

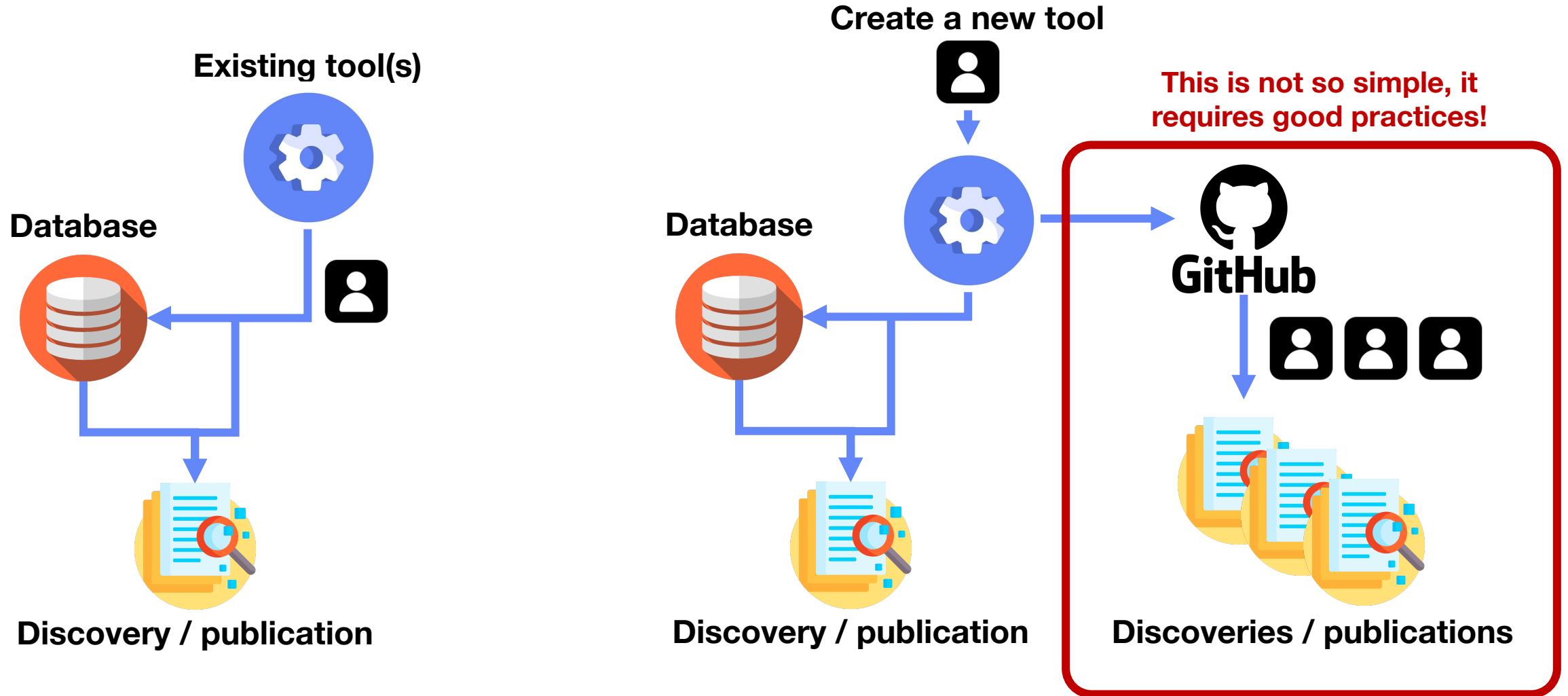
Link: menti.com/ale44temte5w

Good practices

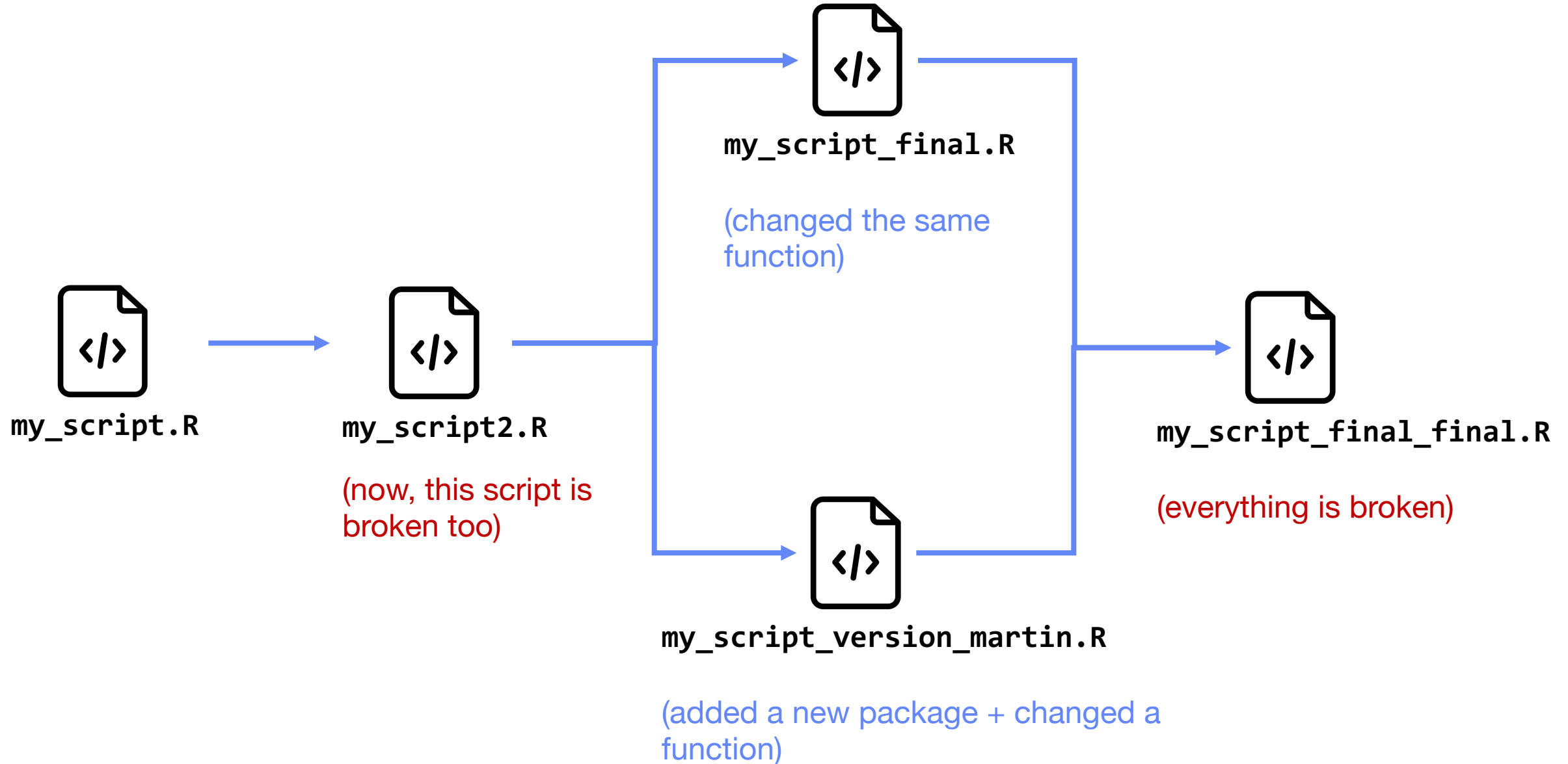
Nathalie Lehmann and Quentin Blampey



Different types of bioinformatics projects



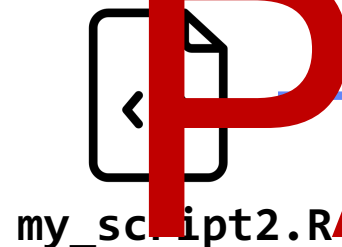
Why using a versioning tool (e.g., Git)



Why using a versioning tool (e.g., Git)

Which file did I use for this article figure? I need to change it!

