

现代化软件开发与可持续的集成方法?



s p e c

A agile ecosystem for earthquake sciences

难度：一般

时长：30分钟

知识点

- 1、 DevOps开发理念vs瀑布式开发， 敏捷式开发工具；
- 2、 Python与Anaconda集成环境
- 3、 Github仓库的使用， 管理你的代码
- 4、 CI/CD的实现方法
- 5、 地学家软件包计划， 如何参与Geoist软件包开发

具备可支持交付能力的软件研发与部署

- 系统如何研发?

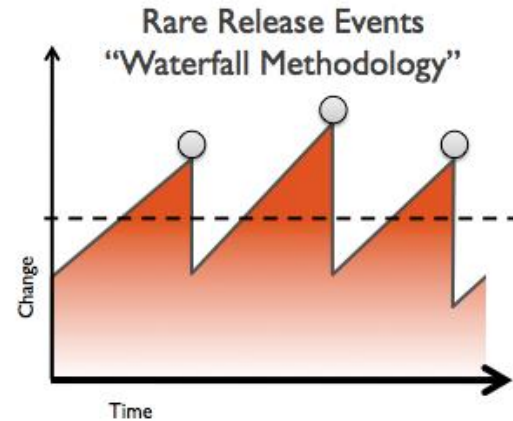
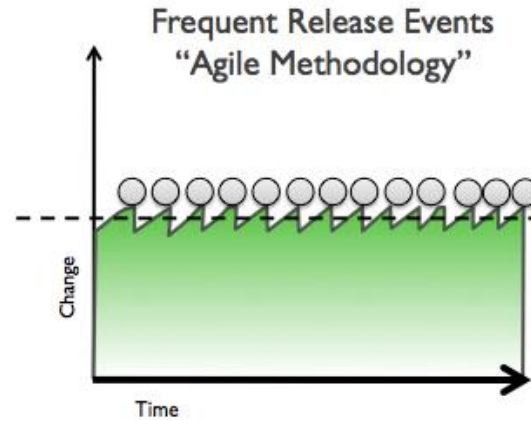
瀑布vs敏捷, 基于云部署

- 采用什么模式开发?

研究所牵头、业务单位参与

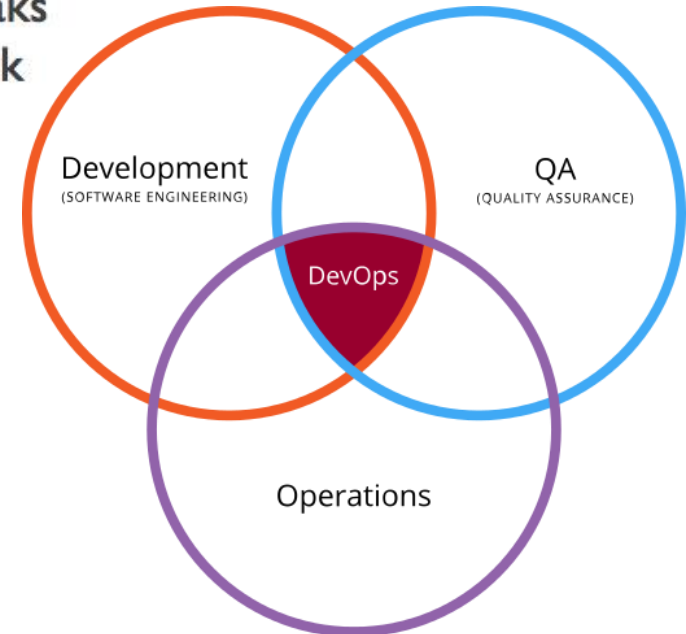
- 怎么让用户使用?

