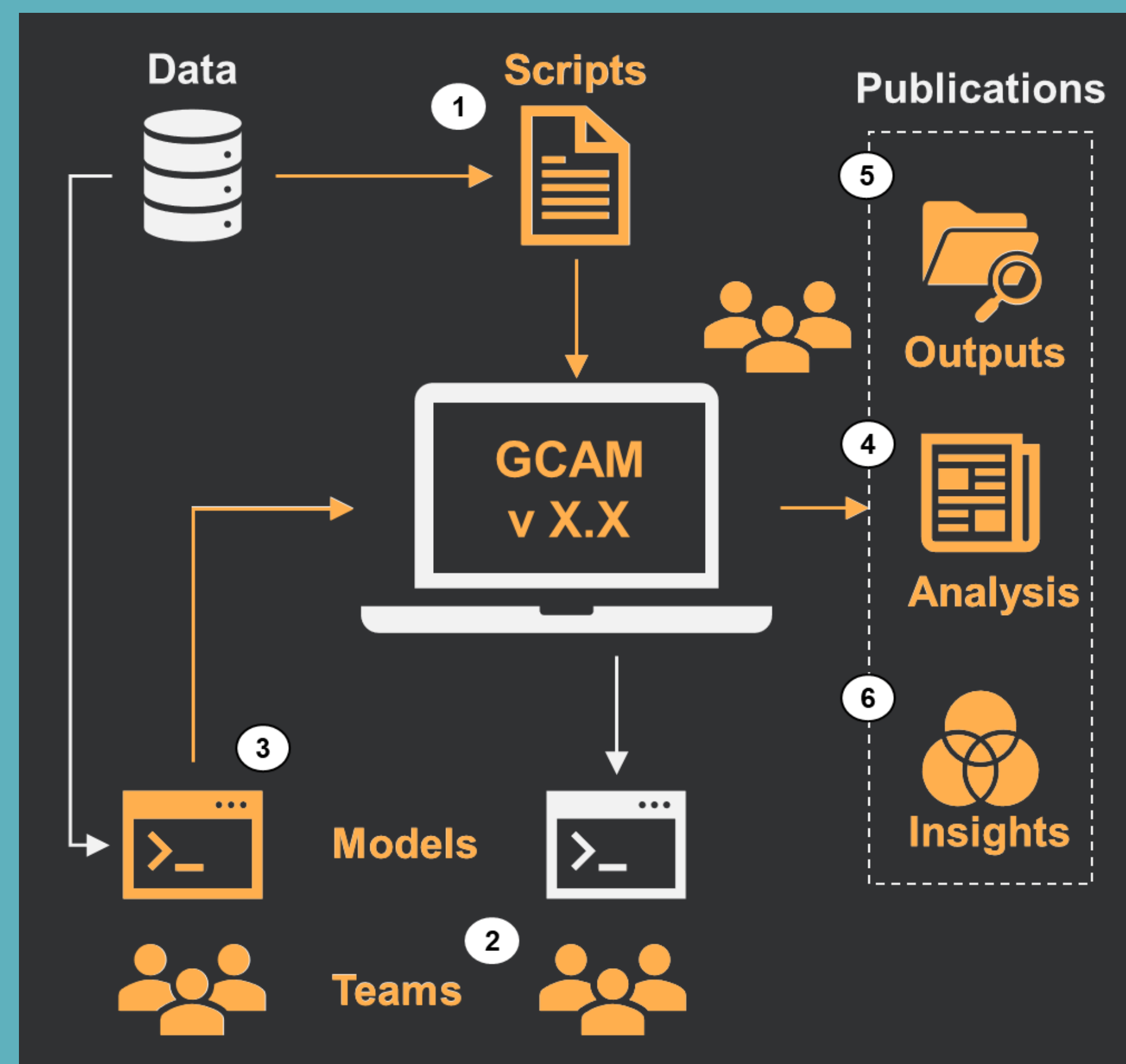# GCIMS – Integration: Reproducible, robust, and scalable workflows for interoperable human-Earth system modeling

Zarrar Khan*, Chris Vernon, Isaac Thompson, and Pralit Patel

***zarrar.khan@pnnl.gov**

Motivation

Issue		Solution
❶	Stranded Scripts	Packaged Models
❷	Stranded Knowledge	Software Map Documentation
❸	Stranded Model Versions	Software Health CI/CD
❹	Stranded Analysis	Reproducible Workflows
❺	Stranded Datasets	Data Management Minting DOIs
❻	Stranded Insights	Dashboard Persistent Insights

