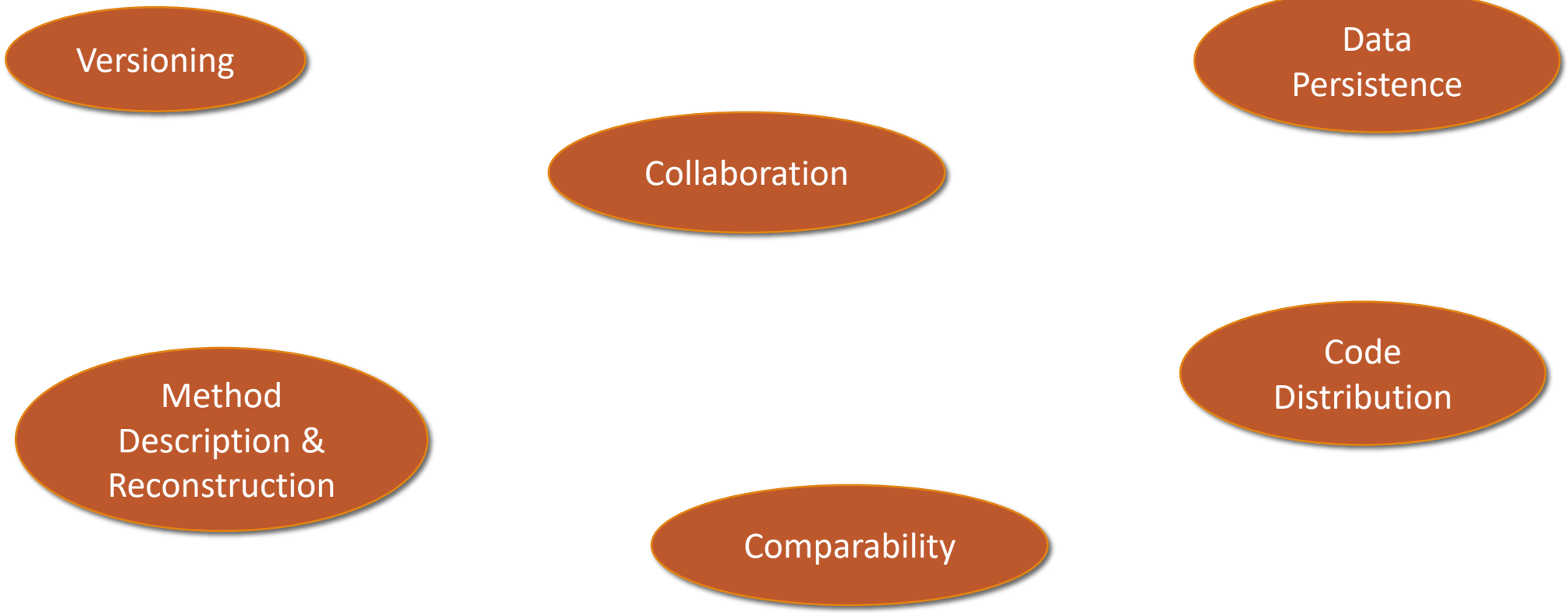
MMRS & Reproducibility

MULTIMODAL REMOTE SENSING

BENJAMIN KELLENBERGER {@GEO.UZH.CH}

Reproducibility?



We work with data (a lot)

- We produce a lot of code
- We experiment with several parametrizations
- We often use the same data

We need to keep track of it all!

We need to be consistent among papers!

We need to become comparable too!

