# Python 资源大全

资源非常多，第一次打开可能会卡顿，推荐用电脑打开，保存下来方便以后查阅

🤫点击「官网」可以打开链接哦

### 包管理

管理包和依赖的工具。

* pip：Python 包和依赖关系管理工具。[官网](https://pip.pypa.io/)
* pip-tools：保证 Python 包依赖关系更新的一组工具。[官网](https://github.com/nvie/pip-tools)
* pipenv：Pyhton 官方推荐的新一代包管理工具。[官网](https://github.com/pypa/pipenv)
* conda：跨平台，Python 二进制包管理工具。[官网](https://github.com/conda/conda/)
* Curdling：管理 Python 包的命令行工具。[官网](http://clarete.li/curdling/)
* wheel：Python 分发的新标准，意在取代 eggs。[官网](http://pythonwheels.com/)

### 交互式解析器

交互式 Python 解析器。

* IPython：功能丰富的工具，非常有效的使用交互式 Python。[官网](https://github.com/ipython/ipython)
* [bpython](http://hao.jobbole.com/bpython/)：界面丰富的 Python 解析器。[官网](http://bpython-interpreter.org/)
* ptpython：高级交互式 Python 解析器， 构建于 [python-prompt-toolkit](https://github.com/jonathanslenders/python-prompt-toolkit) 之上。[官网](https://github.com/jonathanslenders/ptpython)

### 日期和时间

操作日期和时间的类库。

* arrow：更好的 Python 日期时间操作类库。[官网](https://github.com/crsmithdev/arrow)
* Chronyk：Python 3 的类库，用于解析手写格式的时间和日期。[官网](https://github.com/KoffeinFlummi/Chronyk)
* dateutil：Python datetime 模块的扩展。[官网](https://pypi.python.org/pypi/python-dateutil)
* delorean：解决 Python 中有关日期处理的棘手问题的库。[官网](https://github.com/myusuf3/delorean/)
* maya：人性化的时间处理库。[官网](https://github.com/kennethreitz/maya)
* moment：一个用来处理时间和日期的 Python 库。灵感来自于 Moment.js。[官网](https://github.com/zachwill/moment)
* pendulum：一个比 arrow 更具有明确的，可预测的行为的时间操作库。[官网](https://github.com/sdispater/pendulum)
* PyTime：一个简单易用的 Python 模块，用于通过字符串来操作日期/时间。[官网](https://github.com/shinux/PyTime)
* pytz：现代以及历史版本的世界时区定义。将时区数据库引入 Python。[官网](https://launchpad.net/pytz)
* when.py：提供用户友好的函数来帮助用户进行常用的日期和时间操作。[官网](https://github.com/dirn/When.py)

### 文本处理

用于解析和操作文本的库。

* 通用
  + [chardet](http://hao.jobbole.com/chardet/)：字符编码检测器，兼容 Python2 和 Python3。[官网](https://github.com/chardet/chardet)
  + difflib：(Python 标准库)帮助我们进行差异化比较。[官网](https://docs.python.org/2/library/difflib.html)
  + ftfy：让 Unicode 文本更完整更连贯。[官网](https://github.com/LuminosoInsight/python-ftfy)
  + fuzzywuzzy：模糊字符串匹配。[官网](https://github.com/seatgeek/fuzzywuzzy)
  + Levenshtein：快速计算编辑距离以及字符串的相似度。[官网](https://github.com/ztane/python-Levenshtein/)
  + pangu.py：在中日韩语字符和数字字母之间添加空格。[官网](https://github.com/vinta/pangu.py)
  + pypinyin：汉字拼音转换工具 Python 版。[官网](https://github.com/mozillazg/python-pinyin)
  + shortuuid：一个生成器库，用以生成简洁的，明白的，URL 安全的 UUID。[官网](https://github.com/stochastic-technologies/shortuuid)
  + [simplejson](https://github.com/simplejson/simplejson)：Python 的 JSON 编码、解码器。[官网](https://simplejson.readthedocs.io/en/latest/)
  + unidecode：Unicode 文本的 ASCII 转换形式 。[官网](https://pypi.python.org/pypi/Unidecode)
  + uniout：打印可读的字符，而不是转义的字符串。[官网](https://github.com/moskytw/uniout)
  + xpinyin：一个用于把汉字转换为拼音的库。[官网](https://github.com/lxneng/xpinyin)
  + yfiglet-figlet：[pyfiglet -figlet](https://github.com/pwaller/pyfiglet) 的 Python 实现。
  + flashtext: 一个高效的文本查找替换库。[官网](https://github.com/vi3k6i5/flashtext)
* Slug 化
  + awesome-slugify：一个 Python slug 化库，可以保持 Unicode。[官网](https://github.com/dimka665/awesome-slugify)
  + python-slugify：Python slug 化库，可以把 unicode 转化为 ASCII。[官网](https://github.com/un33k/python-slugify)
  + unicode-slugify：一个 slug 工具，可以生成 unicode slugs ,需要依赖 Django 。[官网](https://github.com/mozilla/unicode-slugify)
* 解析器
  + phonenumbers：解析，格式化，储存，验证电话号码。[官网](https://github.com/daviddrysdale/python-phonenumbers)
  + PLY：lex 和 yacc 解析工具的 Python 实现。[官网](http://www.dabeaz.com/ply/)
  + Pygments：通用语法高亮工具。[官网](http://pygments.org/)
  + pyparsing：生成通用解析器的框架。[官网](http://pyparsing.wikispaces.com/)
  + python-nameparser：把一个人名分解为几个独立的部分。[官网](https://github.com/derek73/python-nameparser)
  + python-user-agents：浏览器 user agent 解析器。[官网](https://github.com/selwin/python-user-agents)
  + sqlparse：一个无验证的 SQL 解析器。[官网](https://sqlparse.readthedocs.org/en/latest/)

