

Jupyter Notebooks und die Integration in eine Publikationsplattform

Klaus Thoden, Max-Planck-Institut für Wissenschaftsgeschichte
Kolloquium Wissensinfrastruktur, Universitätsbibliothek Bielefeld

25. Januar 2019

@Edition_OA

GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung

EOA-Entwicklungsteam

○ Samuel Gfrörer

○ Klaus Thoden

○ Malte Vogl



About EOA

localhost:8002/?edit_off=true

eoa edition open access

[About EOA](#) [What is Open Access?](#) [Participate](#) [Projects](#) [Editorial Board](#)

Short Introduction to Edition Open Access

The Edition Open Access (EOA) platform was founded to bring together publication initiatives seeking to disseminate the results of scholarly work in a format that combines traditional publications with the digital medium. It currently hosts the open-access publications of the "Max Planck Research Library for the History and Development of Knowledge" (MPRL) and "Edition Open Sources" (EOS). EOA is open to host other open access initiatives similar in conception and spirit, in accordance with the [Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities](#), which was launched by the Max Planck Society in 2003. By combining the advantages of traditional publications and the digital medium, the platform offers a new way of publishing research and of studying historical topics or current issues in relation to primary materials that are otherwise not easily available. The publications are directed at scholars and students of various disciplines, as well as at a broader public interested in how science and


Projects

localhost:8002/projects/

eoaa

edition
open access


[About EOA](#) [What is Open Access?](#) [Participate](#) [Projects](#) [Editorial Board](#)



Edition Open Sources

Edition Open Sources (EOS) pioneers a new paradigm in publishing, devoted to open access and high-

[↓ More](#)

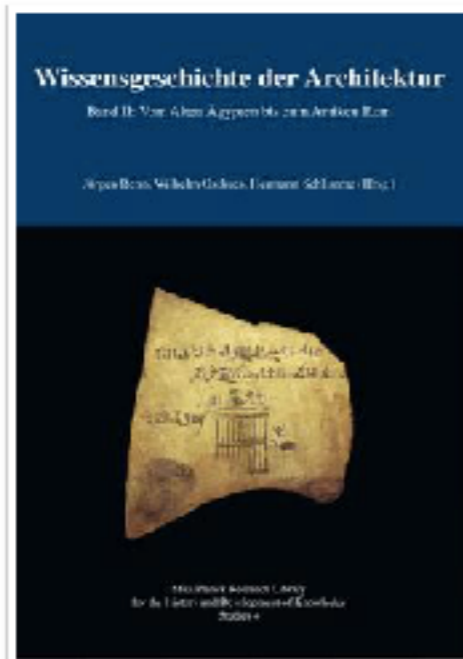


Max Planck Research Library
for the History and
Development of Knowledge

MPRL

The Max Planck Research Library for the History and Development of Knowledge, in accordance with the

[↓ More](#)



ISBN: 978-3-945551-03-4

Print: 37,14 € | 448 p.

Print on Demand: apubli.de

Booksellers can order directly via email at

verlagservice@sigloch.de

Publication Date: Dec. 2, 2014

Studies 4


Wissensgeschichte der Architektur

Band II: Vom Alten Ägypten bis zum Antiken Rom


Jürgen Renn, Wilhelm Osthuus, Hermann Schlimme (Hrsg.)

Submitted by: Horst Eredenkamp, Daniel T. Potts

Click on the entries in the Table of Contents below to download individual chapters.

 Read online

 Download PDF (18.4 MB)

 Download Ebook (42.0 MB)

This publication is licensed under a Creative Commons Attribution-Non Commercial Share Alike 3.0 Germany ([cc by-nc-sa 3.0](http://creativecommons.org/licenses/by-nc-sa/3.0/de/)) License.



Description

Table of Contents

Index of Names

Index of Locations

Die Geschichte des Bauens basiert bis weit in die Moderne auf praxisnahen Wissenstraditionen der Handwerker, Baumeister und Architekten. Solches „handlungsimplicite Wissen“ ermöglichte die Monumentalbauten der frühen Hochkulturen, die modellgebende Architektur und Infrastrukturtechnologie der griechischen und römischen Antike, die Sakral- und Wehrbauten des Mittelalters sowie die riskanten und innovativen Bauprojekte der Renaissance.

Das den großen Bauleistungen der Vergangenheit zugrunde liegende Wissen und seine Entwicklung ist Gegenstand der hier vorgelegten Wissensgeschichte der Architektur, die gemeinsam von der Bibliotheca Hertziana (Max-Planck-Institut für Kunstgeschichte), in Bonn, und dem Max-Planck-Institut für Wissensgeschichte, in Berlin, initiiert wurde. Die



2.3 Parte Guelfa, Karte der Straßen, für deren Instandhaltung die Gemeinde von ‚S. Iacopo a la Zambuca‘ (Sambuca) verpflichtet war und *strada maestra* (im Bild dunkelgrau), 1580–1595 (© Archivio di Stato di Firenze, Regalia del Principato, vol. 1, c.178, mit Erlaubnis des Ministero dei Beni e delle Attività culturali e del Turismo, Reproduktion nicht gestattet). ↑

- 👁 📌 Zu den konkret ausgeführten Straßenbauarbeiten gibt es für das 13. und 14. Jahrhundert praktisch keine Quellen, da die Aufzeichnungen der Ufficiali di Torre und der Parte Guelfa (s. u.) durch einen Brand 1566 zum großen Teil verloren gegangen sind. Ab dem späteren 16. Jahrhundert und insbesondere für die Regierungszeiten von Francesco I (1574–87) und Ferdinando I (1587–1609) lässt sich hingegen nachvollziehen, wie die Bauarbeiten im Einzelnen durchgeführt wurden und welche Arbeitsteilung angestrebt wurde.¹²¹ Im Jahre 1582 wurden etwa Maurer und Handlanger für Arbeiten an der Straße zwischen Florenz und Siena bezahlt. Der *capomastro* Lorenzo Vestrucci schlug der Gemeinde Settignano in einem Gutachten vor, ein schadhaftes Straßennflaster *per mano di buono maestro* reparieren zu lassen.