Data Persistence

Types of Datasets:

Proprietary

✓ Tailored to research
✓ Sometimes the only option

✗ Hardly repeatable setting
✗ Rarely distributed
✗ No comparability

Data Persistence

Types of Datasets:

Public, custom-built

✓ Easily obtainable
✓ Often well-defined
✓ Open: comparative work possible

✗ No separate test dataset

Examples

Proprietary

Data Persistence

Types of Datasets:

Public,
evaluation server

Public, custom-built

Proprietary

✓ Publicly available
✓ Well-defined
✓ Separate, hidden test set
✓ Online evaluation server: you get your accuracy scores only

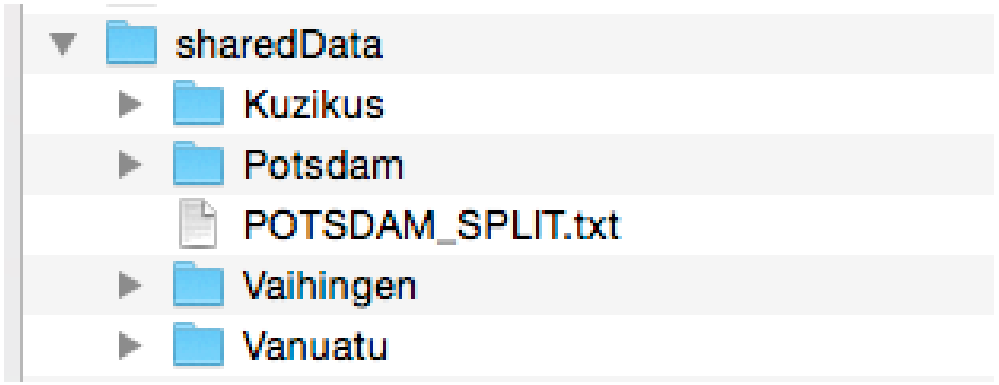
✗ Extensive & expensive for host

Example

Data Persistence

At MMRS:

- We have an internal data server
- We maintain publicly available datasets: examples
- We are also occasionally involved in public contests (with evaluation servers)



Method Reconstruction

We want our *methods* to be reproducible, too!

This involves:

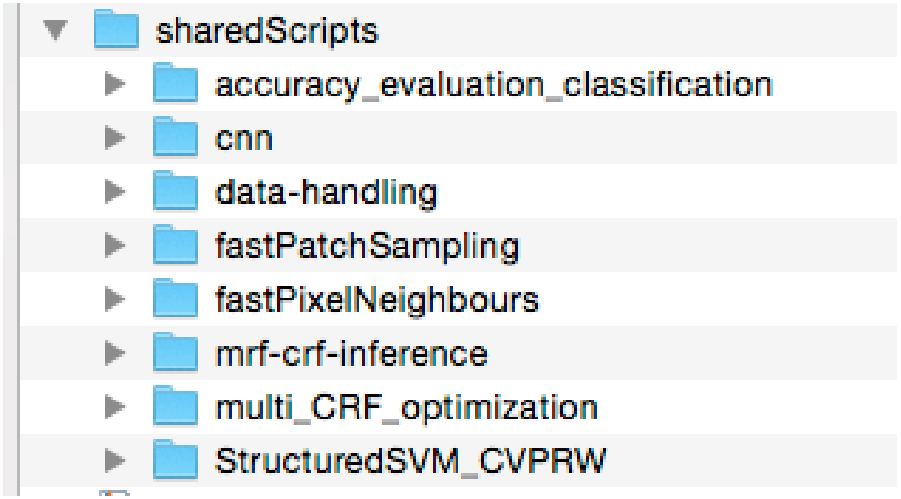
- Detailed description in publication
- A *lot* of code, including (intuitive) examples

At MMRS, we do exactly that!

- All details in paper
- Distribution of code:
  - MMRS webpage
  - Individual researchers‘ webpages: link1, link2, etc.
  - Github: link
- We include parameter settings

Method Reconstruction

In addition: we share useful general-purpose scripts on our server



Versioning

Invaluable paradigm for everything research-related!

- We collaborate
- Our projects evolve as we work on them
- We cannot avoid hardware failures (→ backups!)
- We all make mistakes

Versioning

Idea:

- Relevant files lie on a server (e.g. scripts, LaTeX documents, etc.)

1. Every collaborator has a local copy he works on
2. Once finished, they submit their changes to the server
3. Every (incremental) update is saved and revertible if needed
4. Conflicts are easily resolvable

At MMRS:

- Paper collaboration: SVN
- Code: Github
- With outer world: collaborative systems (e.g. Overleaf, ShareLaTeX)