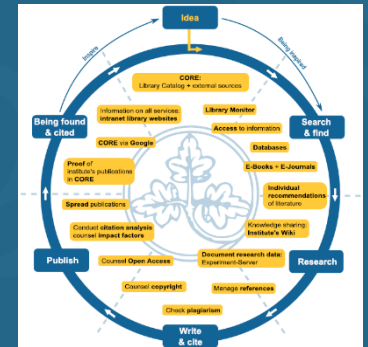


# Das Forschungsdatenmanagementsystem am MPI Gemeinschaftsgüter in Bonn

*Regina Goldschmitt, Dr. Robert Günther, Seyhun Saral, Marius Sonnenschein*  
MPI zur Erforschung von Gemeinschaftsgütern  
{goldschmitt, guenther, sonnenschein, saral}@coll.mpg.de



MPG-Forschungsdatenmanagement-Workshop  
am 19. April 2018 in München

# MPI zur Erforschung von Gemeinschaftsgüter

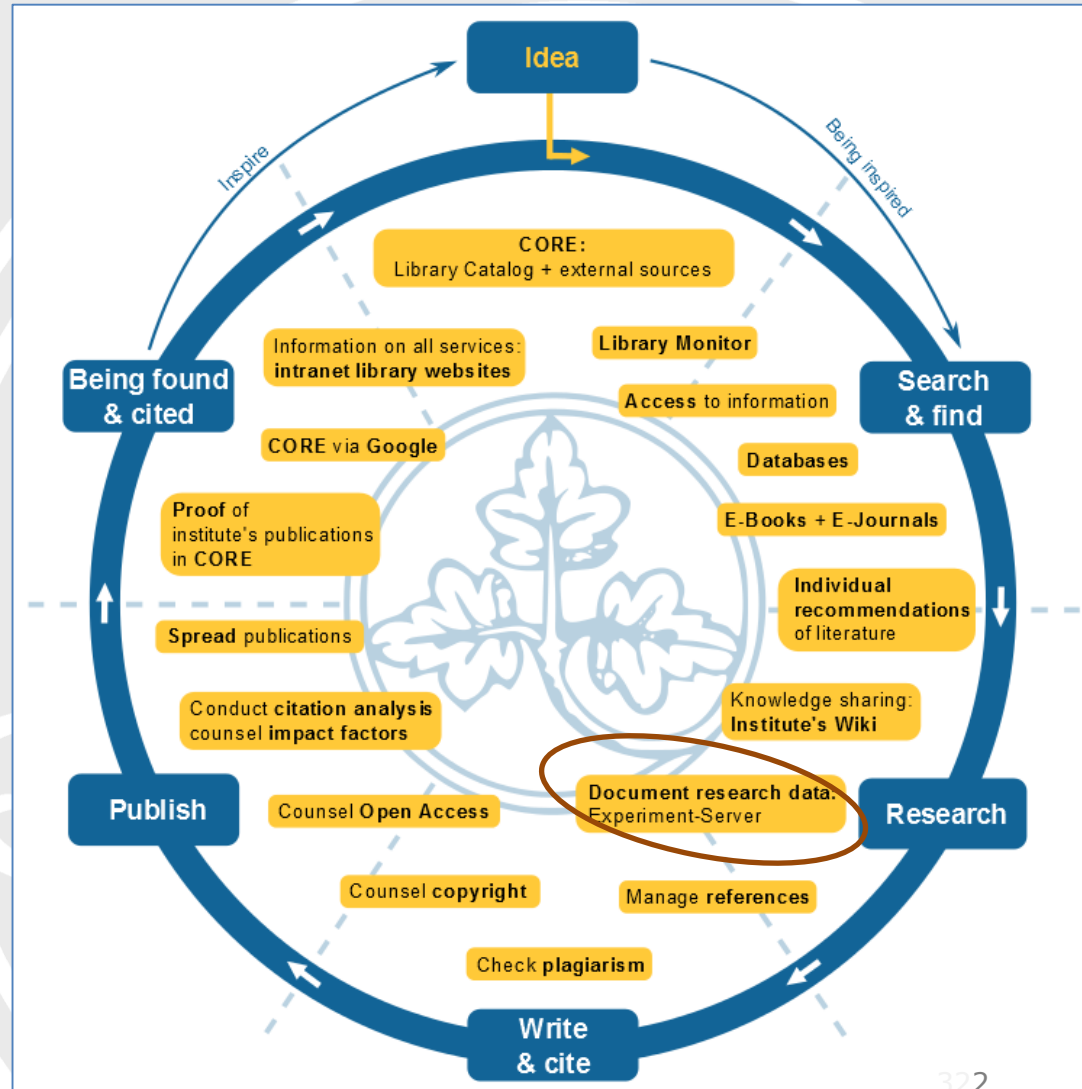
- 2 Abteilungen + 3 Nachwuchsgruppen
- Inter-/multidisziplinär:
  - Recht
  - Ökonomie
  - Psychologie
  - Soziologie
- **Fachspezifische kulturelle Unterschiede bzgl. :**
  - **Publikations- u. Rezeptionsverhalten**
  - **Aufgeschlossenheit für Open Science**
- Experimente im Labor + Feld
- Eyetracking, Videos, Interviews, Online-Surveys, Programmcodes, Dokumente, Statistiken
- Dateigrößen: 10-100 MG, bis TB pro Experiment
- Aktuell: 20 GB / 108 Projekten
- Seit März 2017 produktiv



