

User-Facing Tutorials as Code: Reproducible and Reliable Tutorials with CI/CD

Date: 12 August 2026

Presented by: Brigitta Sipőcz (Caltech/IPAC)

(The slides are available via the link in the page's sidebar.)

Q: What languages do you recommend for the jupyter notebooks? Python, Julia, R?

A: My experience has only been with Python, but any language Jupyter supports should be fine. The choice of language is probably better governed by the languages used in the projects for which you're creating tutorials and what languages you expect your users to be using.

Q: Any comments on workflow using Marimo notebooks?

A: I have not tried it, but I believe that the same approach should work with Marimo notebooks. In my projects, we've chosen to work in the Jupyter ecosystem because it has a large community and is open. Marimo is backed by a company, which some may consider an advantage, but we were concerned about what happens if the company folds or chooses to discontinue the product.

Q: I wanted to get a hands-on or otherwise example on building Unit Tests (and if possible Integration Tests) for a new project adopting best CI/CD practices via the GitHub ecosystem.

A: This topic is really outside the scope of this presentation. But these days, you can find many examples of projects with suites of unit and/or integration tests and an working CI setup. Consider checking out the repositories of some of your favorite software projects. There are also a variety of other resources available about testing

and CI. You might check out resources mentioned on the [BSSw.io](https://bssw.io) website (<https://bssw.io/items?topic=continuous-integration-testing>) or just search more generally. Some basic advice would be to start simple – write one test and get it automated as a CI process. Then you can add more tests, other testing environments, etc.

Q: Historically, investments in reproducibility infrastructure have been underfunded and insufficiently incentivized, at least for communities I know. In particular, publishers have not demanded reproducible results for publication. Does your community do a better job at reproducibility? Your efforts seem to indicate the value of reproducibility by virtue of the investments you can make.

A: I think we're making progress, but slowly. Astronomy is perhaps in a better position than some other fields right now because it is smaller and more dependent on shared facilities and data. In astronomy, expectations for sharing and reproducibility are being led by the practitioners and journals and others are following along.

Q: Sorry if I missed this during the talk, but to what extent are the notebooks being tested? Does the CI just test that all cells run with no issue, or does it go deeper such as testing correctness of cell outputs?

A: At present, the testing ensures that the code in the cell runs without throwing unexpected errors. In principle, it should be possible to extract the code and run it in a setting where you can also check the output it produces, but we're not currently doing that in the tutorials I'm involved with.