



MAX PLANCK  
digital library

# Trust Global and Collaborate Local to Changing for Consistency

## Re-Integrating an Open Data Repository for the Max Planck Society

OR2022 24x7 Track, 8<sup>th</sup> June 2022, Denver

Yves Vincent Grossmann, [grossmann@mpdl.mpg.de](mailto:grossmann@mpdl.mpg.de), <https://orcid.org/0000-0002-2880-8947>

Michael Franke, [franke@mpdl.mpg.de](mailto:franke@mpdl.mpg.de), <https://orcid.org/0000-0002-2661-8242>

## What are the next seven minutes about?



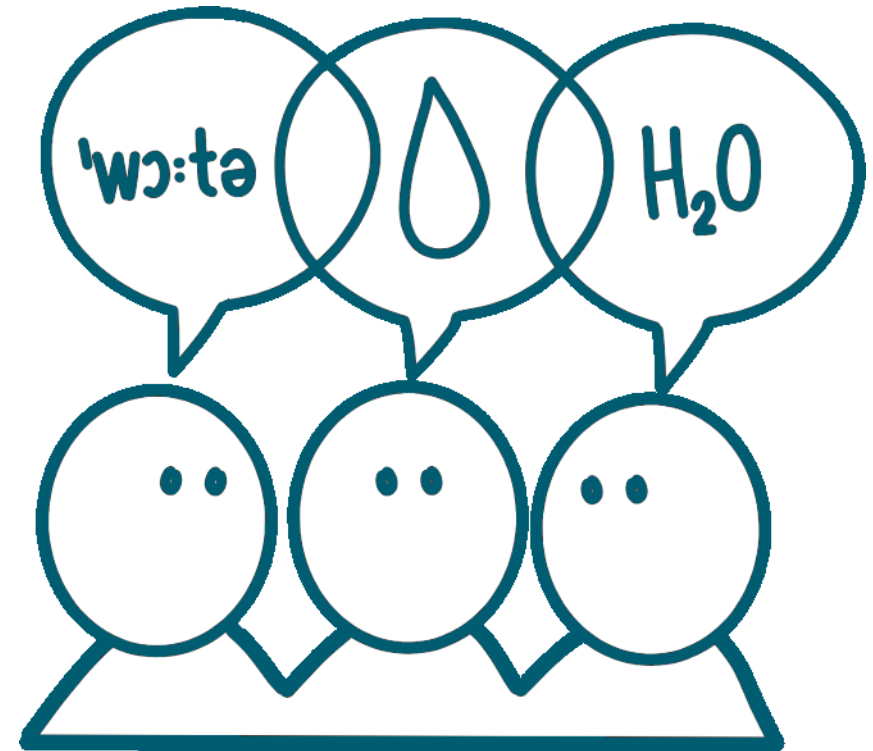
- <https://edmond.mpdl.mpg.de>
- Open data repository for the Max Planck Society

### I am:

- Research Data Management Officer at the Max Planck Digital Library
- Service Manager for Edmond
- Product Owner in the Edmond Software Development Team

## Let's talk about

- Changing demands
- Switching an operating software
- Communication strategy
- Build trust
- Local collaboration





# I. Status quo ante

# What means “local” in the Max Planck Society?

- 86 research institutes
- also Italy, Luxembourg, Netherlands and US (Florida)
- Max Planck Digital Library (central service unit)
- Local libraries at the institutes
- Three main computing facilities
- Administrative Headquarter
- Around 24.000 employees
- Budget of 2 Billion € in 2020
- 29 Nobel laureates