Keywords, Persons and Locations

Francesco I.: [1](#)

Ferdinando I.: [1](#) [2](#)

Vestrucci, Lorenzo: [1](#)

Florenz: [1](#) [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) [8](#) [9](#) [10](#) [11](#) [12](#) [13](#) [14](#) [15](#) [16](#) [17](#) [18](#) [19](#) [20](#) [21](#) [22](#) [23](#) [24](#) [25](#)
[26](#) [27](#) [28](#) [29](#) [30](#) [31](#) [32](#) [33](#) [34](#) [35](#) [36](#) [37](#) [38](#) [39](#) [40](#) [41](#) [42](#) [43](#) [44](#) [45](#) [46](#) [47](#) [48](#) [49](#)
[50](#) [51](#) [52](#) [53](#) [54](#) [55](#) [56](#) [57](#) [58](#) [59](#) [60](#) [61](#) [62](#) [63](#) [64](#) [65](#) [66](#) [67](#) [68](#) [69](#) [70](#) [71](#) [72](#) [73](#)
[74](#) [75](#) [76](#) [77](#) [78](#) [79](#) [80](#) [81](#) [82](#) [83](#) [84](#) [85](#) [86](#) [87](#) [88](#) [89](#)

Siena: [1](#) [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) [8](#) [9](#) [10](#)

Bologna: [1](#) [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) [8](#) [9](#) [10](#) [11](#)

Nipozzano: [1](#)

Raticosa: [1](#) [2](#)

Filigare: [1](#) [2](#)

Firenzuola: [1](#)

Scaricalasino: [1](#)

Romagna: [1](#) [2](#)

de von ,S.
(im Bild
ato, vol. 1,
Turismo,
↑

und 14.
Torre und
gegangen
zeiten von
hingegen
nd welche
aurer und
zahlt. Der
Gutachten

Nachvollziehbarkeit von Forschungsergebnissen

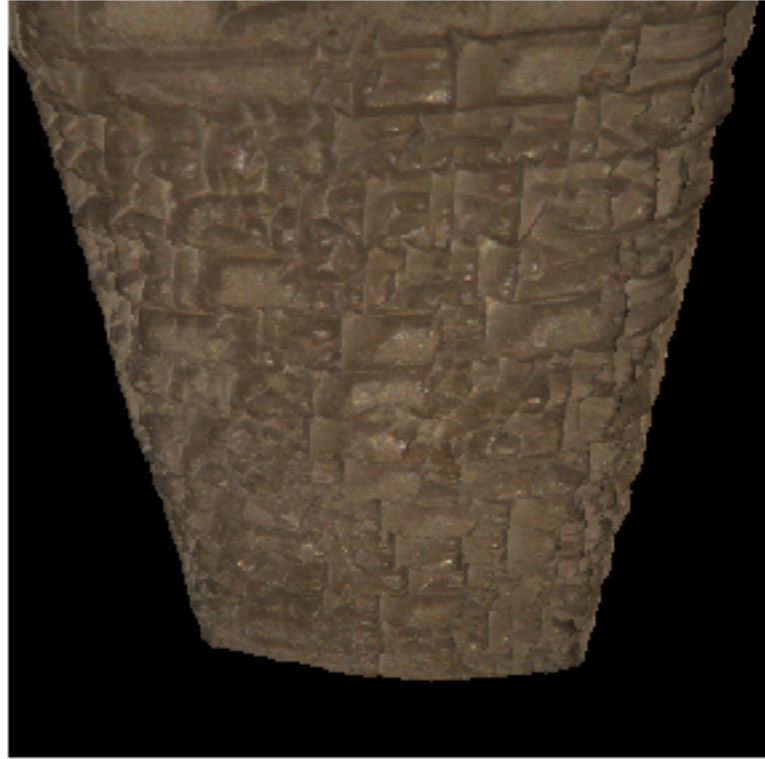
- Grundpfeiler wissenschaftlichen Arbeitens
 - Angabe von Quellen
 - Beschreibung der eingesetzten Methoden
- Einsatz von Computern ermöglicht vielfache Arten der Dokumentation

HS 0106 (size: 94.20MB) ([Download](#) - [meshlabJS](#))



Die Qualität der Ansicht der Scans im Browser ist häufig ungenügend. Zum besseren Lesen empfiehlt sich der [Download](#) und eine Ansicht z.B. mit MeshLab ([externer Link](#)).

[reset camera](#) [center camera](#) [flip y](#) [flip y-z](#)
PointSize: 1.00 [help](#)



Maße: 74,0 X 46,0 X 22,0
Bibl.: Urp.

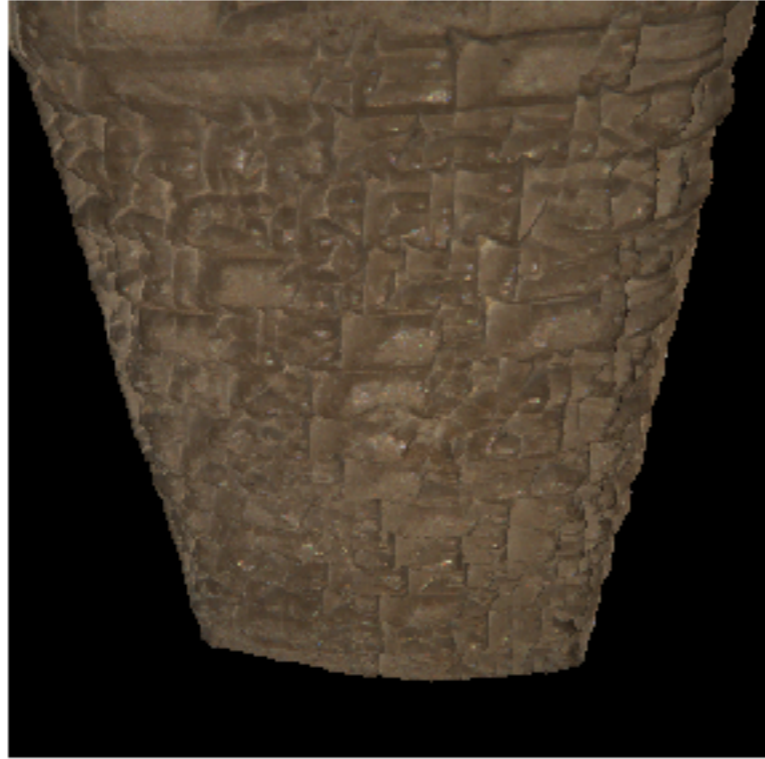
3D-Scan, Screenshot von [Clay Tablet HS0106](#)

HS 0106 (size: 94.20MB) ([Download - meshlabJS](#))



