

Replicability and data/algorithm storage and availability

3rd Workshop on Replication in Extended Reality (WoR XR)

Florian Weidner

Lancaster University 

 IEEE  IEEE COMPUTER SOCIETY  vgtc



ISMAR
23 SYD
Oct 16 - 20



Outline for the next 15 minutes

- **Why** publish data and algorithms?
- **What** could (should) we store?
- **Where** to store?
 - Zenodo
 - OSF
- How do we **document** it?
- Summary

Why publish data?

Reproducibility and replicability. For example:

- Prototype shared: test exactly the same.
- Data shared: calculate exactly the same
- Everything: do exactly the same.

But also:

- Comparison with new findings easier (statistics)
- Making improvements is easier (iterations)

What could (should?) we store?

- *Simulation data*
- *Data from surveys, questionnaires and interviews*
- *Audio-visual data*
- *Software, code, scripts*
- *Measurement data (raw + processed)*
- *Analysis data*
- *...etc.*

Where to store?

- „Popular“: osf.io, zenodo.org, [IEEEDataPort](https://ieee-dataport.org)
- Some institutions have platforms
- Criteria:
 - Offers DOI
 - promises long-term storage
 - Stable funding
 - (Free)



IEEE*DataPort*[™]



Zenodo

All about storage.



ISMAR
23 SYD
Oct 16 - 20



Zenodo UI

- Signup / login
 - Create new upload
 - Get DOI
 - Fill in metadata
 - Upload files
 - Review
 - Publish
- ➔ straightforward

■ ■ ■ Digital Object Identifier^{*}

Do you already have a DOI for this upload? ☒ Yes ☐ No

Copy/paste your existing DOI here...

A DOI allows your upload to be easily and unambiguously cited. Example: 10.1234/foo.bar

Files

Storage available 0 out of 100 files 0 bytes out of 50.00 GB

Drag and drop files - or - Upload files

Basic information

Recommended information

Funding

Alternate identifiers

Related works

References

Publishing information

Conference

Draft

Save draft Preview

Publish

Visibility^{*}

Files only

Public Restricted

Public
The record and files are publicly accessible.

Options

Apply an embargo
Record or files protection must be restricted to apply an embargo.

Example

Published February 25, 2021 | Version 1.0.0

SoftwareOpen

Edit

New version

Share

52
EYES VIEWS

8
DOWNLOADS

Show more details

Versions

Version 1.0.0Feb 25, 202110.5281/zenodo.4562268

Cite all versions? You can cite all versions by using the DOI [10.5281/zenodo.4562267](https://doi.org/10.5281/zenodo.4562267). This DOI represents all versions, and will always resolve to the latest one. [Read more.](#)

External resources

Indexed in

OpenAIRE

Communities

This record is not included in any communities yet.

S3D Dashboard

Weidner, Florian¹

Show affiliations

This data acts as a reference for the Dissertation titled "S3D Dashboard: Exploring Depth on Large Interactive Dashboards".

Files

Readme.md

S3D Dashboard

Author

Florian Weidner

About

This data acts as a reference for the Dissertation titled "S3D Dashboard: Exploring Depth on Large Interactive Dashboards".

Contents

It contains four core elements:

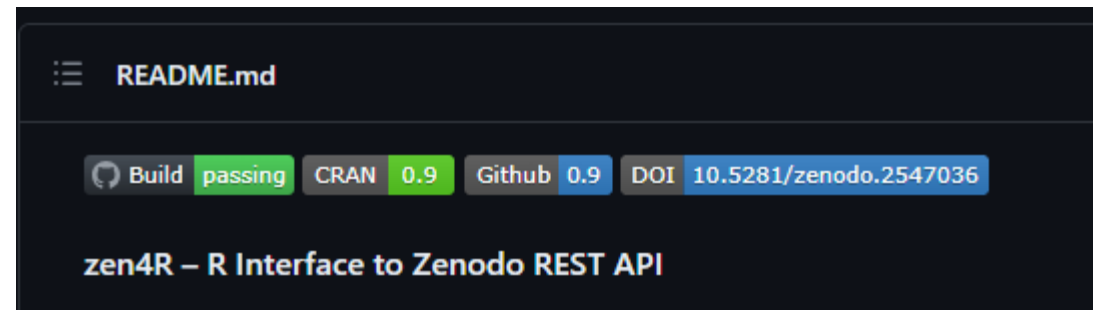
1. The car mock-up. An CAD model (including fbx and Blender files) of a car mock-up designed to provide a large stereoscopic 3D dashboard using rear projection.

Files (1.1 GB)

Name	Size	Download all
------	------	--------------

Comments on Zenodo

- Free
- Funded by CERN (long-term perspective)
- No anonymous links; collaborators can be modified
- Versioning possible
- 50GB
- Links to github release
 - auto-update on new release
 - Nice github Zenodo badge



<https://github.com/eblondel/zen4R>

OSF (Open Science Framework)

It's all about (pre-)registration, project, and resources: lifecycle

TIL: "OSF is maintained and developed by the Center For Open Science (COS)"