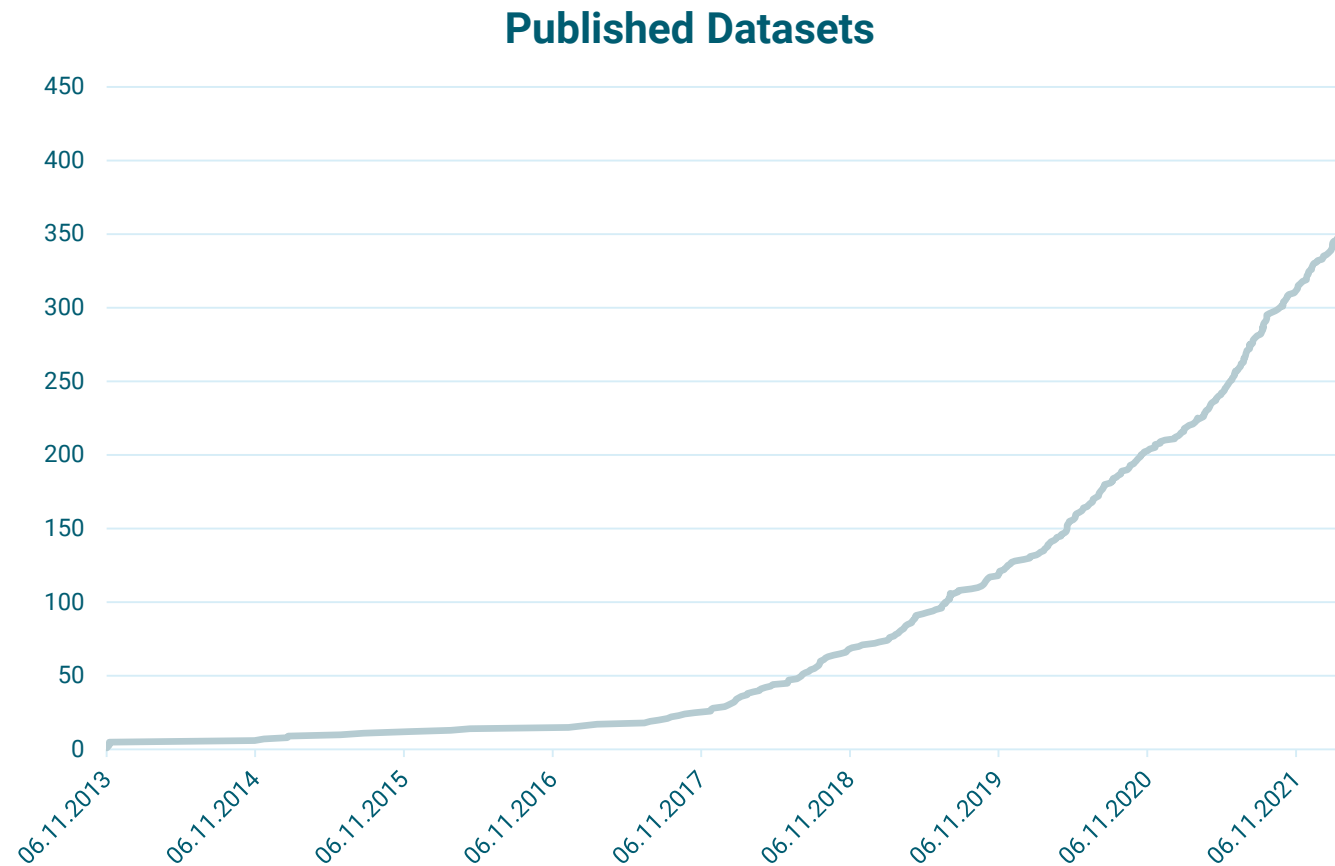
[https://www.mpg.de/zahlen\\_fakten](https://www.mpg.de/zahlen_fakten)  
<https://www.mpg.de/preise/nobelpreis>



[https://www.mpg.de/institute\\_karte](https://www.mpg.de/institute_karte)

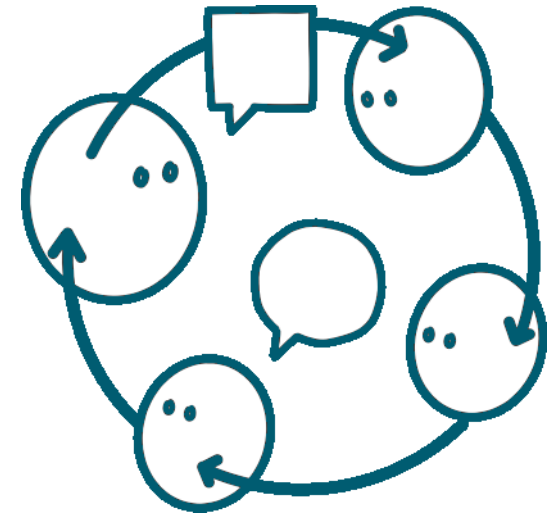
## Edmond's status quo ante February 2022

