst, line 92); [backlink](#)

Unknown interpreted text role "doc".

Extract the data from web pages using XPath.

`:doc:`topics/shell``

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs] index.rst, line 95); [backlink](#)

Unknown interpreted text role "doc".

Test your extraction code in an interactive environment.

`:doc:`topics/items``

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs] index.rst, line 98); [backlink](#)

Unknown interpreted text role "doc".

Define the data you want to scrape.

`:doc:`topics/loaders``

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs] index.rst, line 101); [backlink](#)

Unknown interpreted text role "doc".

Populate your items with the extracted data.

`:doc:`topics/item-pipeline``

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs] index.rst, line 104); [backlink](#)

Unknown interpreted text role "doc".

Post-process and store your scraped data.

`:doc:`topics/feed-exports``

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs] index.rst, line 107); [backlink](#)

Unknown interpreted text role "doc".

Output your scraped data using different formats and storages.

`:doc:`topics/request-response``

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs] index.rst, line 110); [backlink](#)

Unknown interpreted text role "doc".

Understand the classes used to represent HTTP requests and responses.

`:doc:`topics/link-extractors``

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs] index.rst, line 113); [backlink](#)

Unknown interpreted text role "doc".

Convenient classes to extract links to follow from pages.

`:doc:`topics/settings``

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-

resources\scrapy-master\docs\[scrapy-master] [docs] index.rst, line 116); [backlink](#)

Unknown interpreted text role "doc".

Learn how to configure Scrapy and see all `ref` available settings `<topics-settings-ref>`.

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs] index.rst, line 116); [backlink](#)

Unknown interpreted text role "ref".

`:doc:topics/exceptions``

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs] index.rst, line 120); [backlink](#)

Unknown interpreted text role "doc".

See all available exceptions and their meaning.

Built-in services

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs] index.rst, line 125)

Unknown directive type "toctree".

```
.. toctree::
   :caption: Built-in services
   :hidden:

   topics/logging
   topics/stats
   topics/email
   topics/telnetconsole
   topics/webservice
```

`:doc:topics/logging``

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs] index.rst, line 136); [backlink](#)

Unknown interpreted text role "doc".

Learn how to use Python's builtin logging on Scrapy.

`:doc:topics/stats``

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs] index.rst, line 139); [backlink](#)

Unknown interpreted text role "doc".

Collect statistics about your scraping crawler.

`:doc:topics/email``

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs] index.rst, line 142); [backlink](#)

Unknown interpreted text role "doc".

Send email notifications when certain events occur.

`:doc:topics/telnetconsole``

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs] index.rst, line 145); [backlink](#)

Unknown interpreted text role "doc".

Inspect a running crawler using a built-in Python console.

:doc:`topics/webservice`

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs]index.rst, line 149); [backlink](#)

Unknown interpreted text role "doc".

Monitor and control a crawler using a web service.

Solving specific problems

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs]index.rst, line 154)

Unknown directive type "toctree".

```
.. toctree::
   :caption: Solving specific problems
   :hidden:

  faq
   topics/debug
   topics/contracts
   topics/practices
   topics/broad-crawls
   topics/developer-tools
   topics/dynamic-content
   topics/leaks
   topics/media-pipeline
   topics/deploy
   topics/autothrottle
   topics/benchmarking
   topics/jobs
   topics/coroutines
   topics/asyncio
```

:doc:`faq`

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs]index.rst, line 175); [backlink](#)

Unknown interpreted text role "doc".

Get answers to most frequently asked questions.

:doc:`topics/debug`

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs]index.rst, line 178); [backlink](#)

Unknown interpreted text role "doc".

Learn how to debug common problems of your Scrapy spider.

:doc:`topics/contracts`

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs]index.rst, line 181); [backlink](#)

Unknown interpreted text role "doc".

Learn how to use contracts for testing your spiders.

:doc:`topics/practices`

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-

resources\scrapy-master\docs\[scrapy-master] [docs] index.rst, line 184); [backlink](#)

Unknown interpreted text role "doc".

Get familiar with some Scrapy common practices.

`:doc:`topics/broad-crawls``

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs] index.rst, line 187); [backlink](#)

Unknown interpreted text role "doc".

Tune Scrapy for crawling a lot domains in parallel.

`:doc:`topics/developer-tools``

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs] index.rst, line 190); [backlink](#)

Unknown interpreted text role "doc".