ISMAR
23 SYD
Oct 16 - 20

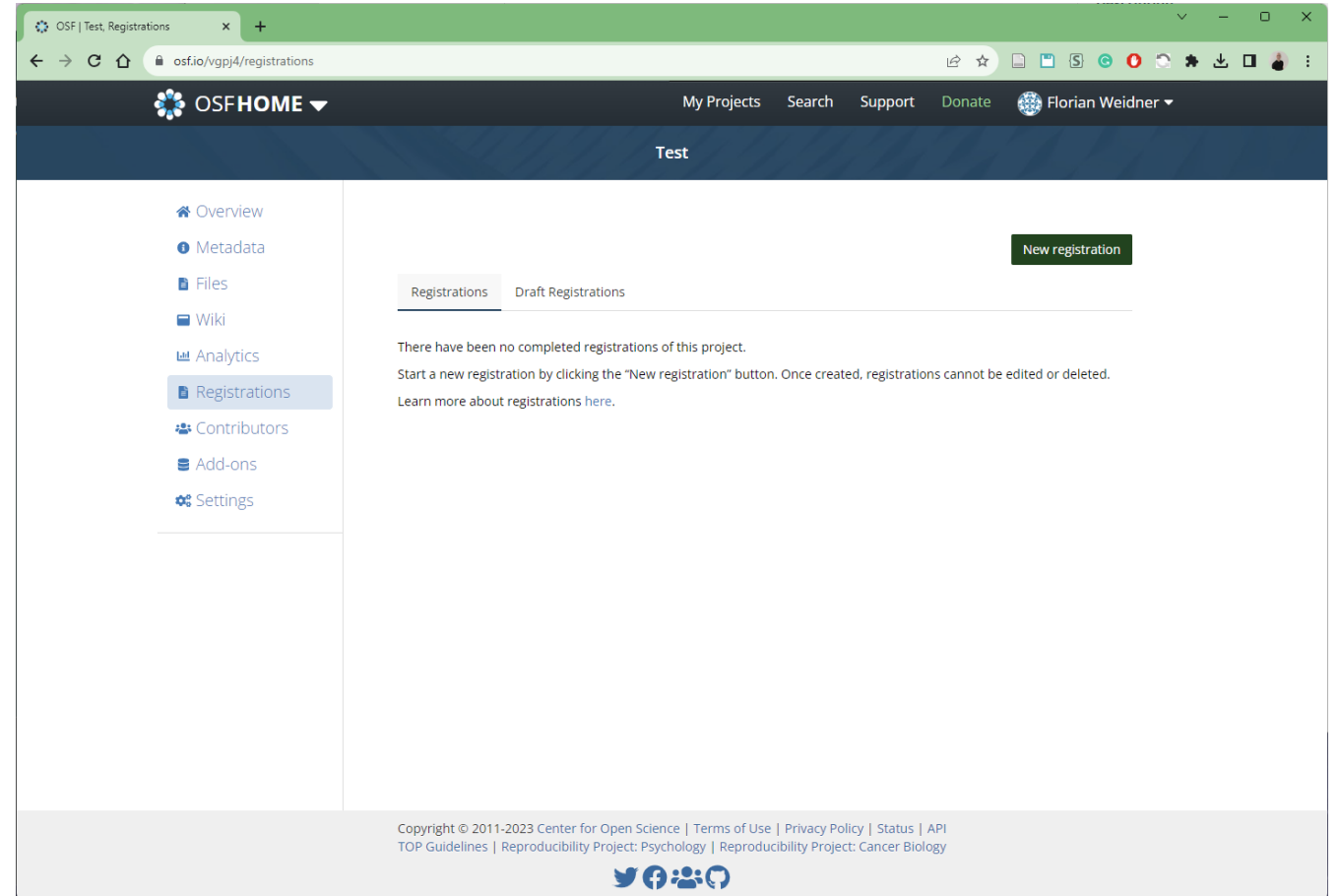


Registration + Project (+ Preprint)

- Registration is the start
- Registration creates project
- Project is for data and collaboration
 - Is flexible
 - Own storage or external
 - History

