SDSC

EPFL ETH zürich

Reproducible & collaborative data science with the RENKU platform

renku

Oksana Riba Grognuz

June 2021

Outline

What is Renku

Why is Renku different

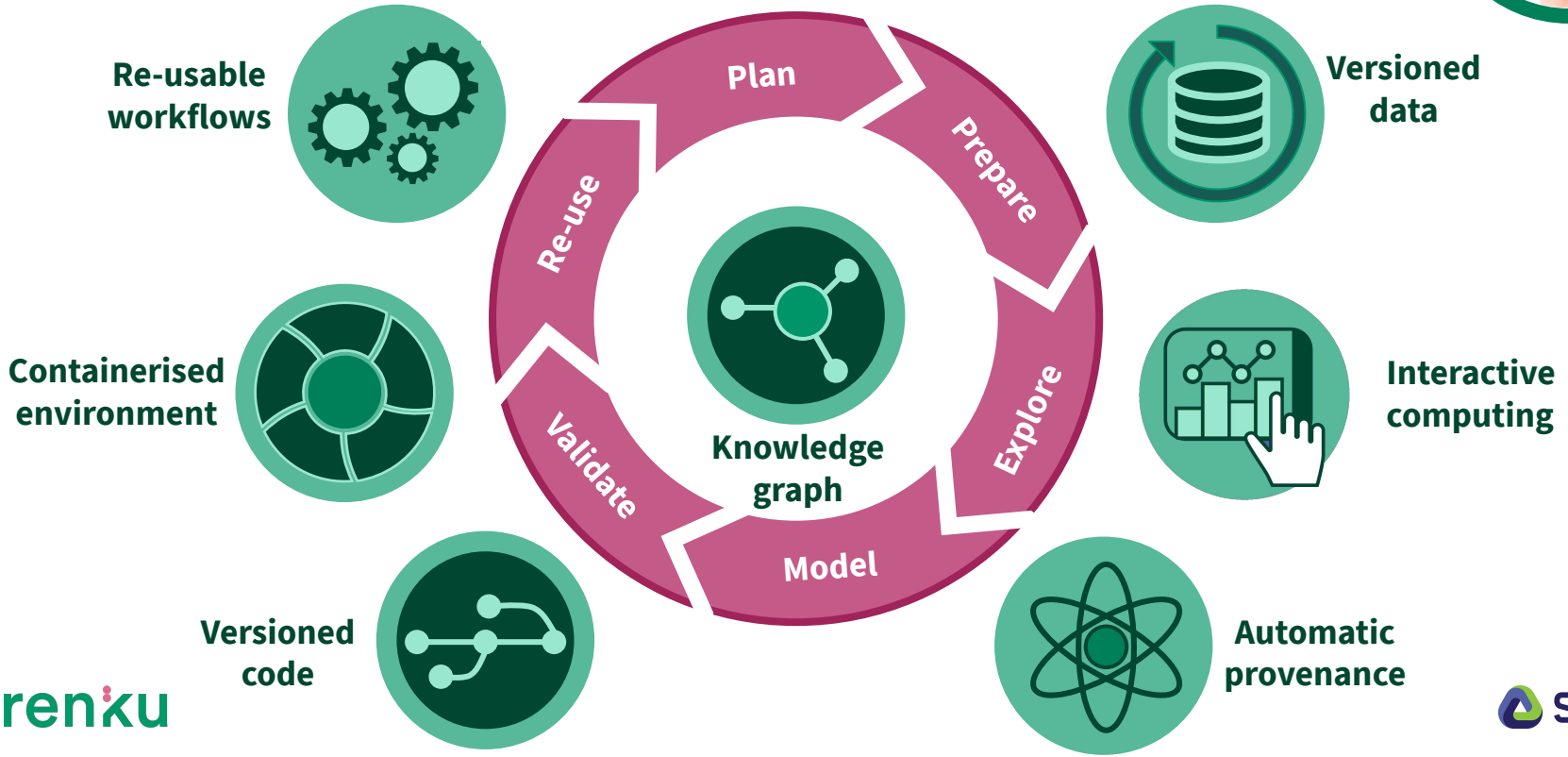
How Renku can help

Where can we get Renku

What is Renku ?