对接移动互联网、企业微信



s p e c

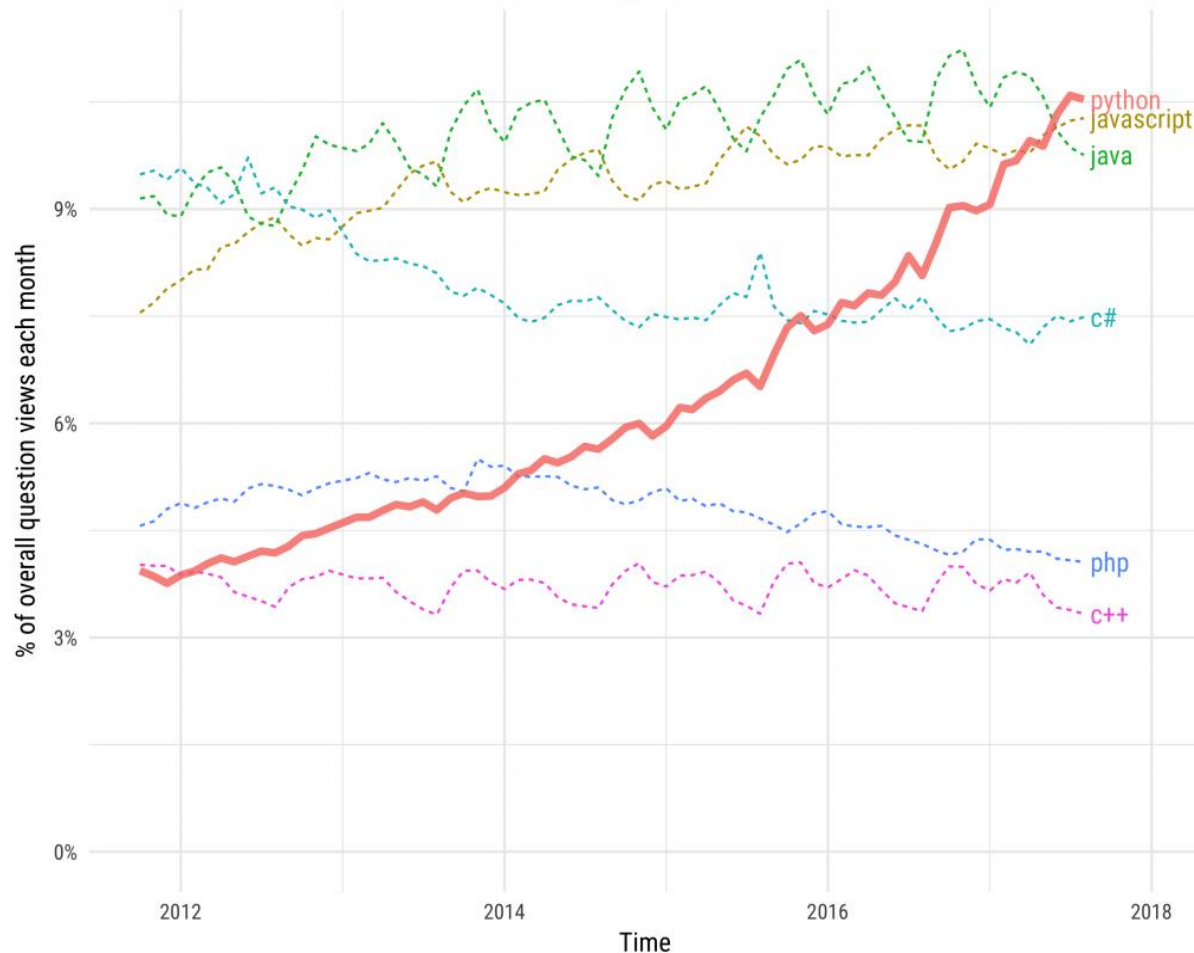
A agile ecosystem for earthquake sciences



