

## **Proposal for Digital life sciences open call EOSC - Open Science in Phytolith Research**

**Contributors:** Emma Karoune, Juan José García-Granero, Carla Lancelotti and Marco Madella, Javier Ruiz-Pérez.

### **Under their call I think we fit into this section:**

Data curation, handling and integration for publishing into the cloud.

- Collection, processing, standardisation, FAIRification, and storage of existing data.

### **Experts will be looking for projects which demonstrate:**

- Scientific excellence and impact
- Potential to make available new valuable resources (data/services/methods) within the EOSC and for the scientific communities represented by the Life Science Research Infrastructures.
- Feasibility of the project within the EOSC-Life runtime (the project must be completed and reported before February 2023 and start within 2 months after project acceptance).
- Availability of resources within the project team: people power, computational & financial resources.
- Availability of resources within the EOSC-Life technical expert teams: if EOSC-Life expert support is required in your project, please contact the relevant experts before submitting your project application to ensure that the requested assistance is available.
- Sustainability after the end of EOSC-Life. Projects should have a clear plan for their sustainability beyond the duration of the funding period (e.g.: data/software/project output management plan).

**These are all the categories in the online application and the details that they have given for each one.**

### **Research Project Title:**

Increasing the FAIRness of phytolith data

### **Institutions to be involved in project and PICs:**

Historic England - 987006041

Spanish National Research Council - 999991722

Universitat Pompeu Fabra - 999867077

**Scientific background of the project (4000 characters) - give details about the scientific question you are addressing:**

Phytoliths are silica-bodies that form within plant cells during the lifetime of the plant (Madella & Lancelotti 2012). Monosilicic acid ( $\text{H}_4\text{SiO}_4$ ) in groundwater is absorbed by the plant roots and is eventually deposited as solid silica dioxide ( $\text{SiO}_2$ ) in and between plant cells forming distinctive shapes (morphotypes), which can be used to identify plant taxa to different taxonomic levels. When the plant dies, phytoliths become deposited naturally in the soil or, if the plant has been exploited by humans, in archaeological layers. These inorganic plant remains are preserved over long time periods and can therefore be readily found in archaeological and palaeoecological samples (see Hodson *et al.* 2020 for examples of recent phytolith research). Consequently, they are used to address research questions in combination with, or instead of, other plant remains that do not survive as well.

Phytoliths are being used for an ever-expanding list of research topics and applications. In recent years, there has been a substantial increase in research, and this has resulted in an upsurge in publications. Archaeological phytoliths are examined from deposits using a variety of methods such as the sub-sampling of bulk samples taken for flotation, column samples from exposed sections after excavation and increasingly a combination of soil micromorphology and then sub-sampling of the same column for sediment processing (Madella & Lancelotti 2012). They are also extracted from ecofacts and artifacts such as dental calculus, coprolites, pottery and grinding stones. Applications in palaeoecology include reconstructing palaeoenvironments and palaeoclimates from core and soil samples, and the newer applications of carbon-14 dating and isotopic analysis of phytoliths are also being increasingly explored.

With increased research comes the difficulty of how to standardise methods, share data and validate analyses. The stage of innovation currently seen in phytolith research is producing a wide range of data, both quantitative and qualitative, from archaeological, palaeoecological and methodological studies. Phytolith data falls into several categories - observational, experimental and computational data. These data must be made FAIR so that other researchers can review, adapt, apply and collate their colleagues' research enabling phytolith research to become more sustainable.

Our ultimate goal is the FAIRification of all phytolith data produced by all researchers whether they are botanists, archaeologists or palaeoecologists. At this stage, however, we need to build the foundations to move towards this goal. This project is the first important step towards more standardised, FAIRer and available data, which will be achieved by a) assessing the FAIRness of existing datasets, b) making clear recommendations to change current data sharing practices for existing and future

datasets and c) introducing these recommendations to the phytolith community by modelling data sharing best practice.

**Technical background (4000 characters)- give details about technical challenges that will be addressed in your project and what work has already been completed:**

The International Phytolith Society (IPS) has started working towards the standardisation of practices in phytolith research for a few years now. This includes important work concerning the standardisation of nomenclature (ICPN 1.0, Madella *et al.* 2005; ICPN 2.0, ICPT 2019) and morphometric methods (Ball *et al.* 2016).