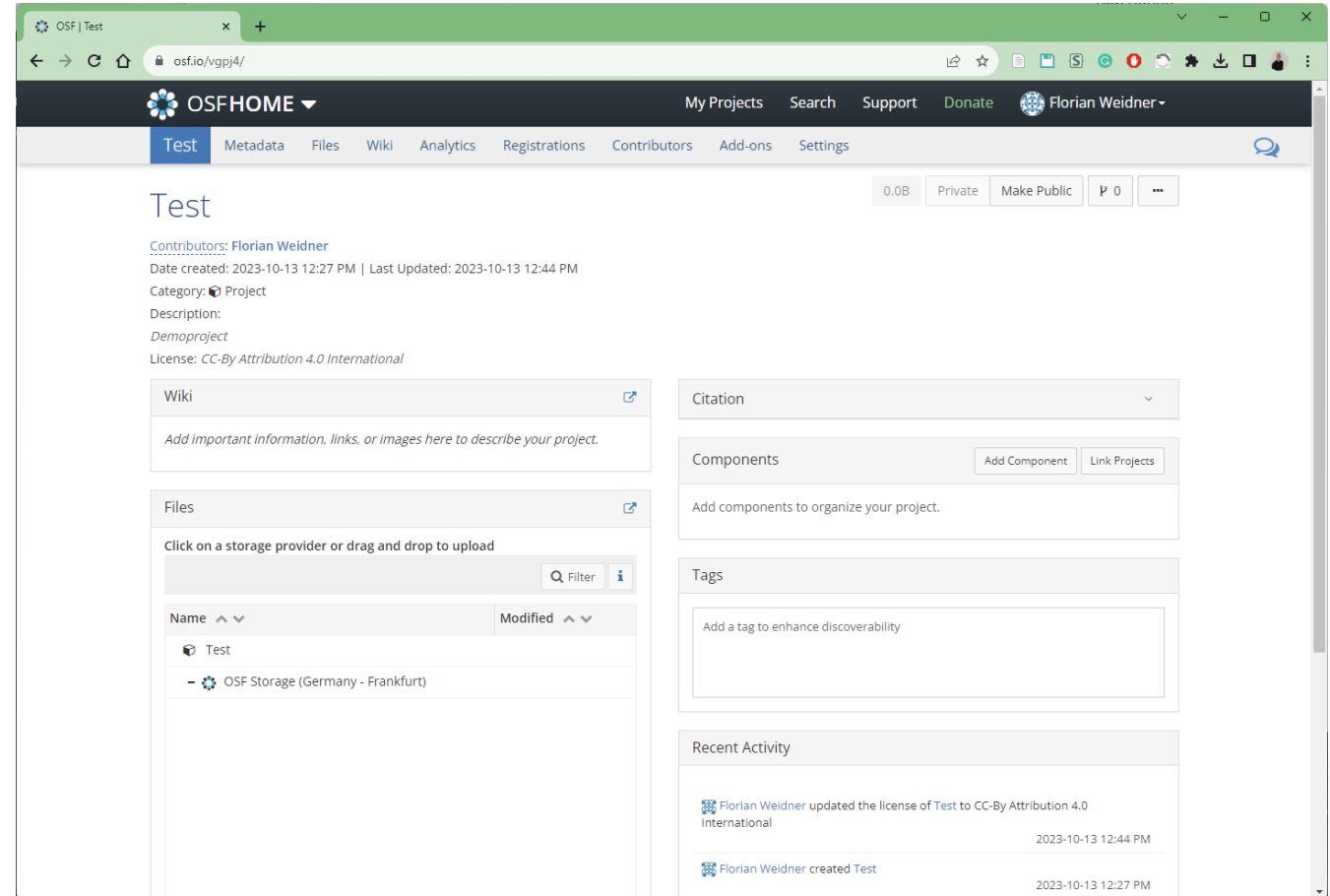
OSF Workflow

Create new registration



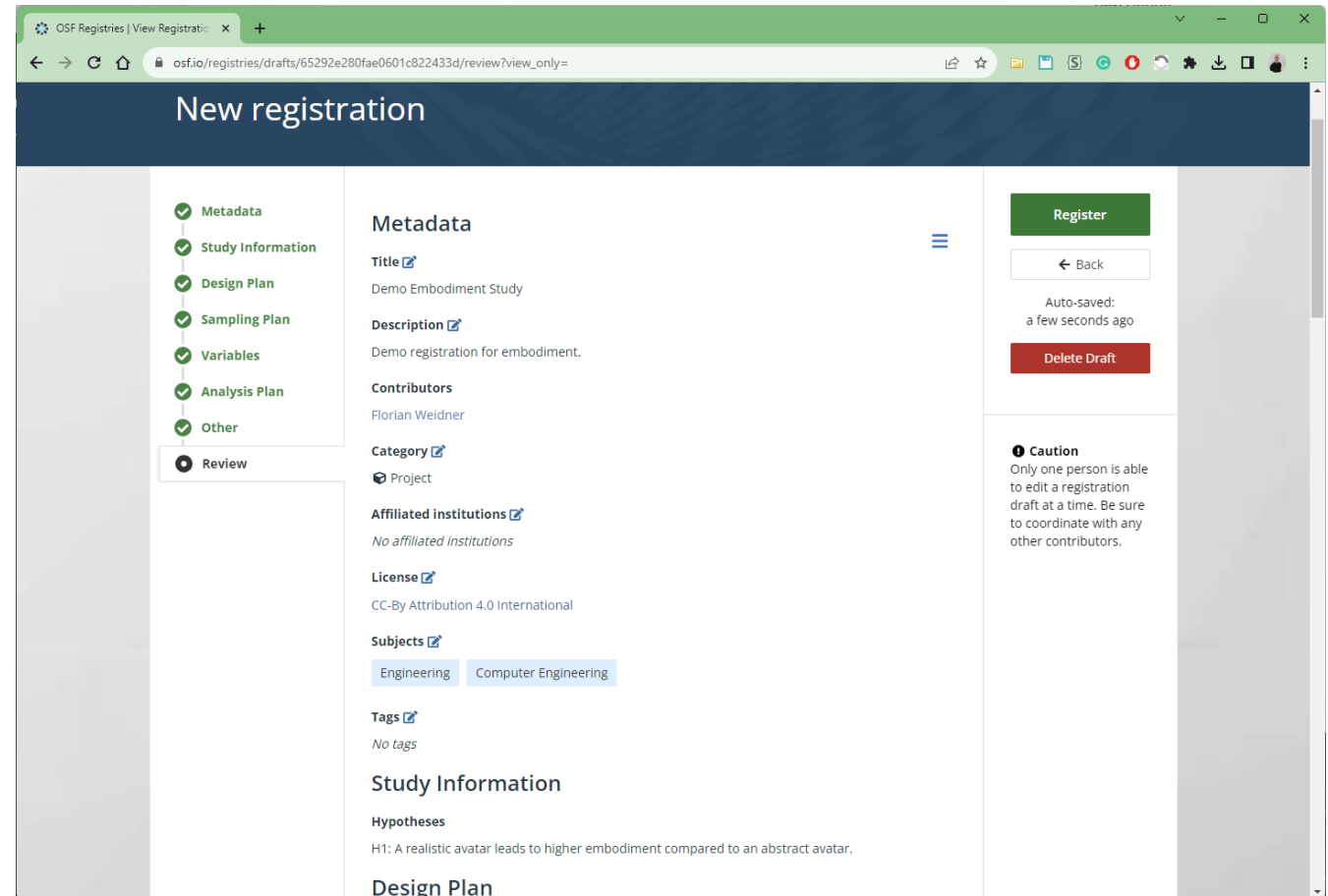
OSF Workflow

Creates new project



OSF Workflow

Fill out all info and
register (updatable)

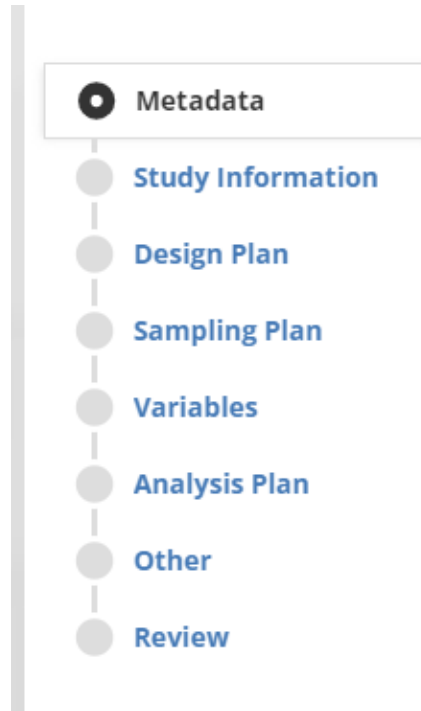


The screenshot shows the 'New registration' form in the OSF Registries interface. The browser address bar shows the URL: `osf.io/registries/drafts/65292e280fae0601c822433d/review?view_only=`. The form is divided into several sections:

- Left sidebar:** A vertical list of steps with checkmarks: Metadata, Study Information, Design Plan, Sampling Plan, Variables, Analysis Plan, Other, and Review (which is currently selected).
- Metadata section:**
 - Title:** Demo Embodiment Study
 - Description:** Demo registration for embodiment.
 - Contributors:** Florian Weidner
 - Category:** Project
 - Affiliated institutions:** No affiliated institutions
 - License:** CC-BY Attribution 4.0 International
 - Subjects:** Engineering, Computer Engineering
 - Tags:** No tags
- Study Information section:**
 - Hypotheses:** H1: A realistic avatar leads to higher embodiment compared to an abstract avatar.
- Design Plan section:** (The content is partially visible at the bottom of the form).
- Right sidebar:**
 - Register button:** A green button labeled 'Register'.
 - Back button:** A button labeled '← Back'.
 - Auto-saved status:** A message indicating 'Auto-saved: a few seconds ago'.
 - Delete Draft button:** A red button labeled 'Delete Draft'.
 - Caution message:** A warning icon followed by the text: 'Only one person is able to edit a registration draft at a time. Be sure to coordinate with any other contributors.'

OSF Workflow: Pre-registration

- General information
- Hypothesis
- Design plan
 - Study type
 - Blinding
 - Study design
 - Randomization
- Sampling plan
 - Existing data*
 - Explanation of existing data
 - Data collection procedure*
 - Sample size* + rationale + stopping rule
- Variables
 - Dependent variables
 - Independent variables*
 - Indexing?
- Analysis plan
 - Statistical models*
 - Transformations
 - Inference criteria (e.g., p-value)
 - Data exclusion
 - Missing data
 - Exploratory analysis



Continue your registration by selecting a registration form:

- ☒ **OSF Preregistration 1**
 - ☐ Open-Ended Registration 1
 - ☐ Qualitative Preregistration 1
 - ☐ Secondary Data Preregistration 1
 - ☐ Generalized Systematic Review Registration 1
 - ☐ Registered Report Protocol Preregistration 1
 - ☐ OSF-Standard Pre-Data Collection Registration 1
 - ☐ Preregistration Template from AsPredicted.org 1
 - ☐ Replication Recipe (Brandt et al., 2013): Post-Completion 1
 - ☐ Replication Recipe (Brandt et al., 2014): Pre-Registration 1
 - ☐ Pre-Registration in Social Psychology (van 't Veer & Giner-Sorolla, 2016): Pre-Registration 1
-
- ☐ Registration prior to creation of data ?
 - ☐ Registration prior to any human observation of the data ?
 - ☐ Registration prior to accessing the data ?
 - ☐ Registration prior to analysis of the data ?
 - ☐ Registration following analysis of the data ?