越来越多的人使用Python

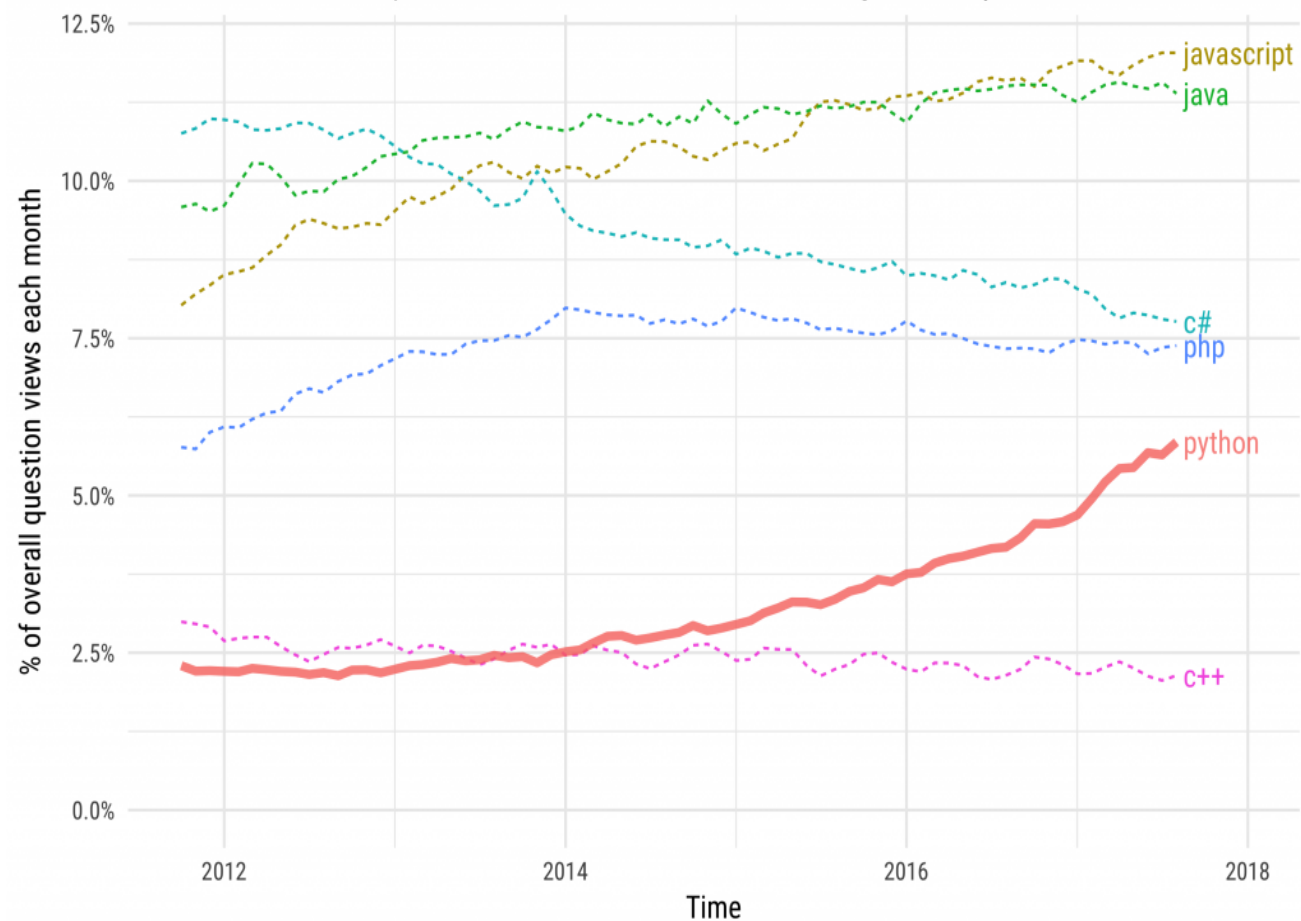
Growth of major programming languages

Based on Stack Overflow question views in World Bank high-income countries



Growth of major programming languages in non-high-income countries

Based on Stack Overflow question views in countries not classified as high-income by the World Bank.



Python的应用场景

- Web开发: Django, Tornado, Flask, ...
- 网络爬虫: Scrapy, Selenium, ...
- 自动化运维: Slatstack, Ansible, ...
- 科学计算: Numpy, Scipy, Matplotlib, ...
- ...
- 许多大型网站就是用Python开发的, 例如YouTube、Instagram, 还有国内的豆瓣。很多大公司, 包括Google、Yahoo等, 甚至NASA都大量地使用Python。

Python做科学计算的优势

- 高质量的科学计算库：
 - 数据分析处理：Numpy, Scipy, Pandas, Statsmodels, ...
 - 可视化：Matplotlib, Seaborn, Bokeh, Plotly, MayaVi2, ...
 - 机器学习：Theano, Tensorflow, SciKit-Learn, Keras, ...
 - 专业相关：ObsPy, Cartopy, ...
- 易学易用
 - 简洁的语法
 - 大量的优质教程
- 良好的生态
 - 广大的用户群
 - 不仅仅是数据处理

安装Python

- 版本选择： Python2? Python3?
 - 目前存在一些项目只支持Python2
 - Python2已经停止开发，未来属于Python3
- 不同发行版安装Python3
 - Ubuntu: `sudo apt-get install python3.6`
 - Archlinux: `sudo pacman -S python3`
 - Centos: 官方源里没有python3, 可以从源码安装 (参考网页 <https://segmentfault.com/a/1190000009922582>)

包管理工具

- pip的安装:

- Ubuntu: `sudo apt install python3-venv python3-pip`

- Archlinux: `sudo pacman -S python-pip`

- Centos参考:

- <https://packaging.python.org/guides/installing-using-linux-tools/>

- anaconda的安装:

- 1. 下载anaconda的安装程序(<https://www.anaconda.com/download/>)

- 2. 执行安装程序: `bash Anaconda-latest-Linux-x86_64.sh`

What is Anaconda?

- Anaconda is a Python distribution that is particularly popular for data analysis and scientific computing
- Open source project developed by Continuum Analytics, Inc.
- Available for Windows, Mac OS X and Linux
- Includes many popular packages: NumPy, SciPy, Matplotlib, Pandas, IPython, Cython Includes
- Spyder, a Python development environment
- Includes conda, a platform-independent package manager

anaconda的使用

- 管理环境：
 - 创建虚拟环境：conda create -n env_name list_of_packages
 - 进入环境：source activate env_name
 - 离开环境：source deactivate
 - 删除环境：conda env remove -n env_name
 - 列出可用环境：conda env list
- 管理包：
 - 安装包：conda install package_name
 - 卸载包：conda remove package_name
 - 帮助：conda help

Ipython

```
Python 3.6.5 (default, May 11 2018, 04:00:52)
Type 'copyright', 'credits' or 'license' for more information
IPython 6.3.1 -- An enhanced Interactive Python. Type '?' for help.

In [1]: import numpy as np

In [2]: np.random.rand(5,5)
Out[2]:
array([[0.13229674, 0.15126078, 0.16357823, 0.51419242, 0.93777158],
       [0.2575341 , 0.15083139, 0.16676434, 0.89683108, 0.71410577],
       [0.90203414, 0.04031882, 0.11572128, 0.99108052, 0.36997237],
       [0.87460682, 0.60470149, 0.11735649, 0.43360977, 0.95160008],
       [0.64384138, 0.8840173 , 0.39482442, 0.16866307, 0.92456145]])

In [3]: np.rand
        np.random
        np.rank
```

Spyder集成开发环境

D:/MyWorks/geoist - Spyder (Python 3.6)

File Edit Search Source Run Debug Consoles Projects Tools View Help

D:\MyWorks\geoist

Project explorer

- geoist
 - .dask
 - binder
 - conda_recipe
 - dataset
 - docs
 - examples
 - extensions
 - geoist
 - __pycache__
 - catqc
 - cattools
 - dataset
 - earthquake
 - gravity
 - gridder
 - inversion
 - others
 - pfm
 - raster
 - snoopy