What is Renku?

Knowledge infrastructure for research life cycle



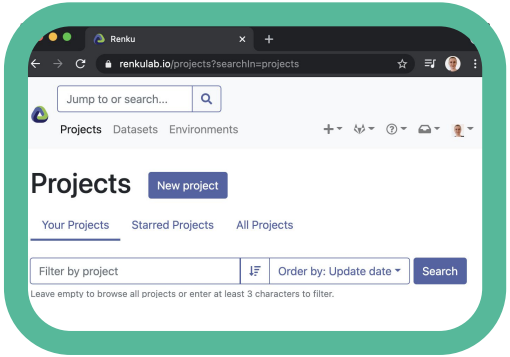
renku

SDSC

Renku is easy

Continue using the tools you are comfortable with

Web App



Interactive sessions

Command line

renku

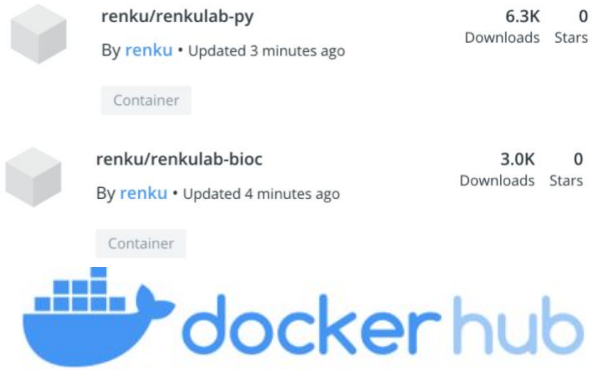
SDSC

Click and use

Maintained stack of Docker images

Web App

Interactive sessions



- **Single-click** project update in the web app
- **No overhead** for local support & operations

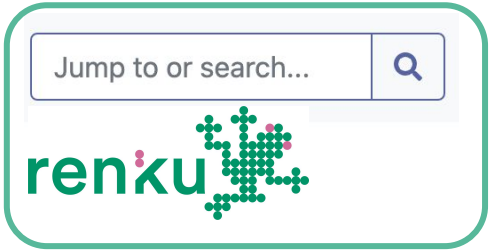
renku

Web App

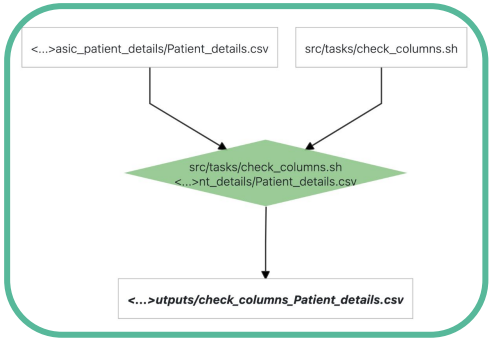
Find, Access, Reuse, Create and Collaborate

Web App

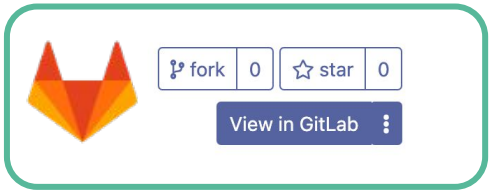
Find & access



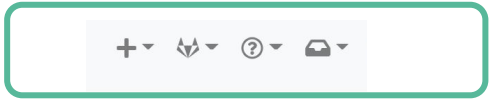
Explore & reuse



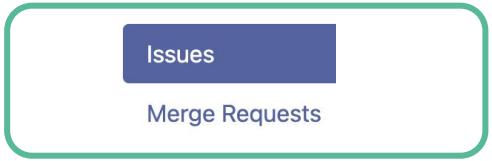
Edit & automate



Create



Collaborate



renku

Key Renku Ingredients

Renku

Clients

UI

Interactive sessions

CLI

Custom services

CORE

Notebooks

Knowledge Graph

API Gateway

Supporting services

GitLab

jupyterhub

Jena

KEYCLOAK

PostgreSQL

Kubernetes

pv

Volumes

pod

Container orchestration

svc

Service management

docker

jupyter

R

CWL

RDF

renku

SDSC

No

vendor or technology lock-in

Except for

docker

renku

SDSC

Why is Renku different?

