

Continous benchmarking of single-cell tools using RENKU

A framework for open and continuous community benchmarking of bioinformatic tools

Project Website: <https://renkulab.io/projects/omnibenchmark>
Source code: <https://renkulab.io/gitlab/omnibenchmark>
License: Apache License, Version 2.0

Almut Luetge^{1,2} Anthony Sonrel^{1,2} Oksana Riba Grognez³ Izaskun Mallona^{1,2} Charlotte Soneson^{2,4} Pierre-Luc Germain^{1,2,5} Federico Marini⁶ Gavin Lee³ Tao Sun³ Christine Choirat³ Rok Rožkar³ Mark D. Robinson^{1,2}

¹ Department of Molecular Life Sciences, University of Zurich, Switzerland
² SIB Swiss Institute of Bioinformatics, Switzerland
³ Swiss Data Science Center, ETH Zurich and EPFL, Switzerland
⁴ Friedrich Miescher Institute, Basel, Switzerland
⁵ D-HEST Institute for Neurosciences, ETH Zürich, Switzerland
⁶ Center for Thrombosis and Hemostasis (CTH), Johannes Gutenberg University Mainz, Germany

BACKGROUND

Benchmarking is critical for the development of performant computational methods and provides important insights for their application. In particular, methods research in **single-cell RNA-seq (scRNA-seq)** provides a perfect use case, where more than 900 tools have been developed in only a few years [1]. Benchmarking efforts are often **not coordinated**, **not extensible** and **not reproducible**.

Current benchmarking approaches have many **limitations**:

- they represent only a **snapshot** of the available methods at a certain time point.
- they become quickly **outdated**, even at the time of a publication.
- they suffer **low comparability**: different procedures, different datasets, different evaluation criteria.
- they lack **standardization**; each benchmarker decides the importance of performance criteria. – all of the above can lead to **different conclusions** among benchmarks made at different time points or at different groups.

CONCEPT

✓ We propose a new **modular** and **extensible** framework based on a free open-source analytic platform, **RENKU**, to offer a continuous and open **community benchmarking system**.

✓ The framework consists of **data**, **method** and **metric** repositories (or ‘modules’) that are connected via a **knowledge graph**.

✓ Performance results can be coupled with an **interactive dashboard** to be explored flexibly by any analyst looking for tool recommendations.

✓ New datasets, methods or metrics can be added by the **community**.

Key features of the developed benchmarking framework:

- Periodic updates to provide the community with the latest recommendations.
- Easily extensible with templates for data, methods or metrics.
- Follows the **FAIR** principles via Docker images, integration with Gitlab and full **provenance**.
- Uses **independent (docker) environments** and flexibly connects and shares modules within and between benchmarks.
- Flexibility to accommodate **various programming languages** used in bioinformatics: R, Python, Bash, Julia, etc.