❶ Packaged Models
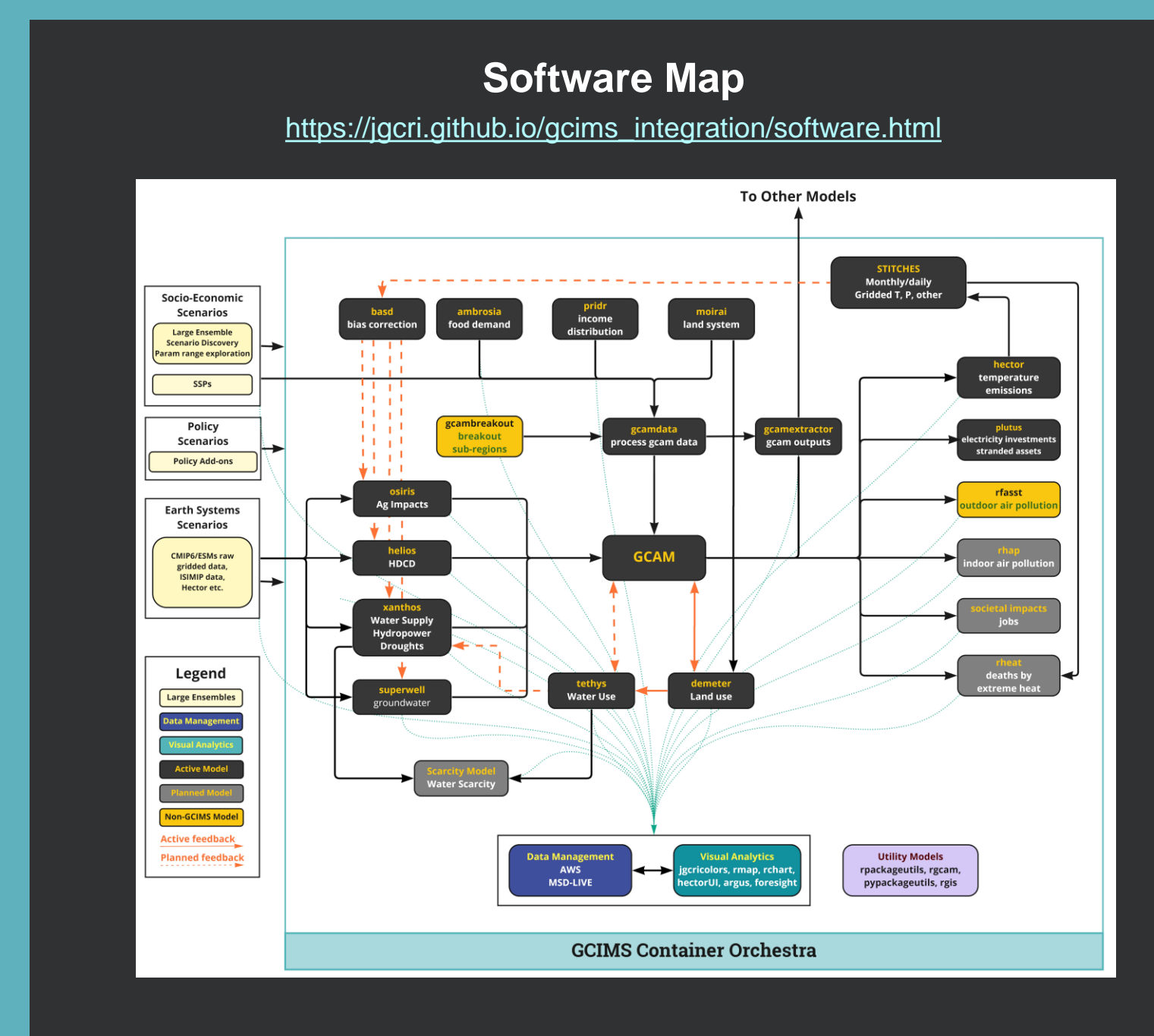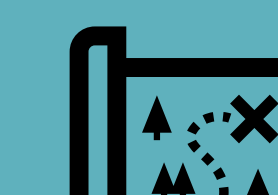
Clear Templates for Packages
https://github.com/JGCRI/pytemplate