Variable explorer

Name	Type	Size	Value
------	------	------	-------

Editor

```
py TsAnomaly.py TsFilter.py TsPSD.py TsResample.py gmtfig.py
1
2 import pandas as pd
3 import pygmt
4
5 fig = pygmt.Figure()
6
7
8 data=pd.read_table('.TAB',sep=",")
9 region = [
10     data['经度'].min() - 2,
11     data['经度'].max() + 2,
12     data['纬度'].min() - 2,
13     data['纬度'].max() + 2,
14 ]
15 fig.coast(shorelines=True, region=region, #[90, 130, 0, 40],
16           projection="M8i", frame=True)
17 fig.plot(x=data['经度'], y=data['纬度'],
18          sizes=0.005 * (2 ** data['Ms']),
19           color="red",
20           style="cc",
21           pen="black")
22 fig.savefig("d:\\test2.png")
23
```

History log

```
history.py
region
runfile('D:/MyProjects/2018-会商技术系统/其他/列装
培训教程/demo/s1e2/gmtfig.py', wdir='D:/
MyProjects/2018-会商技术系统/其他/列装培训教程/
demo/s1e2')
```

IPython console

```
Console 1/A
Python
Type "copyright", "credits" or "license" for more
information.

IPython -- An enhanced Interactive Python.
?      -> Introduction and overview of
IPython's features.
%quickref -> Quick reference.
help    -> Python's own help system.
object? -> Details about 'object', use
'object??' for extra details.
```

```
In [1]:
```

Python console IPython console

Permissions: RW End-of-lines: LF Encoding: UTF-8-GUESSED Line: 2 Column: 1 Memory: 49%

Notebook开发环境



```
window_size = 50
```

```
# parameters for saving data  
res_file = 'despiked.txt'
```

```
In [509]: # Load data  
data=pd.read_csv(pathlib.Path(orig_file),parse_dates=[[0,1]],header=None,delim_whitespace=True,index_col=[0],na_values=na_values)  
data.index.name = 'time'  
data.columns = ['origin_data']
```

```
In [552]: # despiked  
data['despiked'],data['flag'] = tsa.despiked_v2(data['origin_data'].interpolate(),th=thresh_hold)  
ax=data.plot(figsize=(15,12),y=data.columns[:-1])  
ax.set_xlabel('Date')  
ax.set_ylabel('Value')  
plt.grid()  
plt.legend()  
plt.title("The preliminary result by threshold={}".format(thresh_hold),loc='left')
```

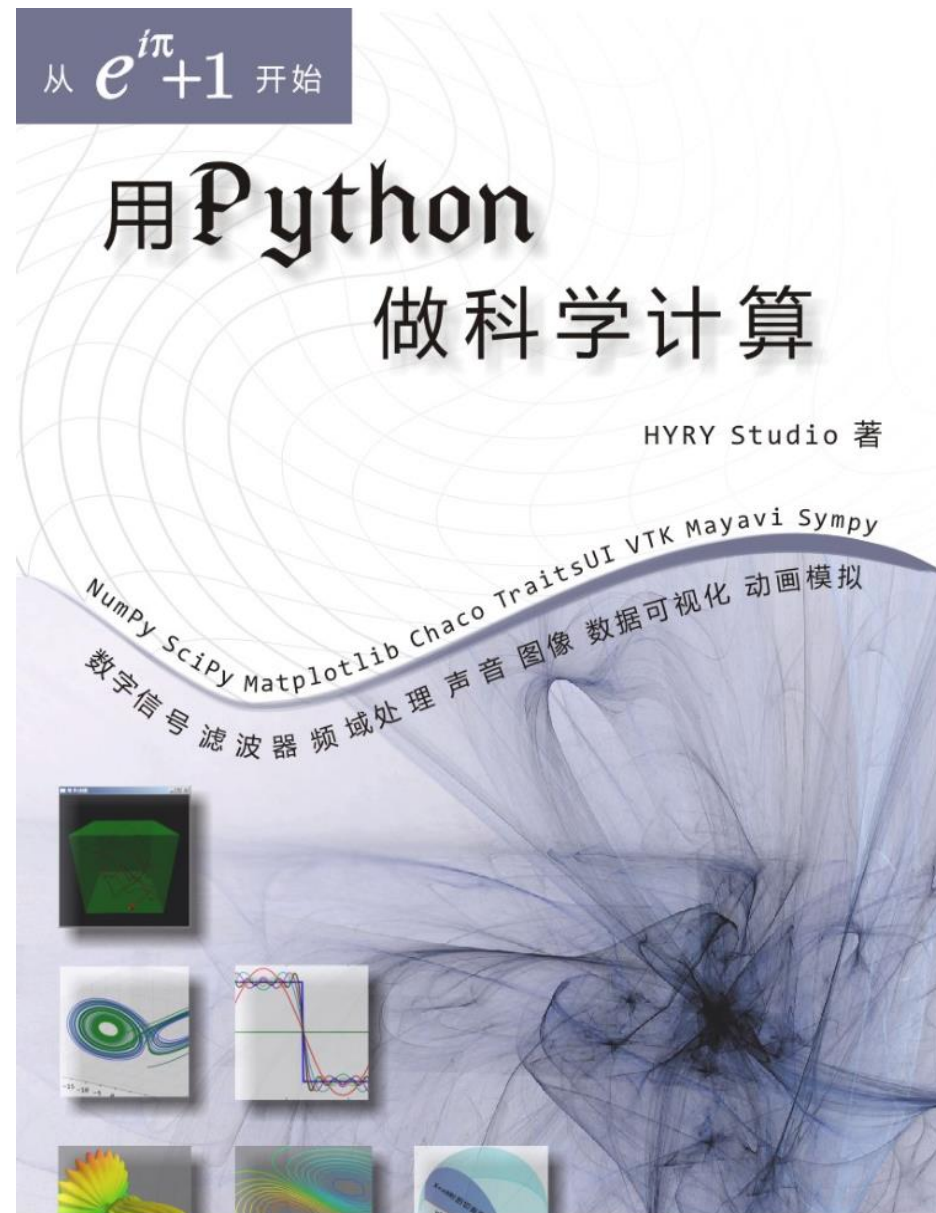
Out[552]: Text(0.0, 1.0, 'The preliminary result by threshold=200.0')

