AstroDB Toolkit: A collaborative data management tool

Kelle Cruz (CUNY Hunter College, NY, USA), **David R. Rodriguez** (STScI, MD, USA),
William J. Cooper (U. of Hertfordshire, UK), Arjun Savel (U of Maryland, MD, USA), and Sherelyn Alejandro (CUNY Hunter College, NY, USA)

The AstroDB Toolkit is a collaborative database management tool intended for small-to-medium sized projects to manage their living datasets. The Toolkit facilitates a novel way of reviewing, updating, and version-controlling the contents of medium-scale databases while still integrating with existing tools. The Toolkit is under active development and not released. The proof-of-concept use case is the SIMPLE Archive, https://simple-bd-archive.org/.

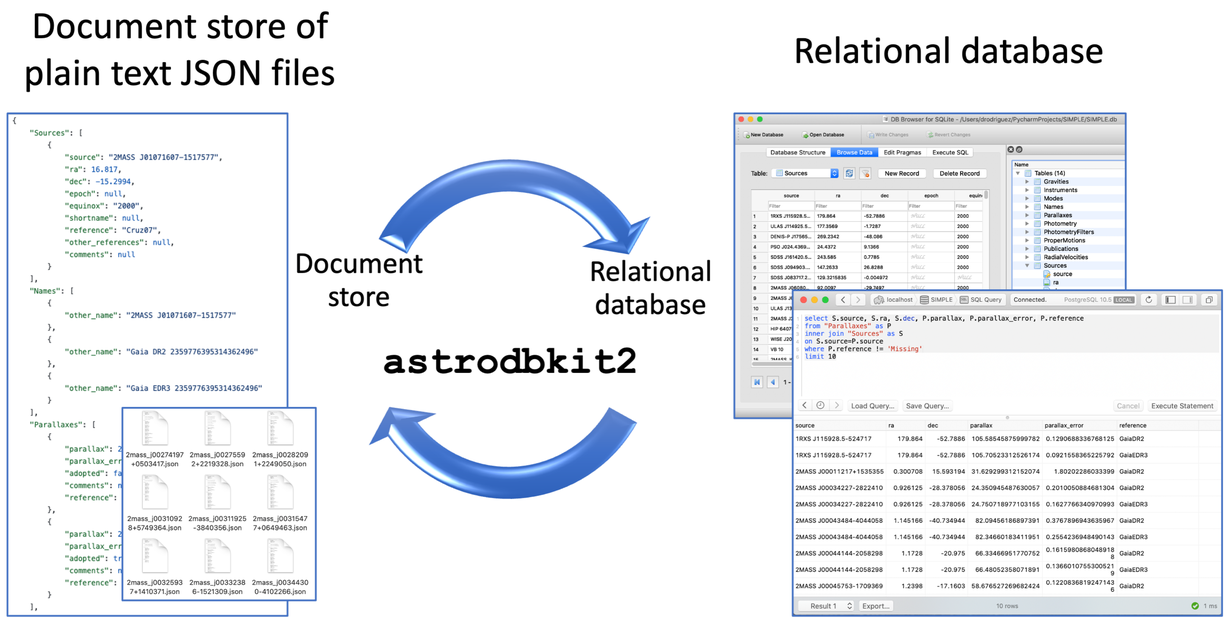
Vision for AstroDB Toolkit

The technical choices that were made in the development of the AstroDB Toolkit follow several design requirements:

- Ability to represent complex data models in a simple fashion,
- Support for datasets up to tens of thousands of sources,
- Collaborative editing which enables community members to modify and maintain data holdings,
- Be usable privately for managing not-yet-public data,
- Interactively explore and visualize holdings via a website,
- Queryable locally via scripting without the need for an internet connection.

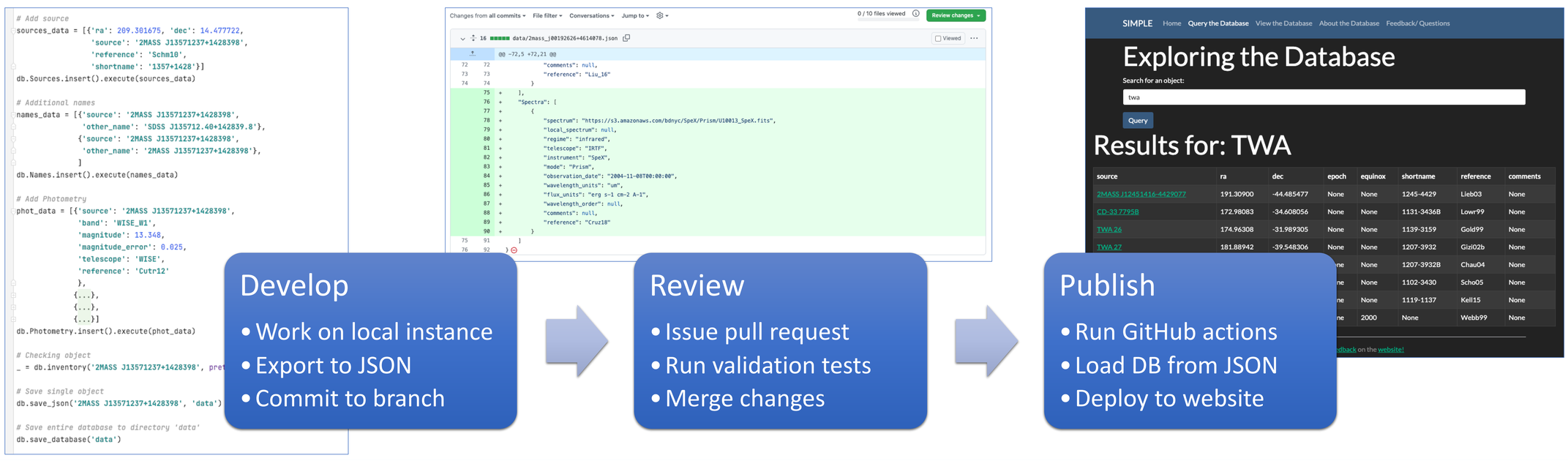
Tool Design for AstroDB Toolkit

Using `astrodbkit2` python package, we convert the database from a document store mode, where individual Sources are stored as separate JSON files (which can in turn be loaded into NoSQL databases like MongoDB), to a relational database such as SQLite or Postgres that can be accessed with standard tools like the DB Browser for SQLite and Postico. JSON files serve as the definitive copy of the database for purposes of version control.



Contribution Workflow

The vision for the AstroDB Toolkit is that anyone can contribute data using a typical GitHub workflow. The chart below illustrates the workflow from inserting new data to a local database instance via Python, review of database modifications in GitHub, and then publishing them to the website.



astrodbtoolkit, SIMPLE-AstroDB/SIMPLE-db/

This material is based upon work supported by the National Science Foundation under Grant No. AST-1909837.