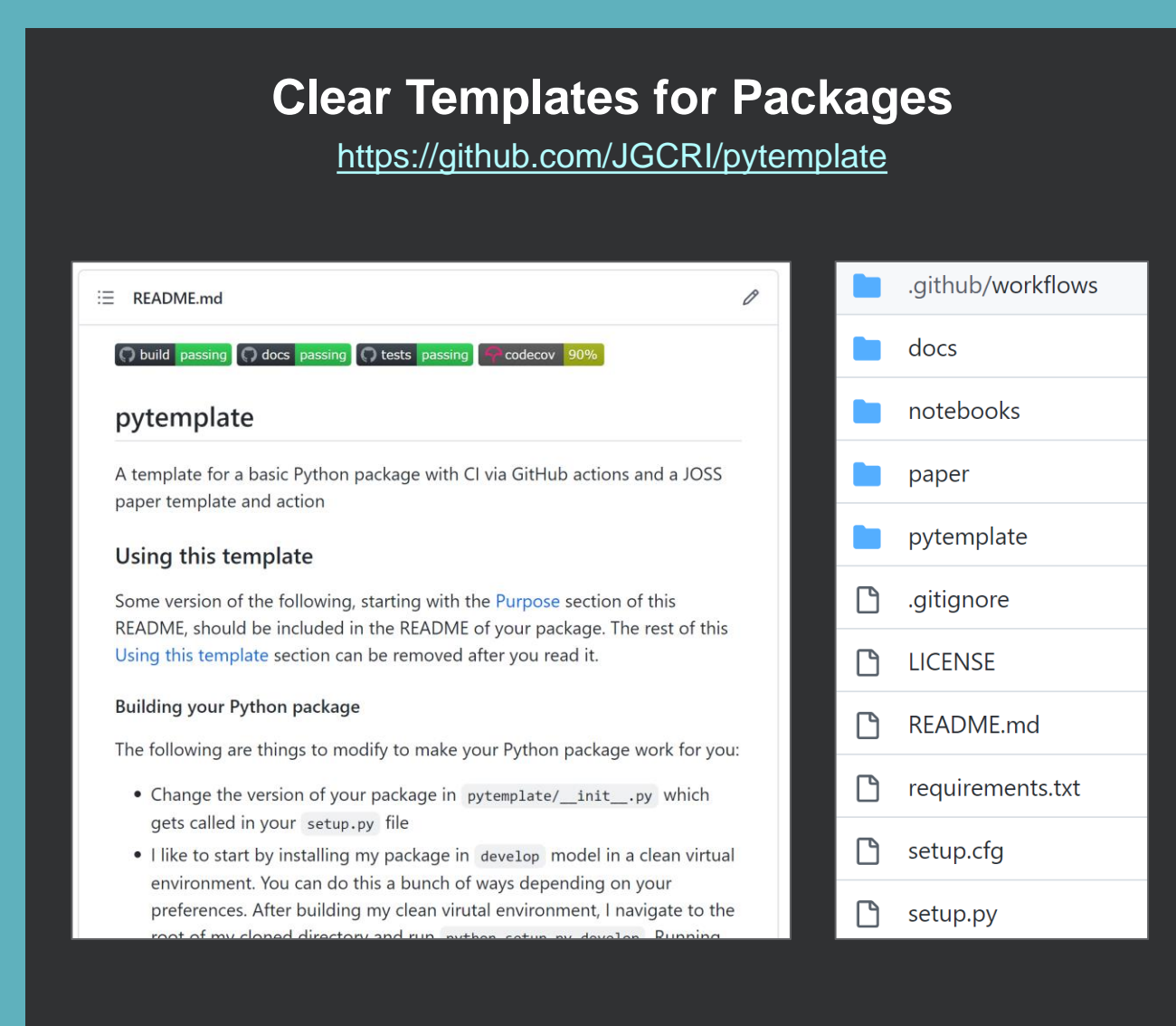
Package Design
- Folder Structure
- Continuous Integration
- Multi OS Build Checks
- Code Testing
- Documentation
- JOSS paper template