Learn how to scrape with your browser's developer tools.

`:doc:`topics/dynamic-content``

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs] index.rst, line 193); [backlink](#)

Unknown interpreted text role "doc".

Read webpage data that is loaded dynamically.

`:doc:`topics/leaks``

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs] index.rst, line 196); [backlink](#)

Unknown interpreted text role "doc".

Learn how to find and get rid of memory leaks in your crawler.

`:doc:`topics/media-pipeline``

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs] index.rst, line 199); [backlink](#)

Unknown interpreted text role "doc".

Download files and/or images associated with your scraped items.

`:doc:`topics/deploy``

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs] index.rst, line 202); [backlink](#)

Unknown interpreted text role "doc".

Deploying your Scrapy spiders and run them in a remote server.

`:doc:`topics/autothrottle``

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs] index.rst, line 205); [backlink](#)

Unknown interpreted text role "doc".

Adjust crawl rate dynamically based on load.

`:doc:`topics/benchmarking``

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs]index.rst, line 208); [backlink](#)

Unknown interpreted text role "doc".

Check how Scrapy performs on your hardware.

`:doc:`topics/jobs``

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs]index.rst, line 211); [backlink](#)

Unknown interpreted text role "doc".

Learn how to pause and resume crawls for large spiders.

`:doc:`topics/coroutines``

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs]index.rst, line 214); [backlink](#)

Unknown interpreted text role "doc".

Use the `:ref:`coroutine syntax <async>``.

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs]index.rst, line 214); [backlink](#)

Unknown interpreted text role "ref".

`:doc:`topics/asyncio``

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs]index.rst, line 217); [backlink](#)

Unknown interpreted text role "doc".

Use `:mod:`asyncio`` and `:mod:`asyncio``-powered libraries.

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs]index.rst, line 217); [backlink](#)

Unknown interpreted text role "mod".

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs]index.rst, line 217); [backlink](#)

Unknown interpreted text role "mod".

Extending Scrapy

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs]index.rst, line 224)

Unknown directive type "toctree".

```
.. toctree::
   :caption: Extending Scrapy
   :hidden:

   topics/architecture
   topics/downloader-middleware
   topics/spider-middleware
   topics/extensions
   topics/api
   topics/signals
   topics/scheduler
   topics/exporters
```

:doc:`topics/architecture`

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs] index.rst, line 239); [backlink](#)

Unknown interpreted text role "doc".

Understand the Scrapy architecture.

:doc:`topics/downloader-middleware`

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs] index.rst, line 242); [backlink](#)

Unknown interpreted text role "doc".

Customize how pages get requested and downloaded.

:doc:`topics/spider-middleware`

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs] index.rst, line 245); [backlink](#)

Unknown interpreted text role "doc".

Customize the input and output of your spiders.

:doc:`topics/extensions`

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs] index.rst, line 248); [backlink](#)

Unknown interpreted text role "doc".

Extend Scrapy with your custom functionality

:doc:`topics/api`

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs] index.rst, line 251); [backlink](#)

Unknown interpreted text role "doc".

Use it on extensions and middlewares to extend Scrapy functionality

:doc:`topics/signals`

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs] index.rst, line 254); [backlink](#)

Unknown interpreted text role "doc".

See all available signals and how to work with them.

:doc:`topics/scheduler`

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs] index.rst, line 257); [backlink](#)

Unknown interpreted text role "doc".

Understand the scheduler component.

:doc:`topics/exporters`

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs] index.rst, line 261); [backlink](#)

Unknown interpreted text role "doc".

Quickly export your scraped items to a file (XML, CSV, etc).

All the rest

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs]index.rst, line 266)

Unknown directive type "toctree".

```
.. toctree::
   :caption: All the rest
   :hidden:

   news
   contributing
   versioning
```

:doc:`news`

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs]index.rst, line 275); [backlink](#)

Unknown interpreted text role "doc".

See what has changed in recent Scrapy versions.

:doc:`contributing`

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs]index.rst, line 278); [backlink](#)

Unknown interpreted text role "doc".

Learn how to contribute to the Scrapy project.

:doc:`versioning`

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs]index.rst, line 280); [backlink](#)

Unknown interpreted text role "doc".

Understand Scrapy versioning and API stability.

Docutils System Messages

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\[scrapy-master] [docs]index.rst, line 5); [backlink](#)

Undefined substitution referenced: "version".