# MPI zur Erforschung von Gemeinschaftsgüter

Motto der Bibliothek:

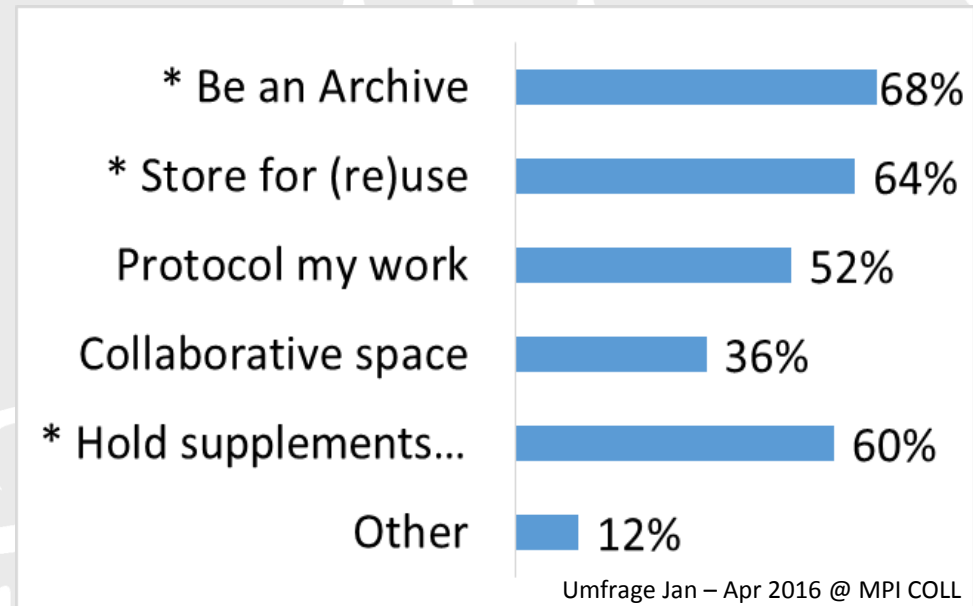
Unsere Nutzer beim **gesamten** Forschungsprozess unterstützen



- Kulturelles Problem
- Daten nicht in der Cloud, sondern Inhouse
- Größte Herausforderung: Zieldefinition / Policy
- Hardware: 200 GB Festplatte, 8 GB RAM, 2.40 GHz CPU
- Programmierkenntnisse: Python, JavaScript, Java
- Aufwand: ca. 1,5 FTE für 2 Jahre
- Starke Orientierung am Workflow der Wissenschaftler
- Schlüssel zum Erfolg: Kommunikation & Zusammenarbeit von Wissenschaft & Bibliothek & IT & Verwaltung

# Die Ziele

- **Infrastruktur** implementieren
- Umsetzung der Vorgaben der **MPG-Regeln zur guten wissenschaftlichen Praxis**
- "Weil sich die Forscher das wünschen" ;-)



- **Strukturierte Ablage und Versionskontrolle**
- **Kollaboratives Arbeiten für Institutsmitarbeiter**

- **Keine** Publikationsplattform
  - ⇒ Journal, Zenodo, Dataverse, OSF, Datorium, SowiDataNet, PsycData, ...
  - ⇒ Wissenschaftler entscheidet, ob, wann, wo und welche Daten veröffentlicht werden
- **Keine** formale Langzeitarchivierung
  - ⇒ Archivierung bei KEEPER

- Auftrag von GD/Institutsleitung: **Ablage-Tool**

- ⇒ Regeln der guten Wiss. Praxis!