- Around 350 published datasets with 3.5 TB and ~200.000 files
- Using [imeji](#) as operating software
  - In-house developed software
  - More for research images less for research data
  - Problems with huge amount of files and file sizes

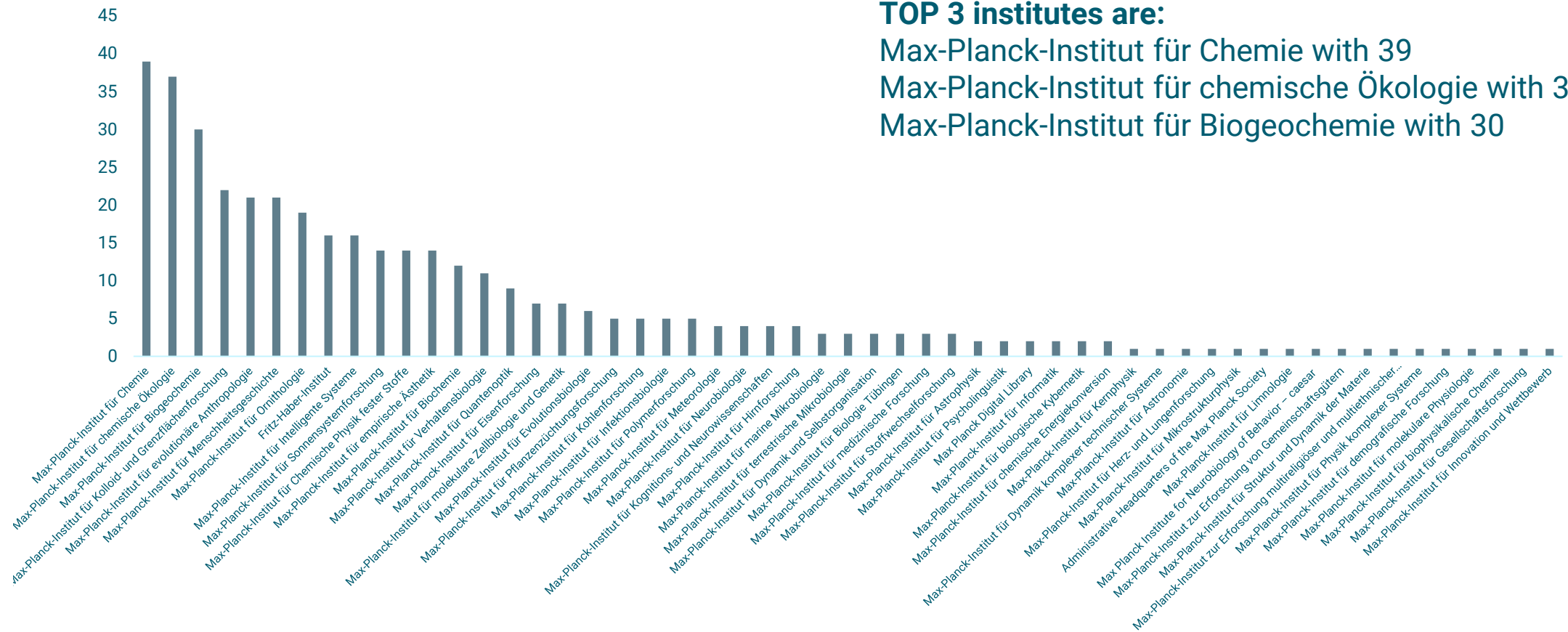


## Different Demands within the Max Planck Society

- Many different subjects: From astrophysics to legal history
- Wide variety of staff use the repository:
  - Different career levels (from PhD students to institute directors)
  - Researchers, colleagues from information and technology, ...
- Demands do not remain stable over time, they (obviously) change