### 特殊文本格式处理

一些用来解析和操作特殊文本格式的库。

* 通用
  + tablib：一个用来处理中表格数据的模块。[官网](https://github.com/kennethreitz/tablib)
* Office
  + Marmir：把输入的 Python 数据结构转换为电子表单。[官网](https://github.com/brianray/mm)
  + openpyxl：一个用来读写 Excel 2010 xlsx/xlsm/xltx/xltm 文件的库。[官网](https://openpyxl.readthedocs.org/en/latest/)
  + pyexcel：一个提供统一 API，用来读写，操作 Excel 文件的库。[官网](https://github.com/pyexcel/pyexcel)
  + python-docx：读取，查询以及修改 Microsoft Word 2007/2008 docx 文件。[官网](https://github.com/python-openxml/python-docx)
  + relatorio：模板化 OpenDocument 文件。[官网](http://relatorio.tryton.org/)
  + unoconv：在 LibreOffice/OpenOffice 支持的任意文件格式之间进行转换。[官网](https://github.com/dagwieers/unoconv)
  + XlsxWriter：一个用于创建 Excel .xlsx 文件的 Python 模块。[官网](https://xlsxwriter.readthedocs.org/en/latest/)
  + xlwings：一个使得在 Excel 中方便调用 Python 的库（反之亦然），基于 BSD 协议。[官网](http://xlwings.org/)
  + [xlwt](http://hao.jobbole.com/xlwt/)：读写 Excel 文件的数据和格式信息。[官网](https://github.com/python-excel/xlwt) / [xlrd](https://github.com/python-excel/xlrd)
* PDF
  + PDFMiner：一个用于从 PDF 文档中抽取信息的工具。[官网](https://github.com/euske/pdfminer)
  + PyPDF2：一个可以分割，合并和转换 PDF 页面的库。[官网](https://github.com/mstamy2/PyPDF2)
  + ReportLab：快速创建富文本 PDF 文档。[官网](http://www.reportlab.com/opensource/)
* Markdown
  + Mistune：快速并且功能齐全的纯 Python 实现的 Markdown 解析器。[官网](https://github.com/lepture/mistune)
  + Python-Markdown：John Gruber’s Markdown 的 Python 版实现。[官网](https://github.com/waylan/Python-Markdown)
  + Python-Markdown2：纯 Python 实现的 Markdown 解析器，比 Python-Markdown 更快，更准确，可扩展。[官网](https://github.com/trentm/python-markdown2)
* YAML
  + PyYAML：Python 版本的 YAML 解析器。[官网](http://pyyaml.org/)
* CSV
  + csvkit：用于转换和操作 CSV 的工具。[官网](https://github.com/wireservice/csvkit)
* Archive
  + unp：一个用来方便解包归档文件的命令行工具。[官网](https://github.com/mitsuhiko/unp)

### 自然语言处理

用来处理人类语言的库。

* [NLTK](http://hao.jobbole.com/nltk/)：一个先进的平台，用以构建处理人类语言数据的 Python 程序。[官网](http://www.nltk.org/)
* jieba：中文分词工具。[官网](https://github.com/fxsjy/jieba)
* langid.py：独立的语言识别系统。[官网](https://github.com/saffsd/langid.py)
* Pattern：Python 网络信息挖掘模块。[官网](http://www.clips.ua.ac.be/pattern)
* SnowNLP：一个用来处理中文文本的库。[官网](https://github.com/isnowfy/snownlp)
* TextBlob：为进行普通自然语言处理任务提供一致的 API。[官网](http://textblob.readthedocs.org/en/latest/)
* TextGrocery：一简单高效的短文本分类工具，基于 LibLinear 和 Jieba。[官网](https://github.com/2shou/TextGrocery)
* thulac:清华大学自然语言处理与社会人文计算实验室研制推出的一套中文词法分析工具包[官网](https://github.com/thunlp/THULAC-Python)