- ⇒ Lokale Implementation (Datensicherheit)

- ⇒ Prozessorientiertes FDM-Tool

- 2 Jahre: 1 wissenschaftlicher Projektkoordinator, 1 Programmierer

- Frühjahr 2016 bis Frühjahr 2018

- Enge Kooperation von Wissenschaft, IT, Verwaltung und Bibliothek

- Bibliothek hat Federführung



# Auswahl des Systems

- Zur Auswahl standen
  - Dataverse
  - Zenodo
  - OSF
    - OS => lokale Anpassungen
    - Anwender-Empfehlung
    - Prozessorientierung
    - Flexibilität
- Vorstellung der Idee in Wissenschaftler-Runde
- MPG-FDM Workshop Ende Juni 2016



# Die Lösung: OSF (Open Science Framework)

**OSF has a rich feature set that enables efficient, organized research projects.**



## **Organize Your Projects.**

Your handy OSF project dashboard helps you keep all of your research and data at your fingertips.



## **Manage complex projects easily.**

Add data, contributors, links to other projects, citations, make certain parts public or private, and even create preprints.



## **Document your process.**

Each project's wiki you can keep a running history of all the work you've done.



## **Add components or folders easily.**

Projects can quickly grow, but OSF's structured approach keeps them clear and organized. Add folders, files, or additional research data.



## **Version Control? Got it.**

Analysis, code or paper drafts will probably change over time, and it is important to keep track of all these different versions. OSF makes it easy by automating version control for files with the same name.



## **Register your work.**

Registrations create time stamped, read-only versions of a project. This means that the files within the main project can continue to change, but the files in the registration will not.



## **Share your work easily.**

OSF allows you to share what you want, when you want. Everything on OSF is private by default. We also give you fine grain control over which aspects of the project are made public.



## **Add-ons extend the power of OSF.**

You can connect other tools you use to OSF--like Dropbox, Github, or Figshare--so you can follow your current workflow in a centralized environment.