❷ Software Map

Software Map
https://jgcri.github.io/gcims_integration/software.html

Track down and Map
- Data sources
- Existing models
- Planned models
- Inter-linkages

❸ Software Health

Overall System Health Monitoring
https://jgcri.github.io/gcims_integration/software.html

Overall System Monitoring
- Continuous Integration
- Multi OS Build Checks
- Code Testing
- Code Coverage
- Documentation

❹ Reproducible Workflows

Reproducible Workflows
https://jgcri.github.io/khan-etal_2022_tethysSSPRCP/metarepo.html
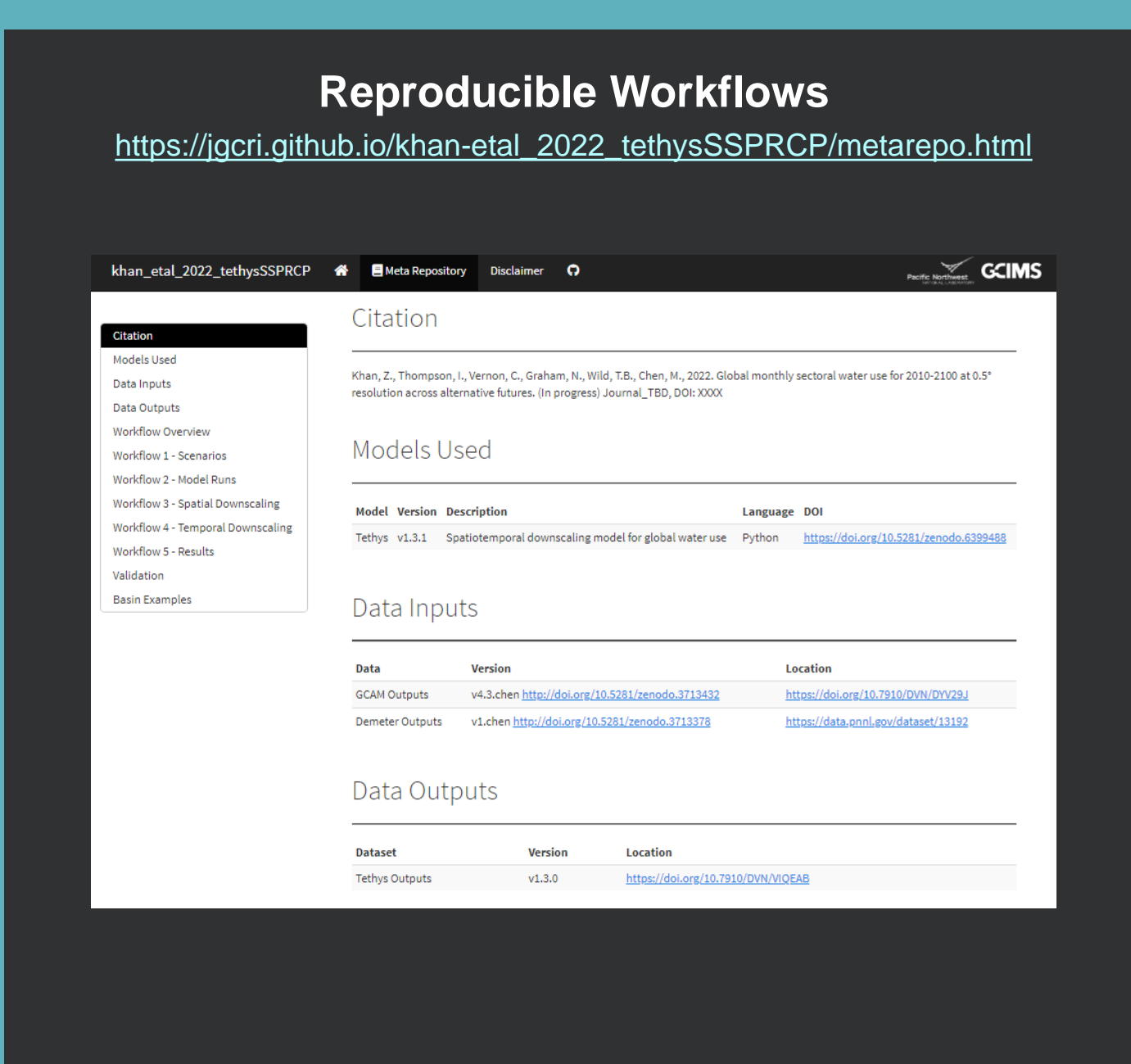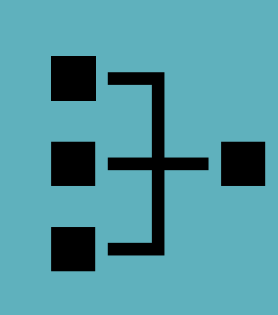
Meta-repo templates
- Inputs with DOI
- Outputs with DOI
- Workflows
- Validation
- Results