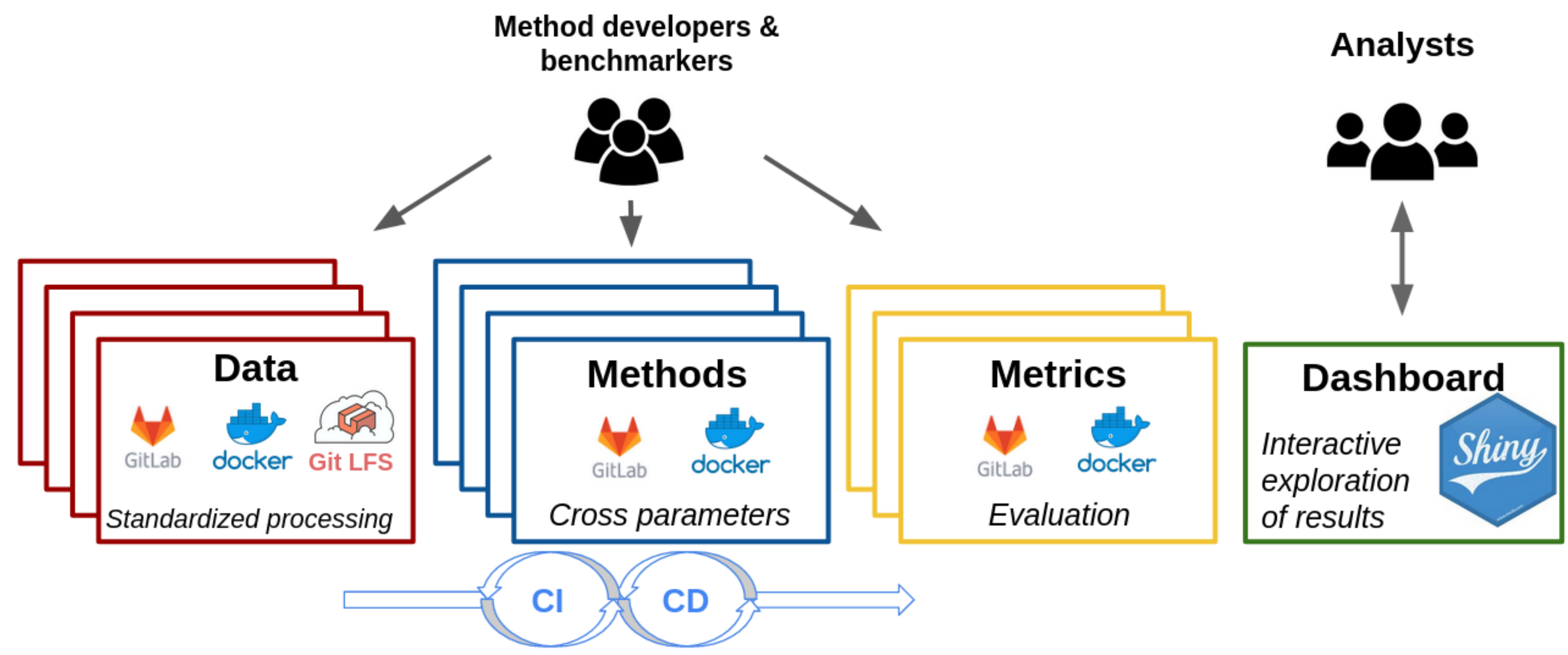


Figure 1: Overview of the proposed benchmarking framework on the RENKU platform. Each step of the benchmark consists of a set of repositories (a.k.a “modules”) that perform a coordinated task (e.g. standardized data processing). New datasets, methods or metrics can be easily integrated into the framework to trigger a continuous delivery to the analysts with up-to-date recommendations. Each component of the framework is automatically tracked by the knowledge graph of RENKU, utilizing Docker images to provide fully reproducible results.

PROTOTYPE

We are currently building a prototype (Figure 1) for community-based benchmarking of single-cell batch correction methods.

Our prototype consists of:

- Two datasets:
 - i. one with 2370 genes x 3613 cells separated in 2 batches; and,
 - ii. one with 1401 genes x 13575 cells separated in 3 batches [2, 3].
- Standardized preprocessing consisting of normalization, highly-variable genes selection, dimensionality reduction.
- An integration using the MNN method [4] (Figure 2).
- Evaluation of batch-effect correction using the LISI metric [5] and cms score [3].
- Exploring the results using the interactive browser bettr [6] (Figure 2).