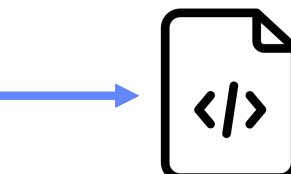
I deleted something in the last version, I can't find it anymore!



(now, this script is broken too)

Which package versions did I use? I can't run my code anymore!

Why are the results different? I need to update it slightly for the revision



(ok, now it seems to work)

Why using a versioning tool (e.g., Git)




+

Source code hub



Writing readable and understandable code

What is this function doing? What is “data”? To what the arguments correspond? What are the possible values? What is the result?



```
def run(data, c=7, a=13.2, m="dense"):  
    ... # some code  
    return results
```

Provide meaningful function AND variable names (self-understandable code)
Document your input types and available values
Describe precisely the output

Writing readable and understandable code

If you can: try to write an online documentation

Sopa

Home

Getting Started

Tutorials >


CLI

API >

Frequently asked questions

Cite us

Spatial-omics pipeline and analysis



Built on top of [SpatialData](#), Sopa enables processing and analyses of image-based spatial omics using a standard data structure and output. We currently support the following technologies: Xenium, MERSCOPE, CosMX, PhenoCycler, MACSIMA, Hyperion. Sopa was designed for generability and low memory consumption on large images (scales to 1TB+ images).

<https://gustaveroussy.github.io/sopa/>

Maintaining a package over time

Time



numpy 1.5.3

numpy 1.5.4

numpy 1.6.0

numpy 2.0.0

numpy 2.0.1

pandas 1.2.7

pandas 1.2.8

pandas 1.3.0

pandas 1.3.1

pandas 2.0.0

pytorch 1.0.1

pytorch 1.0.2

pytorch 1.0.3

pytorch 1.1.0

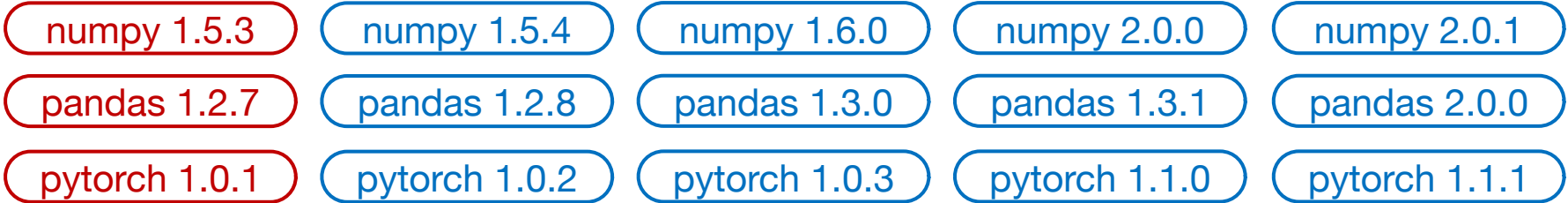
pytorch 1.1.1

Maintaining a package over time

Publish a package

my_package 1.0.0

Time →



Deprecated versions

Maintaining a package over time

Publish a package

my_package 1.0.0

Time →

numpy 1.5.3

numpy 1.5.4

numpy 1.6.0

numpy 2.0.0

numpy 2.0.1

pandas 1.2.7

pandas 1.2.8

pandas 1.3.0

pandas 1.3.1

pandas 2.0.0

pytorch 1.0.1

pytorch 1.0.2

pytorch 1.0.3

pytorch 1.1.0

pytorch 1.1.1

Deprecated versions

Maintaining a package over time

If you don't maintain your code, it will eventually become deprecated!