Containerization
- Scalablity
- Cloud/On-prem/hybrid

❺ Data Management

Persistent Data Management
https://data.msdlive.org/
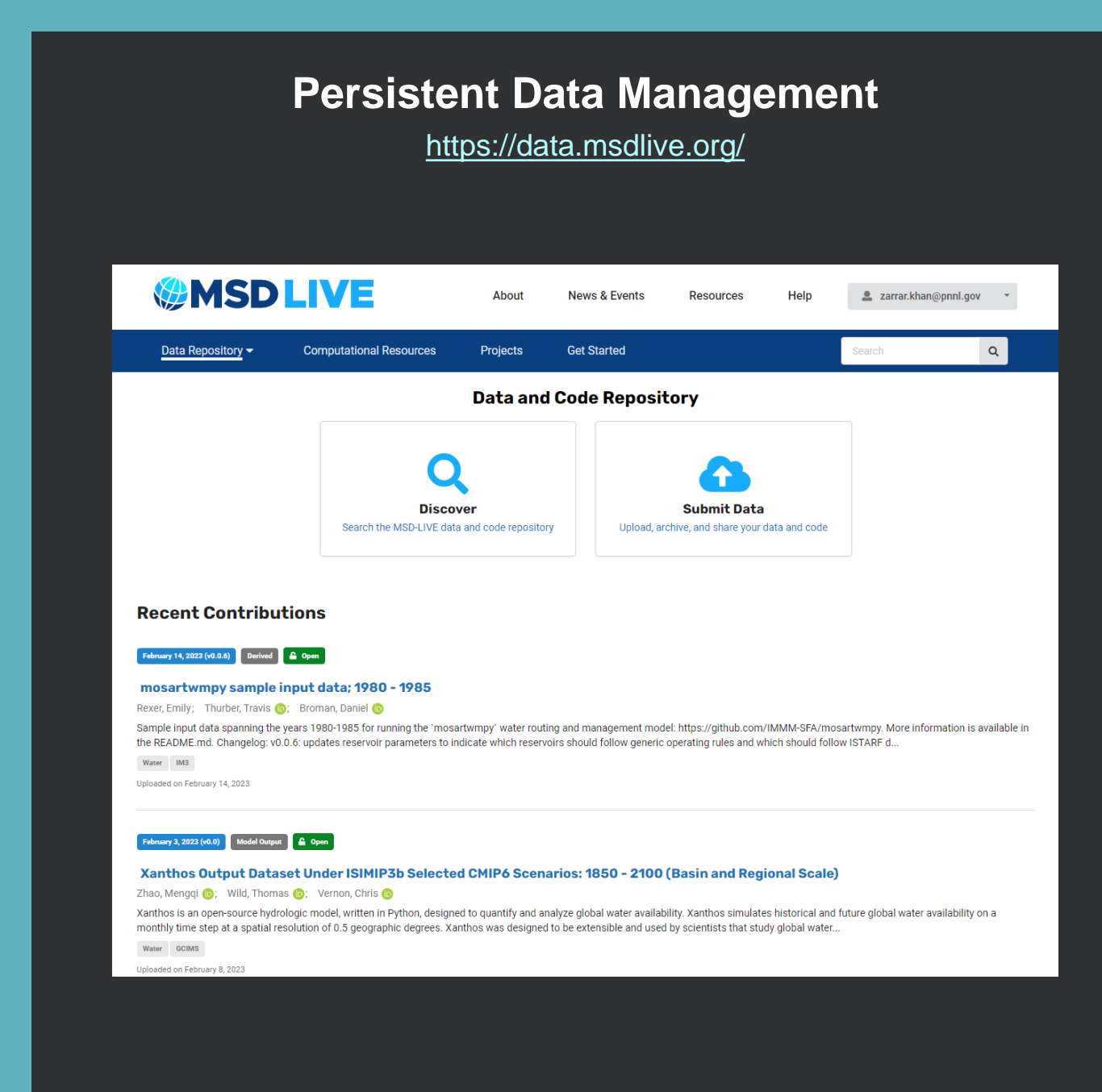
FAIR
- Findability
- Accessibility
- Interoperability
- Reusability