From other

Feature-rich collaborative platforms

Workflow

Environment

Data

Code

CODE OCEAN

renku

gigantum

Pachyderm

Terra

DOMINO DATA LAB

WHOLE TALE

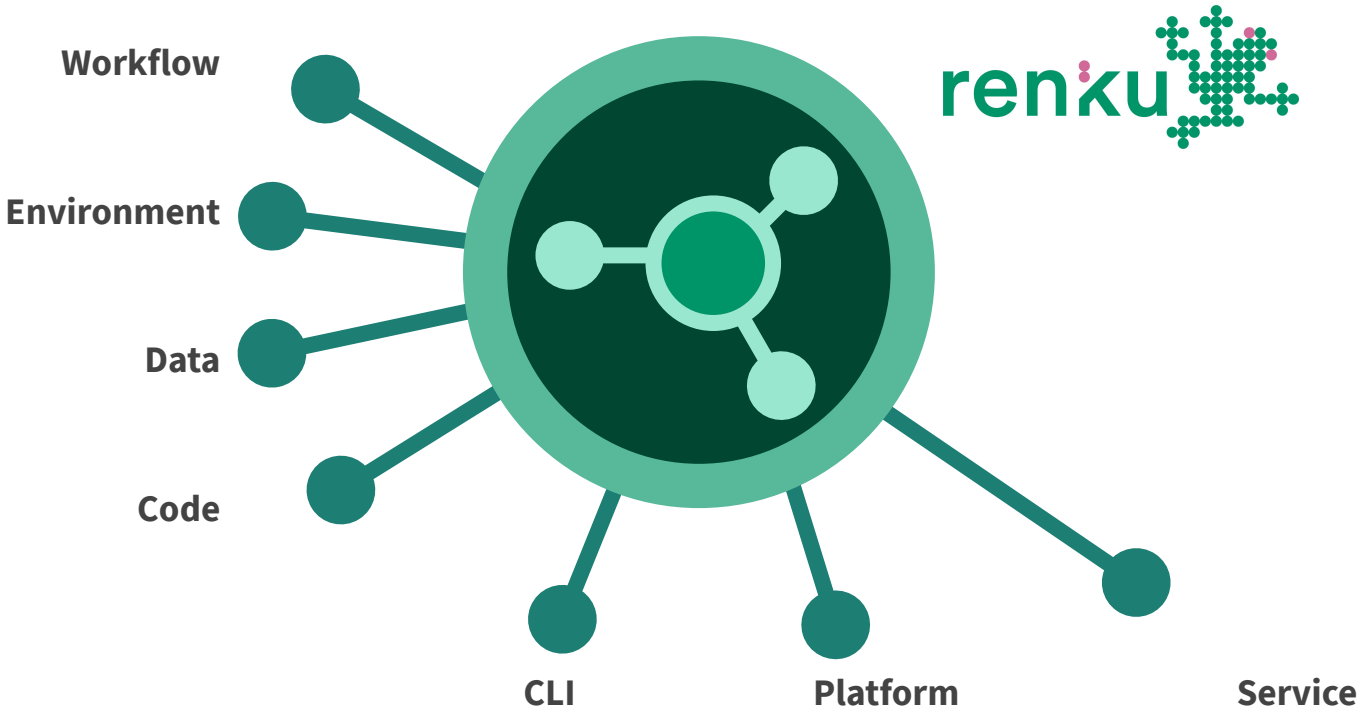
renku CLI

Platform

Service

SDSC

Renku Knowledge Graph



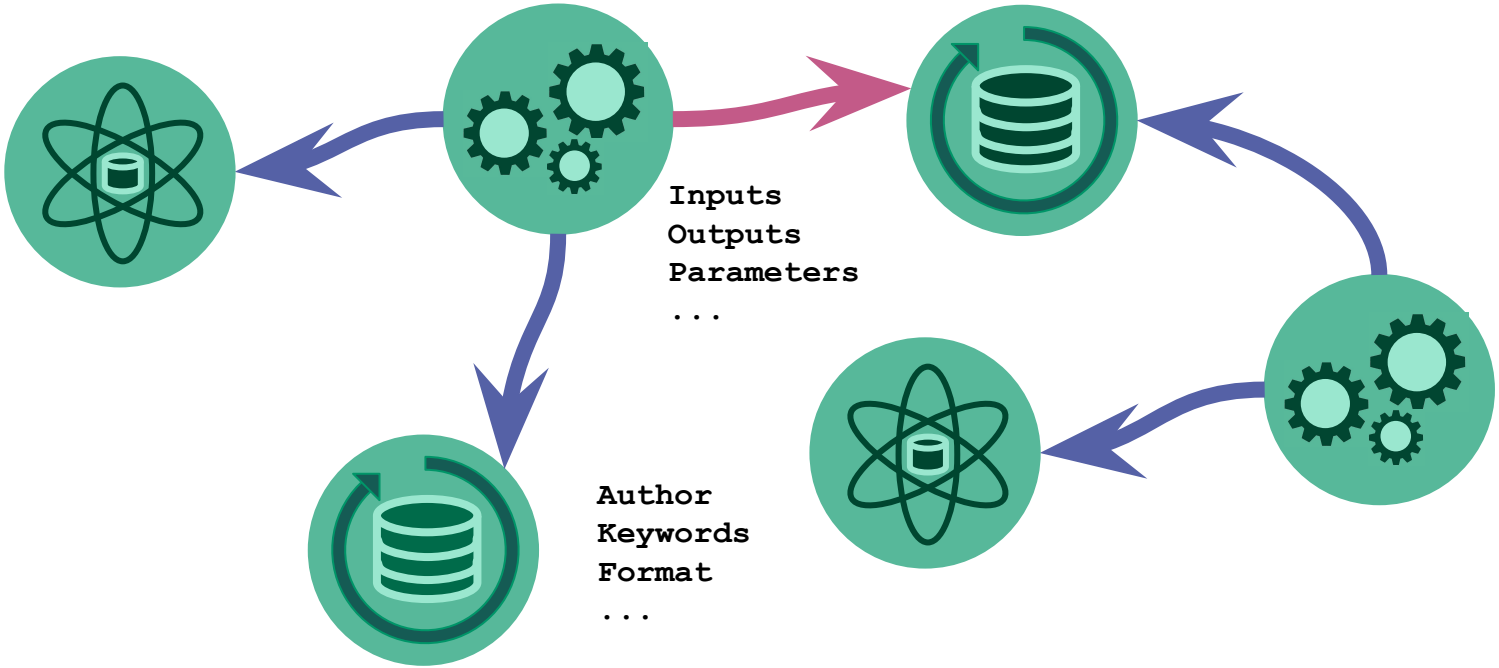
renku

renku

SDSC

Renku Knowledge Graph

Connects versioned digital objects



Uses

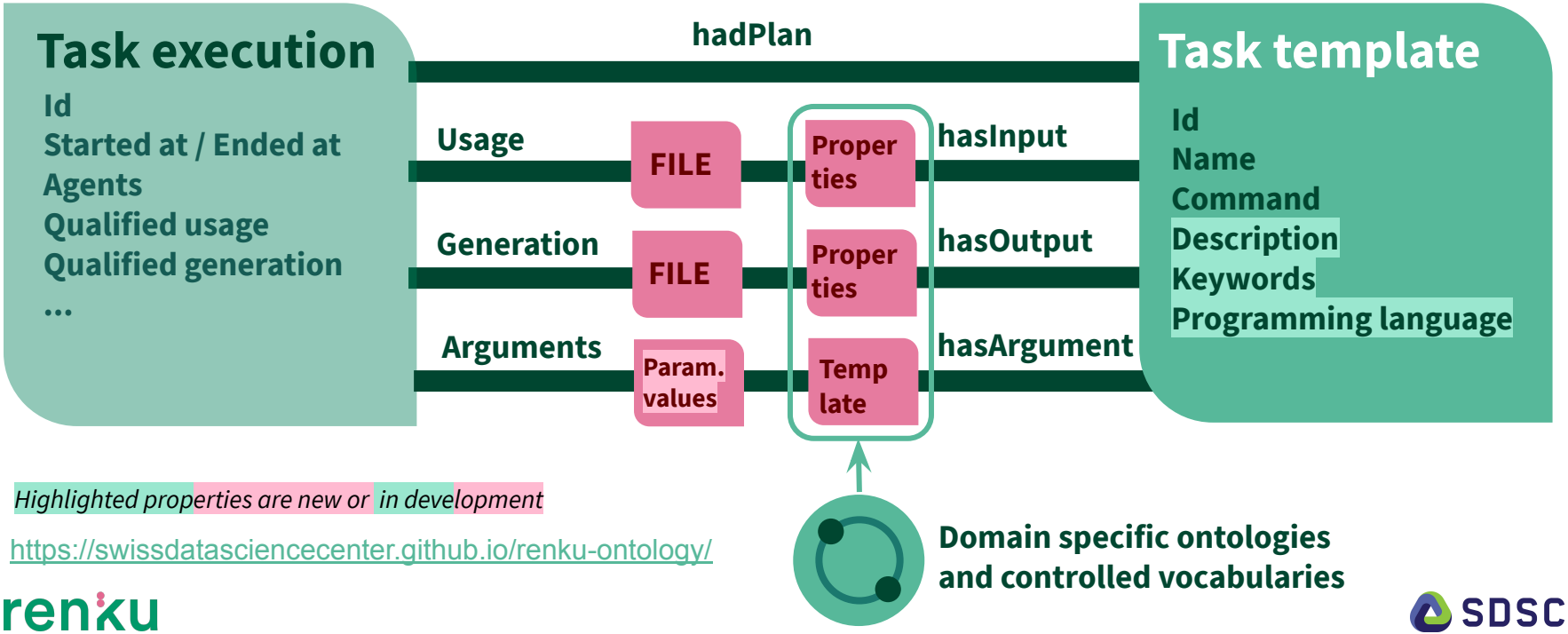
Creates

renku

SDSC

Powerful Provenance Ontology

Based on open standards



Highlighted properties are new or in development

https://swissdatasciencecenter.github.io/renku-ontology/

renku

SDSC

Coupled to Dataset Properties

Based on open standards

Highlighted properties are new or in development

Dataset
- Id
- Agents
- Name
- Description
- Keywords
- wasDerivedFrom
- Url
- Persistent identifier
- Domain types & properties

usedIn

hasPart

Digital document
- Id
- Date created
- Date updated
- Name
- Url
- External
- Checksum
- Encoding format
- Additional type

usedIn

Project
- Id
- Agents
- Date created
- Name
- Schemas
- Template
- Description
- Keywords

atLocation

FILE entity

Domain specific ontologies and controlled vocabularies

renku

SDSC

How Renku helps

Renku brings

Faster value from exploratory projects

Get value faster

By automatically tracking all (failed or not) exploratory avenues across projects and leveraging this knowledge