### 文档

用以生成项目文档的库。

* [Sphinx](http://hao.jobbole.com/sphinx/)：Python 文档生成器。[官网](http://www.sphinx-doc.org/en/latest/)
* awesome-sphinxdoc：[官网](https://github.com/yoloseem/awesome-sphinxdoc)
* MkDocs：对 Markdown 友好的文档生成器。[官网](http://www.mkdocs.org/)
* pdoc：一个可以替换 Epydoc 的库，可以自动生成 Python 库的 API 文档。[官网](https://github.com/BurntSushi/pdoc)
* Pycco：文学编程（literate-programming）风格的文档生成器。[官网](https://github.com/pycco-docs/pycco)
* readthedocs：一个基于 Sphinx/MkDocs 的在线文档托管系统，对开源项目免费开放使用。[官网](https://github.com/rtfd/readthedocs.org/)

### 下载器

用来进行下载的库.

* s3cmd：一个用来管理 Amazon S3 和 CloudFront 的命令行工具。[官网](https://github.com/s3tools/s3cmd)
* s4cmd：超级 S3 命令行工具，性能更加强劲。[官网](https://github.com/bloomreach/s4cmd)
* you-get：一个 YouTube/Youku/Niconico 视频下载器，使用 Python3 编写。[官网](https://www.soimort.org/you-get/)
* youtube-dl：一个小巧的命令行程序，用来下载 YouTube 视频。[官网](http://rg3.github.io/youtube-dl/)

### 图像处理

用来操作图像的库.

* [pillow](http://hao.jobbole.com/pillow/)：Pillow 是一个更加易用版的 [PIL](http://www.pythonware.com/products/pil/)。[官网](http://pillow.readthedocs.org/en/latest/)
* hmap：图像直方图映射。[官网](https://github.com/rossgoodwin/hmap)
* imgSeek：一个使用视觉相似性搜索一组图片集合的项目。[官网](https://sourceforge.net/projects/imgseek/)
* nude.py：裸体检测。[官网](https://github.com/hhatto/nude.py)
* pyBarcode：不借助 PIL 库在 Python 程序中生成条形码。[官网](https://pythonhosted.org/pyBarcode/)
* pygram：类似 Instagram 的图像滤镜。[官网](https://github.com/ajkumar25/pygram)
* python-qrcode：一个纯 Python 实现的二维码生成器。[官网](https://github.com/lincolnloop/python-qrcode)
* Quads：基于四叉树的计算机艺术。[官网](https://github.com/fogleman/Quads)
* scikit-image：一个用于（科学）图像处理的 Python 库。[官网](http://scikit-image.org/)
* thumbor：一个小型图像服务，具有剪裁，尺寸重设和翻转功能。[官网](https://github.com/thumbor/thumbor)
* wand：[MagickWand](http://www.imagemagick.org/script/magick-wand.php)的 Python 绑定。MagickWand 是 ImageMagick 的 C API 。[官网](https://github.com/dahlia/wand)
* face\_recognition：简单易用的 python 人脸识别库。[官网](https://github.com/ageitgey/face_recognition)

### OCR

光学字符识别库。

* pyocr：Tesseract 和 Cuneiform 的一个封装(wrapper)。[官网](https://github.com/jflesch/pyocr)
* [pytesseract](http://hao.jobbole.com/pytesseract/)：[Google Tesseract OCR](https://github.com/tesseract-ocr) 的另一个封装(wrapper)。[官网](https://github.com/madmaze/pytesseract)
* python-tesseract：[Google Tesseract OCR](https://github.com/tesseract-ocr) 的一个包装类。

