

## Riana Minocher: `/eco_minocher/EARS/3_digging`

---

The folder `/eco_minocher/EARS/2_scraped` contains a list of sub-folders, each representing one publication.

In order to get an idea of what was done during 'scraping', read the document **EARS\_WORKFLOW\_SCRAPING** in `/eco_minocher/EARS`.

Every folder in `/eco_minocher/EARS/2_scraped` should contain three files, a **BIB**, a **PDF**, and a **YAML** file.

The **goal of 'digging'** is to obtain all materials necessary to reproduce the results presented in a given publication, namely data and analysis code.

***Your job is to download materials, e.g. datasets, additional information and scientific code used for analysis, that may be published alongside a research article.***

Some material will be downloadable from the publisher's web page.

Sometimes, links will be provided in the publication, leading to an external web page (e.g., GitHub) from where material can be downloaded.

In some cases, it will be necessary to contact the author(s) of the publication and ask them to send the required materials.

---

## PREPARATION

Be aware that more publications will be scraped while you'll be digging. In order to avoid conflicts, order the subfolders inside `/eco_minocher/EARS/2_scraped` from oldest to newest. This way, we are hoping to avoid a situation where you'll try to start the digging process on a folder that is still being scraped!

Inside `/eco_minocher/EARS/3_digging`, create a subfolder with your own name. Start the folder name with a "0\_".

---

### 1) FROM SCRAPED TO DIGGING

Inside `/eco_minocher/EARS/scraped`, choose the 'oldest' folder and drag it into your personal folder in `/eco_minocher/EARS/3_digging`.

In the folder you've just moved, open the YAML. Enter your name into the field 'digger\_name'. *Consult the variable list for additional instructions!*

Go to [www.doi.org](http://www.doi.org). Copy the 'doi' from the YAML and paste it into the search bar on the right.

It should take you to the online version of the publication. Open the corresponding

PDF and make sure that it is indeed the correct publication.

---

## 2) SUPPLEMENTARY MATERIALS

Search on the publisher's website for the publication's supplementary materials

--> there could be a link to "Supplementary Materials" in a sidebar on the right or the left of the page

--> if not, scroll down to the bottom of the article, and search before/ after/ between acknowledgements and references

--> in some cases, you will have to click on another tab on the page such as "Figures and Data"

If there is supplementary materials available for download:

- go to `/eco_minocher/EARS/3_digging/0_templates`
- copy the folder "**supplement**" and paste it into the publication folder you are currently working on
- download all supplementary materials available online and store them in the "supplement" folder
- in the YAML, put "yes" under `supplement_available`

*Consult the variable list for additional instructions!*

---

## 3) MATERIALS STATEMENT

NOTE: *by definition, 'material' is referring to data and/or analysis code.*

The data used for the study and the code written to analyse this data may or may not be available online.

Search for a materials statement in the pdf **and** the online version of the article. (If you have found a solid materials statement in either one, you do NOT have to search the respective other.)

Do this by searching for the following keywords (one by one):

Code

Available

Provided

Data

Analysis

Material

Reproduce

Results

The material statement might be for example:

"The data and code to reproduce our results is available online at [link to website]"

Or

“Data and code for analysis is provided in supplementary materials”

Or

“Data is available on request”

Etc.

Copy paste the statement into "materials\_statement".

For "materials\_statement\_cat", put the statement into one of the following categories:

--> If there is no statement found put: **none**

--> If the statement suggests the materials are available online for download, please put: **available**

--> If the statement mentions contacting an author, please put: **request**

--> If the statement suggests there are 'any' materials available online for download, but there are indications that it is neither data nor analysis code (e.g. "program code"), please put: **it's complicated**

NOTE: put the statement into a category based on the content of the statement alone - regardless of whether you've downloaded supplementary material or not!

*Consult the variable list for additional instructions!*

---

#### 4) DATA AND ANALYSIS CODE, ADDITIONAL MATERIAL LINKS

If the material statement provides a link to download additional materials, follow it. Furthermore, look for links to download additional materials in the PDF and on the web page.

If there is ANY additional material available for download from an external link:

- go to **/eco\_minocher/EARS/3\_digging/0\_templates**

- copy and paste template folder "materials\_link" (a folder ONLY for external materials)

- download all materials available online and store them in the "materials\_link" folder (do not move materials already stored in the "supplement" folder!!)

- in the YAML, insert the hyperlink where the materials were downloaded from under "materials\_link"

Downloading materials from GitHub:

- click on "Clone or download"

- choose option "Download zip"

Once you have downloaded all materials available online, move the entire publication folder to:

**/eco\_minocher/EARS/4\_flagging**

---

#### 5) REQUESTING MATERIALS FROM AUTHORS

**TO BE ADDED...**

---