

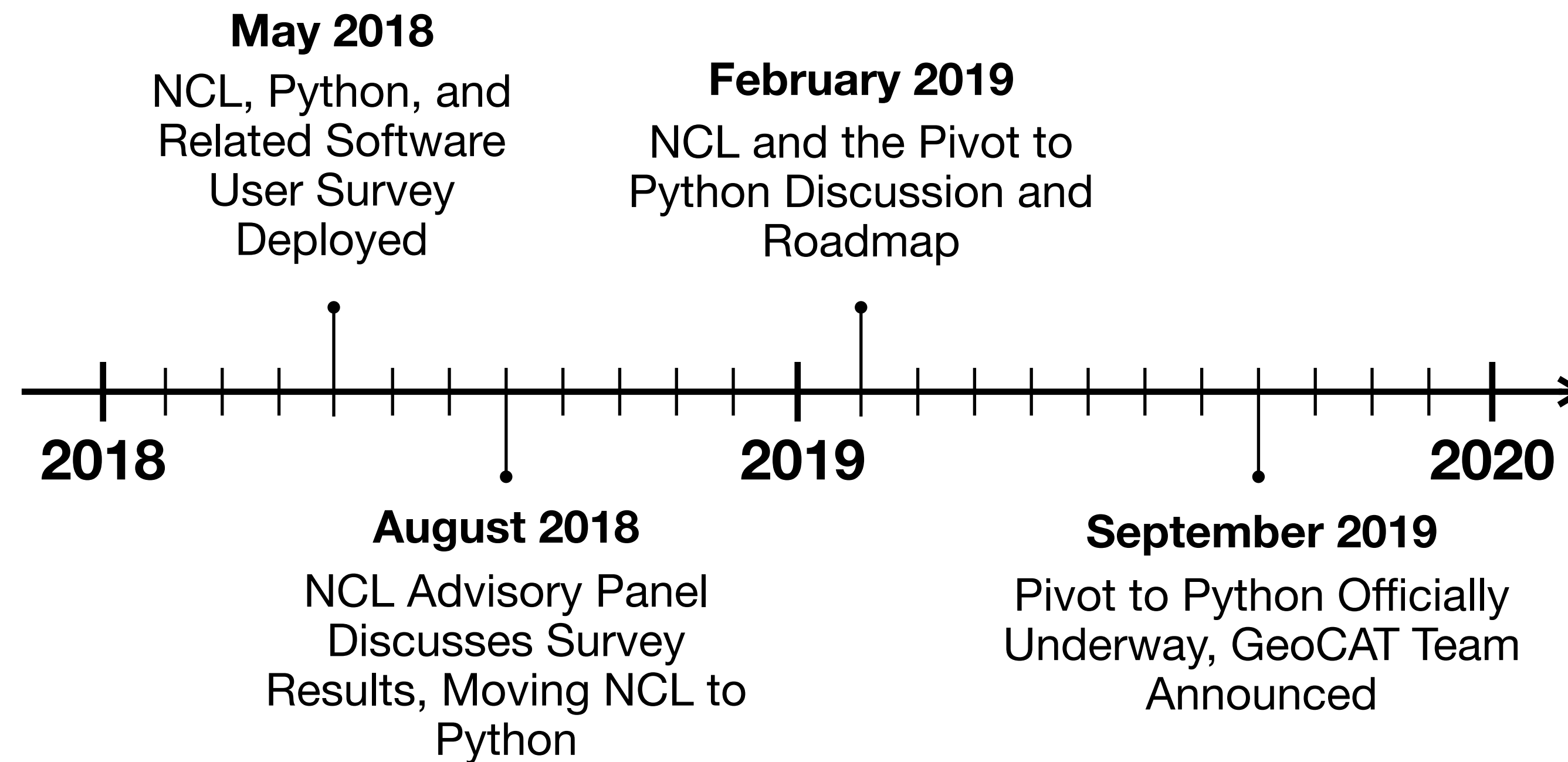
How to write new, old code

Implementing legacy NCL routines in the python ecosystem

Anissa Zacharias | March 4, 2024



History of Pivot to Python





NCL and the Pivot to Python

Discussion and Roadmap

January 2019

Mary Haley
Rick Brownrigg
Kevin Hallock
Bill Ladwig
John Clyne (supervisor)

Visualization and Analysis Systems Technologies
National Center for Atmospheric Research

 NCAR is sponsored by  National Science Foundation

Any opinions, findings and conclusions or recommendations expressed in this material do not necessarily reflect the views of the National Science Foundation.

What is GeoCAT?

Geoscience Community Analysis Toolkit



GeoCAT

“GeoCAT aims to create scalable data analysis and visualization tools for Earth System Science data to serve the geosciences community in the scientific Python ecosystem.”

