Madagascar Open-Source Software Project

Sergey Fomel
3D deep-water reservoir model by Jim Jennings

Tools for Large-Scale Data Analysis
Channel making recipe

- Madagascar modules are filter programs connected into data analysis workflows by Unix pipes and scripts.

- Preferred form of scripts: computational recipes in SCons/Python.

```python
# Make sand fraction grid

def make_sand (grid_par=None, geo_par=None, sand_par=None):
    geo_par_copy = geo_par.copy()
    geo_par_copy['as_height'] = geo_par['as_height'] - geo_par['as_height0']

    sand_par_copy = sand_par.copy()
    sand_par_copy['na_sand'] = sand_par['na_sand'] - sand_par['na_sand0']

    if (grid_par['nz'] > 1):
        Flow ('z_rel', 'aggrade_grid depth_grid layer_00_mask',
            'math a=${SOURCES[0]} d=${SOURCES[1]} m=${SOURCES[2]}
            output="m*(x3-a)/d" |
            clip2 upper=0
        ')
    else:
        Flow ('z_rel', 'aggrade_grid depth_grid layer_00_mask',
            'math a=${SOURCES[0]} d=${SOURCES[1]} m=${SOURCES[2]}
            output="m*(%z-g-a)/d" |
            clip2 upper=0
     ')

    Flow ('pos_shift', 'pos_grid wtop_grid skew_grid',
        'math p=${SOURCES[0]} w=${SOURCES[1]} s=${SOURCES[2]}
        output="p/w+%g+s/2" |
    ')
```

Figure 10. Depth migration of the 3-D synthetic test data. (a) Migrated data. (b) Migrated diffractions.
from rsf.proj import *
from rsf.recipes.beg import server as private

Fetch('zovol.HH','jim',private)
Flow('zovol','zovol.HH','dd form=native | window min1=1.7 max1=2.2')

Fetch('zovel.HH','jim',private)
Flow('zovel','zovel.HH','dd form=native')

def plot3(title):
    return ''
    byte gainpanel=all |
grey3 frame1=80 frame2=220 frame3=140
title="\%s"
label1=Time unit1=s label2=In-line unit2=m label3=Cross-line unit3=m
'' % title

zplot = ''
byte gainpanel=all pclip=100 bias=2297 allpos=y |
grey3 wanttitle=n labelrot=n flat=y
frame1=140 frame2=220 frame3=140
point1=0.85 point2=0.90 Xscreenratio=0.7 Xscreenht=9
Pham, N., Fomel, S. and Dunlap, D., 2019. Automatic channel detection using deep learning: *Interpretation*
Reproducible Research Portal

- Accumulation of knowledge in the form of research papers
- Papers are linked to reproducible examples (figures) and computational recipes
- Reproducibility is maintained by **continuous integration**
Pyramid of Knowledge

- Implement
- Test
- Share
Papers

Workflows

Programs

Pyramid of Knowledge
Pyramid of Knowledge

~2,000 Programs

~1,200 Workflows

~300 Papers

Python

C

LATeX
Lessons from the Madagascar project

Reproducibility in itself is not the goal.
The main beneficiary is the author.
Each computation is a test.
Reproducibility requires maintenance.
Maintenance requires an open community.

In a Nutshell, Madagascar...

...has had 15,011 commits made by 144 contributors representing 1,371,770 lines of code.

...is mostly written in C with an average number of source code comments.

...has a well established, mature codebase maintained by a large development team with increasing Y-O-Y commits.

...took an estimated 382 years of effort (COCOMO) starting with its first commit in May, 2003.

https://www.openhub.net/p/m8r
Contributors

Schools and Workshops
Madagascar Project

- Tools for large-scale data analysis
- Reproducible research portal
- Open-source community

http://ahay.org/
Thank you

Questions?