R and Python Package Collection and Filtering

Robert McLay

December 15, 2022

XALT: Outline

- ▶ XALT can record package usage by R and Python (optionally Matlab)
- ▶ This talk centers on Python with a little on others
- ▶ Special code is required to use allow for tracking

R tracking requires a special package to turn on tracking

- ► Package developed at IU by James McCombs and Michael Scott
- ► I need to get one of them on documentation on how to install it.

TACC

Python tracking

▶ I developed XALT's xalt_sitecustomize.py (installed as sitecustomize.py)

▶ This was with help from Riccardo Murri (with an intro from Kenneth Hoste!)

▶ To track python imports add $XALT_DIR/site_packages to $PYTHONPATH

▶ This puts XALT's sitecustomize.py which is ALWAYS read by every python execution

XALT package tracking interface

► All package tracking "hooks" calls the program xalt_record_pkg to interface with XALT

► It can be found in $XALT_DIR/libexec

```
../xalt_record_pkg program          <PROGRAM> \
                   xalt_run_uuid     <UUID>     \
                   package_name      <NAME>      \
                   package_version   <VERSION>  \
                   package_path      <PATH>
```

Future Topics?

- ► How XALT "hijack" the link process to watermark user programs.
- ► This hijacking is not required for execution tracking
- ► Next Meeting will be on July 22, 2022 at 10:00 am U.S. Central (15:00 UTC)
- ► This talk will be given by my colleague: Amit Ruhela.