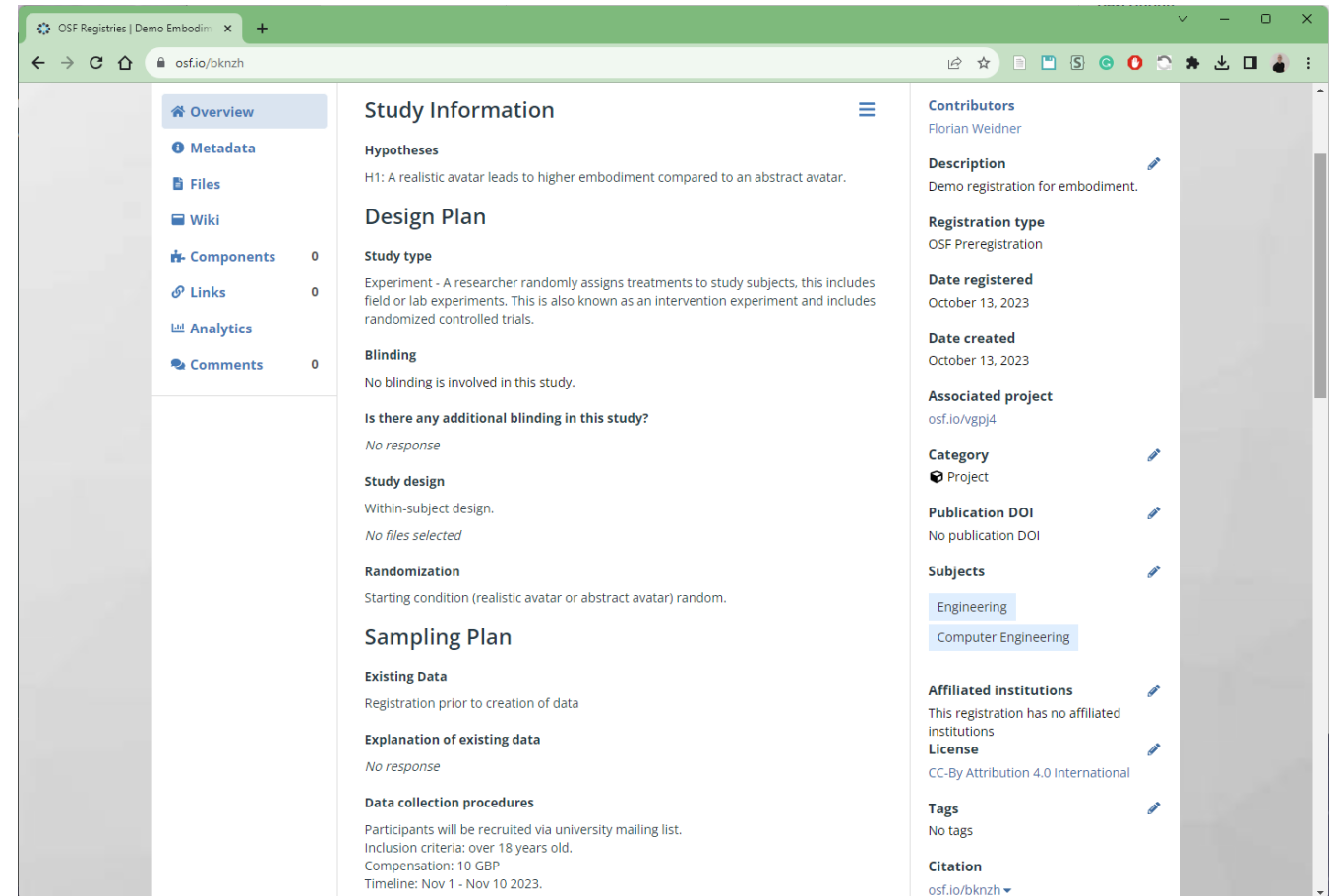
Study design *

Describe your study design. The key is to be as detailed as is necessary given the specific parameters of the design. There may be some overlap between this question and the following questions. That is OK, as long as sufficient detail is given in one of the areas to provide all of the requested information. Examples include two-group, factorial, randomized block, and repeated measures. Is it a between (unpaired), within-subject (paired), or mixed design? Describe any counterbalancing required.

OSF Workflow

Anonymous link to registration possible:
https://osf.io/bknzh/?view_only=bda587ab846f416482460924e019bf91

<https://tinyurl.com/worxr1>



The screenshot shows the OSF Registries page for a study titled "Demo Embodiment". The page is divided into three main sections: Overview, Study Information, and Contributors. The Overview section on the left lists various components like Metadata, Files, Wiki, Components, Links, Analytics, and Comments. The Study Information section in the center provides details about the study, including Hypotheses, Design Plan, Study type, Blinding, Study design, Randomization, Sampling Plan, Existing Data, Explanation of existing data, and Data collection procedures. The Contributors section on the right lists the study's contributors, description, registration type, date registered, date created, associated project, category, publication DOI, subjects, affiliated institutions, license, tags, and citation.

Overview

- Metadata
- Files
- Wiki
- Components 0
- Links 0
- Analytics
- Comments 0

Study Information

Hypotheses
H1: A realistic avatar leads to higher embodiment compared to an abstract avatar.

Design Plan

Study type
Experiment - A researcher randomly assigns treatments to study subjects, this includes field or lab experiments. This is also known as an intervention experiment and includes randomized controlled trials.

Blinding
No blinding is involved in this study.

Is there any additional blinding in this study?
No response

Study design
Within-subject design.
No files selected

Randomization
Starting condition (realistic avatar or abstract avatar) random.

Sampling Plan

Existing Data
Registration prior to creation of data

Explanation of existing data
No response

Data collection procedures
Participants will be recruited via university mailing list.
Inclusion criteria: over 18 years old.
Compensation: 10 GBP
Timeline: Nov 1 - Nov 10 2023.

Contributors
Florian Weidner

Description
Demo registration for embodiment.

