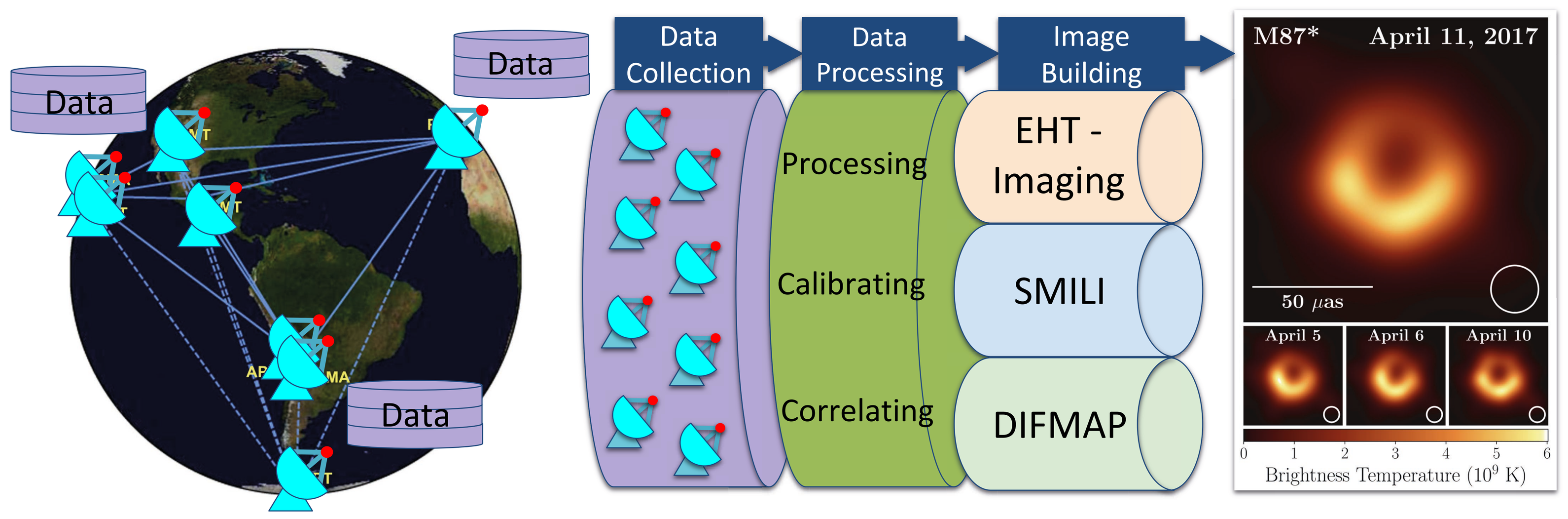
Lessons Learned from Reproducing Images from the First M87 Event Horizon Telescope (EHT)

C. Ketron*, J. Leonard*, B. Roachell*, R. Patel*, R. White†, S. Caino-Lores*, N. Tan*, P. Miles‡, K. Vahi†, E. Deelman†, D. Brown‡, and M. Taufer
U. Tennessee Knoxville*, U. Southern California†, and Syracuse U.‡
Emails: *taufer@utk.edu, †deelman@isi.edu, ‡dabrown@syr.edu

Event Horizon Telescope (EHT) Pipeline



The process the EHT project used to create the black hole images for each day builds on these steps: (i) Eight telescopes collect data from around the world; (ii) The data is processed, calibrated, and correlated before being passed to three separate imaging workflows or pipelines; and (iii) Images from each pipeline are combined into a single black hole image.