## **Project analytics.**

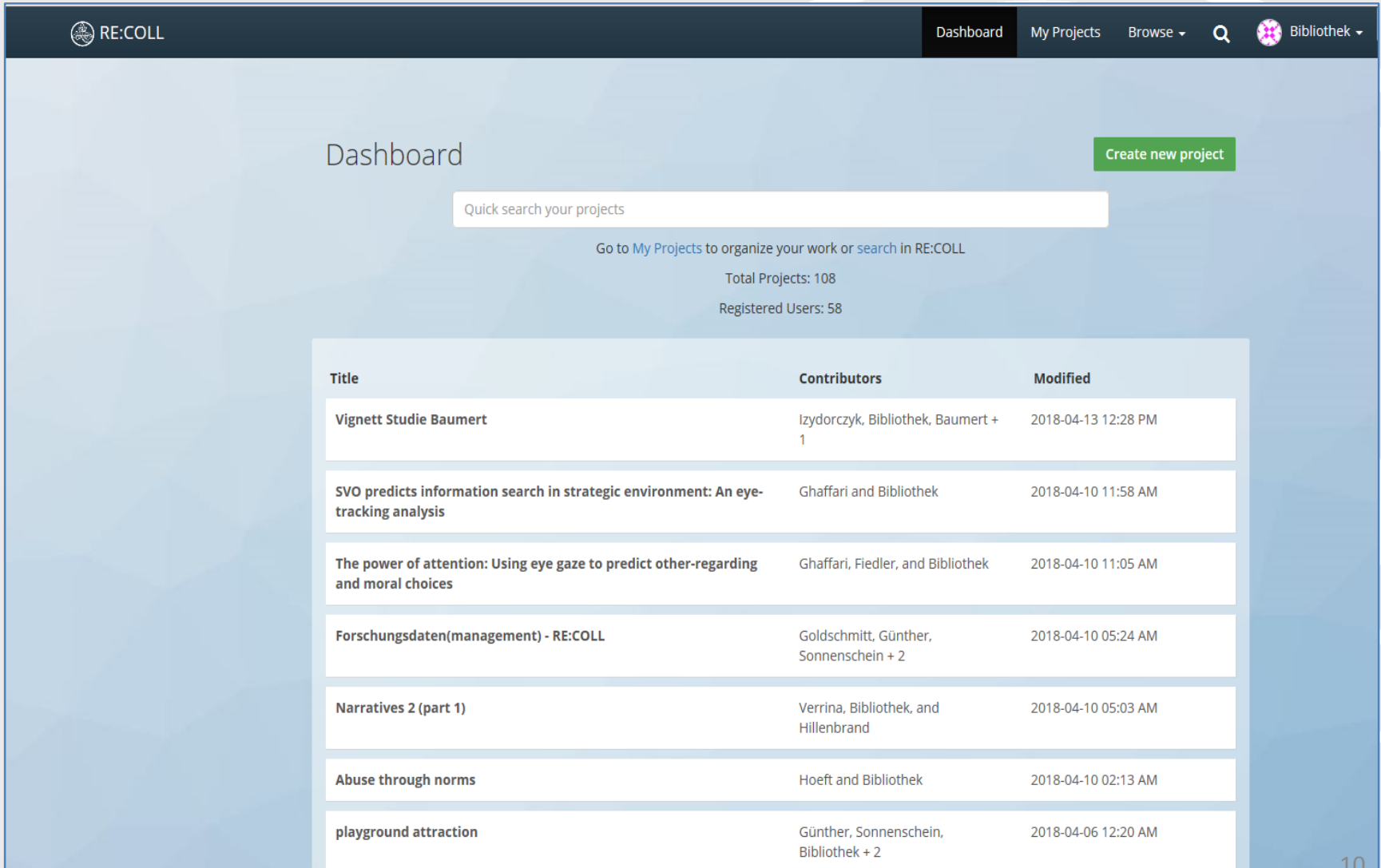
OSF stores project usage information to help you measure the impact of your work. We give you insights about impact beyond citation rates so you can determine how many people are accessing and using your research materials.

<https://osf.io>

<https://github.com/centerforopenscience>

# Die Lösung: OSF lokal = RE:COLL

to recall data, to recall something



The screenshot shows the RE:COLL dashboard interface. At the top, there is a navigation bar with the RE:COLL logo, a search icon, and a user profile icon labeled 'Bibliothek'. The main content area features a 'Dashboard' heading, a search bar, and a 'Create new project' button. Below this, there are statistics for 'Total Projects: 108' and 'Registered Users: 58'. A table lists several projects with their titles, contributors, and modification dates.

Title	Contributors	Modified
Vignett Studie Baumert	Izidorczyk, Bibliothek, Baumert + 1	2018-04-13 12:28 PM
SVO predicts information search in strategic environment: An eye-tracking analysis	Ghaffari and Bibliothek	2018-04-10 11:58 AM
The power of attention: Using eye gaze to predict other-regarding and moral choices	Ghaffari, Fiedler, and Bibliothek	2018-04-10 11:05 AM
Forschungsdaten(management) - RE:COLL	Goldschmitt, Günther, Sonnenschein + 2	2018-04-10 05:24 AM
Narratives 2 (part 1)	Verrina, Bibliothek, and Hillenbrand	2018-04-10 05:03 AM
Abuse through norms	Hoelt and Bibliothek	2018-04-10 02:13 AM
playground attraction	Günther, Sonnenschein, Bibliothek + 2	2018-04-06 12:20 AM

https://re.coll.mpg.de (nur intern)

# Geld für Experimente



<http://severa-ernst.de/wordpress/wp-content/uploads/2012/09/Buchhaltung.jpg>

# NUDGING: Antrag für Experimentiergelder

RE:COLL Dashboard My Project

Forschungsdaten(management) - RE:COLL Files Wiki Contributors **Settings**

Project

Coll Project Settings

Wiki

Project Affiliation / Branding

Email Notifications

COLL Project Settings / **Payment Order** → €€€ for Participants

Research Group: -- Select a Research Group --

Period / Duration:

From: 2018-03-25 Until: 2018-03-26

Number of participants:

36

Average expenditure per participant:

20

Payment carried out by: bank transaction

Cancel Save changes

Wiki

Enable the wiki in Forschungsdaten(management) - RE:COLL.