The preliminary result by threshold=200.0



Python入门书籍

- Learn Python in 10 minutes
(<https://www.stavros.io/tutorials/python>)
- 《A Byte of Python》
- 《用Python做科学计算》



numpy/scipy/matplotlib

- numpy: 提供ndarray对象以及基本的数学运算;
- scipy: 常用的数据处理算法, 如数值积分、插值、数字信号处理、线性代数、最优化等;
- matplotlib: 绘制2-d图形。
 - matplotlib gallery: <https://matplotlib.org/gallery.html>
- 一个例子...

pandas

- pandas is an open source, BSD-licensed library providing high-performance, easy-to-use **data structures and data analysis tools** for the Python programming language.
- data structure: 一维Series, 二维DataFrame
- analysis tools: 基本算术运算、数据统计信息、Category类型离散化、查改增删等
- 一个例子...



PyGMT

dev

Search docs

GETTING STARTED

Overview

Installing

Making your first figure

Gallery

USER GUIDE

Frames, ticks, titles, and labels

Projections

Coastlines and borders

Plotting data points

REFERENCE DOCUMENTATION

API Reference

GETTING HELP AND CONTRIBUTING

[Gitter chatroom.](#)

About

PyGMT is a library for processing geospatial and geophysical data and making publication quality maps and figures. It provides a Pythonic interface for the [Generic Mapping Tools \(GMT\)](#), a command-line program widely used in the Earth Sciences.

We rely heavily on new features currently being implemented in GMT. In particular, a new *modern execution mode* that greatly simplifies figure creation. **These features are not available in the 5.4 version of GMT.** They will be part of the future 6.0 release of GMT predicted for early 2019.

Project goals

- Make GMT more accessible to new users.
- Build a Pythonic API for GMT.
- Interface with the GMT C API directly using ctypes (no system calls).
- Support for rich display in the Jupyter notebook.
- Integration with the Scipy stack: numpy.ndarray or pandas.DataFrame for data tables and xarray.DataArray for grids.

obspy

- 文档: <http://docs.obspy.org/contents.html>
- obspy gellery: <http://docs.obspy.org/gallery.html>

- ObsPy Tutorial

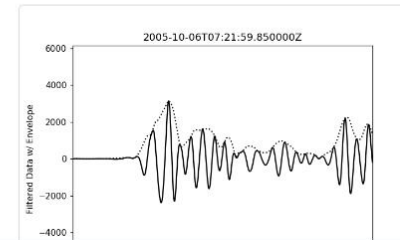
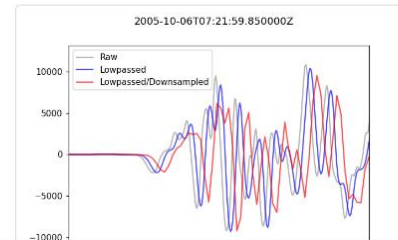
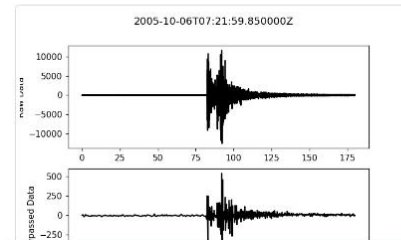
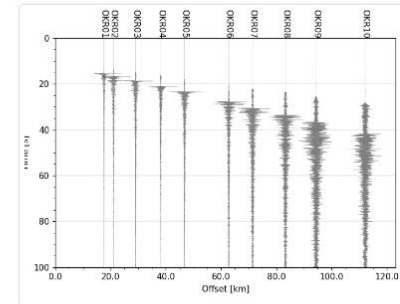
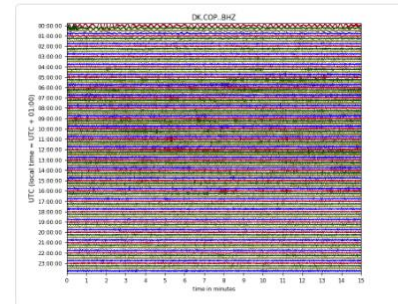
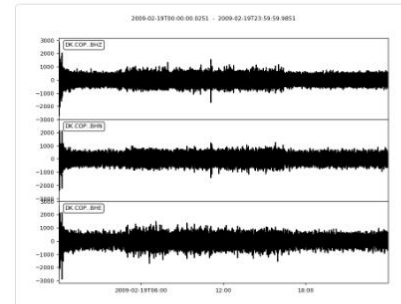
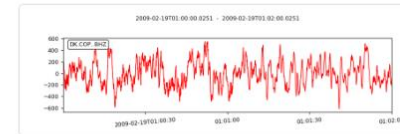
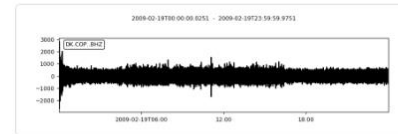
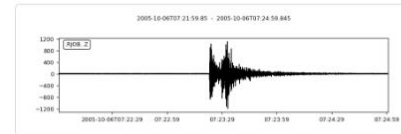
- Introduction to ObsPy