# Published Datasets in Edmond by Max Planck Institutes



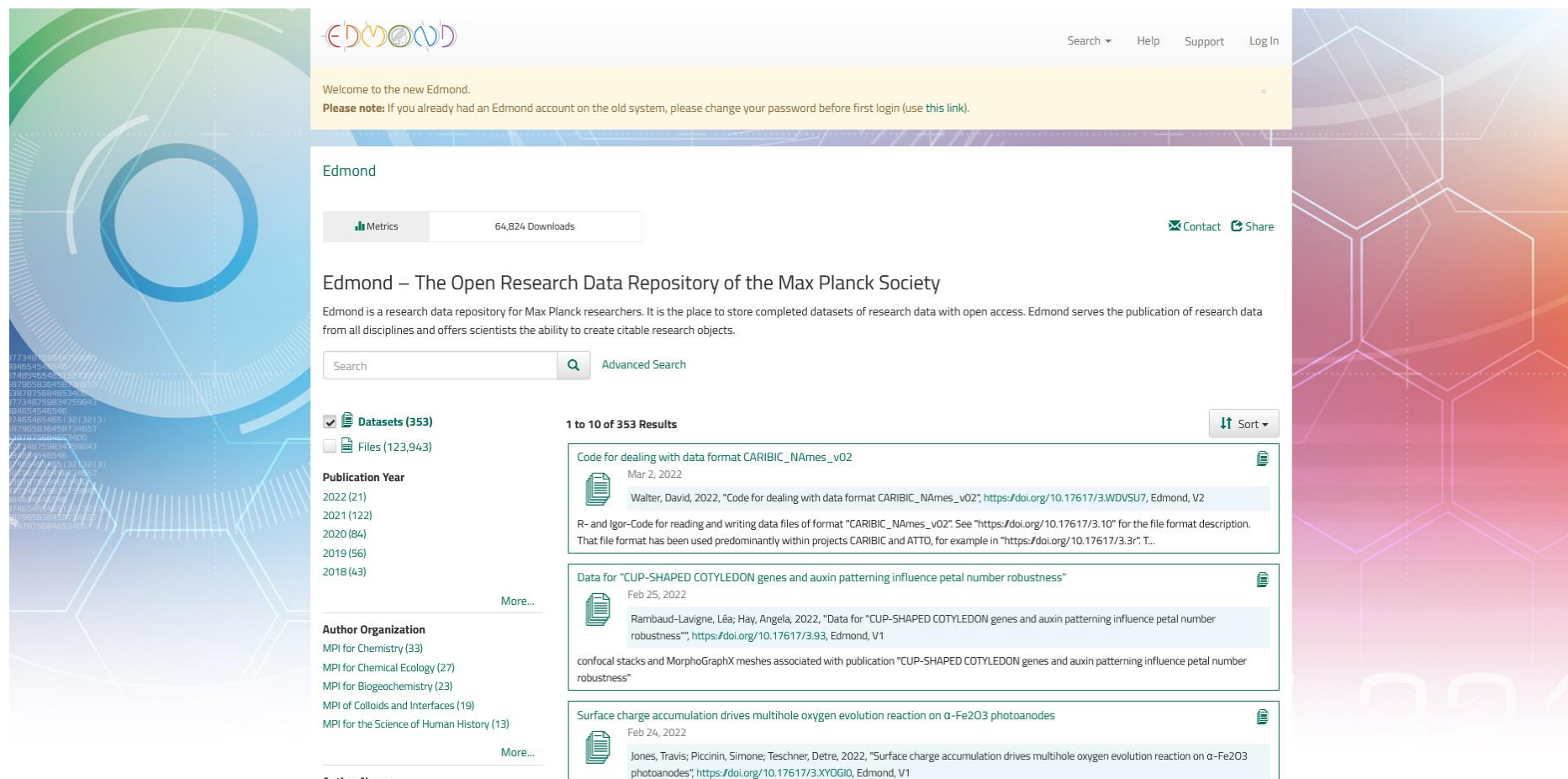


## II. Status quo

## Challenge: Switch to another Operating Software

- Decision to converse Edmond's Software in 2021 and after eight years
- Fulfill different types of demands:
  - Rich metadata schema for datasets
  - PID triangle: DOI + ORCID + ROR
  - Different user roles
  - Identity Access Management → usage of the MPG Single Sign-On
  - Versioning on dataset level
  - And even more...
- Discussion: In-house development vs. use of an existing software
  - Result: Dataverse addressed best all the different needs