To allow all RE:COLL users to edit the wiki, Forschungsdaten(management) - RE:COLL must be published.

Project Affiliation / Branding

Projects can be affiliated with institutions that have created OSF for Institutions accounts. This allows:

- institutional logos to be displayed on public projects
- public projects to be discoverable on specific institutional landing pages
- single sign-on to the OSF with institutional credentials
- FAQ

Email Notifications

These notification settings only apply to you. They do NOT affect any other contributor on this project.

- Forschungsdaten(management) - RE:COLL
  - Comments added
  - Files updated

Möchten Sie „PaymentOrder.pdf“ (53,3 KB) von „re.coll.mpg.de“ öffnen oder speichern? Öffnen Speichern

Max-Planck-Institut zur Erforschung von Gemeinschaftsgütern  
Max Planck Institute for Research on Collective Goods



## PAYMENT ORDER

for money used in experiments

Applicant	Regina Goldschmitt
Research unit	
Exact name of experiment	Forschungsdaten(management) - RE:COLL
Abbreviation of experiment (acronym)	RE:COLL
RE:COLL Project ID <small>(mandatory; please do not change)</small>	9ust5
Period / Duration	2018-03-25 - 2018-03-26
Number of participants	36
Average expenditure per participant **	20.0
Sum in EUR	720.0

\*\*Expenses per participant cannot be more than 25% of the calculated average. If they are more than 25%, we require an additional explanation. If necessary, please talk to Ms Morgenstern about the expenditure before the experiment starts.

Payment carried out by

bank transaction

Cheques and small-change denominations over 1,500 euro have to be declared to the bank in advance. Please send a brief e-mail with the details (when/ which denominations etc.) to the accounting department!

Brief description of experiment theme and its objectives:

Vorträge, Aufsätze, Protokolle im Zusammenhang mit der Einführung von RE:COLL als Nachfolger des Experimentsservers EXP. Vorschläge für eine FDM-Richtlinie.

Bonn, \_\_\_\_\_ Signature of applicant: \_\_\_\_\_

Bonn, \_\_\_\_\_ Signature of Director or  
Team leader for approval: \_\_\_\_\_

(To be filled in by administration) ↓

Factually correct:	
Established by:	
Payment ordered:	
Account:	
Cost unit / Project:	

# Orientierung am Workflow

RE:COLL Dashboard My Projects Browse Q Bibliothek Send Feedback

The visible and hidden costs of control... Files Contributors Settings

## The visible and hidden costs of control under delegation

Private P 0

**Contributors:** Claudia Cerrone, Philip Brookings  
**Date created:** 2018-03-05 11:21 AM | **Last Updated:** 2018-03-05 02:26 PM  
**Abbreviation of experiment (acronym):** delegation  
**Description / Abstract:** When delegating a decision to an agent, principals may grant limited discretion (e.g. by allowing the agents to choose from a subset of options) in order to avoid that agents make suboptimal choices. But agents can react to the limited discretion received by reducing their performance. This project explores whether agents' response to principals' delegation decisions are driven by reciprocity, and whether principals, anticipating that, will retain full control or grant full discretion, rather than opting for limited discretion.

€€€ for Participants

**Files**

Name	Modified
The visible and hidden costs of control under delegation	
- COLL Storage	
- 0. General	
- COLL Storage	
- 1. Raw Data	
- COLL Storage	
- 2. Prepared - Analyzed Data	
- COLL Storage	
- 3. Scripts used	
- COLL Storage	
- 4. Instructions	
- COLL Storage	
- 5. Publication	
- COLL Storage	

**Citation** re.coll.mpg.de/ukwx

**Components**

- 0. General  
Cerrone  
2 contributions
- 1. Raw Data  
Cerrone  
2 contributions
- 2. Prepared - Analyzed Data  
Cerrone  
2 contributions
- 3. Scripts used  
Cerrone  
2 contributions
- 4. Instructions  
Cerrone  
2 contributions
- 5. Publication  
Cerrone  
2 contributions