- 1. Python Introduction for Seismologists
 - 2. UTCDateTime
 - 3. Reading Seismograms
 - 4. Waveform Plotting Tutorial
 - 5. Retrieving Data from Data Centers
 - 6. Filtering Seismograms
 - 7. Downsampling Seismograms
 - 8. Merging Seismograms
 - 9. Beamforming - FK Analysis
 - 10. Seismogram Envelopes
 - 11. Plotting Spectrograms
 - 12. Trigger/Picker Tutorial
 - 13. Poles and Zeros, Frequency Response
 - 14. Seismometer Correction/Simulation
 - 15. Clone an Existing Dataless SEED File
 - 16. Export Seismograms to MATLAB
 - 17. Export Seismograms to ASCII
 - 18. Anything to MiniSEED
 - 19. Beachball Plot
 - 20. Basemap Plots
 - 21. Interfacing R from Python
 - 22. Coordinate Conversions
 - 23. Hierarchical Clustering
 - 24. Visualizing Probabilistic Power Spectral Densities
 - 25. Array Response Function
 - 26. Continuous Wavelet Transform
 - 27. Time Frequency Misfit
 - 28. Visualize Data Availability of Local Waveform Archive
 - 29. Travel Time and Ray Path Plotting
 - 30. Cross Correlation Pick Correction
 - 31. Handling custom defined tags in QuakeML and the ObsPy Catalog/Event framework
 - 32. Handling custom defined tags in StationXML with the ObsPy Inventory
 - 33. Creating a StationXML file from Scratch
 - 34. Connecting to a SeedLink Server

- Advanced Exercise

Gallery

Click on any image to see full size image and source code



Geoist项目

面向地球科学原型化研究的开源Python程序库

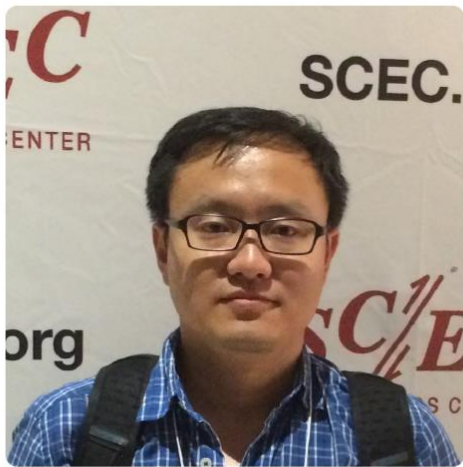


build passing codecov 15% docs passing launch binder Azure Pipelines succeeded license MIT

- [English Documentation](#)

主要功能

- 重磁位场数据分析与处理
- 地震重力平差和数据处理
- 地球物理反演方法
- 大地测量数据同化和建模
- 地震目录分析
- 时间序列分析和异常检测



Steve Chen

igp-gravity

对一切未知技术都有兴趣

Edit bio

CEA-IGP

Beijing, China

<https://igp-gravity.github.io/>

Overview **Repositories 9** Stars 7 Followers 5 Following 6

Find a repository...