Our Mission

- *Published projects often have obstacles in reproducibility:*
  - *Minimal documentation describing codes' compilation & execution*
  - *Missing information of used heterogeneous computing resources*
  - *Complex data pre- & post-processes*
- We study the reproducibility of the Event Horizon Telescope Collaboration's (EHTC) project on the computed M87 black hole shadow image
- We generate images similar to the original images for each pipeline & reproduced plots of baseline telescope data
- Our solution supports understanding, analyzing, and reusing of the published scientific results

Reproducing Science as a Team

Team:

- Reproduction of EHT project by team of undergraduate students from the University of Tennessee (UT) and the University of Southern California (USC)
- Data validated by Ria Patel (UT) & Rebecca White (USC)
- Image-building algorithms validated by Ross Ketron (UT - *EHT-Imaging algorithm*), Jacob Leonard (UT - *SMILI algorithm*), & Brandan Roachell (UT - *DIFMAP algorithm*)
- Reproducibility of work is done with the same data sets and workflows used for the original project

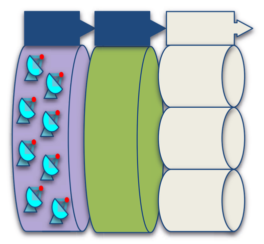
Challenges:

- Missing scripts to plot baseline telescope data
- Package dependencies not documented
- Installation across different architectures with dependency issues
- Missing detailed image post-processing steps to fully reproduce released figures
- Distributed knowledge, codes, and tools across groups of the original project

Opportunities:

- Interact with the scientists who worked on the original project
- Learn astrophysics theory and concepts
- Strengthen problem solving skills in a cohort setting
- Understand the impact and applicability of reproducible research

Data Validation



Type of data:

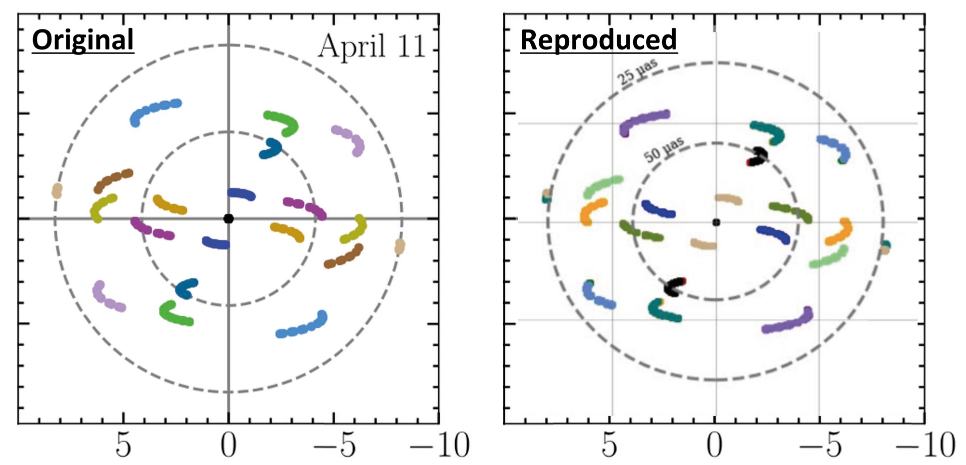
- Timestamps, telescope locations, and radio frequencies
- Time Stamp: April 5, 6, 7, 10, & 11 in 2017
- Each day has data for low & high telescope frequencies

Eight telescopes (table on the right) are located across the earth and work in concert as a huge single telescope

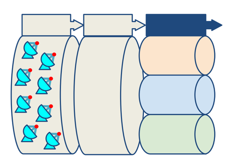
Data is collected in a distributed way and processed at a single location