Registration type
OSF Preregistration

Date registered
October 13, 2023

Date created
October 13, 2023

Associated project
osf.io/vgpj4

Category
Project

Publication DOI
No publication DOI

Subjects
Engineering
Computer Engineering

Affiliated institutions
This registration has no affiliated institutions

License
CC-BY Attribution 4.0 International

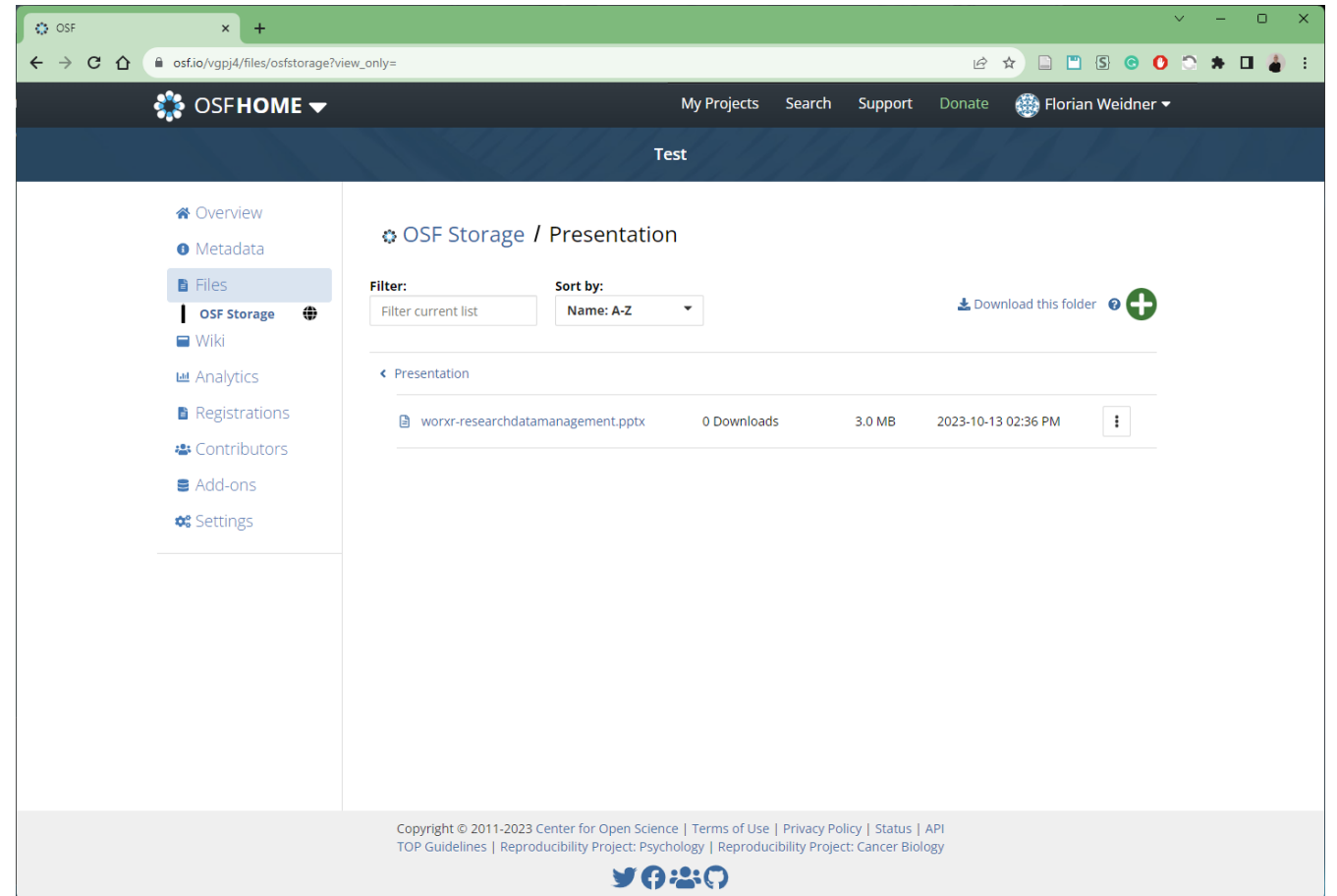
Tags
No tags

Citation
osf.io/bknzh

... do study ...

OSF Workflow

Add some data.



OSF Workflow

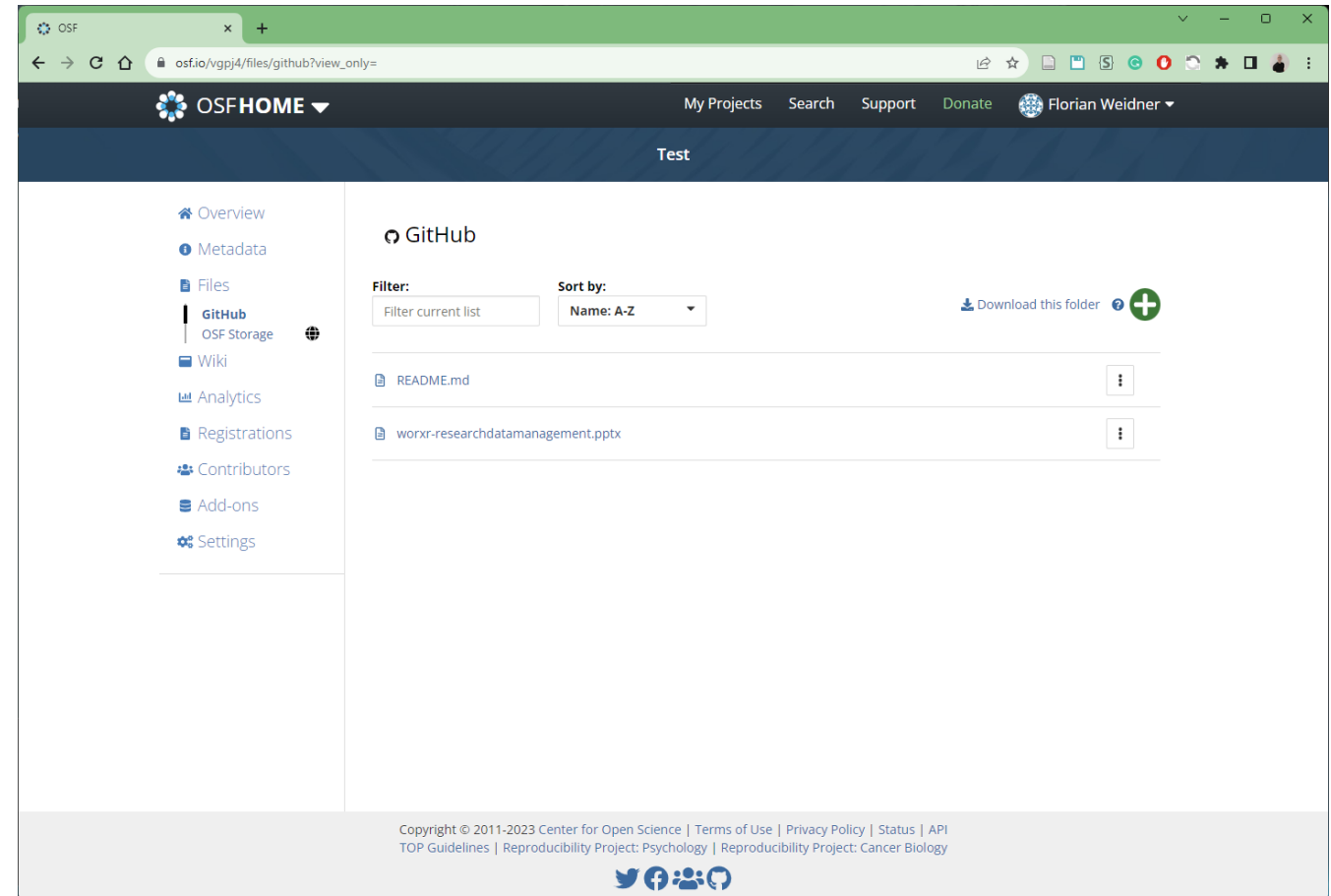
- Github + OSF
- No file mirror!