Type: All

Language: All

New

igp-gravity.github.io

小G创新俱乐部, pages技术, Hexo建站, Next模板

HTML ★ 4 Updated 19 hours ago



test.github.io

我的个人写作至readthedocs自动发布

my

HTML ★ 3 Updated 6 days ago

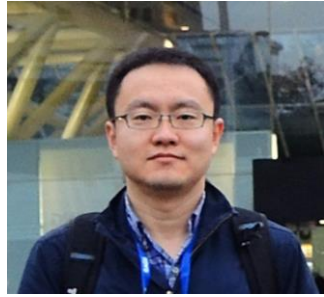


detrend_example

Python GNU General Public License v3.0 Updated on 10 Aug



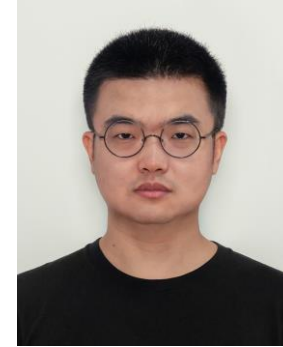
Team at IGP-CEA



Dr. Chen
Geodynamics and Geodesy



Dr. Li
Geodesy



Dr. Han
Geodesy

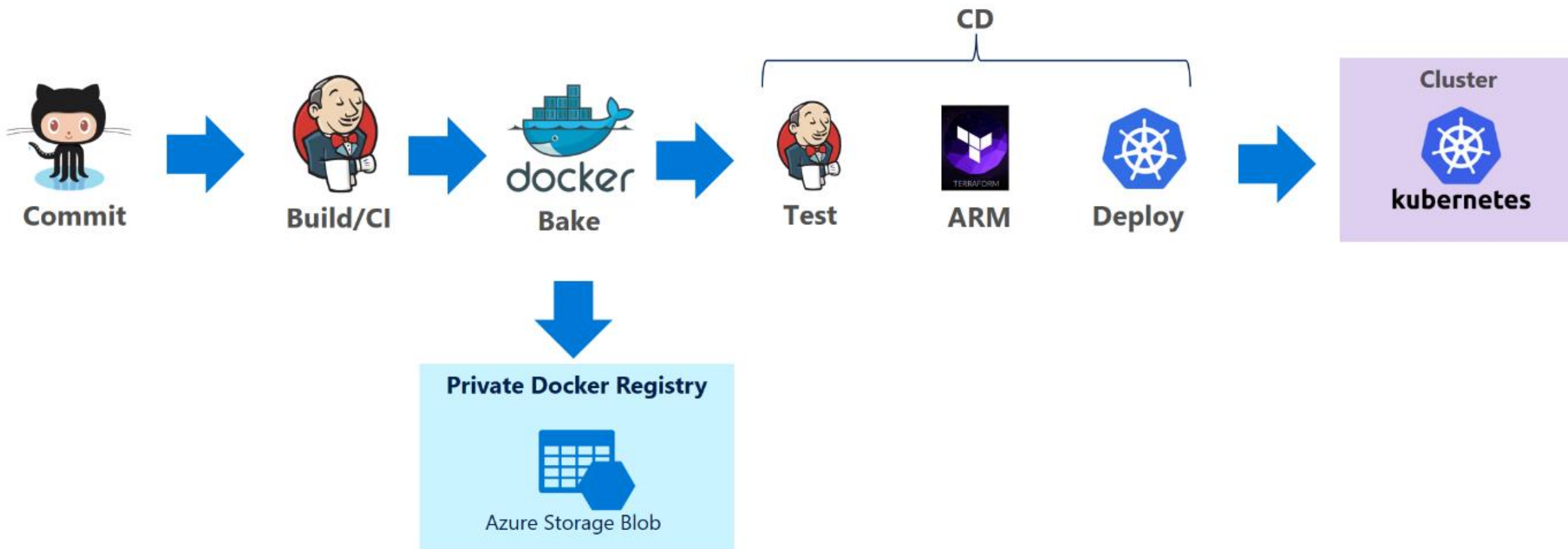


Dr. Zhao
Geophysical modelling



Dr. Zhang
Geodynamical modelling

现代化软件开发于部署流程





geoist

Overview

Boards

Repos

Pipelines

Builds

Releases

#20190112.1: add NOAA geomag model

All logs

Triggered 1月12日 at 下午9:29 for steveshichen igp-gravity/geoist master fd0d998

Logs Summary Tests

Test Python36 Succeeded

Publish Job Succeeded

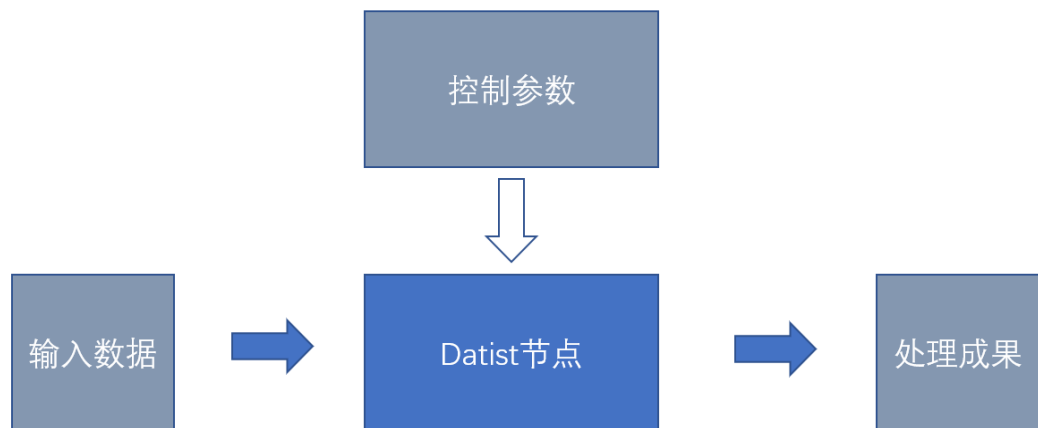