There has also been encouragement to adopt a common set of best practices to aid standardisation (Zurro *et al.* 2016) throughout the phytolith research lifecycle as it is clear that current published work presents a wide range of approaches to data analysis, presentation and interpretation. Further standardisation will enable broader phytolith research beyond technique-orientated work, permitting greater opportunity for its application to inform on past cultures and their strategies of plant resource exploitation as well as the dynamics related to climate change and anthropic-driven environmental modifications.

However, a recent review of open science practices in phytolith research (Karoune 2020 and [pre-print Karoune 2020](https://osf.io/fa7q3/) <https://osf.io/fa7q3/>) found the adoption of these approaches is lacking. Data sharing was found to be needing considerable improvement (53% of articles shared some form of data) and the majority of data shared was not reusable (96%). This was due to these articles not providing raw data in a supplementary file or in an open repository.

Much of the data was found to not be accessible because it exists behind journal paywalls (only 14% of articles were open access). Very little phytolith data is currently deposited in open repositories; only one article out of 341 in the study shared raw data in an open online repository.

It was also found that adequate metadata was not included in these articles. Methods were not clearly written as full protocols and adoption of the standard nomenclature was only 51%. Other identification indicators such as photos were commonly included but often did not encompass the whole breadth of phytolith morphotypes identified, which results in lack of confidence in the identifications made in the study.

In terms of FAIRification, the phytolith community is taking its first steps at the beginning of the journey. The assessment of open science practices conducted by the PI (Karoune 2020 and [pre-print Karoune 2020](https://osf.io/fa7q3/)) and the subsequent project that has been developed during the Open Life Science mentoring and cohort-based training programme have led to increased awareness and interest in open science in phytolith research and related disciplines. Talks, blogs and the formation of a working group has resulted in the initiation of this project. The work conducted so far, in a sense, has completed the “preparing” and “training” phases, as proposed in

the FAIRification processes set out in David *et al.* (2020). However, more training in open science skills and, in particular, training related to implementing the FAIR principles is still needed within the phytolith community.

This project will go further into the pre-FAIRifying phase in which a collaborative strategy will be built from the results of the detailed assessment of phytolith data. There will, however, be challenges in creating community consensus on the recommendations developed in this project. Therefore, the presentation of empirical data concerning the potential reuse of existing phytolith data and providing solutions to issues such as the standardisation of methods and nomenclature will be crucial for community adoption.

Skills training in open science tools must be addressed to make phytolith researchers confident to access online repositories and work on collaborative projects. This work has already begun through collaboration with *The Turing Way* to create chapters related to collaborative working and particularly accessible content for onboarding novice Github users. Developing training materials for the FAIRification of phytolith data and commonly used tools will also help to overcome this challenge.

**Planned work (4000 characters)- include the key deliverables of the planned work with their approximate timings (months) after project initiation:**

Initial steps towards the FAIRification of phytolith data will be conducted through three work packages (WPs):

#### **WP1. FAIR assessment of existing data (months 1-6)**

Published phytolith datasets from South America and Europe will be assessed to establish the breadth of phytolith data (including archaeological, palaeoecological and modern botanical studies). This will initially be from articles published between 2016 to 2020, but earlier data will be collected if needed to reach data saturation (Saunders *et al.* 2017). These two regions have been selected because different phytolith ‘traditions’ exist among researchers working in Europe and South America. Most European phytolith studies use the standardised nomenclature (ICPN 1.0), whereas South American researchers have developed their own naming criteria. In spite of this, the overall principle of the methods used for processing phytolith samples is the same (the separation of phytoliths from other particles in sediments or modern plants), as well as the research questions addressed and the types of data analysis that is conducted on primary phytolith data, thus ensuring the comparability of the datasets.

The FAIR assessment will be conducted by establishing a subject-specific FAIR assessment tool, similar to the [FAIR enough?](#) checklist, using a google form to standardise data collection.

**Key deliverables:**

- A Github repository of data from published articles and FAIR assessment datasets as csv files.
- A research compendium in the Github repository and archived in Zenodo.
- Data paper to improve findability and therefore facilitate the reproducibility of the study. APC costs for all articles published within this project will be covered by one of the participating institutions.

**WP2. Community building and training for FAIRification (throughout the project)**

At the outset of the project, a survey will be sent out to our research community to gain insight into how researchers see their own data sharing practice. This will gather information about current data sharing practices and opinions on why more open practices are not used.