### 音频

用来操作音频的库

* audiolazy：Python 的数字信号处理包。[官网](https://github.com/danilobellini/audiolazy)
* audioread：交叉库 (GStreamer + Core Audio + MAD + FFmpeg) 音频解码。[官网](https://github.com/beetbox/audioread)
* beets：一个音乐库管理工具及 [MusicBrainz](https://musicbrainz.org/) 标签添加工具。[官网](http://beets.io/)
* dejavu：音频指纹提取和识别。[官网](https://github.com/worldveil/dejavu)
* [django-elastic-transcoder](http://hao.jobbole.com/django-elastic-transcoder/)：Django + [Amazon Elastic Transcoder](http://aws.amazon.com/elastictranscoder/)。[官网](https://github.com/StreetVoice/django-elastic-transcoder)
* eyeD3：一个用来操作音频文件的工具，具体来讲就是包含 ID3 元信息的 MP3 文件。[官网](http://eyed3.nicfit.net/)
* id3reader：一个用来读取 MP3 元数据的 Python 模块。[官网](http://nedbatchelder.com/code/modules/id3reader.py)
* m3u8：一个用来解析 m3u8 文件的模块。[官网](https://github.com/globocom/m3u8)
* mutagen：一个用来处理音频元数据的 Python 模块。[官网](https://bitbucket.org/lazka/mutagen)
* pydub：通过简单、简洁的高层接口来操作音频文件。[官网](https://github.com/jiaaro/pydub)
* pyechonest：[Echo Nest](http://developer.echonest.com/) API 的 Python 客户端。[官网](https://github.com/echonest/pyechonest)
* talkbox：一个用来处理演讲/信号的 Python 库。[官网](http://scikits.appspot.com/talkbox)
* TimeSide：开源 web 音频处理框架。[官网](https://github.com/Parisson/TimeSide)
* tinytag：一个用来读取 MP3, OGG, FLAC 以及 Wave 文件音乐元数据的库。[官网](https://github.com/devsnd/tinytag)
* mingus：一个高级音乐理论和曲谱包，支持 MIDI 文件和回放功能。[官网](http://bspaans.github.io/python-mingus/)

### Video

用来操作视频和 GIF 的库。

* moviepy：一个用来进行基于脚本的视频编辑模块，适用于多种格式，包括动图 GIFs。[官网](http://zulko.github.io/moviepy/)
* scikit-video：SciPy 视频处理常用程序。[官网](https://github.com/aizvorski/scikit-video)

### 地理位置

地理编码地址以及用来处理经纬度的库。

* GeoDjango：世界级地理图形 web 框架。[官网](https://docs.djangoproject.com/en/dev/ref/contrib/gis/)
* GeoIP：MaxMind GeoIP Legacy 数据库的 Python API。[官网](https://github.com/maxmind/geoip-api-python)
* geojson：GeoJSON 的 Python 绑定及工具。[官网](https://github.com/frewsxcv/python-geojson)
* geopy：Python 地址编码工具箱。[官网](https://github.com/geopy/geopy)
* pygeoip：纯 Python GeoIP API。[官网](https://github.com/appliedsec/pygeoip)
* django-countries：一个 Django 应用程序，提供用于表格的国家选择功能，国旗图标静态文件以及模型中的国家字段。[官网](https://github.com/SmileyChris/django-countries)

### 数据库

Python 实现的数据库。

* pickleDB：一个简单，轻量级键值储存数据库。[官网](https://pythonhosted.org/pickleDB/)
* PipelineDB：流式 SQL 数据库。[官网](https://www.pipelinedb.com/)
* TinyDB：一个微型的，面向文档型数据库。[官网](https://github.com/msiemens/tinydb)
* ZODB：一个 Python 原生对象数据库。一个键值和对象图数据库。[官网](http://www.zodb.org/en/latest/)

### 数据库驱动

用来连接和操作数据库的库。

* MySQL：
  + [awesome-mysql](http://shlomi-noach.github.io/awesome-mysql/) 系列aiomysql：基于 asyncio 的异步 MySQL 数据库操作库。[官网](https://github.com/aio-libs/aiomysql)
  + mysql-python：Python 的 MySQL 数据库连接器。[官网](https://sourceforge.net/projects/mysql-python/)
  + ysqlclient：[mysql-python](https://github.com/PyMySQL/mysqlclient-python) 分支，支持 Python 3。
  + oursql：一个更好的 MySQL 连接器，支持原生预编译指令和 BLOBs。[官网](https://pythonhosted.org/oursql/)
  + PyMySQL：纯 Python MySQL 驱动，兼容 mysql-python。[官网](https://github.com/PyMySQL/PyMySQL)
* PostgreSQL
  + psycopg2：Python 中最流行的 PostgreSQL 适配器。[官网](http://initd.org/psycopg/)
  + queries：psycopg2 库的封装，用来和 PostgreSQL 进行交互。[官网](https://github.com/gmr/queries)
  + txpostgres：基于 Twisted 的异步 PostgreSQL 驱动。[官网](http://txpostgres.readthedocs.org/en/latest/)
* 其他关系型数据库
  + apsw：另一个 Python SQLite 封装。[官网](http://rogerbinns.github.io/apsw/)
  + dataset：在数据库中存储 Python 字典
  + pymssql：一个简单的 Microsoft SQL Server 数据库接口。[官网](http://www.pymssql.org/en/latest/)
* NoSQL 数据库
  + asyncio-redis：基于 asyncio 的 redis 客户端 (PEP 3156)。[官网](https://github.com/jonathanslenders/asyncio-redis)
  + cassandra-python-driver：Cassandra 的 Python 驱动。[官网](https://github.com/datastax/python-driver)
  + HappyBase：一个为 Apache HBase 设计的，对开发者友好的库。[官网](http://happybase.readthedocs.org/en/latest/)
  + Plyvel：一个快速且功能丰富的 LevelDB 的 Python 接口。[官网](https://plyvel.readthedocs.org/en/latest/)
  + py2neo：Neo4j restful 接口的 Python 封装客户端。[官网](http://py2neo.org/2.0/)
  + pycassa：Cassandra 的 Python Thrift 驱动。[官网](https://github.com/pycassa/pycassa)
  + PyMongo：MongoDB 的官方 Python 客户端。[官网](https://docs.mongodb.org/ecosystem/drivers/python/)
  + redis-py：Redis 的 Python 客户端。[官网](https://github.com/andymccurdy/redis-py)
  + telephus：基于 Twisted 的 Cassandra 客户端。[官网](https://github.com/driftx/Telephus)
  + txRedis：基于 Twisted 的 Redis 客户端。[官网](https://github.com/deldotdr/txRedis)