GitHub Add-on Terms

Function	Status
Permissions	Making an OSF project public or private is independent of making a GitHub repo public or private. The OSF does not alter the permissions of linked GitHub repos.
View / download file versions	GitHub files and their versions can be viewed/downloaded via OSF.
Add / update files	Adding/updating files in the project via OSF will be reflected in GitHub.
Delete files	Files deleted via OSF will be deleted in GitHub.
Logs	GitHub dynamically updates OSF logs when files are modified outside the OSF. Changes to GitHub repos made before the repo is linked to the OSF will not be reflected in OSF logs.
Forking	Forking a project or component does not copy Github authorization unless the user forking the project is the same user who authorized the Github add-on in the source project being forked.
Registering	GitHub content will be registered, but version history will not be copied to the registration.

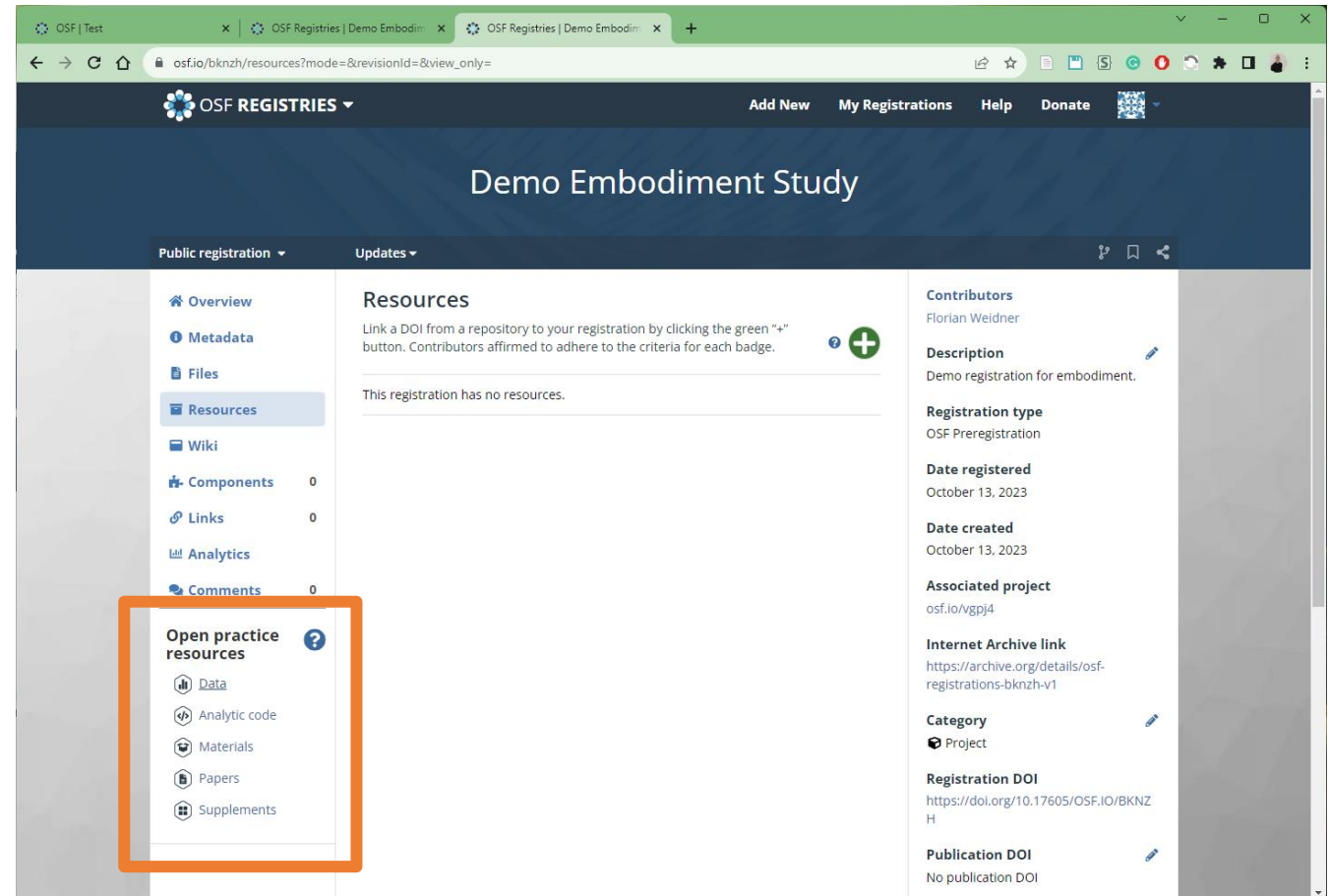
- This add-on connects your OSF project to an external service. Use of this service is bound by its terms and conditions. The OSF is not responsible for the service or for your use thereof.
- This add-on allows you to store files using an external service. Files added to this add-on are not stored within the OSF.