This survey will give the option to the authors of the existing datasets to make their data open access. This also gives us the opportunity to work with them to improve these datasets in terms of their FAIRness. This could take the same approach as the '*Bring your own data workshop*' at the University of Cambridge that aimed to improve data management through training. These improved datasets will then be used as case studies to raise awareness of the need for planning data management and provide resources for more training in our community. Training our community in open science skills, such as Github, will help to aid accessibility to open science tools in general but also help to advance the implementation of future FAIR data criteria.

**Key deliverables:**

- Survey data will be stored in the Github repository/Zenodo.
- Conference paper at the International Meeting for Phytolith Research (September 2021).
- User support resources will be created to demonstrate the process of changing existing datasets into FAIRer datasets in terms of minimum information and best practice. This will include a FAIR assessment tool and training materials - all archived on Zenodo.

**WP3. Development of guidelines for existing datasets and future FAIR phytolith data (months 7-12)**

The FAIR assessment of existing datasets will be used to propose criteria for minimum information requirements for data sharing and potential areas of reuse that consider the limitations of using these data. These constraints will be used to suggest a best practice criteria for the implementation of future FAIR phytolith data.

**Key deliverable:**

- A peer-reviewed open-access research article to share findings of the FAIR assessment and the new data sharing criteria.

**Which of the following tools/workflow management systems/registries/infrastructure do you plan to use for your project?**

This initial project will establish a repository using Github linked to Zenodo for long-term archiving of the datasets gathered from the two geographic regions, the datasets created from the FAIR assessment of the phytolith data and the community survey data. All research outputs will be deposited in this repository and archived as part of the research compendium.

Github will also be used as a project workflow management system making use of the kanban project boards for planning, issues and pull requests to work on collaborative documentation and the creation of a webpage to advertise and promote the project to the phytolith community.

RStudio will be used for data analysis and visualisation. All code and descriptions of data analysis will be stored in the research compendium.

Although it is beyond the aim (and feasibility) of this project, our future aim is to establish a sustainable open-access phytolith repository to share software, data analysis tools, datasets and user support resources. This could potentially be created through an existing data repository such as Pangaea.

**Provide Key features of the data resource/services that you will be using (4000 characters) - give details of data types, dimensions, data models, ontologies, current storage and hardware containerisation, curation, software and licenses, sensitive data or relevant ethical aspects:****Data types:**

The data we are capturing comes from a wide range of analyses: archaeological, palaeoecological, botanical and methodological data. This can be qualitative and quantitative.

Survey data will be anonymised, so no personal data will be collected conforming to GDPR regulations.

The FAIR assessment tool will collect data concerning:

- What data is being published - raw data, absolute counts, percentages, morphotype groupings.

- What format the data is provided in and its location - tables in text, pdf, excel, csv, supplementary files or open repository.
- What system is being used for standardisation of nomenclature - ICPN 1.0, ICPN 2.0 or a different region- or lab-specific nomenclature system.
- What other metadata is provided and is being used for the interpretation of the data - this could be the sampling and processing methods, counting method, taphonomy indicators such as diagenesis and pH, information about dating the deposits, geographical information, botanical information about the plants being studied.
- What statistical analysis this data is being used for - this determines what data is needed in the first place and how others might be able to reuse the data.

### **Ontologies/Standardisation:**

The extent to which standardisation is being currently used for phytolith datasets and the extent to which standardisation can be achieved in the future is a key part of this project.

Nomenclatures currently vary for phytolith morphotypes - ICPN 1.0 is currently used by some researchers although many have developed their own region-specific naming criteria. There is a new version of ICPN (2.0) but there has been criticism of this new system as it has made substantial changes to the previous version. In future steps towards FAIRification, there may be a need to be inventive to overcome this particular problem such as creating a system of conversion between the commonly used nomenclatures to allow researchers to reuse datasets in spite of the lack of consistent nomenclature.

Processing methods are another area in which standardisation needs to occur. There are currently many different methods in use and we do not know the extent to which this influences the phytolith data produced. Therefore, recording which method is used and also encouraging researchers to publish full protocols may be the extent of standardisation possible. To address this issue, members of this project (CL & MM) are also involved in a parallel project with the NSF-US to assess the interoperability of data arising from different processing methods.

Other areas in need of standardisation are plant species names, geographic names and dates from archaeological and palaeoecological deposits. These issues will be addressed in the survey sent out to the community and the extent of variation will be assessed in the existing datasets.