Test Python36

Agent: Hosted Agent

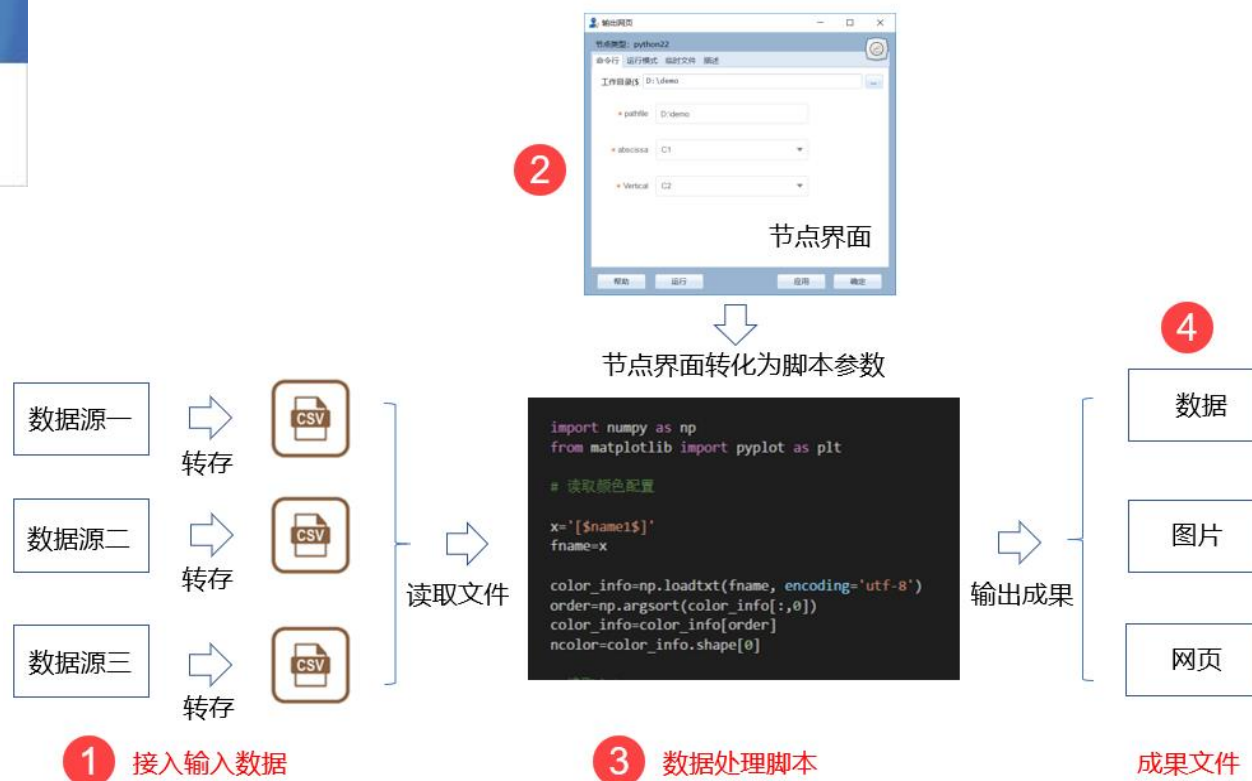
Started: 2019/1/12 下午9:29:24

8m 3s

Prepare job · succeeded	<1s
Initialize Agent · succeeded	<1s
Initialize job · succeeded	1s
Get sources · succeeded	11s
UsePythonVersion · succeeded	2s
CondaEnvironment · succeeded	7m 19s
pytest · succeeded	24s
PublishTestResults · succeeded	1s

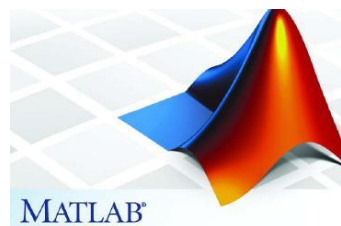


自定义节点数据处理过程



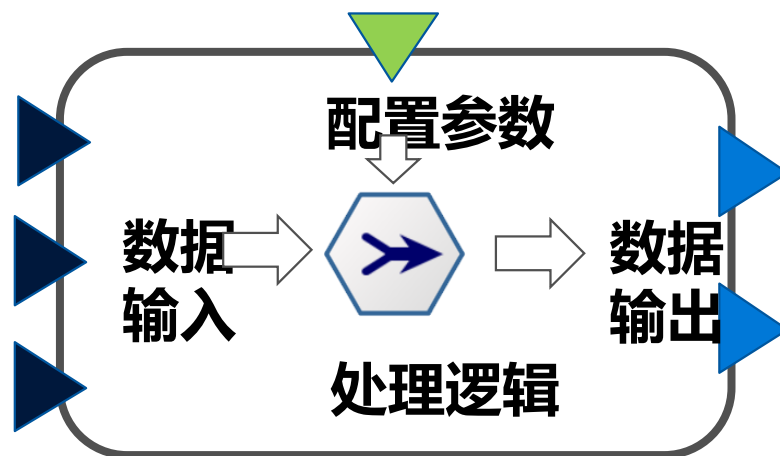
扩展节点的意义

除系统提供的一组常用的数据分析处理算法外，还提供强大的**算法集成能力**，通过简单易用开放接口可以和R、Python、Matlab、DLL算法库等常用的算法库和用户编写的算法库**无缝集成**，可以借助这些成熟的算法大大扩展系统的专业数据处理算法能力。比如分析地震目录数据常用的“ssEDA”“ssBase”等。



节点配置和节点解耦，通过API形式调用，配置界面可以是WPF，也可以是网页格式。

输入数据和数据表解耦，输入数据可能是一个json文件，也可以是一个数据表，也可能是消息队列



输出数据和数据表解耦，输出数据可能是一个json文件，也可以是一个数据表，也可能是消息队列

微服务式节点模型图

用流程使数据畅通，构建预报业务新生态



s p e c

A agile ecosystem for earthquake sciences