Recent Activity

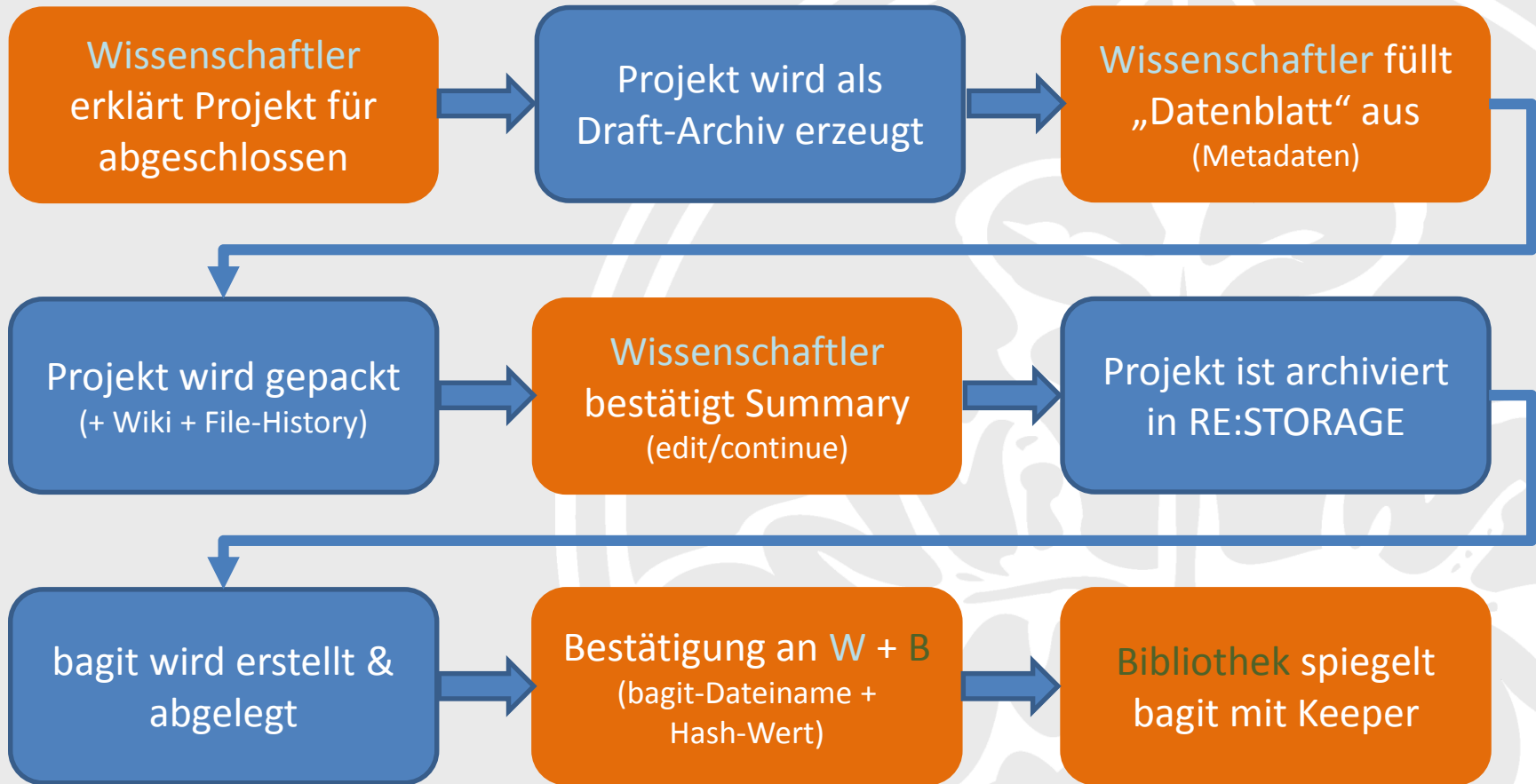
- Einarbeitung in Quellcode
- nicht benötigte Komponenten entfernen:  
Google Drive, Dropbox, figshare, OneDrive usw.
- Anpassung und Erweiterungen:
  - Anbindung an LDAP
  - Anpassung des Datenspeichers
  - Verschlüsselte Verbindung mit TLS
  - Einbau Feedback Button
  - Automatisches Anlegen der „fünf Ordner“



## Wichtige lokale Anpassungen

- Zugriff über EZ-Proxy
- Anzeige aller Projekte
- Erzeugen der RE:COLL Payment Order
- Bibliothek darf immer „Lesen“
- Zusätzliche Berechtigungsstufe P.I.
- Aus „fünf Ordner“ werden „fünf Components“
- Workflow Einfrieren des Status und Archivierung als zip