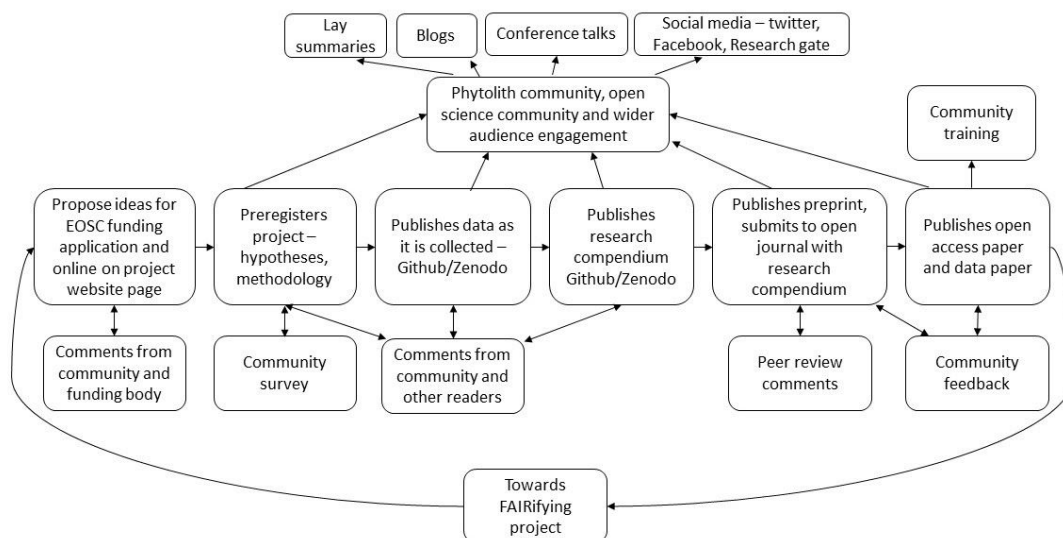
**Software:** It is the aim of the project to use free-to-use software and tools to enable complete accessibility of the community to the outputs produced.

**Licenses:** All new data will be published using a CC0 license. Existing datasets will also be published using CC0, unless permission is not gained from the original authors and journal publishers. All documentation will be given a CC BY 4.0 licence.

Phytolith data is not considered to be sensitive data and the ethical use of existing data only needs to be considered in terms of the original author's own copyright and license restrictions. However, every effort will be made to contact the authors of the datasets used in this study to be able to discuss licenses for the data and hopefully gain permission for CC0 licenses.

**Key features of your workflow (if relevant to your project)(4000 characters) - description of the steps to be executed in your workflow, ontologies, containerisation, workflow engine, software and license, registry, repository. Is your workflow fully established? Is it cloud ready? Describe the limitations and improvements to be done (data input/output, runs only on my computer, user authentication/authorisation, compute resources, amount of human intervention. licensing):**

See workflow diagram attached to the application.



**Compute and data resources required (4000 characters) - What compute and data resources are required for your project? Be explicit with the requirements, e.g. CPUhrs, RAM, GBhr storage, Network, ancillary services (data transfer, workflow execution) etc. Describe the sensitivity of the data that will be processed by your project. For each required resource please state whether you would request access through EOSC-Life or whether you have access in-house or from another source:**



Computing hardware and locally used software such as Microsoft office will be supplied by the Host Institution (Historic England). Data at Historic England is stored in their network and backed up daily. There will also be access to journal publications through Historic England and the other participant institutions.

Any other software and tools used in the project will be open and free to use to increase the accessibility of the research outputs created in the project. These will include the use of Google drive for collaborative data collection and documentation writing, RStudio for data analysis and visualisation, and Github for an online data and metadata repository, project management and for creating a project webpage. All outputs will be archived on Zenodo.

**Technical expertise in your team (4000 characters) - Please list which members of the project team will be responsible for which aspect of the proposed work and detail their expertise.:**

Project members will include Emma Karoune (PI, Historic England), Javier Ruiz-Pérez (Universitat Pompeu Fabra), Juan José García-Granero (Spanish National Research Council), Carla Lancelotti (Universitat Pompeu Fabra) and Marco Madella (Universitat Pompeu Fabra).

It is requested that this grant will cover the salary costs of the PI (Emma Karoune) and a post-doctoral research assistant (Javier Ruiz-Pérez), both on 0.6 FTE. This is well within the maximum cost of the grant. The rest of the team will be carrying out this research within their current roles at no cost to the project. Employing two researchers will enable a larger team of researchers for this collaborative project; both bringing with them unique skills to aid the completion of the project.