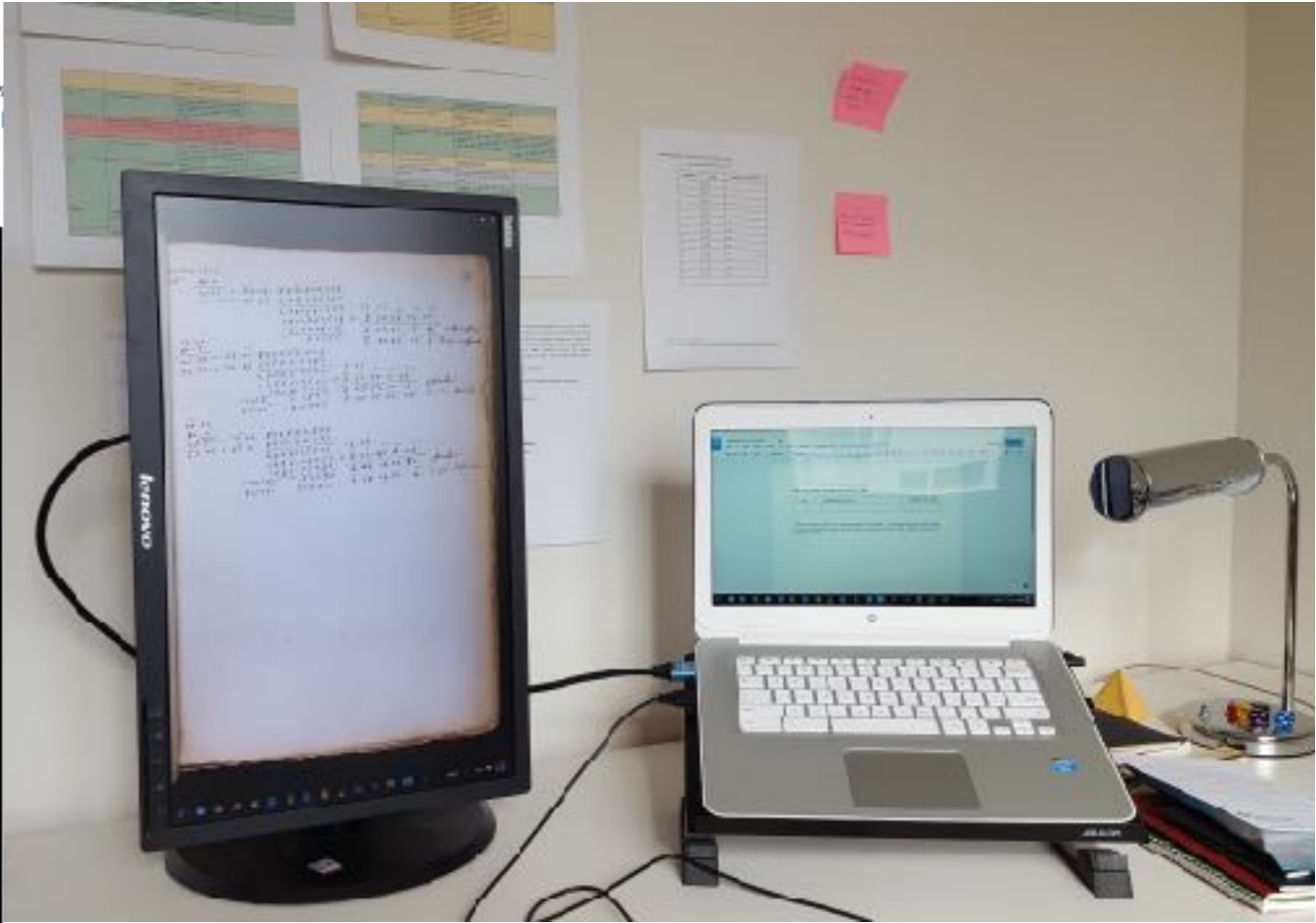
Die Qualität der Ansicht der Scans im Browser ist häufig ungenügend. Zum besseren Lesen empfiehlt sich der Download und eine Ansicht z.B. mit MeshLab (externer)

reset camera center camera flip y flip y-z
PointSize:1.00



Maße: 74,0 X 46,0 X 22,0
Bibl.: Urp.

3D-Scan, Screenshot von [Clay Tablet HS0106](#)



Arbeitsplatz von Robert Goulding, beim Bearbeiten der Manuskripte von [Thomas Harriot](#)

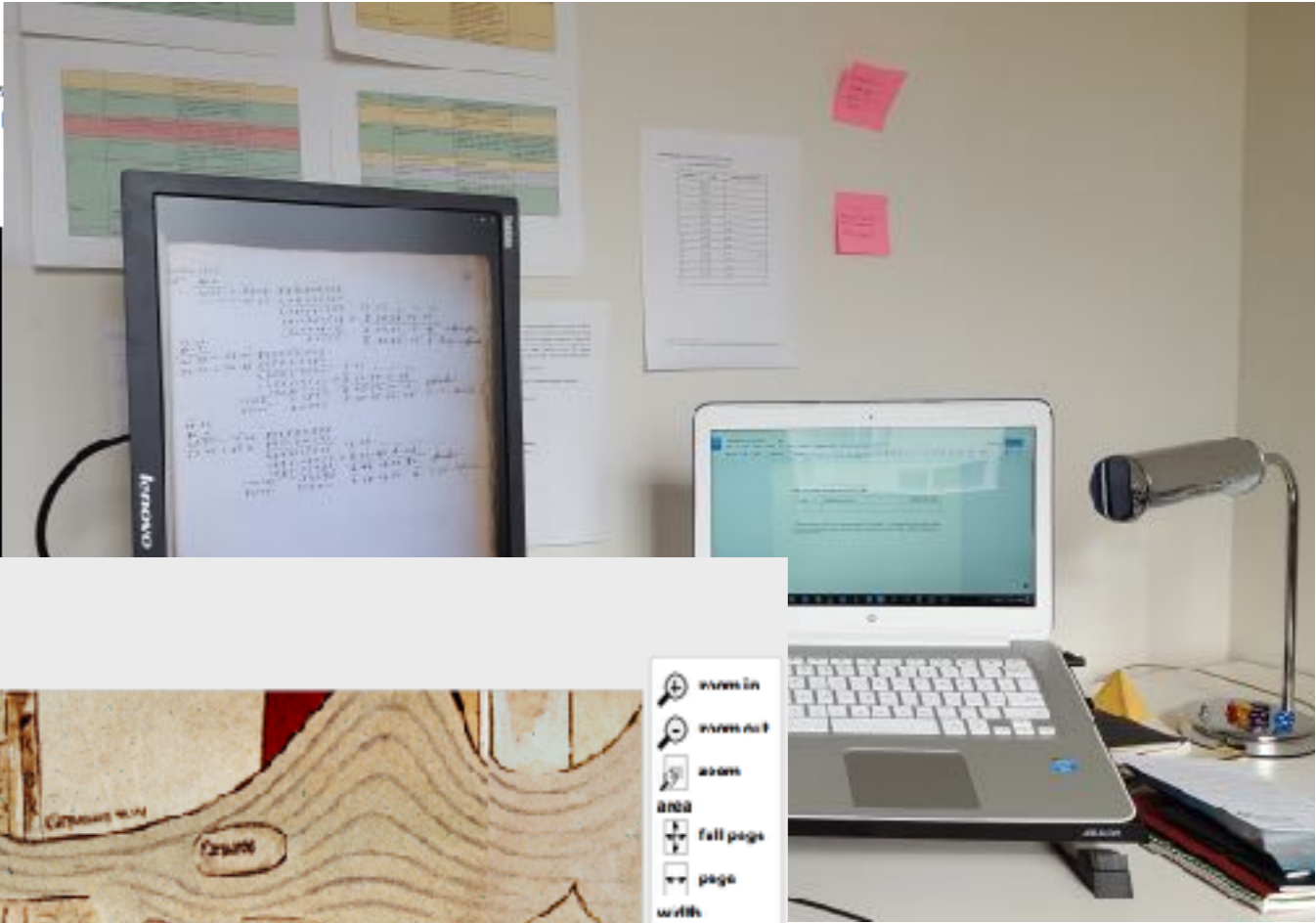
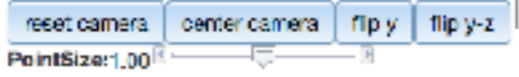
HS 0106 (size: 94.20MB) ([Download - meshlabJS](#))



Maße: 74,0 X 46,0 X 22,0
Bibl.: Urp.