GeoCAT-comp

- Computational component of the GeoCAT project
- Provides implementations of computational functions for operating on geosciences data
- Many (but not all) functions from NCL

GeoCAT-examples & GeoCAT-viz

- Visualization component of GeoCAT project
- GeoCAT-examples provides a gallery containing visualization examples from many plotting categories of geosciences data
- GeoCAT-viz provides tools to help plot data, including convenience and plotting functions

Project Raijin



- Part of the collaboration team developing uxarray, an Xarray-styled package for reading and directly operating on unstructured grid datasets following UGRID conventions
- Community-owned, sustainable, scalable tools on unstructured climate and global weather data

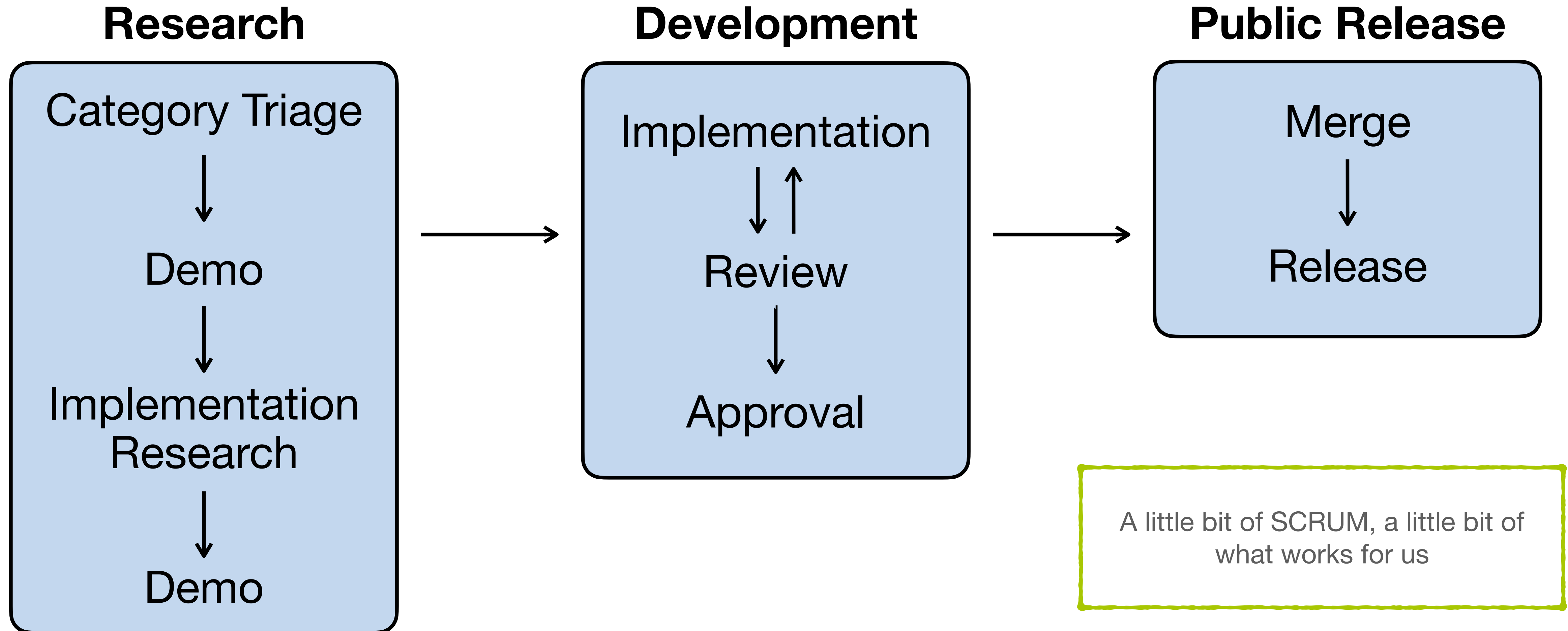
Python Tools



Python Package Development Tools



Lifecycle of Implementation



Some “old code” problems

Reproducibility challenges: how close is close enough?

Things that can impact numerical reproducibility

- Floating point
- Resolution of intermediate results
- Differences in underlying math libraries
- Optimization
- Differences across processors and compilers
- ... many other things

CAM (Community Atmospheric Model) example

$A(I) + B + TOL$

where

- TOL was very small and positive
- $A(I)$ and B could be large

Compiler evaluated this as $A(I) + (B + TOL)$

Hoisted constant $B + TOL$ out of the loop

TOL got rounded away....

Code the Future

© 2013, Intel Corporation. All rights reserved. Intel and the Intel logo are trademarks of Intel Corporation in the U.S. and/or other countries. *Other names and brands may be claimed as the property of others.