# Software Migration: From imeji to Dataverse



The screenshot displays the Edmond website, which is the Open Research Data Repository of the Max Planck Society. The interface includes a search bar, navigation links (Search, Help, Support, Log In), and a welcome message. The main content area shows search results for "Code for dealing with data format CARIBIC\_NAmes\_v02".

**Edmond**

Welcome to the new Edmond.  
**Please note:** If you already had an Edmond account on the old system, please change your password before first login (use [this link](#)).

**Edmond – The Open Research Data Repository of the Max Planck Society**

Edmond is a research data repository for Max Planck researchers. It is the place to store completed datasets of research data with open access. Edmond serves the publication of research data from all disciplines and offers scientists the ability to create citable research objects.

**Search**  **Advanced Search**

**Metrics** 64,824 Downloads **Contact** **Share**

**Publication Year**

- 2022 (21)
- 2021 (122)
- 2020 (84)
- 2019 (56)
- 2018 (43)

**Author Organization**

- MPI for Chemistry (33)
- MPI for Chemical Ecology (27)
- MPI for Biogeochemistry (23)
- MPI of Colloids and Interfaces (19)
- MPI for the Science of Human History (13)

**1 to 10 of 353 Results**

**Code for dealing with data format CARIBIC\_NAmes\_v02**  
 Mar 2, 2022  
 Walter, David, 2022, "Code for dealing with data format CARIBIC\_NAmes\_v02"; <https://doi.org/10.17617/3.WDV5U7>, Edmond, V2  
 R- and Igor-Code for reading and writing data files of format "CARIBIC\_NAmes\_v02". See "<https://doi.org/10.17617/3.10>" for the file format description. That file format has been used predominantly within projects CARIBIC and ATTO, for example in "<https://doi.org/10.17617/3.3r>". T...

**Data for "CUP-SHAPED COTYLEDON genes and auxin patterning influence petal number robustness"**  
 Feb 25, 2022  
 Rambaud-Lavigne, Léa; Hay, Angela, 2022, "Data for "CUP-SHAPED COTYLEDON genes and auxin patterning influence petal number robustness"; <https://doi.org/10.17617/3.93>, Edmond, V1  
 confocal stacks and MorphoGraphX meshes associated with publication "CUP-SHAPED COTYLEDON genes and auxin patterning influence petal number robustness"

**Surface charge accumulation drives multihole oxygen evolution reaction on  $\alpha$ -Fe<sub>2</sub>O<sub>3</sub> photoanodes**  
 Feb 24, 2022  
 Jones, Travis; Piccinin, Simone; Teschner, Detre, 2022, "Surface charge accumulation drives multihole oxygen evolution reaction on  $\alpha$ -Fe<sub>2</sub>O<sub>3</sub> photoanodes"; <https://doi.org/10.17617/3.XYOGIO>, Edmond, V1

Current main page Edmond on Dataverse

# Migration to a New Operating Software

## Data Migration

- all research data has been migrated
- all local user accounts have been migrated
- all collection IDs and DOIs still refer to the published research data



## New Functionalities

- Single Sign-On for the Max Planck Society
- direct upload to object storage
- version control
- responsive user interface
- ...