Publish a package

my_package 1.0.0

Time →

numpy 1.5.3

numpy 1.5.4

numpy 1.6.0

numpy 2.0.0

numpy 2.0.1

pandas 1.2.7

pandas 1.2.8

pandas 1.3.0

pandas 1.3.1

pandas 2.0.0

pytorch 1.0.1

pytorch 1.0.2

pytorch 1.0.3

pytorch 1.1.0

pytorch 1.1.1

Deprecated versions

Backward compatibility

my_package 1.0.0

```
def annotate_cell_types(X: np.ndarray, d: dict):  
    ...
```

New version

my_package 1.1.0

```
def get_cell_types(X: np.ndarray, markers: dict):  
    ...
```

It will break everything! E.g. Seurat

Solution 1:

```
def annotate_cell_types(X: np.ndarray, d: dict):  
    warning("This function is deprecated and will be removed in Jan. 2025")  
  
    return get_cell_types(X, markers=d)
```

Backward compatibility

Solution 2: Release notes

v0.2.2

[0.2.2] - 2024-08-07

Major

- New disk format for shapes using `GeoParquet` (the change is backward compatible) [#542](#)

Minor

- Add `return_background` as argument to `get_centroids` and `get_element_instances` [#621](#)
- Ability to save data using older disk formats [#542](#)

Fixed

- Circles validation now checks for inf or nan radii [#653](#)
- Bug with table name in torch dataset [#654](#) [@LLehner](#)

Testing and continuous integration (CI)



Which change broke what? And when? On which OS? On which use case?

Testing and continuous integration (CI)

Testing = write some code that tests if everything is working

```
===== test session starts =====  
platform linux2 -- Python 2.7.3 -- py-1.4.20 -- pytest-2.5.2 -- /usr/bin/python  
collected 9 items  
  
test/test_basic_integers.c:14: test_some_integers() PASSED  
test/test_basic_integers.c:15: test_some_integers() PASSED  
test/test_basic_integers.c:21: test_more_integers() FAILED  
test/test_basic_integers.c:22: test_more_integers() FAILED  
test/test_basic_strings.c:16: test_some_strings() PASSED  
test/test_basic_strings.c:17: test_some_strings() PASSED  
test/test_basic_strings.c:26: test_more_strings() FAILED  
test/test_basic_strings.c:27: test_more_strings() FAILED  
test/test_basic_strings.c:28: test_more_strings() PASSED
```

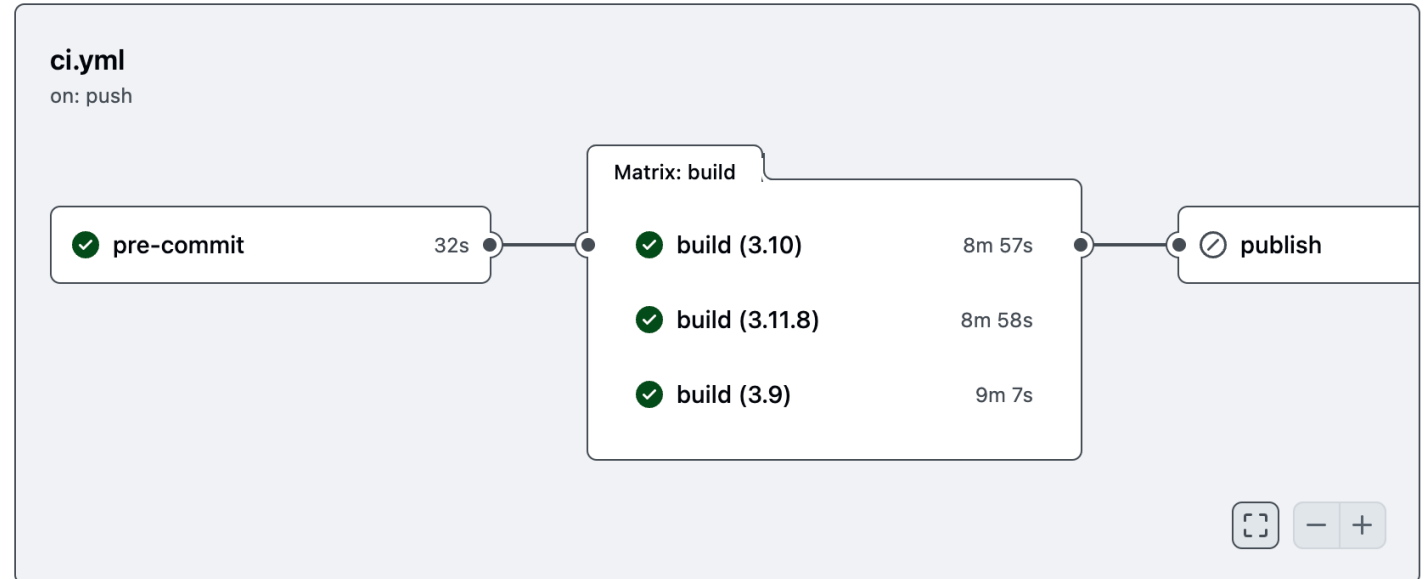
It tells you when some **specific use cases** are broken