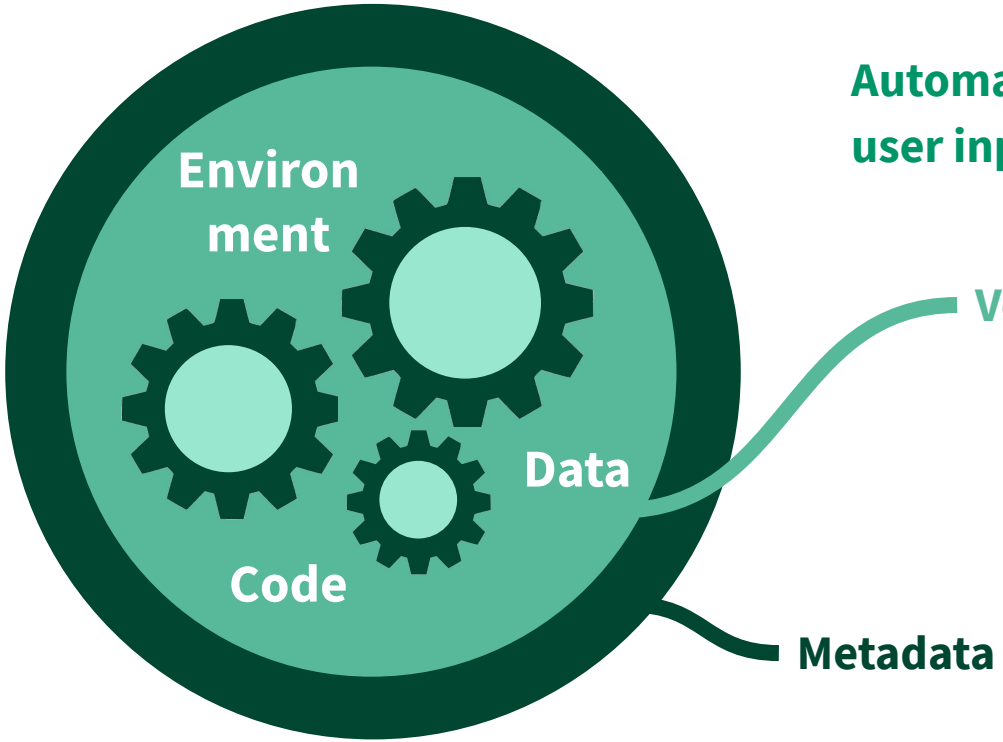
Knowledge graph

renku

SDSC

Renku workflows are

Findable, reusable and reproducible



Automatic workflow generation from user input

Versioned components

