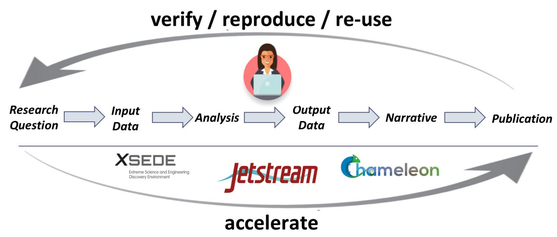
Publishing Transparent & Reproducible Computational Research with Whole Tale

Kyle Chard, Niall Gaffney, Mihael Hategan, Matthew B. Jones, Kacper Kowalik, Sara Lambert, Bertram Ludäscher, Timothy McPhillips, Jarek Nabrzyski, Victoria Stodden, Ian Taylor, Thomas Thelen, Matthew Turk, Craig Willis

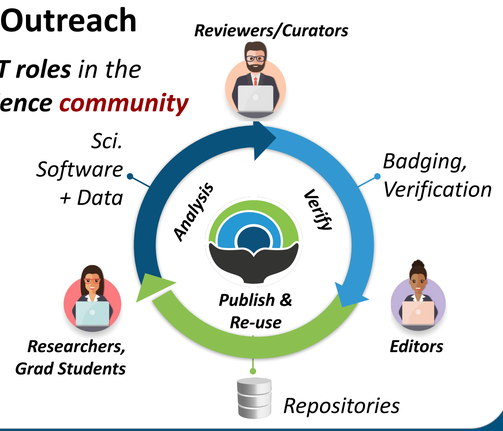
What is **Whole Tale**?

- NSF-funded **Data Infrastructure Building Blocks** (*DIBBs*) project*
- Open source platform to support **computational transparency** & **reproducibility**
- Simplifies the process of **creating, publishing & verifying** computational research artifacts (which we call **tales**)
- **Whole Tale = Infrastructure + Education + Outreach**

Whole Tale (WT)** from the **user's perspective



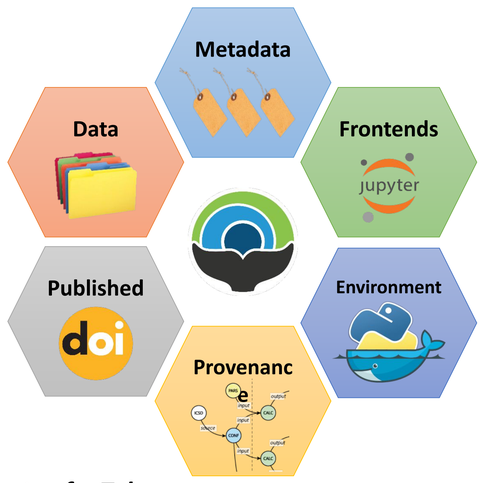
WT roles** in the **science community



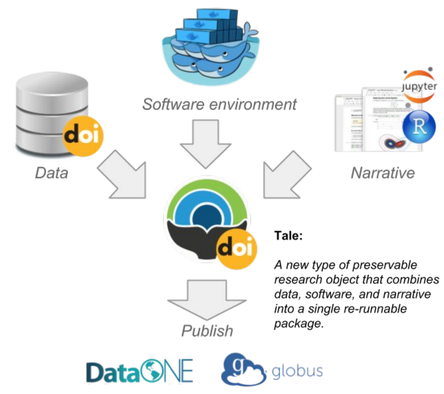
*#1541450: Merging Science & Cyberinfrastructure Pathways: The Whole Tale

What's **in** a Tale?

- **Tale** = executable **research object** that captures **data**, **code**, and complete **software environment** along with the (traditional) science **narrative**
- Standards-based tale **format**: BagIt-RO, JSON-LD metadata, …

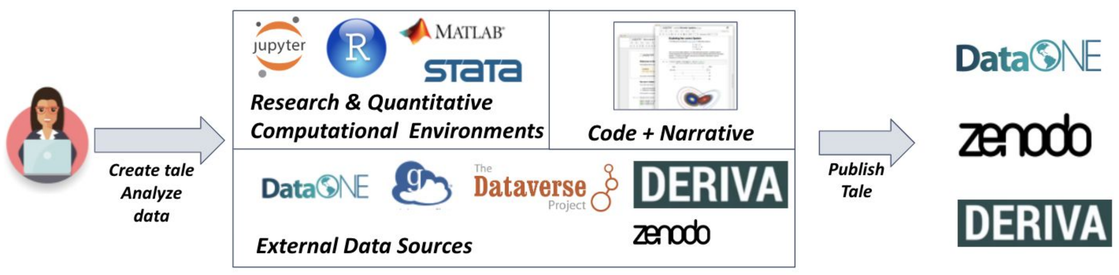


***Elements** of a Tale*

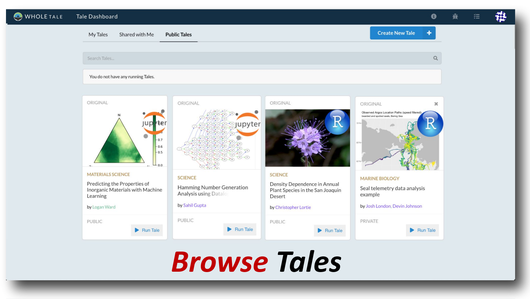


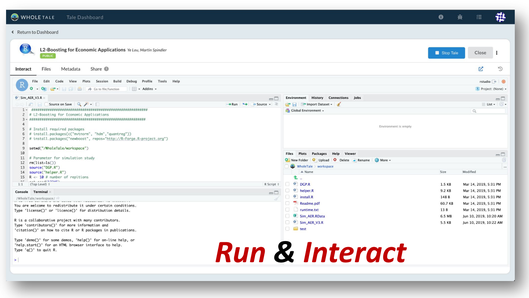
***Publishing** a Tale to DataONE, Globus, Zenodo, …*