When everything is “**passing**”, you’ll be more confident that your code is stable

Testing and continuous integration (CI)

.yaml file

```
on:  
  push:  
    branches: [master]  
  
jobs:  
  build:  
    runs-on: ubuntu-latest  
  
    - name: Install dependencies  
      run: pip install -e .  
  
    - name: Tests  
      run: poetry run pytest  
  
    - name: Deploy doc  
      run: mkdocs gh-deploy
```



<https://github.com/gustaveroussy/sopa/actions/runs/11442161900>

Other keywords to dive deeper

- Precise error handling (e.g., using `assert`)
- PEP style guide
- Formatting and linting (e.g., `black`, `flake8`, `ruff`, ...)
- `pre-commit` tool
- Typing

Contributing to open-source

scverse / spatialdata

Code Issues 152 Pull requests 10 Discussions Actions Projects 1


Beta Give feedback

is:issue state:open Labels Milestones New issue

| Open 152 | Closed 225 | Author | Labels | Projects | Milestones | Assignees | ... |
|--|------------|---------------|---|----------|------------|-----------|-----|
| Detach SpatialData object from its Zarr store | | Marius1311 | | | | 0 | |
| Dask dataframe vs Polars | | gouinK | | | | 0 | |
| Expose interpolation kwargs for dask_image.ndinterp.affine_transform() | | LucaMarconato | | | | 0 | |
| Performance improvement of query and aggregate with cuspatial | | LucaMarconato | method: aggregation 12 34 method: query performance 🚀 | | | 0 | |
| Respect xarray coordinates for ImageContainers | | conradfoo | method: transforms 1 | | | 4 | |

Contributing to open-source

Open


 conradfoo opened 3 weeks ago edited by conradfoo · Edits · ...

Is your feature request related to a problem? Please describe.
Currently, if I parse an xarray DataArray into an image container, the coordinates are overwritten with pixel coordinates even if I do not specify a transformation.

Describe the solution you'd like
Transfer over the xarray coordinates from the initial xarray into the spatialdata object

Describe alternatives you've considered
It is possible (but annoying) to construct a transformation object from my coordinate list and feed that into the constructor.

Additional context
I'm using spatialdata-0.2.3
Thanks!



Assignees
No one assigned

Labels
method: transforms

Type
No type


Projects
SpatialData tasks
Status Done

Milestone
No milestone

Relationships
None yet


Development
No branches or pull requests

Notifications

 LucaMarconato 3 weeks ago Member · ...

Hi @conradfoo, thanks for reporting.

We actually want to implement this feature. Tracked here [#308](#) (in "there will be a helper function to convert back and forth between BaseTransformation (and so also NGFFBaseTransformation) and xarray coordinates."). I will increase the priority of this feature request.



You can answer to other people issues (even if it's not your own package)
If you fixed your issue, add a comment to explain how you fixed it

Contributing to open-source

<https://scverse.zulipchat.com/>

The screenshot displays the Zulip chat interface. At the top, there's a navigation bar with 'Conversations récentes' and a search bar labeled 'Rechercher'. Below this, a sidebar on the left lists navigation options: 'VUES' (Conversations récentes, Boîte de réception, Fil groupé, Mentions, Réactions, Messages favoris, Brouillons) and 'MESSAGES DIRECTS' (Paul Kiessling, giovanni, Katie Miller, Luca Ma..., Anna Schaar, Mikaela Koutrouli, Notification Bot). Below the sidebar is a 'CANAUUX' section with a list of channels like '# 2024-09 conference munich', '# 2025-xx-xx [org] hackathon owkin', etc.