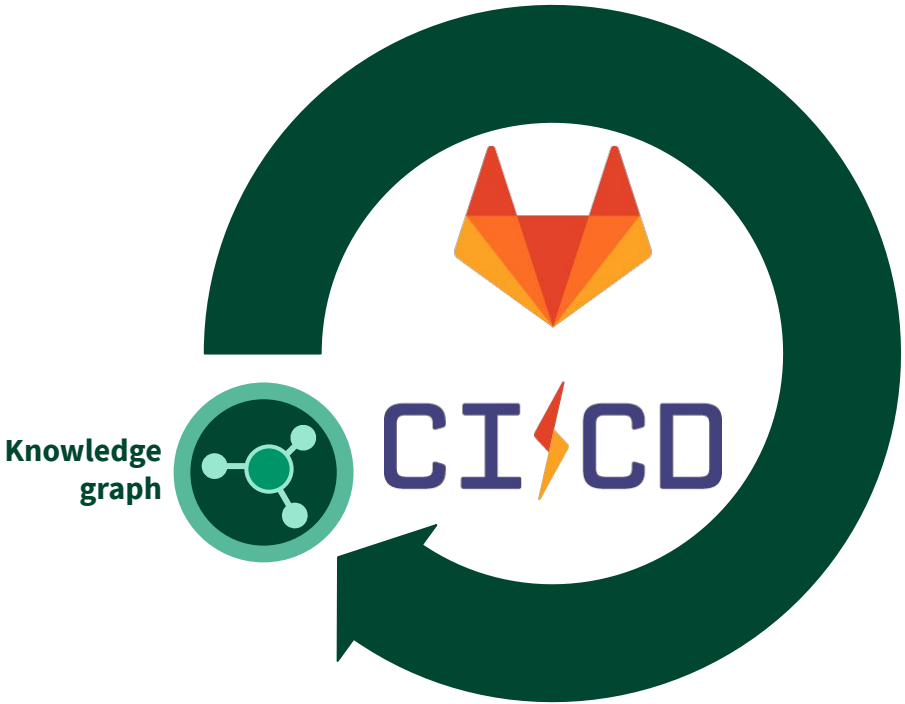
Metadata

renku

SDSC

Renku takes care of

Continuous compliance & self-assessment



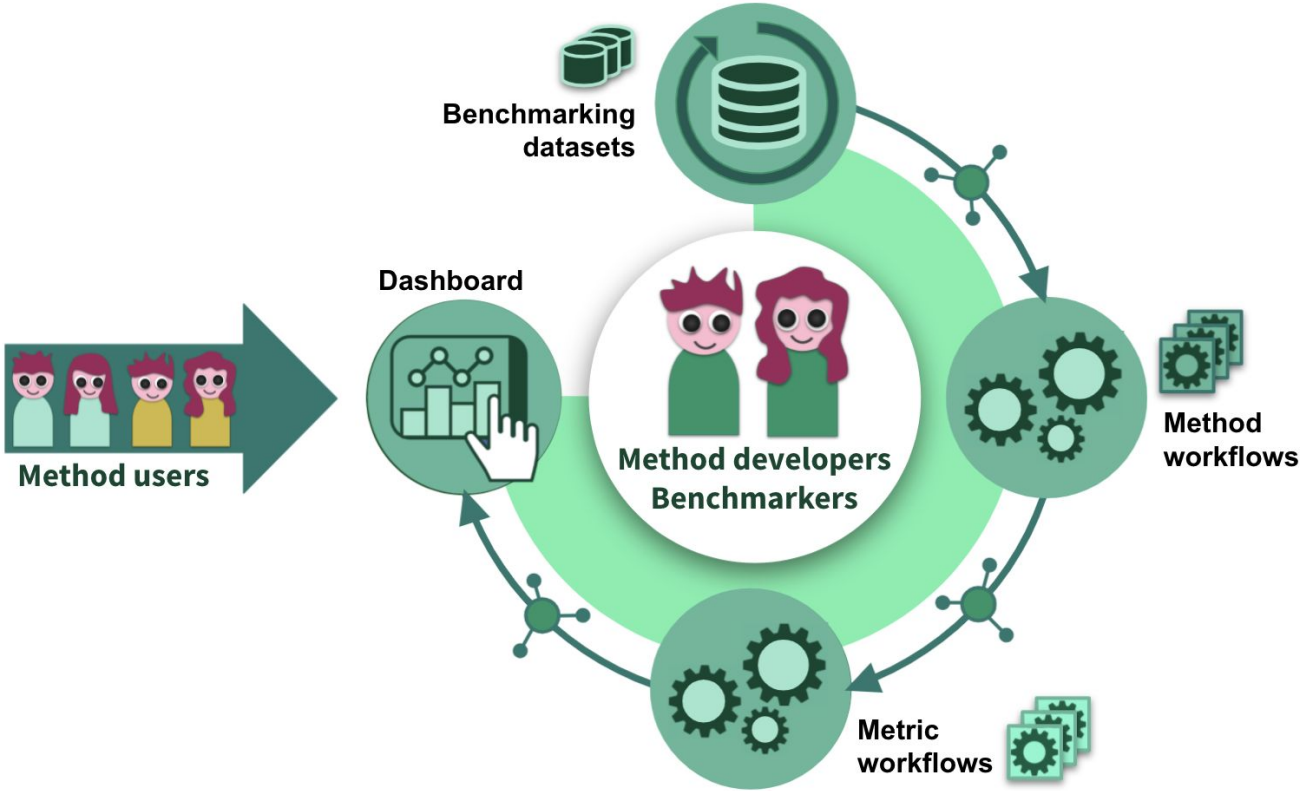
Continuously evaluating the analytical workflows against standards.