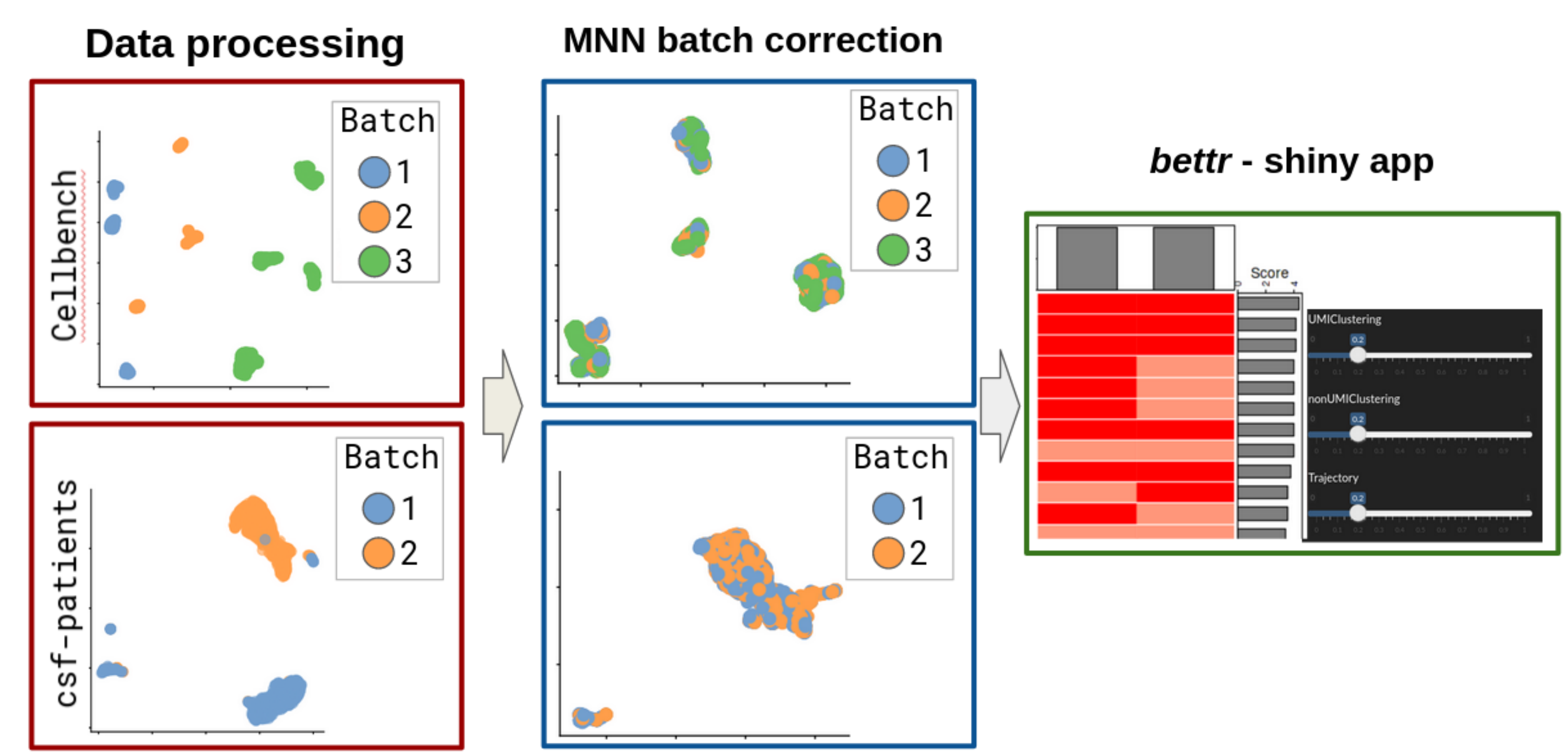


Figure 2: Example of the proposed framework applied to benchmark single cell RNA-seq batch correction methods. Datasets are uniformly processed and passed to a batch-correction method within a separate module. Results are displayed in a shiny app for an interactive exploration of the results. All modules are connected through the RENKU knowledge graph, which allows workflows to be tracked from a set of repositories using different environments and programming languages.



RENKU is an open and collaborative platform that provides a knowledge infrastructure for the entire research life cycle (Figure 3). The platform and its tools are built on top of a stack of open-source components and aims to make data science reproducible.

On **RENKU**:

- Since RENKU is based on cloud computing, fully interactive sessions can be accessed via a the browser, with no local installation needed.
- Versioning and containers ensure precise and reproducible computational environments.
- Datasets and workflows are automatically tracked in a **knowledge graph**, which can be queried from within a project, a group of projects or even across deployments.
- Workflows can be re-run or updated automatically when inputs such as datasets or scripts change.
- CI/CD can be leveraged to automate cumbersome tasks, such as fetching results of a piece of analysis to integrate into a dashboard.