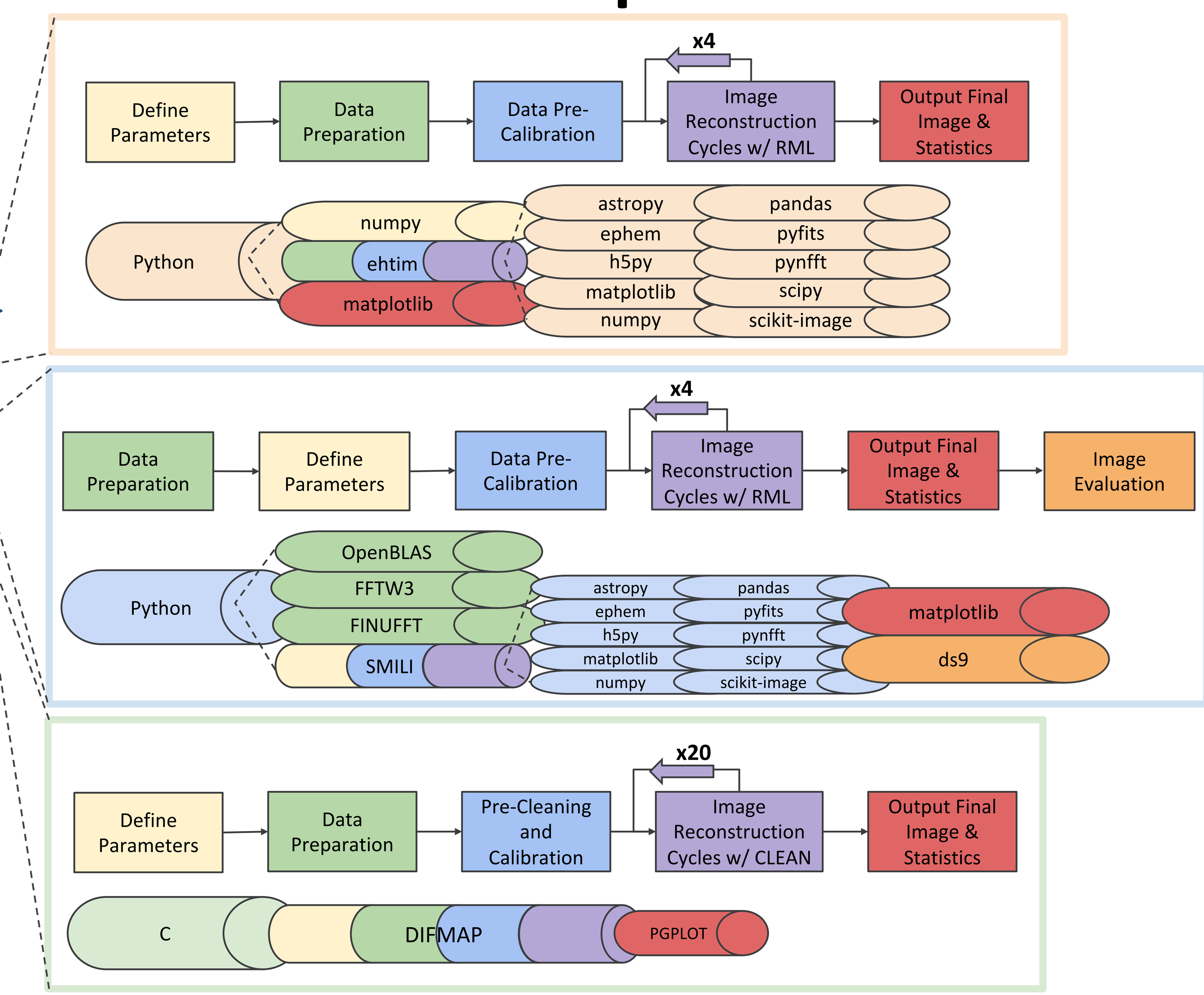
8 EHT Array Telescopes	
Antarctica	SPT - South Pole Telescope
Arizona	SMT - Submillimeter Array
Chile	ALMA - Atacama Large Millimeter/Submillimeter Array APEX - Atacama Pathfinder Experiment Telescope
Hawaii	CMT - James Clerk Maxwell Telescope SMA - Submillimeter Array
Mexico	LMT - Large Millimeter Telescope Alfonso Serrano
Spain	PV 30m - Pico Veleta 30m Telescope

Example of matching between data used for the **original project** (left) and our **reproducibility effort** (right) – same data sets are used