3D-Scan, Screenshot von [Clay Tablet HSO](#)

Die Qualität der Ansicht der Scans im Browser ist häufig ungenügend. Zum besseren Lesen empfiehlt sich der Download und eine Ansicht z.B. mit MeshLab (externer...)



ECHO [Ebstorfer Weltkarte, 1492](#)

Image Thumbnail overview Document information

List of thumbnails

page << 1 of 1 >>

None Thumbnail

Hi-Resolution Scan, Screenshot der [Ebstorfer Weltkarte](#)

kripte von [Thomas Harriot](#)

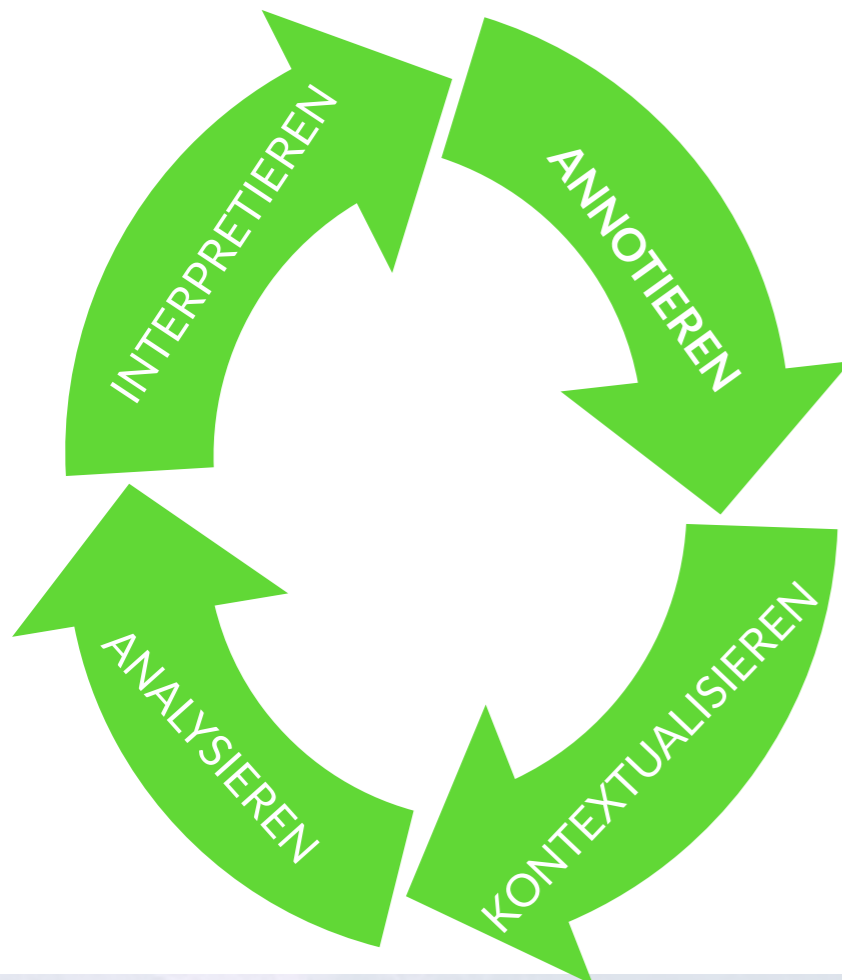
Forschungsaktivitäten

Exploration

Zusammenfassung
existierender Quellen

Korpusbildung

Sampling/Filtern



Publizieren

Forschungsaktivitäten

Exploration

Zusammenfassung
existierender Quellen

Korpusbildung

Sampling/Filtern

INTERPRETIEREN

ANNOTIEREN

ANALYSIEREN

KONTEXTUALISIEREN

Publizieren

Ergebnisse

Forschungsfrage

Hypothesen

Korpus

Sample/Teilkorpus

verfeinerte
Hypothesen

neue Hypothesen

Antworten

Visualisierungen



Forschungsaktivitäten

Exploration

Zusammenfassung
existierender Quellen

Korpusbildung

Sampling/Filtern

INTERPRETIEREN

ANNOTIEREN

ANALYSIEREN

KONTEXTUALISIEREN

Publizieren

Ergebnisse

Forschungsfrage

Hypothesen

Korpus

Sample/Teilkorpus

verfeinerte
Hypothesen

neue Hypothesen

Antworten

Visualisierungen

Wissensgenerierung

Notizen

Quellenübersicht

Bibliographien

Datenbanken

Teilkorpora

Drafts/ Wrap-Ups

Präsentationen

Publikationen



GitHub interface showing the file `localbibliography.bib` in the repository `langsci / 215`. The file is 6459 lines (5236 sloc) and 163 KB. The content is a BibTeX bibliography with two entries:

```
1 % Encoding: UTF-8
2 @article{Aalberse2009,
3     author = {Aalberse, Suzanne and Don, Jan},
4     journal = {Morphology},
5     pages = {3-14},
6     title = {Syncretism in {Dutch} dialects},
7     volume = {19},
8     year = {2009}
9 }
10
11
12 @article{Aalberse2011,
13     author = {Aalberse, Suzanne and Don, Jan},
14     journal = {Morphology},
15     pages = {327-350},
16     title = {Person and number syncretisms in {Dutch}},
17     volume = {21},
18     year = {2011}
19 }
20
```

<https://github.com/langsci/215/blob/master/localbibliography.bib>

Reproduzierbare Forschung

- Dokumentation von genutzten Programmen und ihrer Anwendung
- Genaue Nachvollziehbarkeit der genutzten Datasets

Reproduzierbare Forschung

- Dokumentation von genutzten Programmen und ihrer Anwendung
- Genaue Nachvollziehbarkeit der genutzten Datasets
- Erfordert eine wissenschaftliche Praxis, in der Daten und Software ebenso zitiert werden wie Literatur!