### CMS

内容管理系统

* odoo-cms: 一个开源的，企业级 CMS，基于 odoo。[官网](http://www.odoo.com/)
* django-cms：一个开源的，企业级 CMS，基于 Django。[官网](http://www.django-cms.org/en/)
* djedi-cms：一个轻量级但却非常强大的 Django CMS ，考虑到了插件，内联编辑以及性能。[官网](http://djedi-cms.org/)
* FeinCMS：基于 Django 构建的最先进的内容管理系统之一。[官网](http://www.feincms.org/)
* Kotti：一个高级的，Python 范的 web 应用框架，基于 Pyramid 构建。[官网](http://kotti.pylonsproject.org/)
* Mezzanine：一个强大的，持续的，灵活的内容管理平台。[官网](http://mezzanine.jupo.org/)
* Opps：一个为杂志，报纸网站以及大流量门户网站设计的 CMS 平台，基于 Django。[官网](http://opps.github.io/opps/)
* Plone：一个构建于开源应用服务器 Zope 之上的 CMS。[官网](https://plone.org/)
* Quokka：灵活，可扩展的小型 CMS，基于 Flask 和 MongoDB。[官网](http://quokkaproject.org/)
* [Wagtail](http://hao.jobbole.com/wagtail/)：一个 Django 内容管理系统。[官网](https://wagtail.io/)
* Widgy：最新的 CMS 框架，基于 Django。[官网](https://wid.gy/)

### 队列

处理事件以及任务队列的库。

* celery：一个异步任务队列/作业队列，基于分布式消息传递。[官网](http://www.celeryproject.org/)
* huey：小型多线程任务队列。[官网](https://github.com/coleifer/huey)
* [mrq](http://hao.jobbole.com/mrq/)：Mr. Queue -一个 Python 的分布式 worker 任务队列， 使用 Redis 和 gevent。[官网](https://github.com/pricingassistant/mrq)
* rq：简单的 Python 作业队列。[官网](http://python-rq.org/)
* simpleq：一个简单的，可无限扩张的，基于亚马逊 SQS 的队列。[官网](https://github.com/rdegges/simpleq)

### 搜索

对数据进行索引和执行搜索查询的库和软件。

* django-haystack：Django 模块化搜索。[官网](https://github.com/django-haystack/django-haystack)
* elasticsearch-py：Elasticsearch 的官方底层 Python 客户端。[官网](https://www.elastic.co/guide/en/elasticsearch/client/python-api/current/index.html)
* elasticsearch-dsl-py：Elasticsearch 的官方高级 Python 客户端。[官网](https://github.com/elastic/elasticsearch-dsl-py)
* solrpy：[solr](http://lucene.apache.org/solr/) 的 Python 客户端。[官网](https://github.com/edsu/solrpy)
* Whoosh：一个快速的纯 Python 搜索引擎库。[官网](http://whoosh.readthedocs.org/en/latest/)

### HTML 处理

处理 HTML 和 XML 的库。

* BeautifulSoup：以 Python 风格的方式来对 HTML 或 XML 进行迭代，搜索和修改。[官网](http://www.crummy.com/software/BeautifulSoup/bs4/doc/)
* bleach：一个基于白名单的 HTML 清理和文本链接库。[官网](http://bleach.readthedocs.org/en/latest/)
* cssutils：一个 Python 的 CSS 库。[官网](https://pypi.python.org/pypi/cssutils/)
* html5lib：一个兼容标准的 HTML 文档和片段解析及序列化库。[官网](https://github.com/html5lib/html5lib-python)
* lxml：一个非常快速，简单易用，功能齐全的库，用来处理 HTML 和 XML。[官网](http://lxml.de/)
* MarkupSafe：为 Python 实现 XML/HTML/XHTML 标记安全字符串。[官网](https://github.com/pallets/markupsafe)
* pyquery：一个解析 HTML 的库，类似 jQuery。[官网](https://github.com/gawel/pyquery)
* requests-html：人性化的，Pythonic 的 HTML 解析库。[官网](https://github.com/kennethreitz/requests-html)
* untangle：将 XML 文档转换为 Python 对象，使其可以方便的访问。[官网](https://github.com/stchris/untangle)
* xhtml2pdf：HTML/CSS 转 PDF 工具。[官网](https://github.com/xhtml2pdf/xhtml2pdf)
* xmltodict：像处理 JSON 一样处理 XML。[官网](https://github.com/martinblech/xmltodict)