Workflows and their Implementation

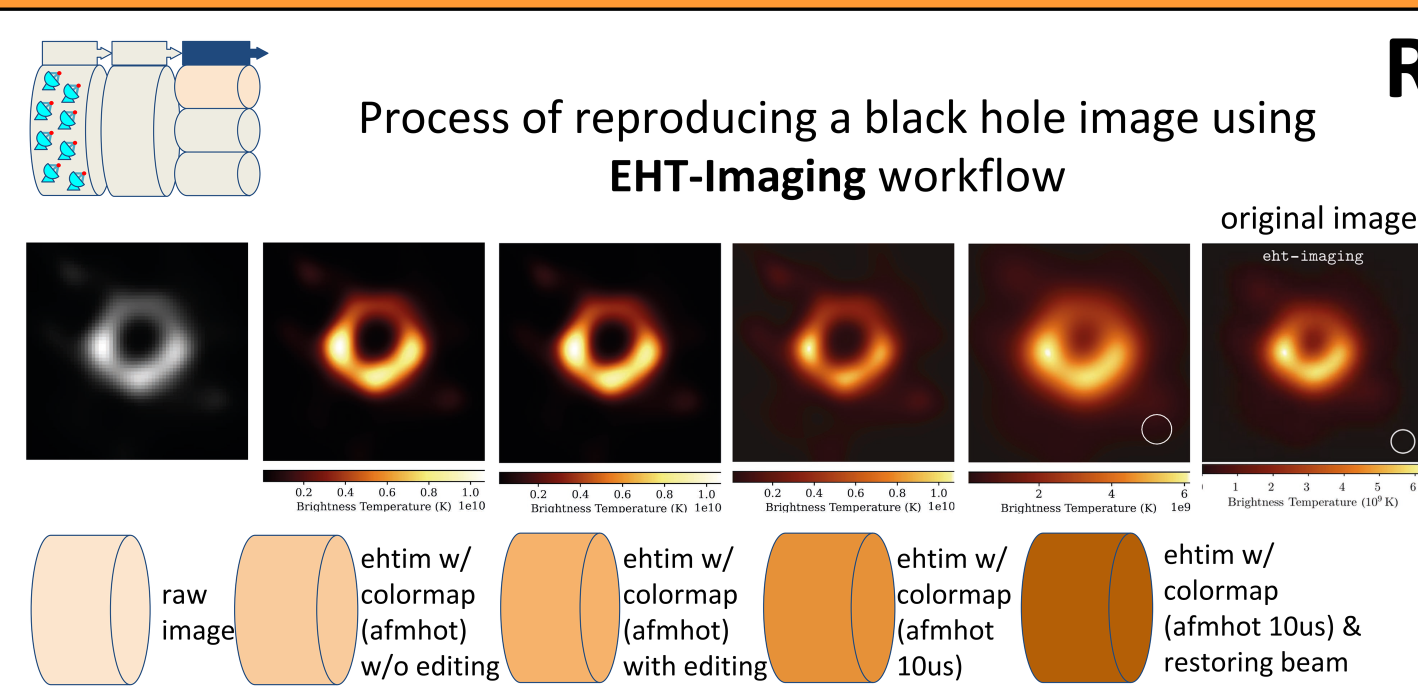




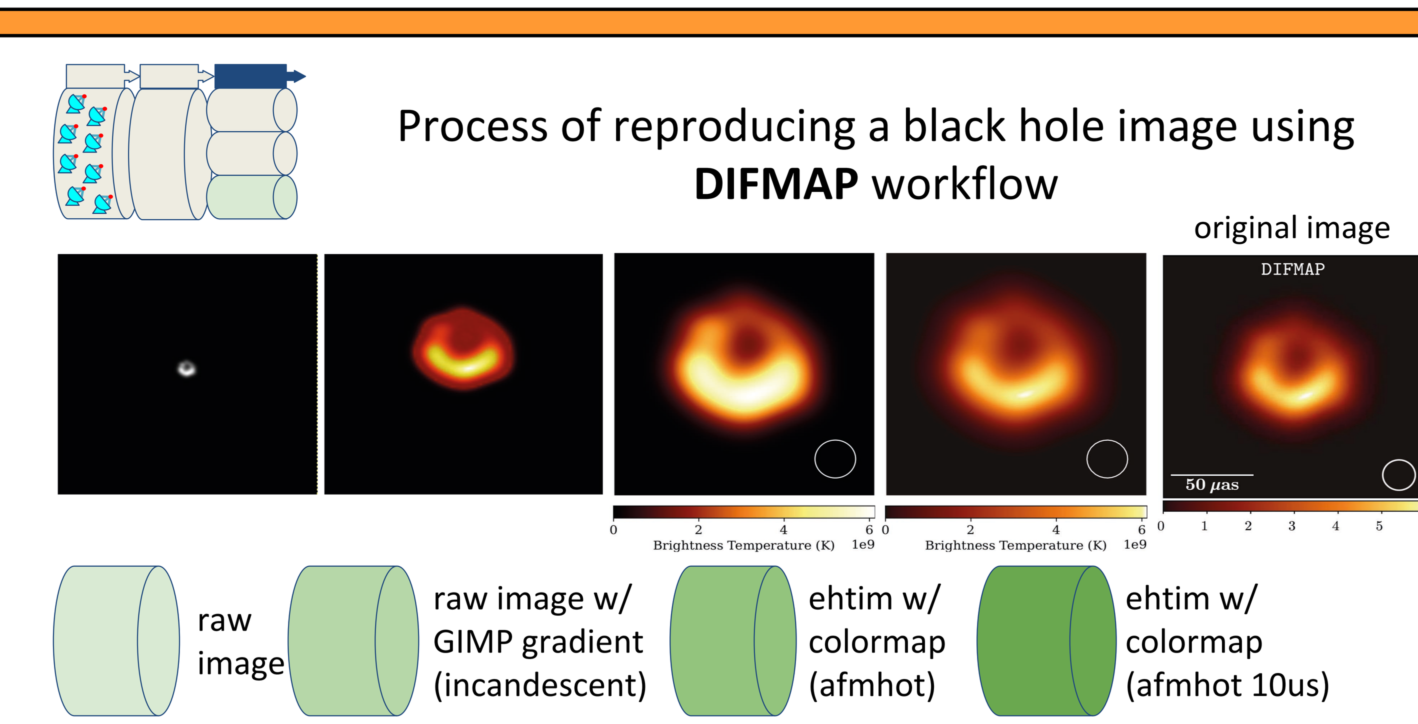
High-level workflow operations (top) and implementation including programming languages, libraries, and tools (bottom) for each one of the three methods -- EHT-Imaging (pick box), SMILI (blue box), and DIFMAP (green box)