Zitierweise von Tools

- Autor_in des Tools schlägt Zitierweise vor
 - üblicherweise ein wissenschaftlicher Artikel

Zitierweise von Tools

- Autor_in des Tools schlägt Zitierweise vor
- üblicherweise ein wissenschaftlicher Artikel

Publications

If you use HeidelTime, please cite the appropriate paper (in general, this would be the journal paper [4]; if you use HeidelTime with automatically created resources, please cite paper [10]; if you use HeidelTime for temponym tagging, please cite paper [11]):

1. Strötgen, Gertz: HeidelTime: High Quality Rule-based Extraction and Normalization of Temporal Expressions. SemEval'10. [pdf](#) [bibtex](#)
2. Strötgen, Gertz: Temporal Tagging on Different Domains: Challenges, Strategies, and Gold Standards. LREC'12. [pdf](#) [bibtex](#)

<https://github.com/HeidelTime/heideltime#publications>

Zitierweise von Tools

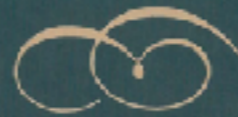
- Link zum Quellcode
- Angabe der Versionsnummer
- Dokumentation der verwendeten Parameter
- Link zu Quelldateien und Ergebnissen

Zitierweise von Tools

- Link zum Quellcode
- Angabe der Versionsnummer
- Dokumentation der verwendeten Parameter
- Link zu Quelldateien und Ergebnissen

```
$ java -jar de.unihd.dbs.heideltime.standalone.jar \  
Ulysses.234m1.0.xml -dct 1904-06-16 -t COLLOQUIAL -v_
```

«Literate Programming»



Donald E. Knuth

<Emphatic declarations 1>;

emphic; array [size] of *small*; *large*; *beauty*; *real*;

<True confessions 10>;

for *readers* (*human*) do write (*web*);

while *programming* = *art* do

 begin *incr* (*pleasure*); *decr* (*bugs*); *incr* (*portability*);

incr (*maintainability*); *incr* (*quality*); *incr* (*solitary*);

 end (*happily ever after*)

This code is used in theory and practice.

I believe that the time is ripe for significantly better documentation of programs, and that we can best achieve this by considering programs to be works of literature. Hence, my title:
»Literate Programming«

-Donald E. Knuth



Code is more often read than written.

-Guido van Rossum



Anwendungsgebiete

- Programmierung und Dokumentation von LaTeX-Klassen

L^AT_EX

Anwendungsgebiete

- Programmierung und Dokumentation von LaTeX-Klassen

L^AT_EX

- Dokumentation und Erstellung von XML-Schema in TEI



Literate Programming und interaktive Publikationen

- Interaktive Publikation, die die Methode beschreibt und die Ergebnisse dokumentiert
- Browsergestützte Anwendungen

Literate Programming und interaktive Publikationen

- Interaktive Publikation, die die Methode beschreibt und die Ergebnisse dokumentiert
- Browsergestützte Anwendungen
- Die vorgestellte Entwicklung ist die Integration dieser Notebooks in eine Publikationsplattform

Jupyter Notebooks



The screenshot shows a Jupyter Notebook window titled "Webseiten_als_Quellen.ipyn". The notebook content includes a title "Suchanfrage I: isisCB explore", a paragraph explaining the task, and three code cells. The first cell (index 8) defines a URL, makes a request, and parses the XML. The second cell (index 9) defines a base URL. The third cell (index 10) finds paragraphs containing "Article" or "Book".

```
Webseiten_als_Quellen.ipyn ×
Python 3

Suchanfrage I: isisCB explore

Auf der Webseite für das isisCB explore Tool der "Isis Bibliography of the History of Science" werden Suchergebnisse über eine Link-Liste dargestellt.

In der Liste werden einzelne Ergebnisse in paragraphen sortiert (mit dem tag <p> ). Wir generieren eine Unterliste aller paragraphen die das Word Article oder Book enthalten.

[8]: url2 = 'https://data.isiscb.org/isis/?q=copernicus&models-isisdata.citation&sort_order_citation-publication_date_for_sort&sc
r2 = requests.get(url2)

data2 = r2.text

soup2 = BeautifulSoup(data2, 'xml')

[9]: baseURL = 'https://data.isisch.org/'

Suche alle Paragraphen, die im Text das Wort Article oder Book enthalten.

[10]: plist = soup2.findAll('p')
queListe = [x for x in plist if "Article" in x.text or "Book" in x.text]
***
```

Jupyter Notebooks



Example_3.ipynb

Markdown ▾

Chapter

Title 1

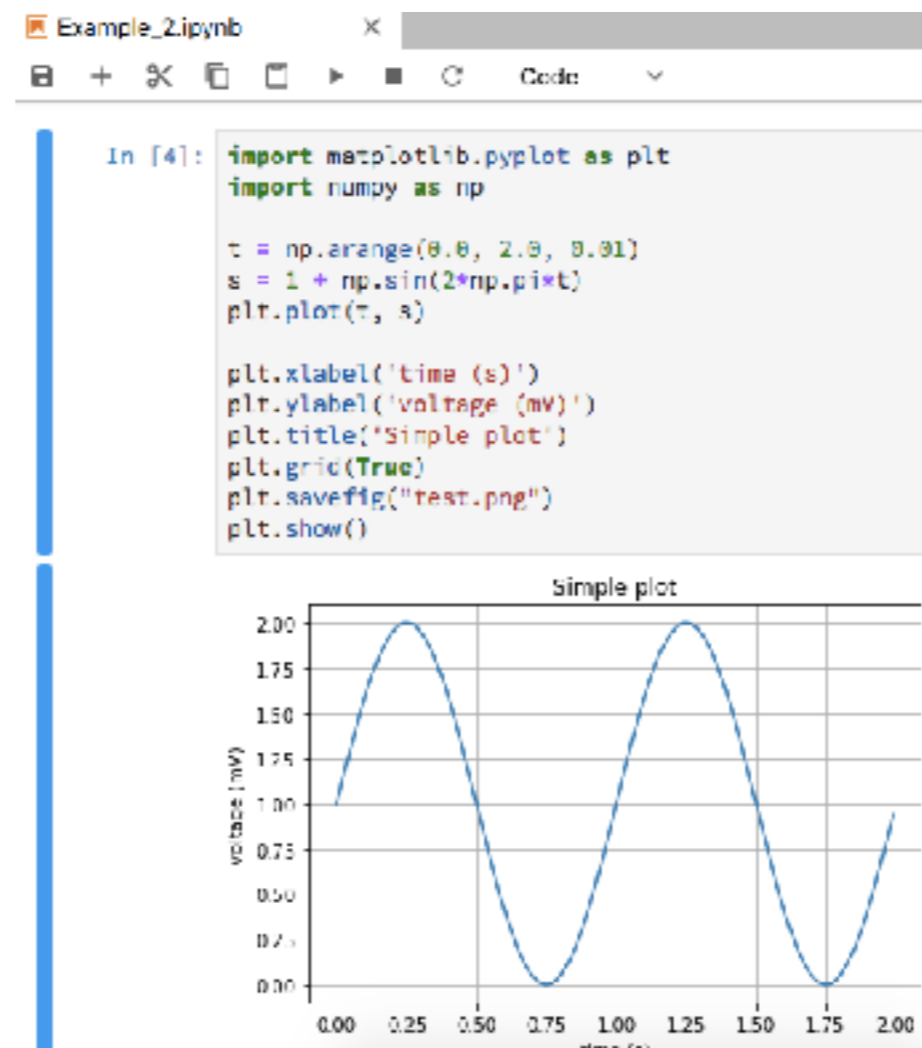
Text has *many* options for **formatting**.

