

# OAI-PMH Workflow

Thorsten Wetzenstein

2023-12-17

ORCID <https://orcid.org/0000-0001-9589-3136>

Loading the necessary libraries

```
library(OAIHarvester)
```

1. Definition of Variables

```
baseurl <- "https://macau.uni-kiel.de/oai"
```

```
spec <- "deepgreen"
```

2. Use OAI-PMH Verb "List Records"

```
records <- OAIHarvester::oaih_list_records(baseurl, set = spec)
```

Harvest all metadata

```
records_data <- records[, "metadata"]
```

3. Transformation of OAI-PMH XML into table format for R. List with character vectors.

Empty Records with the status "deleted" should not be harvested so set length > 0L

```
records_data <- OAIHarvester::oaih_transform(records_data[length(records_data) > 0L])
```

4. Convert to Tibble format, a simple and modern form of a data frame.

```
rec_dat_tib <- as_tibble(records_data)
```

5. Write Data to table

```
WriteXLS::WriteXLS(x = rec_dat_tib, ExcelFileName = "rec_data.xlsx", verbose = "TRUE")
```

6. Load library tidyverse for data wrangling

```
library(tidyverse)
```

link to the complete publication <http://dx.doi.org/10.57892/100-36>



CC-BY 4.0