renku

SDSC

Example: community benchmarking

Modular & extensible framework on Renku



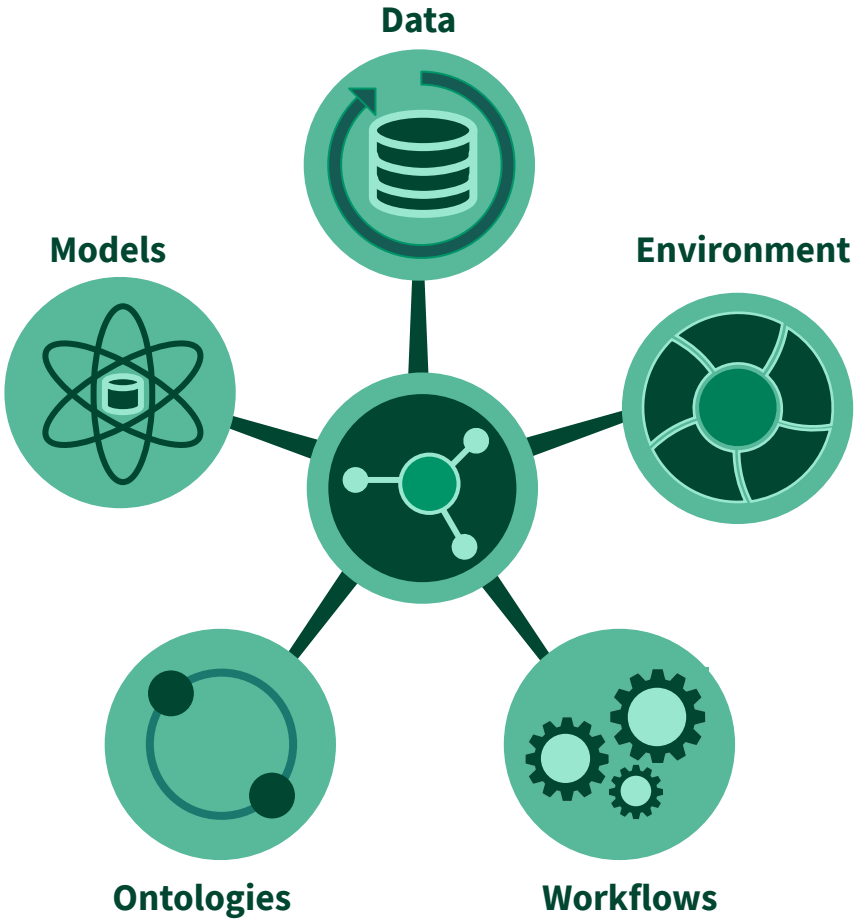
renku

SDSC

Renku Vision

A FAIR Ecosystem

Findable
Accessible
Interoperable
Reusable



Versioned digital objects
Reproducible & reusable
Domain-specific provenance
Portable & interoperable
Based on open standards

renku

SDSC

Where Renku

Where can I access Renku?

Command line + platform + service

```
pipx install renku
```

Command Line

Platform

Open source on GitHub

RenkuLab.io

Service

renku

SDSC

Where can I get Renku?

Custom deployment with personalised features

SaaS

Cloud

In-house

renku

SDSC

THANK YOU & get in touch

Ask questions @ Discourse
https://renku.discourse.group/

Chat about Renku @ Gitter
https://gitter.im/SwissDataScienceCenter/renku

Check out manuals & tutorials
https://renku.readthedocs.io/en/latest/getting_started.html

Request features & report bugs @ GitHub
https://github.com/SwissDataScienceCenter/renku

renku

SDSC