Results

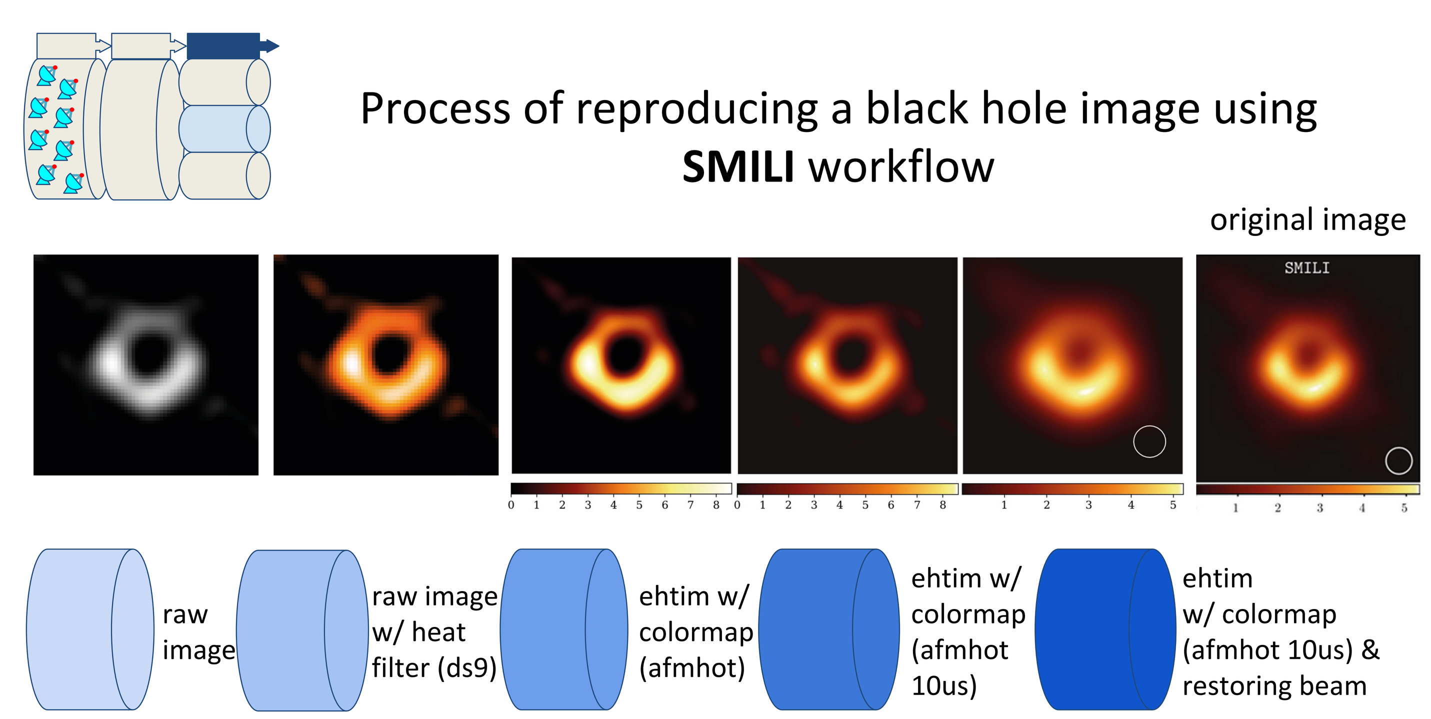
Process of reproducing a black hole image using **EHT-Imaging** workflow



Process of reproducing a black hole image using **DIFMAP** workflow



Process of reproducing a black hole image using **SMILI** workflow



Lessons Learned

- Original work prioritized the science and implementation of the data collection, data processing, and image building
- Our work extends this contribution by validating workflows used and providing detailed documentation allowing non-scientists to more readily execute the workflow(s)
- Our work can enable other scientists to generate new scientific discovery from published findings, open-source methods, and open-access data
- Our accessible, easy-to-use artifacts can enable a diverse group of participants in this exciting research, including junior students, minorities, and underserved communities
- Original researchers from the EHT Collab were contacted for assistance with some aspects of our reproduction and provided valuable insights

Citations:
[1] Event Horizon Telescope Collaboration, "First M87 Event Horizon Telescope Results. I. The Shadow of the Supermassive Black Hole", *The Astrophysical Journal*, vol. 875, L1, 2019. doi:10.3847/2041-8213/ab0ec7.
[2] Event Horizon Telescope Collaboration, "First M87 Event Horizon Telescope Results. IV. Imaging the Central Supermassive Black Hole", *The Astrophysical Journal*, vol. 875, L4, 2019. doi:10.3847/2041-8213/ab0e85.

Funding support:
The work in this poster is funded by the National Science Foundation (NSF) under grants #2041977, #2041901, #2041878, #2028923, and #2028930.
The authors want to thank the scientists who worked on the original project and provided us with valuable insights during this project

Syracuse University

THE UNIVERSITY OF TENNESSEE KNOXVILLE