

ADS is hiring! We are looking for an astrophysicist/project scientist; we also have openings for developers and a DevOps engineer. [Read more and apply here.](#)



## Reproducible, Automated and Portable Computational and Data Science Experimentation Pipelines with Popper

Jimenez, I. ; Maltzahn, C.

Currently, approaches to scientific research require activities that take up much time but do not actually advance our scientific understanding. For example, researchers and students spend countless hours reformatting data and writing code to attempt to reproduce previously published research. What if the scientific community could find a better way to create and publish our workflows, data, and models to minimize the amount of the time spent "reinventing the wheel"? Popper is an NSF and CROSS sponsored protocol and CLI tool for implementing scientific exploration pipelines following a DevOps approach. Popper allows researchers and students to generate work that is easy to reproduce.

Modern open source software (OSS) development communities have created tools that make it easier to manage large codebases, allowing them to deal with high levels of complexity, not only in terms of managing code changes, but with the entire ecosystem that is needed in order to deliver changes to software in an agile, rapidly changing environment. These practices and tools are collectively referred to as DevOps. The Popper Experimentation Protocol repurposes the DevOps practice in the context of scientific explorations so that researchers can leverage existing tools and technologies to maintain and publish scientific analyses that are easy to reproduce. By following Popper, researchers can produce portable, automated and version-controlled experimentation pipelines that are easier to re-execute.

In this talk/poster, we will briefly introduce DevOps and give an overview of best practices. We will then show how these practices can be repurposed for carrying out scientific explorations and illustrate using some examples. We will also walk the audience through the usage of the Popper CLI tool, showing examples from multiple domains such as High Energy Physics, Genomics, and Atmospheric Sciences.

**Publication:** American Geophysical Union, Fall Meeting 2018, abstract #IN53A-04  
**Pub Date:** December 2018  
**Bibcode:** 2018AGUFMIN53A..04J  
**Keywords:** 1908 Cyberinfrastructure; INFORMATICSDE: 1916 Data and information discovery; INFORMATICSDE: 1948 Metadata: Provenance; INFORMATICSDE: 1976 Software tools and services; INFORMATICS

 Feedback/Corrections?