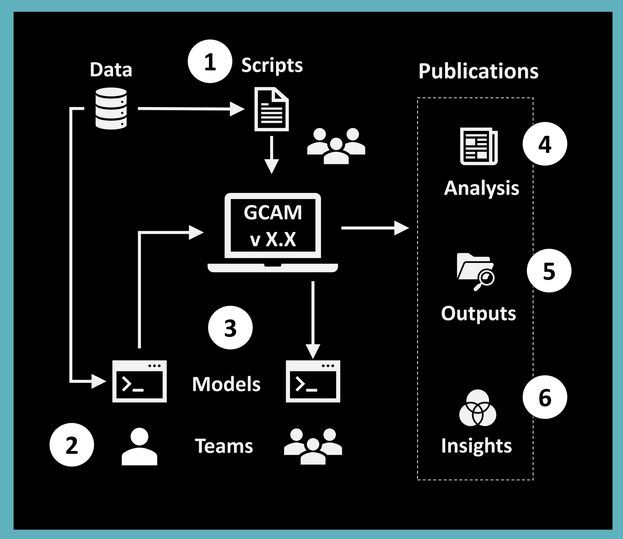
GCIMS - Integration: Reproducible, robust, and scalable workflows for interoperable human-Earth system modeling

Zarrar Khan*, Chris Vernon, Pralit Patel, Isaac Thompson, Benjamin Knight

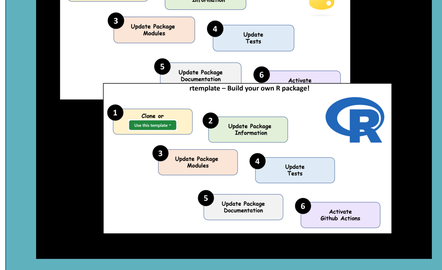
***zarrar.khan@pnl.gov**

Motivation



	Issue	Solution
1	Stranded Scripts	Pakcaged Models
2	Stranded Knowledge	Software Map Documentation
3	Stranded Model Versions	Software Health CI/CD
4	Stranded Analysis	Reproducible Workflows
5	Stranded Datasets	Data Management Miniting DOIs
6	Stranded Insights	Dashboard Persistent Insights

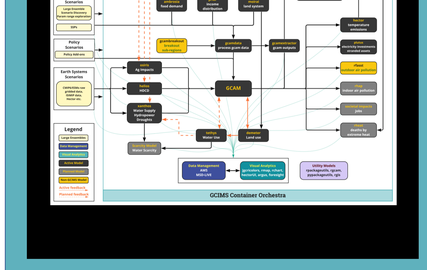
1 Packaged Models



Package Design

- Folder Structure
- Continuous Integration
- Multi OS Build Checks
- Code Testing
- Documentation

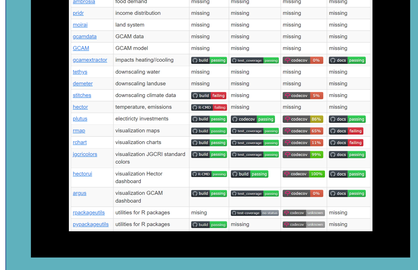
2 Software Map



Track down and Map out

- Data sources
- Existing models
- Planned models
- Inter-linkages

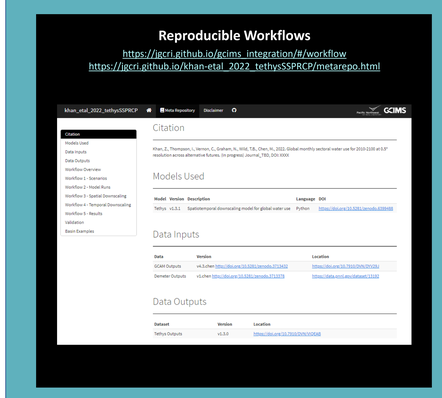
3 Software Health



Overall System Monitoring

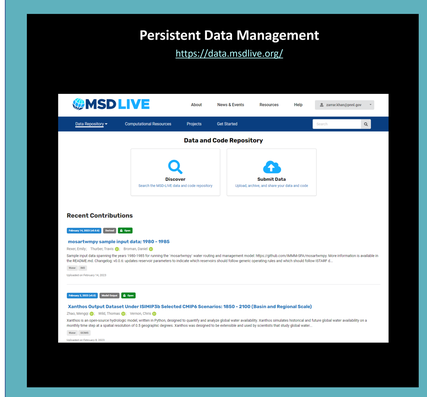
- Continuous Integration
- Multi OS Build Checks
- Code Testing
- Code Coverage
- Documentation

4 Reproducible Workflows



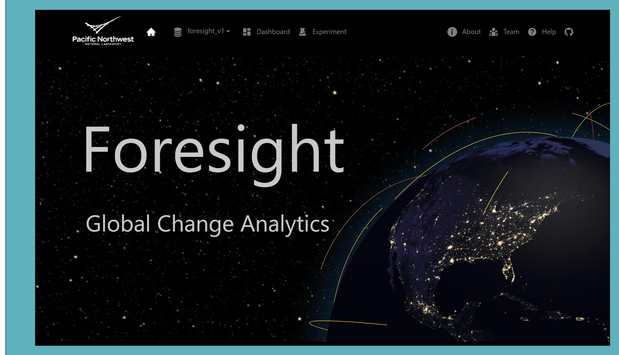
- **Meta-repo templates**
  - Inputs with DOI
  - Outputs with DOI
  - Workflows
  - Validation
  - Results
- **Containerization**
  - Scalable
  - Cloud/On-prem/hybrid

5 Data Management



- **FAIR**
  - Findability
  - Accessibility
  - Interoperability
  - Reusability
- **Key Features**
  - Minted with DOIs
  - Meta-data
  - Storage (MSD, Deception, Cloud)

6 Persistent Insights



Curated Insights

- Persisting key figures across existing & future research
- Persisting key analysis across existing & future research
- Accesibility and reusability for a broad audience

U.S. DEPARTMENT OF ENERGY BATTELLE

PNNL is operated by Battelle for the U.S. Department of Energy

IMS GLOBAL CHANGE INTERSECTORAL MODELING SYSTEM

11/9/2023 | PNNL-SA-XXXXX

www.pnnl.gov