

## # README file for European dataset of FAIR Phytoliths Project

### ## Data collection methodology

List of European countries:

[https://en.wikipedia.org/wiki/List\\_of\\_European\\_countries\\_by\\_area](https://en.wikipedia.org/wiki/List_of_European_countries_by_area)

**Download Publish or Perish software** (<https://harzing.com/resources/publish-or-perish>) - you can run keyword queries on Google Scholar from there, and export the results to .csv.

1. Use Publish or Perish to run a search for keywords 'Phytoliths' and 'Europe'. ✓
2. Save this search as a csv file in the Europe/data-raw/search-data folder on google drive. ✓
3. Start working your way down the search list:
  - a. All articles must be saved in Europe/data-raw/articles folder.
  - b. Our criteria for our study is the article must have primary data, be from the region and published in the date range of 2016-2020 (first published on internet).
  - c. Once you have downloaded an article and started work on that article for the FAIR assessment - change the colour for text in the 'to edit' search doc - green for meets our criteria and red for not meeting it.
  - d. Do not edit the original csv file or the google sheet that says raw.
4. Each time you work on the FAIR assessment - please put notes below with the date, your name and what you have done in your working session. This will make it easier for us to work asynchronously on this task.
5. The R package 'roadoi' (uses unpaywall with R) has been used on our list of articles to find information about the open access status of the articles. Documentation for the data categories in this package can be found here:  
<https://unpaywall.org/data-format>.
6. This information was integrated into the data set after the rest of the data was collected.

### ### Moderation:

- After doing the assessment on the paper, highlight your comments in **yellow** (in this file) any papers in need of moderation. A moderation meeting will then be conducted to discuss this paper and any data categories that have been hard to assess.

### ### Data collection work session details:

#### 15 September 2021 - Emma Karoune

- Used publish or perish to do search of google scholar using key terms above - phytoliths and europe.
- Saved as csv on my local computer and also put in the google drive folder.
- Made a 'raw' and 'to edit' version of this as a google sheet.

## **## Data cleaning**

### **29 September 2021 - Juanjo García-Granero**

- Searched the journal articles (no book chapters) and marked in green (in the 'to edit' version) those matching the chronological (available online between 2016 and 2020) and geographic parameters. For Russian case studies, only those in geographically 'European Russia' (up to the Ural Mountains) were considered (e.g., not the Altai region). Articles in Russian were not considered.
- Uploaded the selected articles to the 'Articles' folder in Gdrive. Got to 39 after checking the first ~300 articles in the list, will finish soon.

### **30 September 2021 - Juanjo García-Granero**

- Finished checking the list and downloaded 11 more articles. I had to check the first 500 entries in the list from publish or perish, if we need to increase the number of articles in the future we'll probably need to consider earlier articles (2011-2015).
- One paper I published in 2021 (but which appeared online in 2020) was not in the list (García-Granero et al. 2020 JAMT). This might be added if one of the selected articles turns out not to be valid.

### **7th October 2021 - Emma Karoune**

- Renamed all articles with first author-etal(if required)-date of publication
- Also checked that all were in our date range (2016-2020 - first available online) and yes they were!

### **8th October 2021 - Emma, Juanjo, Javier and Celine**

- We went through the first article together and filled in the data collection form.
- Issues we need to deal with are:
  - Adding none to what data repository category
  - Making stats category multiple choice by changing to check list
  - Problem with Open access category as some of use get green for unpaywall even though the access is through this is institution access and not showing real green access on a repository - we are going to now check this using an R package that uses the DOI so we need to put the DOI for each paper that we assess into the article search spreadsheet list (spreadsheet taken from publish or perish now on google drive in column L). R package is <https://github.com/ropensci/roadoi>
  - Need to clarify research gate/academia category as hard to do

### **11th October 2021 - Emma and Javier**

- Javier has added closed to Journal-type - as some journals from s america do not give an option to be open access.
- Leave all categories for open access articles to the end so that this can be done with the R package above.

### **15th October 2021 - Emma**

- FAIR assessment for Andriopoulou and Christidis (2020) - need help with this one as a tricky paper to assess: My comments are: This is also about chemical/mineralogy

composition of the phytoliths so could be under taphonomy too? Parr et al. 2001 for dry ashing method also uses Rovner 1971, 1972 - as is looking at multiple extractions methods for plant materials - the methods used would be clearer if they had attached full protocols as supp files or in a repository - how many plants? 1 for each species? Counting method is not clear as it is mostly about the mineral composition but there is mention finding the quantity of phytoliths so counting must have been done?

- and Aseyeva et al 2019 - much more straight forward so all done.
- Added not used to qu 16 for articles that do not use icpn. - changed on both s america and europe forms as well so it is included in the drop down list.

#### 18th October - Juanjo

- Banerjea et al. 2020
- Bejenaru et al. 2016 - does not have a doi! - add any online identifier

#### 20th October - Juanjo

- Benes et al. 2016
- Bohme et al. 2017

#### 29th October - Juanjo

- Borderie et al. 2020
- Burguet-Coca et al. 2020

#### 31st October - Emma

- Cunha et al. 2019
- Dal Corso et al 2018
- Dal Corso et al 2017
- Danu et al 2019

Must check that if it says adapted from --- ref - this must be put down as summary not specified one ref.