# Archivierungsprozess: Einfrieren / Archivieren





# Archivierung

## Erfolgsemail

To: sonnenschein@coll.mpg.de

From: re.bib@coll.mpg.de

Subject: Confirmation: New archive created on RE:CO

Message: Hello Marius Sonnenschein,

You successfully archived the project Components. Yeah!  
Please find a copy of the archived data in the RE:COLL

File name: rz38t09.04.18-10-03-16.zip

SHA256 Hash Value: 76b8511982905ae3c9918575ff43

Sincerely,

Container-Software bagit =  
im Prinzip zip-Archiv mit  
vorgegebener Struktur + Metadaten  
nach Dublin Core

## Components. Yeah!

Contributors: Robert Günther, Marius Sonnenschein

Date created: 2017-07-13 08:16 AM | Last Updated: 2018-04-09 05:03 PM

Abbreviation of experiment (acronym): CY!

Description / Abstract:

...and: The whole thing can be made "public" (for members of RE:COLL)

€€€ for Participants

Wiki

NEW

- Make projects public within the institute.
- export Wiki as PDF
- New projects organized in components.
- Support page with FAQs

Updated

- Payment order form for long abstracts.
- PDF form will be created on the server side.

Fixed

- Rendering of \*.docx files -
- Fixed IE glitches on myProjects page.

More on "Release Notes"  
[Read More](#)

Files

Click on a storage provider or drag and drop to upload

Filter ⓘ

Name	Modified
Components. Yeah!	
- RE:Storage	
rz38t09.04.18-10-03-16.zip	2018-04-09 05:03 PM
- 0. General	
- RE:Storage	
json.json	2018-02-26 10:56 AM
- 1. Raw Data	
- RE:Storage	
Signals.xlsx	2017-11-16 10:10 AM
- 2. Prepared - Analyzed Data	
- RE:Storage	
Live-Universe-Rest.xlsx	2017-11-20 10:01 PM
- 3. Scripts used	
- RE:Storage	
..	

Citation re.coll.mpg.de:5000/rz38t

Components [Add Component](#) [Link Projects](#)

### 0. General

Günther & Sonnenschein  
1 contribution

### 1. Raw Data

Günther & Sonnenschein  
0 contribution

### 2. Prepared - Analyzed Data

Günther & Sonnenschein  
0 contribution

### 3. Scripts used

Günther & Sonnenschein  
0 contribution

### 4. Instructions

Günther & Sonnenschein  
0 contribution

### 5. Publication

Günther & Sonnenschein  
0 contribution

Tags

add a tag to enhance discoverability

Recent Activity

Marius Sonnenschein added file rz38t09.04.18-10-03-16.zip to RE:Storage in Components. Yeah!  
2018-04-09 05:03 PM

Marius Sonnenschein initiated registration of Components. Yeah!

# Beschreibende Elemente bei der Archivierung

Bezeichner	Bemerkung	Beispiel
ID	Bezeichner auf RE:COLL	3fgh8
Title	Titel des Projekts	„Normative Abweichungen in sozialen ...“
P.I.	Projektverantwortliche Person	Planck, Max
Archived on	Datum d. Archivierung	2018-02-07
Date	Datum d. Projektbeginn	2016-11-05
Filename	Name des Containers	2018-02-07_3fgh8.zip
Hash	Prüfsumme nach SHA-2	e3b0c44298fc1c149afbf4c8996fb92427ae41e46 49b934ca495991b7852b855

Anforderungen gemäß Dublin Core, MPI-intern  
Minimum an Metadaten, da archiviert /nicht veröffentlicht wird!

# Serverübersicht

Mail (mail.coll.mpg.de)  
– Benutzer „re.bib“

Gitolite  
– Benutzer „gitolite3“

Postgresql Server  
– Database

OSF (re.coll.mpg.de)  
– Server :443  
– API :8000  
– Celery Worker :5672  
– (Nicht nach Außen geöffnet)

Mongo (re.coll.mpg.de)  
– Wiki  
Elasticsearch :9200

CAS (re.coll.mpg.de)  
– Server :8443

Postgresql (re.coll.mpg.de)  
– Session

MFR (re.coll.mpg.de)  
– Server :7778

Waterbutler (re.coll.mpg.de)  
– Server :7777

# Lessons Learned

- Der Knüller: Antrag auf Experimentiergelder in RE:COLL
- OSF: **offener Quellcode, aber wenig Support**
- OSF: ständige **Weiterentwicklung, die wir nicht mitziehen**
  