OPTIMIZATION NOTICE >

Example

```
do i=1,10000
```

```
    sum = sum + sqrt(values(i))
```

```
end do
```

No optimization: 0.2919665E+08

With optimization: 0.2919677E+08

With advanced instructions: 0.2919678E+08

Misaligned data: 0.2919677E+08

Double precision: 0.29196781789902E+08

Code the Future

© 2013, Intel Corporation. All rights reserved. Intel and the Intel logo are trademarks of Intel Corporation in the U.S. and/or other countries. *Other names and brands may be claimed as the property of others.

OPTIMIZATION NOTICE >

Improving Numerical Reproducibility in C/C++/Fortran, Steve Lionel, SC13

Some “old code” problems

Reproducibility challenges: how close is close enough?

Things that can impact numerical reproducibility

- Floating point
- Resolution of intermediate results
- Differences in underlying math libraries
- Optimization
- Differences across processors and compilers
- ... many other things

Reproducibility decisions

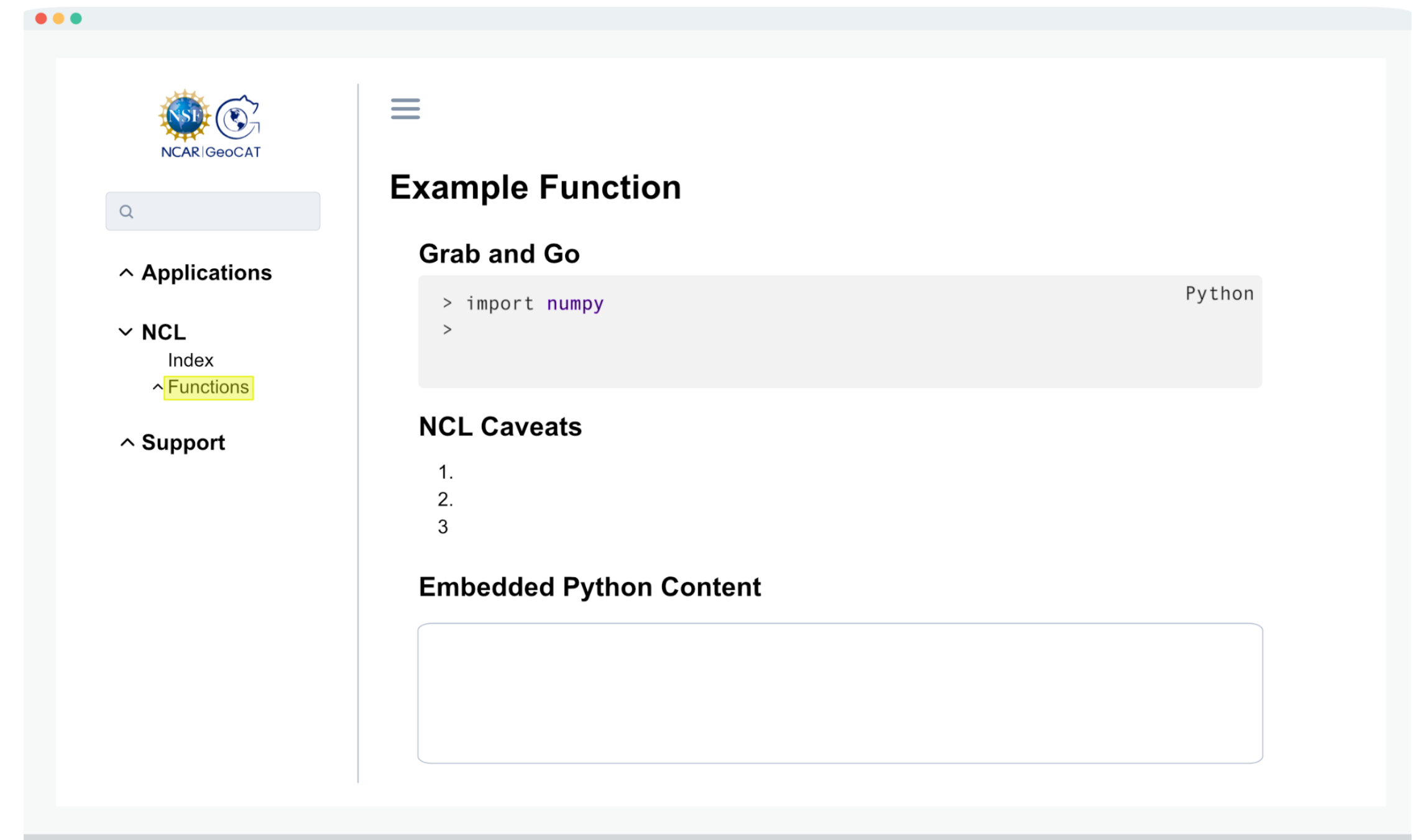
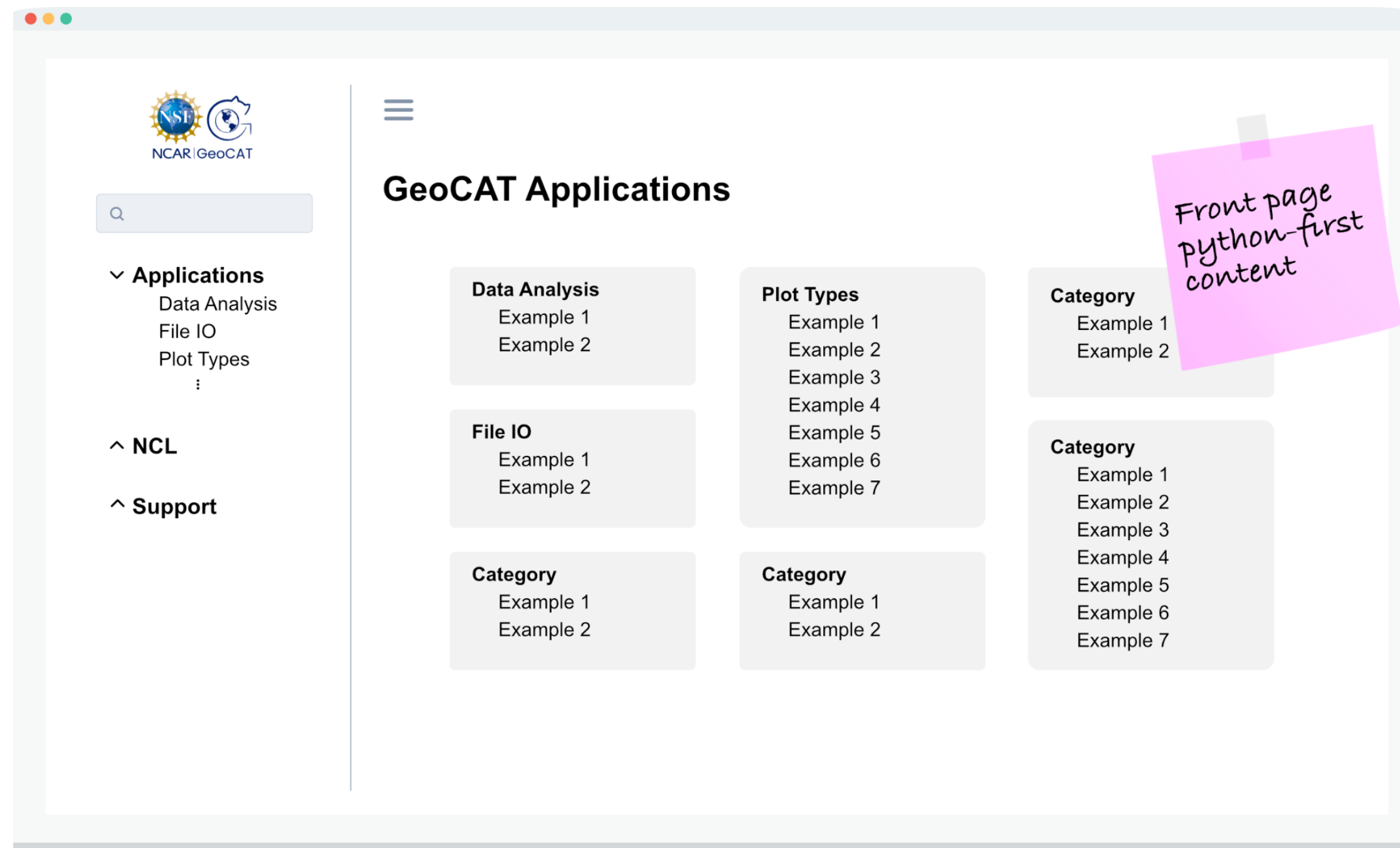
- Should we use more modern or updated algorithms? Should we keep the old ones available?
- What changes do we make to the algorithms to counteract Python’s speed limitations?
- Do we fix bugs from the NCL code?

Other considerations

- What should we reproduce and take ownership of?
- What functionality is better owned by other community packages?
- Domain specific considerations

See Steve Lionel’s SC13 slides, [Improving Numerical Reproducibility in C/C++/Fortran](#)

Case Study: GeoCAT-applications



Requests? Suggestions? Burning Questions?

**We'd legitimately love to hear from you,
you're part of our process**

Feature request form:

<https://forms.gle/6DTo3ELLri4DAGfG8>

Want one on one advice?

<https://ncar.github.io/esds/office-hours/>



my team's email: geocat@ucar.edu
my email: anissaz@ucar.edu
my boss' email: oero@ucar.edu