E.g. `[Link](https://www.mpiwg-berlin.mpg.de)` becomes [Links](https://www.mpiwg-berlin.mpg.de)

Subsection 1

Formulaes can be written in Latex: `$\int_0^{10} x^3 dx$` becomes $\int_0^{10} x^3 dx$

Jupyter Notebooks



Datenstruktur

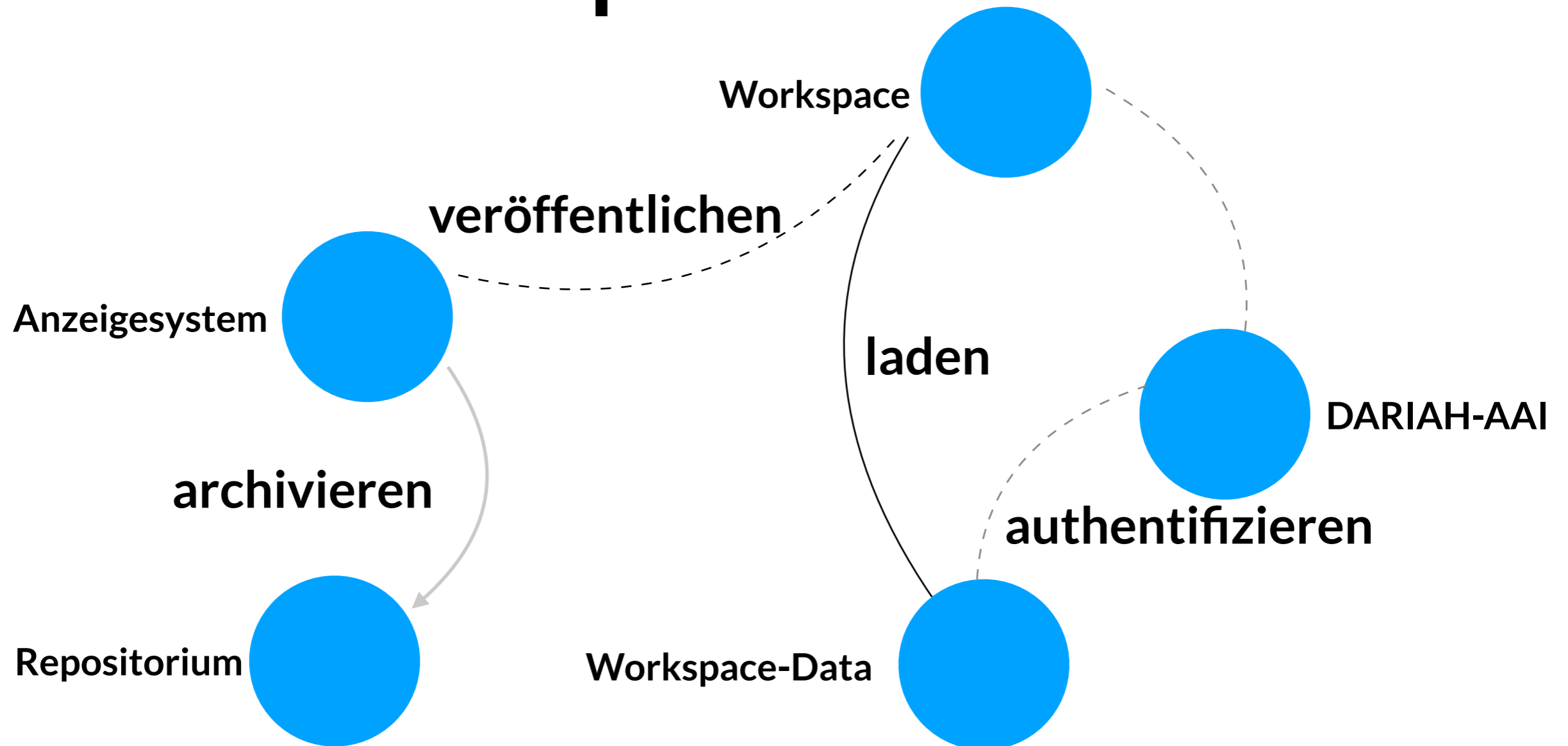
```
publication/  
├── config  
│   └── jupyter_notebook_config.py  
├── data  
│   ├── index.ipynb  
│   └── example_data.csv  
├── Dockerfile  
└── packages  
    └── requirements.txt
```


Archivierbar

Webseiten_als_Quellen.ipyn...X

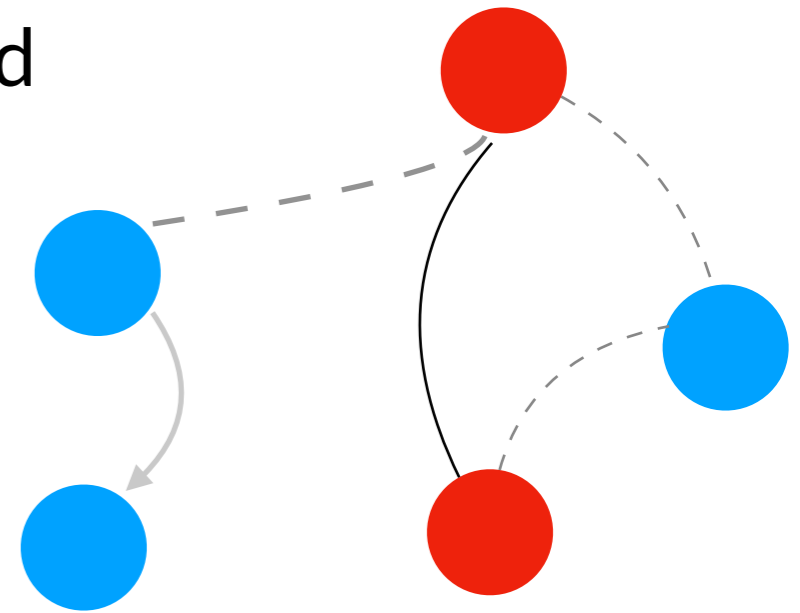
```
1  {
2  "cells": [
3  {
4  "cell_type": "markdown",
5  "metadata": {},
6  "source": [
7  "# Beispiel: Arbeiten mit Webseiten als Quellen\n",
```

