Web 框架、网络爬虫、网络内容提取、模板引擎、数据库、数据可视化、图片处理、文本处理、自然语言处理、机器学习、日志、代码分析等

环境管理

管理 Python 版本和环境的工具

* p：非常简单的交互式 python 版本管理工具。[官网](https://github.com/qw3rtman/p)
* pyenv：简单的 Python 版本管理工具。[官网](https://github.com/yyuu/pyenv)
* Vex：可以在虚拟环境中执行命令。[官网](https://github.com/sashahart/vex)
* virtualenv：创建独立 Python 环境的工具。[官网](https://pypi.python.org/pypi/virtualenv)
* virtualenvwrapper：virtualenv 的一组扩展。[官网](https://pypi.python.org/pypi/virtualenvwrapper)

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/)

包仓库

本地 PyPI 仓库服务和代理。

* [warehouse](https://pypi.org/)：下一代 PyPI。[官网](https://github.com/pypa/warehouse)
* bandersnatch：PyPA 提供的 PyPI 镜像工具。[官网](https://bitbucket.org/pypa/bandersnatch)
* devpi：PyPI 服务和打包/测试/分发工具。[官网](http://doc.devpi.net/)
* localshop：本地 PyPI 服务（自定义包并且自动对 PyPI 镜像）。[官网](https://github.com/mvantellingen/localshop)

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)

网站

* [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)