爬取网络站点的库

* Scrapy：一个快速高级的屏幕爬取及网页采集框架。[官网](http://scrapy.org/)
* cola：一个分布式爬虫框架。[官网](https://github.com/chineking/cola)
* Demiurge：基于 PyQuery 的爬虫微型框架。[官网](https://github.com/matiasb/demiurge)
* feedparser：通用 feed 解析器。[官网](http://pythonhosted.org/feedparser/)
* Grab：站点爬取框架。[官网](http://grablib.org/)
* MechanicalSoup：用于自动和网络站点交互的 Python 库。[官网](https://github.com/hickford/MechanicalSoup)
* portia：Scrapy 可视化爬取。[官网](https://github.com/scrapinghub/portia)
* pyspider：一个强大的爬虫系统。[官网](https://github.com/binux/pyspider)
* RoboBrowser：一个简单的，Python 风格的库，用来浏览网站，而不需要一个独立安装的浏览器。[官网](https://github.com/jmcarp/robobrowser)

### 网页内容提取

用于进行网页内容提取的库。

* Haul：一个可以扩展的图像爬取工具。[官网](https://github.com/vinta/Haul)
* html2text：将 HTML 转换为 Markdown 格式文本。[官网](https://github.com/Alir3z4/html2text)
* lassie：人性化的网页内容检索库。[官网](https://github.com/michaelhelmick/lassie)
* micawber：一个小型网页内容提取库，用来从 URLs 提取富内容。[官网](https://github.com/coleifer/micawber)
* [newspaper](http://hao.jobbole.com/python-newspaper/)：使用 Python 进行新闻提取，文章提取以及内容策展。[官网](https://github.com/codelucas/newspaper)
* opengraph：一个用来解析开放内容协议(Open Graph Protocol)的 Python 模块。[官网](https://github.com/erikriver/opengraph)
* [python-goose](http://hao.jobbole.com/python-goose/)：HTML 内容/文章提取器。[官网](https://github.com/grangier/python-goose)
* python-readability：arc90 公司 readability 工具的 Python 高速端口。[官网](https://github.com/buriy/python-readability)
* sanitize：为杂乱的数据世界带来调理性。[官网](https://github.com/Alir3z4/python-sanitize)
* sumy：一个为文本文件和 HTML 页面进行自动摘要的模块。[官网](https://github.com/miso-belica/sumy)
* textract：从任何格式的文档中提取文本，Word，PowerPoint，PDFs 等等。[官网](https://github.com/deanmalmgren/textract)

### 并发和并行

用以进行并发和并行操作的库。

* multiprocessing：(Python 标准库) 基于进程的“线程”接口。[官网](https://docs.python.org/2/library/multiprocessing.html)
* threading：(Python 标准库)更高层的线程接口。[官网](https://docs.python.org/2/library/threading.html)
* eventlet：支持 WSGI 的异步框架。[官网](http://eventlet.net/)
* gevent：一个基于协程的 Python 网络库，使用 [greenlet](https://github.com/python-greenlet/greenlet)。[官网](http://www.gevent.org/)
* Tomorrow：用于产生异步代码的神奇的装饰器语法实现。[官网](https://github.com/madisonmay/Tomorrow)
* uvloop：在 libuv 之上超快速实现 asyncio 事件循环。[官网](https://github.com/MagicStack/uvloop)