#### 2nd November 2021 - Emma

- Danu et al 2020
- Daura et al. 2019
- Delhon et al 2020
- Delplace et al 2020 - worth moderating as it's another chemical analysis weird one! - look to see if about counting method or not.

#### 9th November - Juanjo

- Dzhafiezova 2020 - Analysis on thin sections, hard to assess 'Processing-method-communication' → go back and check if thin section preparation method is reported
- Égüez et al. 2020
- Field et al. 2018
- García-Granero et al. 2018
- García-Granero et al. 2020
- García-Rivero et al. 2018

### 16th November - Juanjo

- Golyeva and Svirida 2017 - The study uses both modern and palaeoecological samples but the aim is more methodological than palaeoecological, so I have labelled it as 'Modern'
- Goude et al. 2020
- Juhola et al. 2019
- Kaiser et al. 2020 - 'Counting method' not applicable because this is a chemical analysis of phytoliths within plants, not a phytolith analysis.
- Koromila et al. 2018
- Kufel-Diakowska et al. 2019

### 23rd November - Celine

- Wroth et al. 2019
- Wouters et al. 2017
- Vigliature et al. 2019 This study documents phytolith contents with X-Ray spectroscopy and assesses bioremediation possibilities of Miscanthus sinensis for treating contaminated soils/ It is difficult to assess for the FAIR project.

### 24th November - Celine

- Tsartsidou & Kotsakis 2020
- Sumegi et al. 2019

### 25th November - Celine

- Stanc et al. 2020
- Shishlina 2018
- Rusakov et al. 2018
- Puppe et al. 2017 This article aims at studying the different BSi pools influence short-term changes in the water-soluble Si fraction in soils, so I think it is a geochemical type of study. But it can be also be plant physiology (as the phytolith abundance and volume are parts of the study). The counting method is difficult to assess because it is not really about phytolith identification but more about plant and soil interaction.
- Portillo et al. 2017 The article aims at demonstrating the importance of taphonomy in phytolith and pollen studies. The researchers are using archaeological and experimental data to make their point. Maybe this study type is more methodological-taphonomy than archaeological.
- Ponomarenko et al. 2019

### 26th November 2021 - Javier

- Highlighted in green the assessed papers (and papers to be assessed) in file 2021-09-15\_phytoliths-south-america-google-scholar-search\_to-edit and added their DOIs when missing. Obtained open access/repositories information of the papers using the roadio R package (function oadoi\_fetch), which collects the information from Unpaywall. Added the file Unpaywall\_Europe containing the results obtained (both the open access status of the assessed papers and the available open repositories) in FAIR-assessment/Europe/data-raw.

### **2nd December 2021 - Celine**

- Peto et al. 2017
- Parkinson et al. 2018
- Ovchinnikov et al. 2020
- Norstrom et al. 2019

### **9th December 2021 - Celine**

- Meunier et al. 2017
- Merceron et al. 2016, the phytolith counting method is connected to the Dental Microwear Texture Analysis (DMTA), I am not sure if I caught it properly.
- Lisztes-Szabo et al. 2020

### **15th December 2021 - Celine**

- Lisztes-Szabo et al. 2019
- Lippi et al. 2017
- Kurbanova et al. 2018

### **9th January 2022 - Emma**

- Going through papers that needed moderation and checking in line with what we decided in the meeting on 17th Dec 2021.
- Andriopoulou and Christidis (2020) - changes made.
- Delplace et al 2020 - Not applicable added for counting method.

### **21st January 2022 - Juanjo**

- Revisited Dzhanfezova 2020 to assess 'Processing-method-communication' → Not the clearest paper ever, I have put it down as 'Summary of method' in the Excel file
- Kaiser et al. 2020 → I have added 'Not applicable' in the counting method
- Bejenaru et al. 2016 → added the URL to the '2021-09-12\_phytolith-europe-google-scholar-search\_to-edit' file

### **21st January 2022 - Emma**

- Changing open-access-article column to oa\_status from unpaywall R data spreadsheet.
- Kufel-Diakowska, B., Miazga, B., Łucejko, J., Ribechini, E., & Li, W. (2019). - checking into the status of the journal, also looks like it is free open access so diamond rather than gold/green - row 30.
- Kurbanova, F., Makeev, A., Puzanova, T., Kust, P., & Chernov, T. (2018). - this is not a journal article - we need to do one more article to replace this one. - row 51.

### **23rd January 2022 - Emma**

- Add Herrescher et al 2018 to replace other article but it is based on sites in Georgia - are we counting Georgia in Europe? JG says yes ok.
- Added repositories column info.
- Just need to check country for new article - this now ok
- and then row 30 oa-status - this now ok - Open and diamond

#### **4th Feb 2022 - Emma**

- Final checks now done.
- Copy made for archive =  
2022-02-04\_FAIR-Phytoliths-Data-Assessment-Final-Europe-raw
- Do not alter this file now
- The responses spreadsheet should be used for data cleaning and analysis.