- Nach Projektende (5/2018):  
**Lab-Manager Seyhun Ali Saral & Bibliothek**
  
- Hohe Akzeptanz: **108 Projekte, 58 Nutzer**
- **Bewusstsein für Forschungsdaten**



## Ausblick:

- Kulturwandel zu mehr **Open Science**
- Spannungsfeld mit Datenhoheit, **Datensicherheit, Datenschutz**

# Vielen Dank für Ihre Aufmerksamkeit!

Für Fragen stehen wir gerne zur Verfügung:

- Marius Sonnenschein
- Regina Goldschmitt
- Robert Günther
- Seyhun Ali Saral

*{goldschmitt, sonnenschein, guenther, saral}@coll.mpg.de*

Max-Planck-Institut zur Erforschung von Gemeinschaftsgütern

Bibliothek

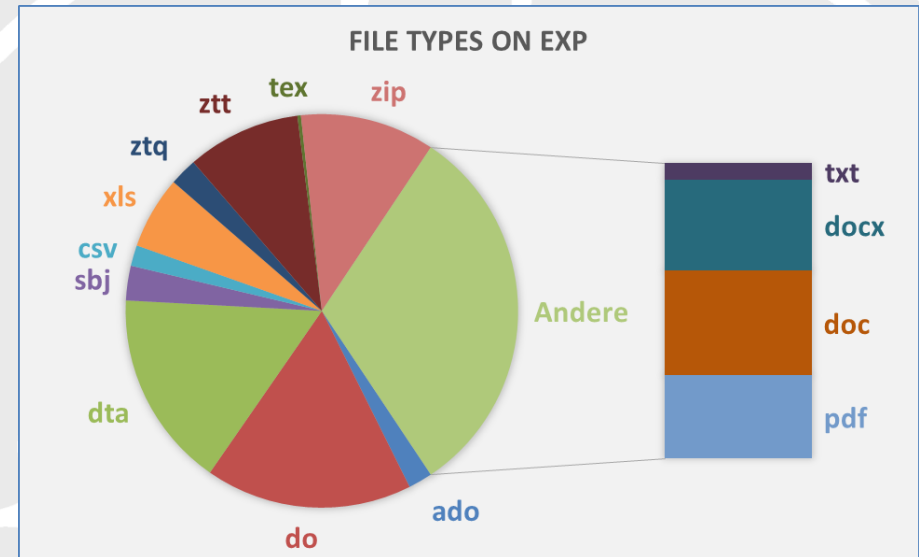
Kurt-Schumacher-Str. 10, 53113 Bonn

<http://www.coll.mpg.de/>

<http://www.coll.mpg.de/content/institute-library>

# Anhang 1: Datenarten und Datenmenge

- Rohdaten für Statistik
- Programmcode
- Dokumente, wie z.B. Manuskripte, Notizen, Instruktionen usw.
- Experimente im Labor, Eyetracking, Videos, Interviews, Feldversuche, ...
- Zunächst: pro Experiment ca. 10 bis 100 MB, bis TB;
- **Produktiv seit März 2017**
- **5 GB/Datei Upload**
- **Aktuell: 20 GB / 108 Projekte**



(Umfrage Jan – Apr 2016 @ MPI COLL)

## Anhang 2: Kerndatensatz Metadaten am MPI COLL

Element	Beschreibung	Element	Beschreibung
Title	Titel des Projekts auf RE:COLL	Format	Mime-type
Creator	Verantwortlicher Wissenschaftler	Identifizier	Interner Bezeichner auf RE:COLL
Contributor	Koautoren (auch externe)	Source	Labor, in dem das Experiment durchgeführt wurde (DecisionLab, BonnEconLab oder ähnliches)
Description	Kurzbeschreibung	Language	Sprache
Subject	Schlüsselwörter	Relation	Beziehung zur Veröffentlichung, als Referenz einzugeben
Publisher	MPI Gemeinschaftsgüter	Project lifespan	Laufzeit des Projekts
Date	Datum der Archiverstellung	Rights	© MPI Gemeinschaftsgüter; kann auf CC-By gesetzt werden
<<Type>>	Studentyp als kontrolliertes Vokabular		