### 测试

进行代码库测试和生成测试数据的库。

* 测试框架
  + unittest：(Python 标准库) 单元测试框架。[官网](https://docs.python.org/2/library/unittest.html)
  + nose：nose 扩展了 unittest 的功能。[官网](https://nose.readthedocs.org/en/latest/)
  + contexts：一个 Python 3.3+ 的 BDD 框架。受到 C# – Machine.Specifications 的启发。[官网](https://github.com/benjamin-hodgson/Contexts)
  + hypothesis：Hypothesis 是一个基于先进的 Quickcheck 风格特性的测试库。[官网](https://github.com/DRMacIver/hypothesis)
  + mamba：Python 的终极测试工具， 拥护 BDD。[官网](http://nestorsalceda.github.io/mamba/)
  + PyAutoGUI：PyAutoGUI 是一个人性化的跨平台 GUI 自动测试模块。[官网](https://github.com/asweigart/pyautogui)
  + pyshould：Should 风格的断言，基于 [PyHamcrest](https://github.com/hamcrest/PyHamcrest)。[官网](https://github.com/drslump/pyshould)
  + pytest：一个成熟的全功能 Python 测试工具。[官网](http://pytest.org/latest/)
  + green：干净，多彩的测试工具。[官网](https://github.com/CleanCut/green)
  + pyvows：BDD 风格的测试工具，受 Vows.js 的启发。[官网](http://heynemann.github.io/pyvows/)
  + Robot Framework：一个通用的自动化测试框架。[官网](https://github.com/robotframework/robotframework)
* Web 测试
  + Selenium：[Selenium](http://www.seleniumhq.org/) WebDriver 的 Python 绑定。[官网](https://pypi.python.org/pypi/selenium)
  + locust：使用 Python 编写的，可扩展的用户加载测试工具。[官网](https://github.com/locustio/locust)
  + sixpack：一个和语言无关的 A/B 测试框架。[官网](https://github.com/seatgeek/sixpack)
  + splinter：开源的 web 应用测试工具。[官网](https://splinter.readthedocs.org/en/latest/)
* Mock 测试
  + mock：(Python 标准库) 一个用于伪造测试的库。[官网](https://docs.python.org/3/library/unittest.mock.html)
  + doublex：Python 的一个功能强大的 doubles  测试框架。[官网](https://pypi.python.org/pypi/doublex)
  + freezegun：通过伪造日期模块来生成不同的时间。[官网](https://github.com/spulec/freezegun)
  + httmock：针对 Python 2.6+ 和 3.2+ 生成 伪造请求的库。[官网](https://github.com/patrys/httmock)
  + httpretty：Python 的 HTTP 请求 mock 工具。[官网](http://falcao.it/HTTPretty/)
  + responses：伪造 Python 中的 requests 库的一个通用库。[官网](https://github.com/getsentry/responses)
  + VCR.py：在你的测试中记录和重放 HTTP 交互。[官网](https://github.com/kevin1024/vcrpy)
* 对象工厂
  + factoryboy：一个 Python 用的测试固件 (test fixtures) 替代库。[官网](https://github.com/rbarrois/factoryboy)
  + mixer：另外一个测试固件 (test fixtures) 替代库，支持 Django, Flask, SQLAlchemy, Peewee 等。[官网](https://github.com/klen/mixer)
  + modelmommy：为 Django 测试创建随机固件。[官网](https://github.com/vandersonmota/modelmommy)
* 代码覆盖率
  + coverage：代码覆盖率测量。[官网](https://pypi.python.org/pypi/coverage)
  + Codecov：一个代码覆盖率测试工具，为开源项目提供免费代码覆盖率测试服务。[官网](https://codecov.io/)
* 伪数据
  + faker：一个 Python 库，用来生成伪数据。[官网](http://www.joke2k.net/faker/)
  + fake2db：伪数据库生成器。[官网](https://github.com/emirozer/fake2db)
  + radar：生成随机的日期/时间。[官网](https://pypi.python.org/pypi/radar)
* 错误处理
  + FuckIt.py：FuckIt.py 使用最先进的技术来保证你的 Python 代码无论对错都能继续运行。[官网](https://github.com/ajalt/fuckitpy)

### 调试工具

用来进行代码调试的库。

* 调试器
  + ipdb：IPython 启用的 [pdb](https://docs.python.org/2/library/pdb.html)。[官网](https://pypi.python.org/pypi/ipdb)
  + pudb：全屏，基于控制台的 Python 调试器。[官网](https://pypi.python.org/pypi/pudb)
  + pyringe：可以在 Python 进程中附加和注入代码的调试器。[官网](https://github.com/google/pyringe)
  + wdb：一个奇异的 web 调试器，通过 WebSockets 工作。[官网](https://github.com/Kozea/wdb)
  + winpdb：一个具有图形用户界面的 Python 调试器，可以进行远程调试，基于 rpdb2。[官网](http://winpdb.org/)
  + django-debug-toolbar：为 Django 显示各种调试信息。[官网](https://github.com/django-debug-toolbar/django-debug-toolbar)
  + django-devserver：一个 Django 运行服务器的替代品。[官网](https://github.com/dcramer/django-devserver)
  + flask-debugtoolbar：django-debug-toolbar 的 flask 版。[官网](https://github.com/mgood/flask-debugtoolbar)
* 性能分析器
  + lineprofiler：逐行性能分析。[官网](https://github.com/rkern/lineprofiler)
  + [Memory Profiler](http://hao.jobbole.com/memory_profiler/)：监控 Python 代码的内存使用。[官网](http://pypi.python.org/pypi/memory_profiler)、[内存](https://github.com/fabianp/memoryprofiler)
  + profiling：一个交互式 Python 性能分析工具。[官网](https://github.com/what-studio/profiling)
* 其他
  + pyelftools：解析和分析 ELF 文件以及 DWARF 调试信息。[官网](https://github.com/eliben/pyelftools)
  + python-statsd：[statsd](https://github.com/etsy/statsd/) 服务器的 Python 客户端。[官网](https://github.com/WoLpH/python-statsd)