# Interactions with Users

## Direct Communication

- interviews with different stakeholders
- internal presentations
- Communication by email and phone

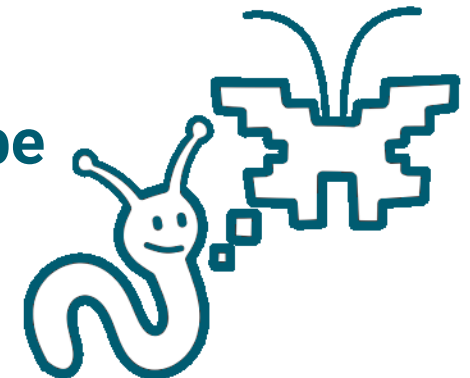
## User Testing

- selected ten users for testing
- collected feedback during tutorials

## Virtual Tutorials

- four virtual tutorials (EN/GER)
- on-demand trainings for institutes or working groups

**Need to understand + try to improve = Getting better, maybe**





### III. Next steps

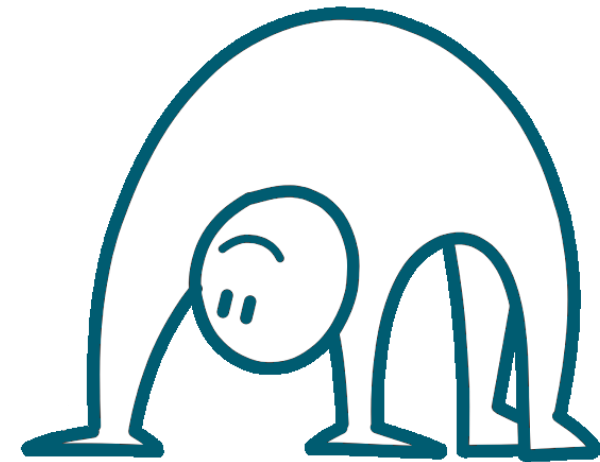
# Building Trust

## Keeping in touch

- Stay connected with all 86 MPG institutes
- Communicate about subject-specific requirements
  - difficult to reflect in an institutional repository

## Improve Edmond

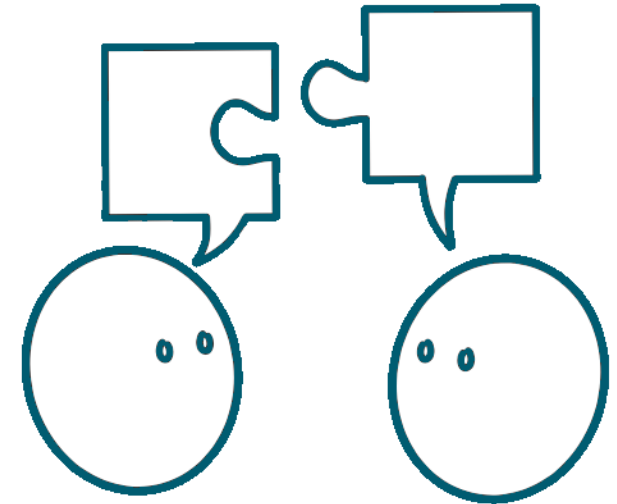
- Provide high-quality support
- Strengthen long-term preservation
- In application process for CoreTrustSeal
- Series of RDM workshops, connected to this topic



# Enable Local Collaboration

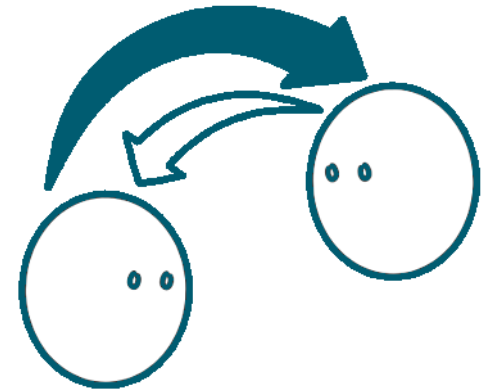
## Failures Possibilities to Learn

- No one is really waiting for your repository
  - see [re3data](#)
- Data sharing does not has a high priority (yet)
- Local collaboration means social interaction, which highly depends on humans
  - If the colleagues don't work together, the best repository is useless
- Talk to the “right” people
  - but who is that?



## Summary Slide

1. Different fields of research → many different demands towards an institutional repository
2. Switching the Operating Software → Challenging, especially the data migration
3. Communication, communication, communication → “looking” over the shoulder of your user is time-consuming but very helpful





**Thanks to the team!**



# Thanks for your attention!

[grossmann@mpdl.mpg.de](mailto:grossmann@mpdl.mpg.de)