The main area shows a list of recent conversations with the following columns: Canal, Sujet, Participants, and Heure. The current conversation is '# spatialdata-dev' with the subject 'datashader aggregation in spatialdata-plot'. Other conversations include 'general' (scverse community meetings), '2025-xx-xx [org] hackathon owkin' (Klckoff meeting), 'repo-management' (onetimesecret), 'random' (Decoupler), '2025-xx-xx [org] hackathon owkin' (hybrid), 'random' (merge announcement and general), 'spatial' (spatialdata query returns region larger than min/max limit), 'scvi-tools' (AnnCollection), and 'spatial' (spatialdata meetings).

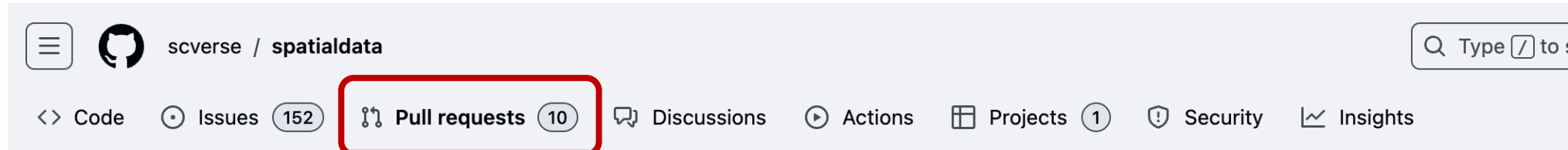
At the bottom, there are three buttons: 'Composer un message', 'Démarrer une nouvelle conversation', and 'Nouveau message direct'.

| Canal | Sujet | Participants | Heure |
|----------------------------------|--|----------------|-----------------|
| # spatialdata-dev | <u>datashader aggregation in spatialdata-plot</u> | 4 participants | il y a 4 heures |
| general | scverse community meetings | 2 participants | il y a 2 jours |
| 2025-xx-xx [org] hackathon owkin | Klckoff meeting | 4 participants | il y a 2 jours |
| repo-management | onetimesecret | 2 participants | il y a 3 jours |
| random | Decoupler | 4 participants | il y a 3 jours |
| 2025-xx-xx [org] hackathon owkin | hybrid | 3 participants | il y a 3 jours |
| random | merge announcement and general | 4 participants | il y a 3 jours |
| spatial | spatialdata query returns region larger than min/max limit | 5 participants | il y a 4 jours |
| scvi-tools | AnnCollection | 2 participants | il y a 5 jours |
| spatial | spatialdata meetings | 2 participants | il y a 7 jours |

Contributing to open-source (going further)

Pull Request example:

<https://github.com/scverse/spatialdata/pull/617>



scverse / spatialdata

Code Issues (152) **Pull requests (10)** Discussions Actions Projects (1) Security Insights

refactor bounding box #617

Merged

LucaMarconato merged 12 commits into `main` from `giovp/refactor_transform` on Jul 12

Conversation (23)

Commits (12)

Checks (6)

Files changed (5)



giovp commented on Jul 5 • edited

Member

I would need this refactor in order to move on with the data loader performance PR. [@LucaMarconato](#) what this PR basically does is unifying the building request code for `DataArray|DataTree` with the already existing `_get_bounding_box_corners_in_intrinsic_coordinates` for `DaskDataFrame|GeoPandasDataFrame`. I *think* this is correct, as in the tests for `spatial_query` pass, but I am not sure that the logic is correct.

I had to disable the tests for the dataloaders but they will be worked out in a separate PR. This PR is ready to review, but I would not merge it for now, and wait for finishing the dataloader improvement work.



Contributing to open-source (going further)

Create your own library:

- Good dependency manager
- Good testing (all OS, many Python/R versions)
- Good documentation
- Release your tool on PyPI / Bioconductor
- Manage your package over time