Komponenten



Komponenten

- Workspace-Data: Ablage für Daten und Programme, Kollaboration
- Workspace: Notebooks schreiben und ausführen

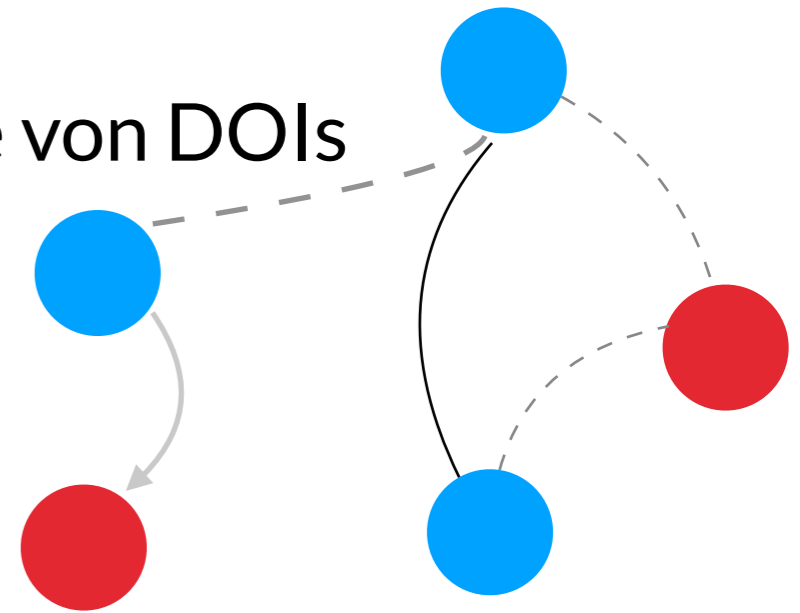


Komponenten

- Anmeldung über DARIAH-AAI



- Publikation in Repository, Vergabe von DOIs




Projects

Projects


localhost:8002/projects/

90%



eoa edition open access


[About EOA](#) [What is Open Access?](#) [Participate](#) [Projects](#) [Editorial Board](#)



Edition Open Sources

Edition Open Sources (EOS) pioneers a new paradigm in publishing, devoted to open access and high-

↓ More



Max Planck Research Library
for the History and
Development of Knowledge

MPRL

The Max Planck Research Library for the History and Development of Knowledge, in accordance with the

↓ More



Edition Open Sources

Edition Open Sources (EOS) pioneers a new paradigm in publishing, devoted to open access and high-

↓ [More](#)

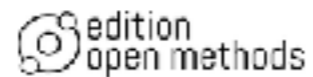


Max Planck Research Library
for the History and
Development of Knowledge

MPRL

The Max Planck Research Library for the History and Development of Knowledge, in accordance with the

↓ [More](#)