### 数据可视化

进行数据可视化的库。 参见: [awesome-javascript](https://github.com/sorrycc/awesome-javascript#data-visualization)。

* matplotlib：一个 Python 2D 绘图库。[官网](http://matplotlib.org/)
* bokeh：用 Python 进行交互式 web 绘图。[官网](https://github.com/bokeh/bokeh)
* ggplot：ggplot2 给 R 提供的 API 的 Python 版本。[官网](https://github.com/yhat/ggplot)
* plotly：协同 Python 和 matplotlib 工作的 web 绘图库。[官网](https://plot.ly/python/)
* pyecharts：基于百度 Echarts 的数据可视化库。[官网](https://github.com/chenjiandongx/pyecharts)
* pygal：一个 Python SVG 图表创建工具。[官网](http://www.pygal.org/en/latest/)
* pygraphviz：Graphviz 的 Python 接口。[官网](https://pypi.python.org/pypi/pygraphviz)
* PyQtGraph：交互式实时 2D/3D/ 图像绘制及科学/工程学组件。[官网](http://www.pyqtgraph.org/)
* SnakeViz：一个基于浏览器的 Python's cProfile 模块输出结果查看工具。[官网](http://jiffyclub.github.io/snakeviz/)
* vincent：把 Python 转换为 Vega 语法的转换工具。[官网](https://github.com/wrobstory/vincent)
* VisPy：基于 OpenGL 的高性能科学可视化工具。[官网](http://vispy.org/)

### 编辑器插件

编辑器和 IDE 的插件

* Emacs
  + Elpy：Emacs Python 开发环境。[官网](https://github.com/jorgenschaefer/elpy)
* Sublime Text
  + SublimeJEDI：一个 Sublime Text 插件，用来使用超赞的自动补全库 Jedi。[官网](https://github.com/srusskih/SublimeJEDI)
  + Anaconda：Anaconda 把你的 Sublime Text 3 变成一个功能齐全的 Python IDE。[官网](https://github.com/DamnWidget/anaconda)
* Vim
  + [YouCompleteMe](http://hao.jobbole.com/youcompleteme/)：引入基于 [Jedi](https://github.com/davidhalter/jedi) 的 Python 自动补全引擎。[官网](https://github.com/Valloric/YouCompleteMe)
  + Jedi-vim：绑定 Vim 和 Jedi 自动补全库对 Python 进行自动补全。[官网](https://github.com/davidhalter/jedi-vim)
  + Python-mode：将 Vim 变成 Python IDE 的一款多合一插件。[官网](https://github.com/klen/python-mode)
* Visual Studio
  + PTVS：Visual Studio 的 Python 工具。[官网](https://github.com/Microsoft/PTVS)

### 集成开发环境

流行的 Python 集成开发环境。

* PyCharm：商业化的 Python IDE ，由 JetBrains 开发。也有免费的社区版提供。[官网](https://www.jetbrains.com/pycharm/)
* LiClipse：基于 Eclipse 的免费多语言 IDE 。使用 PyDev 来支持 Python 。[官网](http://www.liclipse.com/)
* Spyder：开源 Python IDE。[官网](https://github.com/spyder-ide/spyder)

### 自动聊天工具

用于开发聊天机器人的库

* Errbot：最简单和最流行的聊天机器人用来实现自动聊天工具。[官网](http://errbot.io/en/latest/)

### 金融数据

* Tushare ：一个可以提供免费股票、基金、期货、港股等金融数据的 Python 开源数据。[官网](http://tushare.org/)
* Ta-Lib ：金融数据技术分析库，可以依据原始金融数据计算各种技术指标,计算性能比较优异。[官网](https://www.ta-lib.org/)

## 资源

在这里可以找到新的 Python 库。

* [r/Python](https://www.reddit.com/r/python)
* [CoolGithubProjects](https://www.coolgithubprojects.com/)
* [Django Packages](https://www.djangopackages.com/)
* [Full Stack Python](http://www.fullstackpython.com/)
* [Python 3 Wall of Superpowers](http://python3wos.appspot.com/)
* [Python Hackers](http://pythonhackers.com/open-source/)
* [Python ZEEF](https://python.zeef.com/alan.richmond)
* [Trending Python repositories on GitHub today](https://github.com/trending?l=python)
* [PyPI Ranking](http://pypi-ranking.info/alltime)