Cancel Confirm



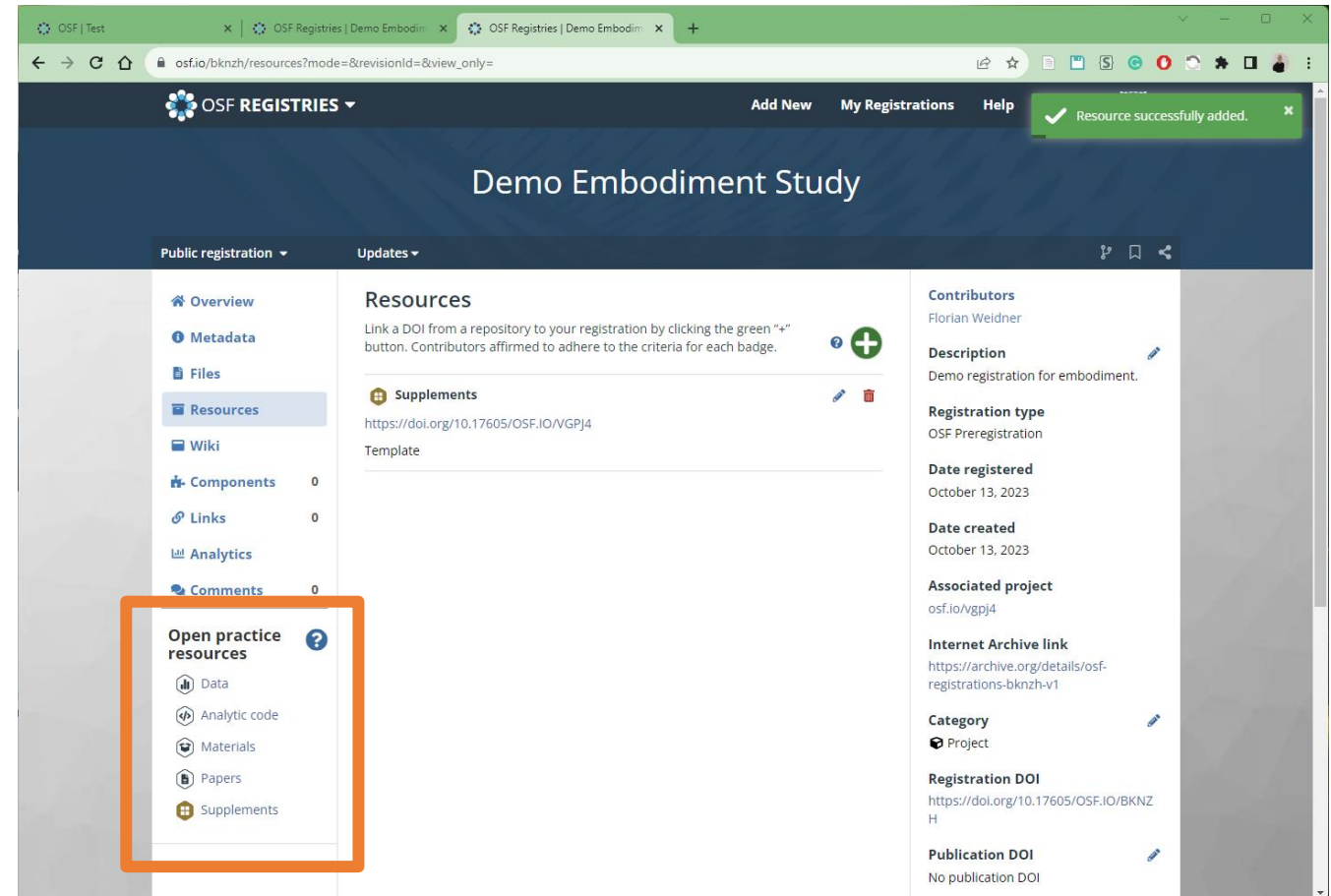
OSF Workflow

Link project to registration via project DOI.



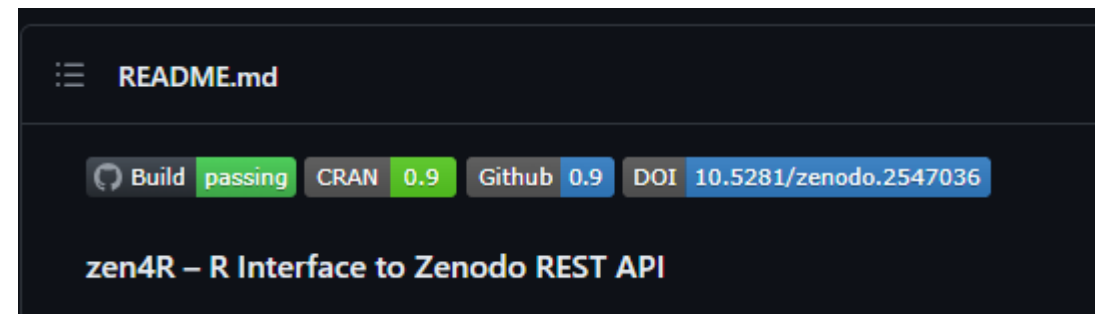
OSF Workflow

- Link project to registration via project DOI 😊
- [1] F. Weidner, "Demo Embodiment Study", 13-Oct-2023. [Online]. Available: <https://osf.io/bknzh>.



OSF comments

- Free
- Foundation well-funded (long-term perspective)
- Preregistration
- Anonymous links
- Versioning possible
- 50GB
- Links to github release
→ no file mirror.
- Covers research lifecycle



<https://github.com/eblondel/zen4R>

Tl;dr: OSF

1. Create and fill out pre-registration.
2. Do study.
3. Create project and upload data, analytics code, materials, papers, and supplemental.
4. Create view-only links (private or not)
5. Add them to manuscript.
6. Be happy 😊

IEEE DataPort

All about datasets.



ISMAR
23 SYD
Oct 16 - 20



IEEE DataPort comments

- Free only for subscribers
- Up to 2TB or 10TB
- Open access for 1.950 USD (sometimes discounts)
- Long-term via IEEE
- Linked to IEEE account
- No preregistration
- DOI
- (designed for competitions)