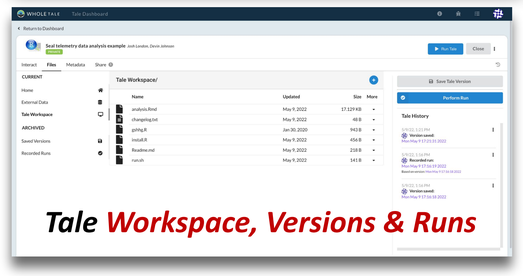
Whole Tale **Platform**

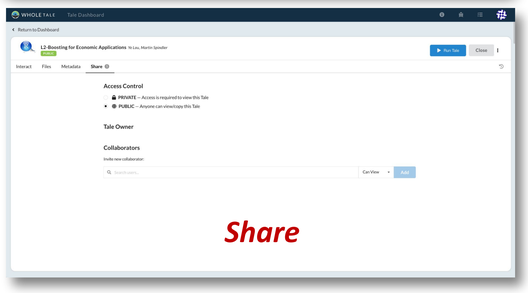


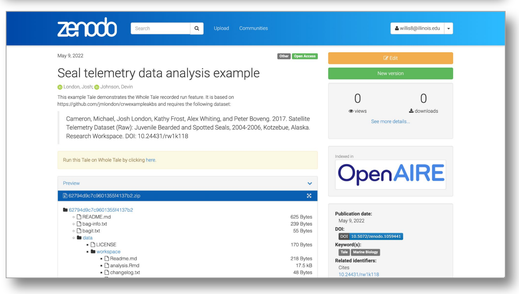
- **Authenticate** using your institutional identity
- **Access** commonly-used **computational environments** including **licensed** software
- Easily **customize** your environment
- Create or upload **your data and code**
- Cite and access externally **published data**
- **Share** with other users
- Deposit to **archival repositories**
- Get a **persistent identifier**



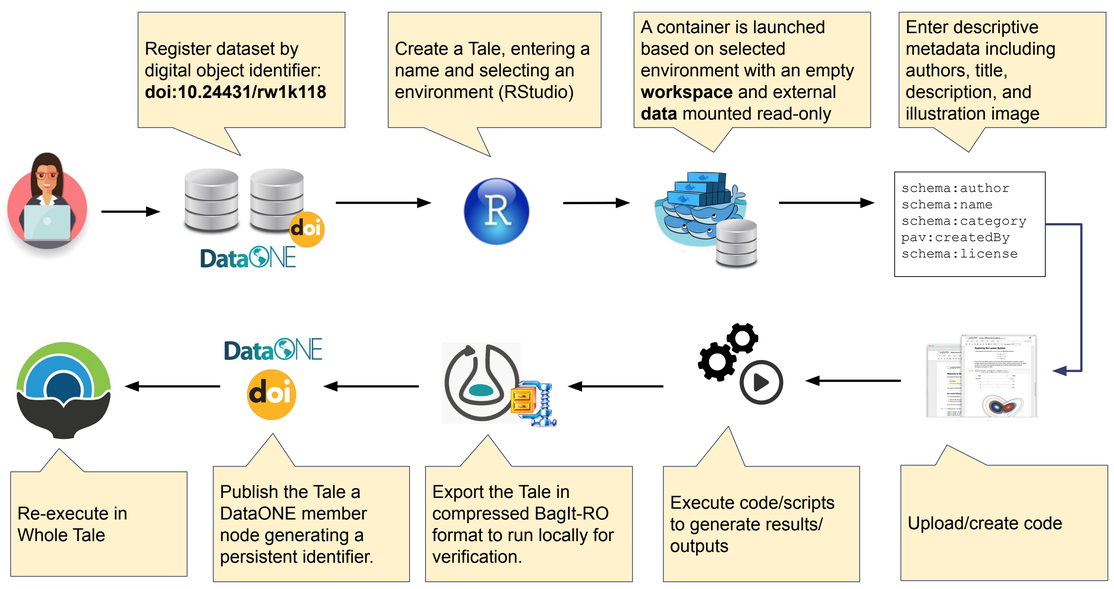








Tale Creation **Workflow**



Key Features

- Web-based platform for interactive and exploratory data-driven and computational- analysis using popular environments (JupyterLab, RStudio, MATLAB, Stata)
- Versioning, sharing, exporting, and publishing
- "Recorded runs" ensure that the specific version of code and data are captured along with outputs/results for publication
- Integration with popular research repositories
- Creates standards-based archival research artifacts

Additional Features (coming soon)

- Publishing container images
- Support for larger-scale computations (memory/cores)
- Improved accessibility/universal design
- Support for VS Code

ILLINOIS — UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN · THE UNIVERSITY OF CHICAGO · UCSB — UNIVERSITY OF CALIFORNIA SANTA BARBARA · TEXAS — The University of Texas at Austin · UNIVERSITY OF NOTRE DAME

https://wholetale.org
https://dashboard.wholetale.org
https://github.com/whole-tale