Key Features
- Minted with DOIs
- Meta-data
- Storage (MSD Live, Deception, Cloud)

❻ Persistent Insights
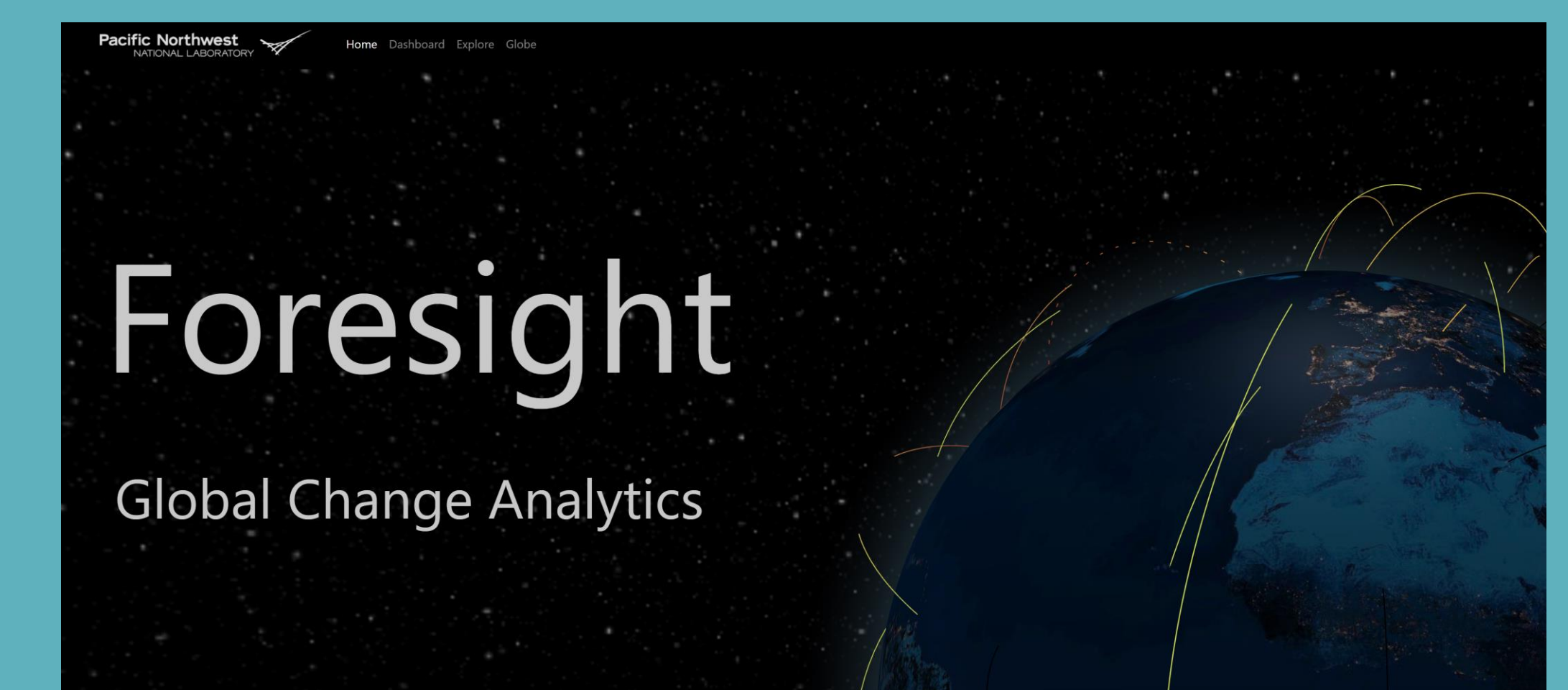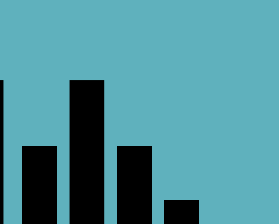
Curating Insights
- Persisting key figures across existing & future research
- Persisting key analysis across existing & future research
- Accessibility and reusability for a broad audience