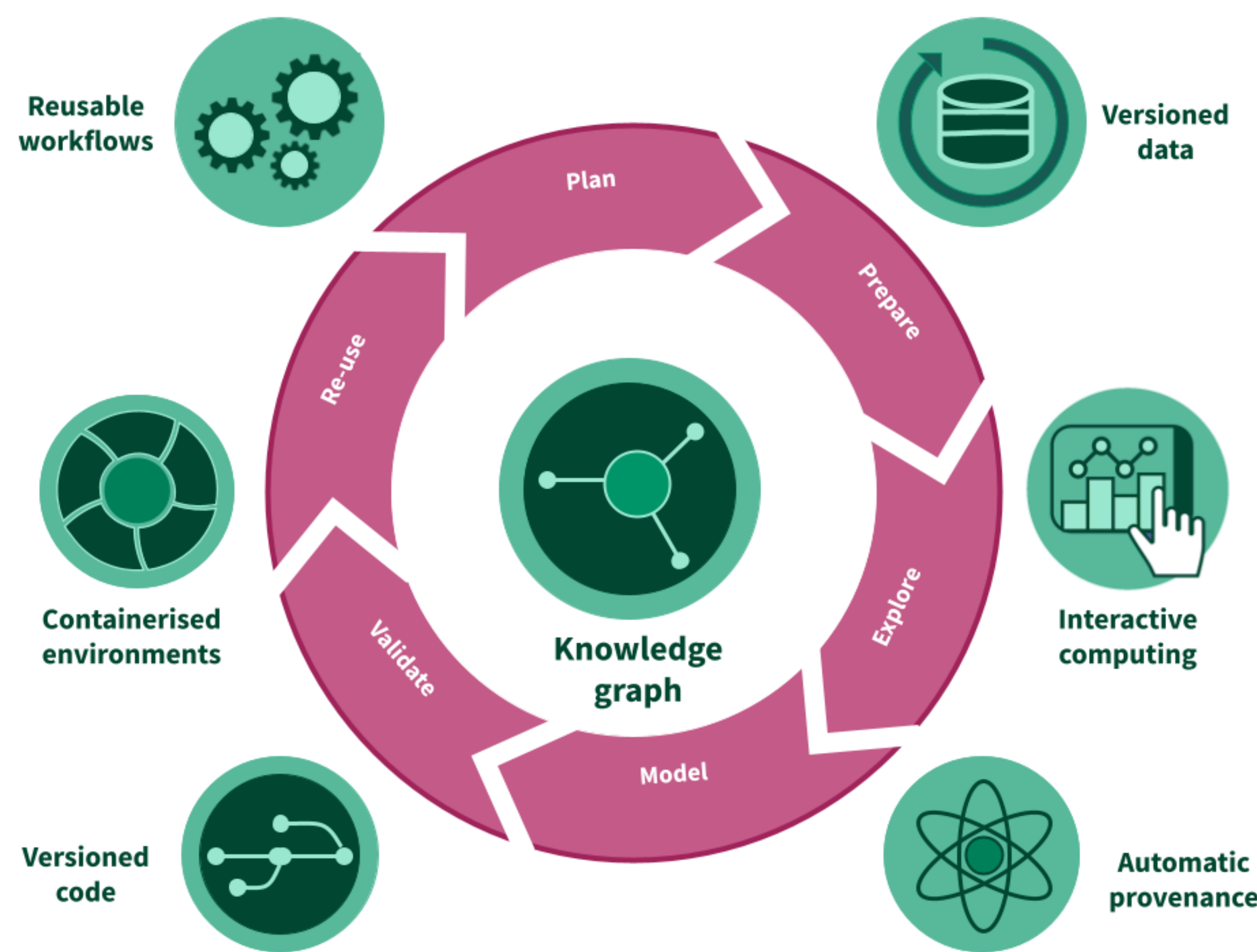


Figure 3: Overview of RENKU components that allow reproducible scientific work. Centered around the knowledge graph, renku provides code versioning, containerized environments, workflows, data tracking and automatic provenance.

References

1. Zappia L, Phipson B, Oshlack A. Exploring the single-cell RNA-seq analysis landscape with the scRNA-tools database. PLOS Computational Biology. 2018;14:e1006245. doi:10.1371/journal.pcbi.1006245.
2. Su S, Tian L, Dong X, Hickey PF, Freytag S, Ritchie ME. CellBench: R/Bioconductor software for comparing single-cell RNA-seq analysis methods. Bioinformatics. 2020;36:2288–90. doi:10.1093/bioinformatics/btz889.
3. Lütge A, Zyporych-Walczak J, Kunzmann UB, Crowell HL, Calini D, Malhotra D, et al. CellMixS: Quantifying and visualizing batch effects in single-cell RNA-seq data. Life Science Alliance. 2021;4. doi:10.26508/lsa.202001004.
4. Haghighi L, Lun ATL, Morgan MD, Marioni JC. Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. Nature Biotechnology. 2018;36:421–7. doi:10.1038/nbt.4091.
5. Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, Wei K, et al. Fast, sensitive and accurate integration of single-cell data with Harmony. Nature Methods. 2019;16:1289–96. doi:10.1038/s41592-019-0619-0.
6. Marini F, Soneson C. Bettr: A better way to explore what is best. 2021. <https://github.com/federicomarini/bettr>. Accessed 19 Jul 2021.