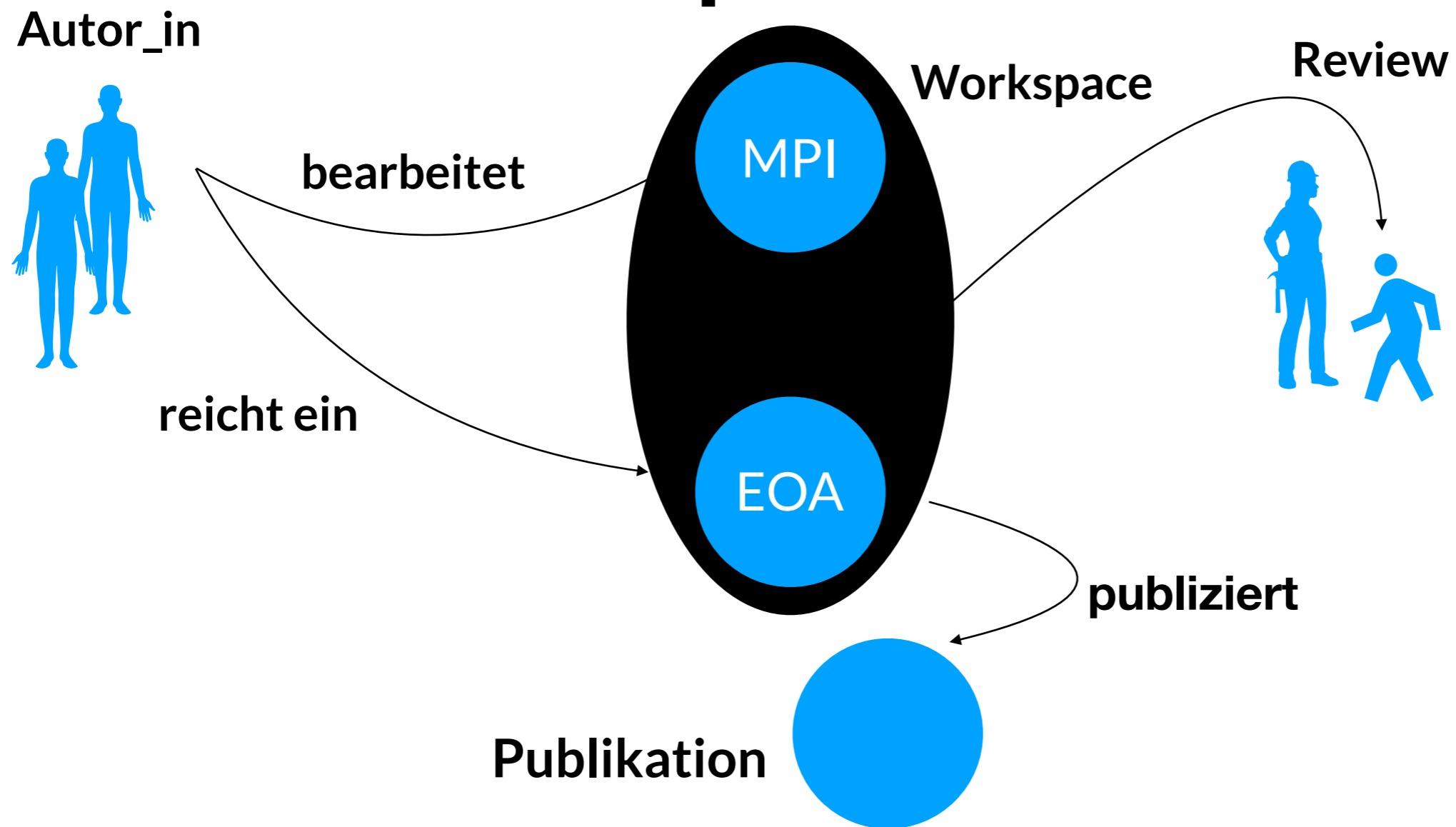


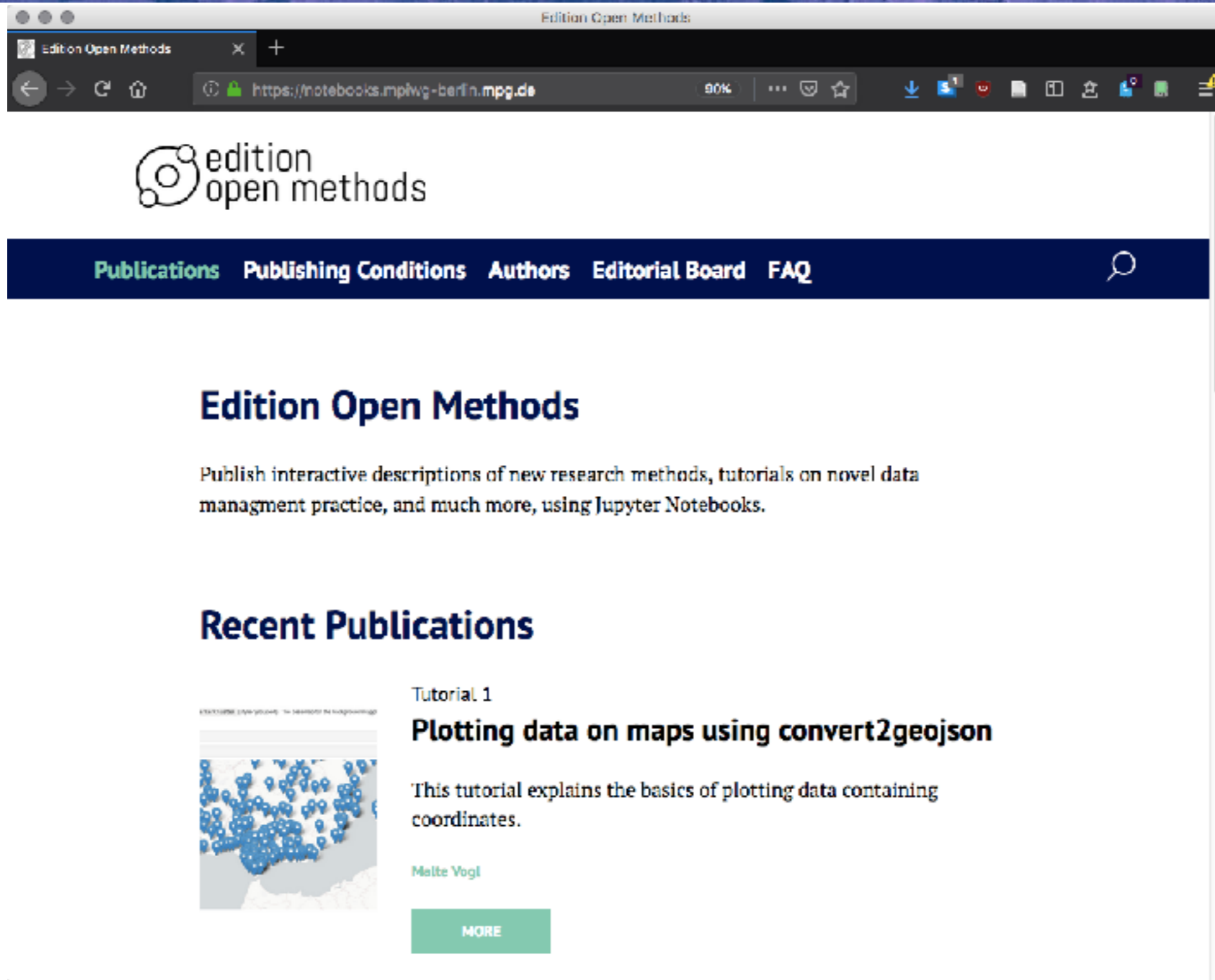
Edition Open Methods

Publish interactive descriptions of new research methods, tutorials on novel data management practice, and

↓ [More](#)

Workspace & EOA





The image shows a browser window with the URL <https://notebooks.mpiwg-berlin.mpg.de>. The page features the logo for 'edition open methods' and a navigation bar with links for 'Publications', 'Publishing Conditions', 'Authors', 'Editorial Board', and 'FAQ'. The main content area is titled 'Edition Open Methods' and includes a sub-header 'Recent Publications'. One publication is highlighted: 'Tutorial 1: Plotting data on maps using convert2geojson' by Melite Vogl. The description states: 'This tutorial explains the basics of plotting data containing coordinates.' A green 'MORE' button is located below the description.

edition
open methods

[Publications](#) [Publishing Conditions](#) [Authors](#) [Editorial Board](#) [FAQ](#)

Edition Open Methods

Publish interactive descriptions of new research methods, tutorials on novel data management practice, and much more, using Jupyter Notebooks.

Recent Publications

Tutorial 1

Plotting data on maps using convert2geojson

This tutorial explains the basics of plotting data containing coordinates.

Melite Vogl

MORE

Zusammenfassung

- Komponenten von Forschungsinfrastrukturen werden weitergenutzt (Repositorien, Authentifizierung)
- Open-Access-Publikationen können mit Code-Elementen angereichert werden
- Engere Verschränkung von Forschungsdaten und deren Interpretation

Zusammenfassung

- Relevanz von Forschungsdaten im Forschungs- und Begutachtungsprozess erhöhen
- Wissenschaftler_innen ermutigen, sowohl Text als auch Daten zu publizieren
- Förderern und anderen Komitees zeigen, dass auch Forschungsdaten eine ernstzunehmende wissenschaftliche Leistung sind

Quellen

- Jupyter Notebook: <https://jupyter.org/>
- Jupyter Hub: <https://github.com/jupyterhub/jupyterhub>
- Nextcloud: <https://nextcloud.com>
- Notebook Beispiele: <https://github.com/jupyter/jupyter/wiki/A-gallery-of-interesting-Jupyter-Notebooks>
- DH-Bibliotheken: <https://github.com/computational-antiquity>
- Programming Historian: <https://programminghistorian.org>

Demo

- ① <https://workspace.mpiwg-berlin.mpg.de>
- ① <https://data.workspace.mpiwg-berlin.mpg.de>
- ① <https://notebooks.mpiwg-berlin.mpg.de>

Vielen Dank

<https://github.molgen.mpg.de/EditionOpenAccess>

<https://gitlab.gwdg.de/EditionOpenAccess>

<http://edition-open-access.de/>

<http://edition-open-sources.org/>

<http://mprl-series.mpg.de/>

ea@mpiwg-berlin.mpg.de

@Edition_OA