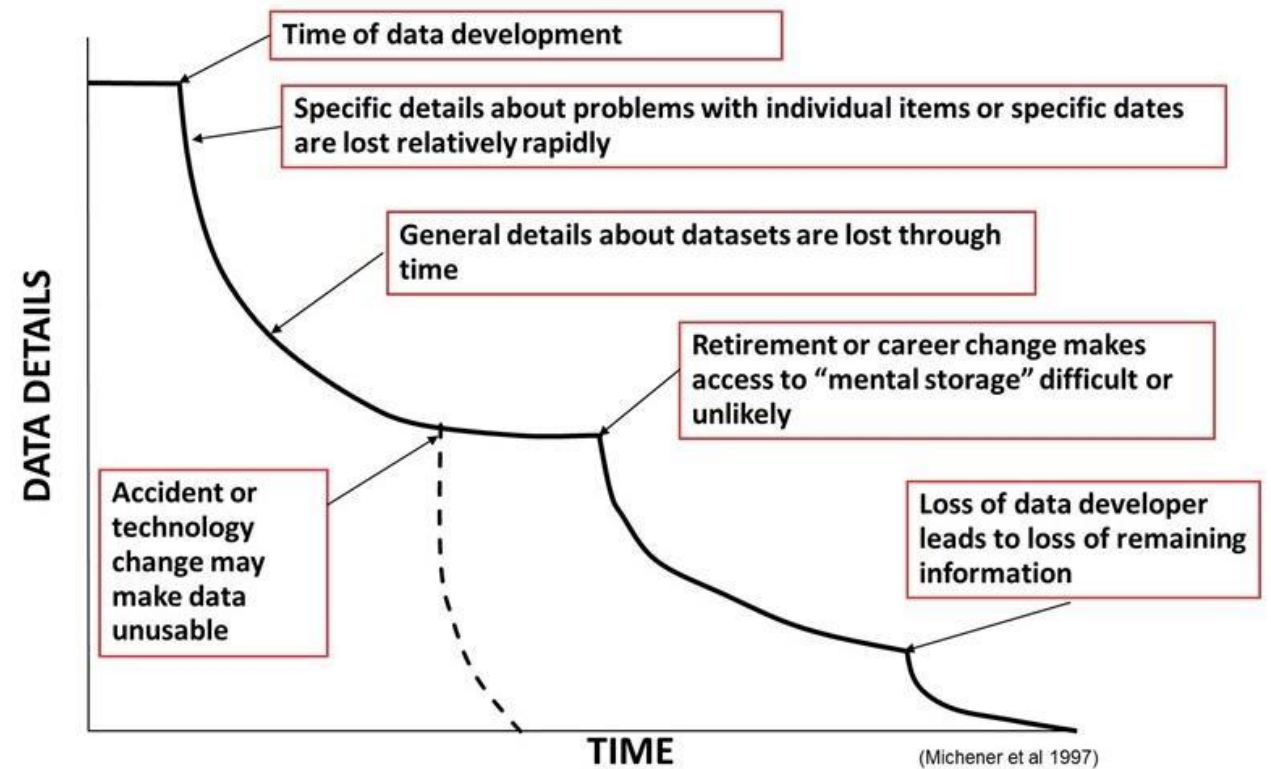
Data documentation



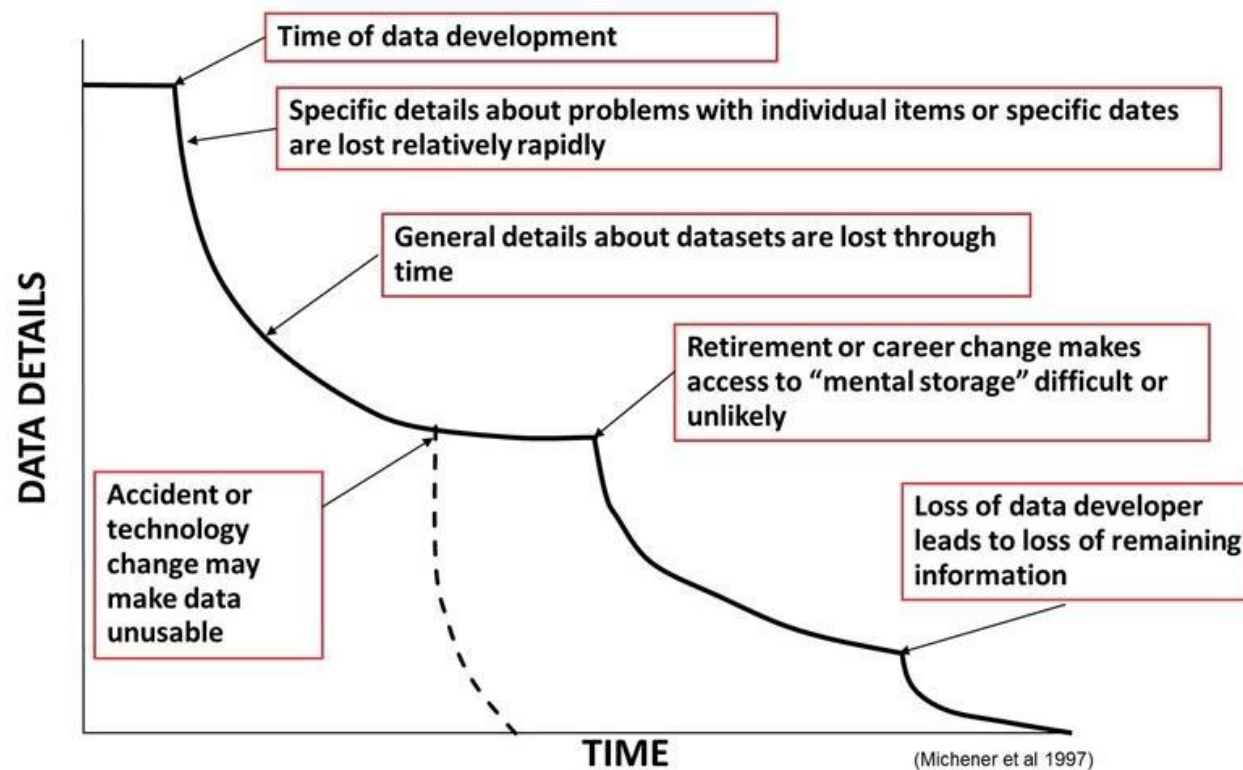
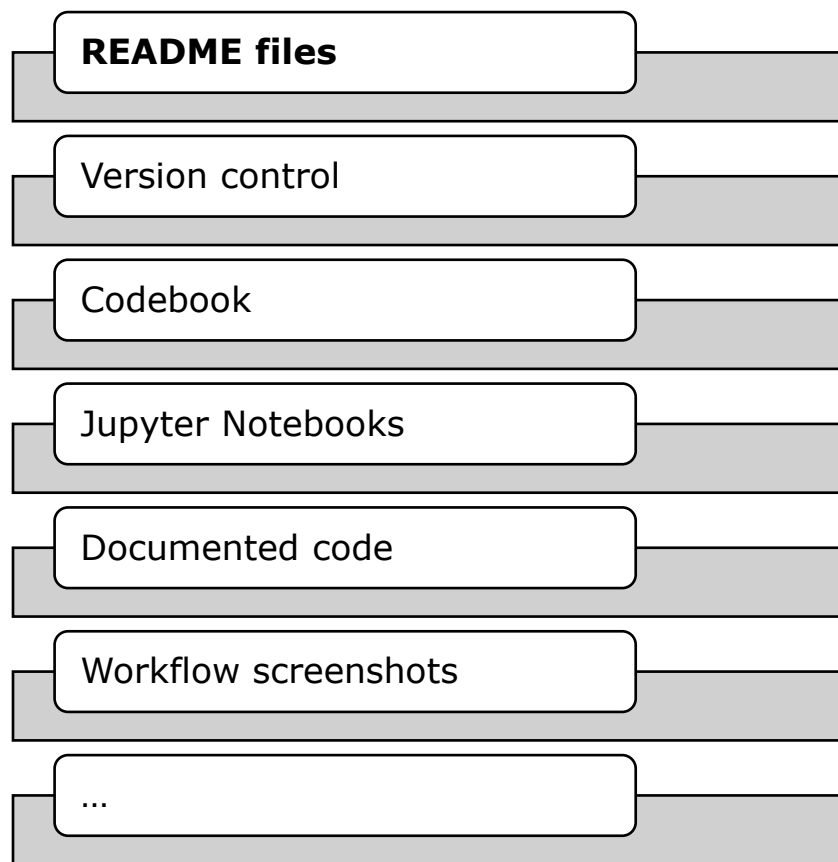
ISMAR
23 SYD
Oct 16 - 20



How do we document it?



How do we document it?



Summary



ISMAR
23 SYD
Oct 16 - 20



IEEE



IEEE
COMPUTER
SOCIETY

IEEE
vgtc

All together?



IEEE*DataPort*[™]



- Soft requirement / extra recognition?
 - Executable + setup instructions
 - Study protocol and self-defined questions
 - Raw data + scripts to calculate results in paper
 - Video
 - Preprint
- OSF, Zenodo, IEEE DataPort?
 - OSF: allrounder, free, pre-registration, 50GB
 - Zenodo: easy via github, free, 50GB, no pre-registration
 - DataPort: large storage, paid

Replicability and data/algorithm storage and availability

3rd Workshop on Replication in Extended Reality (WoR XR)

Florian Weidner, Lancaster University

Thank you! Questions?



Some resources

- <https://www.acm.org/publications/policies/digital-artifacts>
- <https://journals.ieeeauthorcenter.ieee.org/create-your-ieee-journal-article/research-reproducibility/>

FAIR data

- https://en.wikipedia.org/wiki/FAIR_data
- <https://www.go-fair.org/fair-principles/>
- Fair != open