**Emma Karoune** - PhD in Archaeology (macro-botanical and phytolith research), trained in the use of Github for collaborative projects and open science leadership through Open Life Science 2. She is also a trained teacher with extensive experience of planning and running training courses both in person and online. Emma will lead the project repository set up and the FAIR assessment of the two phytolith datasets, and will personally conduct the assessment of phytolith data from Central and northern Europe (WP1). She will work collaboratively with other members of the working group to develop community engagement (WP2) and roll out the data sharing guidelines for existing and future phytolith data (WP3).

**Javier Ruiz-Pérez** - MSc in Terrestrial Ecology and PhD in Archaeology (submitted in December 2020), focused on phytolith analyses in palaeoecological and archaeological contexts in South America. He will conduct the assessment of phytolith data from South America (WP1) and will work collaboratively with other members of the working group to develop community engagement (WP2) and roll out the data sharing guidelines for existing and future phytolith data (WP3).

**Juan José García-Granero** - PhD in Archaeology (macro- and micro-botanical remains, including phytoliths), with experience in the Mediterranean region. He will conduct the assessment of phytolith data from southern Europe (WP1) and will work collaboratively with other members of the working group to develop community engagement (WP2) and roll out the data sharing guidelines for existing and future phytolith data (WP3).

**Carla Lancelotti** - PhD in Archaeobotany (anthracology and phytoliths), with experience both in the Mediterranean and South America. She will collaborate with the other members on the assessment of phytolith data (WP1) and she will help design the survey and analyse data (WP2). Carla is experienced in the use of Github for sharing data and conducting collaborative research and an active user of Rstats in her daily research. She is also collaborating with her Institution to set up the new repository system for sharing primary research data (Dataverse.cat).

**Marco Madella** - Degree in Natural Sciences and PhD in Environmental Archaeology, with a world-wide experience in phytolith studies. He will collaborate with members of the project specifically on assessing phytolith data (WP1) and the development of common guidelines for sharing data arising from phytolith studies (WP3). Marco has extensive experience in databases and especially concerning standardisation of data related to phytoliths. He has been involved in the development of the first standardised nomenclature (ICPN 1.0).

### **Will you or your team benefit from training?**

Although it is the aim of this project to make the outputs as accessible as possible and therefore lower the skill level required to access the data by producing a simple research compendium to be archived in Zenodo. The team would benefit from training to enable the packaging of the computational environment for reproducible phytolith research such as using Binder or containerisation for future progression of this project.

**FAIR data (2000 characters) - Please describe how your project work will improve upon the current situation with regards to FAIR data standards and how you will ensure that your project adheres to FAIR data principles.:**

The majority of phytolith data are currently not inline with FAIR principles. Greater sustainability in terms of Findability, Accessibility and Reusability is possible, although it is likely complete interoperability will be harder due to issues of standardisation of methods and nomenclature. This will limit interoperability, although the creation of a cooperative framework for FAIR data, in terms of minimum requirements and best practice, will be a considerable improvement to the current situation.

At this point, there are likely to be a number of steps towards FAIRification of phytolith data and some aspects of this process will be easier than others. These potential steps forward are summarised in Table 1 ([attached to the application](#)); however, more detail concerning this process will only become clear after the assessment of existing data (WP1) and the analysis of the survey results (WP2).

This particular project's outputs will be made FAIR by creating a research compendia to structure and store all of the data and metadata produced. This will be archived in Zenodo to create persistent identifiers for all research outputs. All software and tools used in the project will be open and free-to-use, therefore improving the accessibility of the workflow. Findability and accessibility will be increased by the publication of a data paper and an open-access research article with a data availability statement. The training resources will conform to the 'ten simple rules for making training materials FAIR' (Garcia *et al.* 2020) and will also be archived in Zenodo.

**Expected impact (2000 characters) - Please include both expected scientific impact and expected impact providing resources for the EOSC and the scientific community.:**

The impact on the phytolith community will be substantial as we are starting from a low point. Creating a new cooperative culture of open science will allow the widespread movement towards the FAIR principles and set a new standard for future research. Changing data sharing practice to more findable and accessible habits, such as using open repositories, will enable reproducible studies that will aid the development of more robust methodologies and allow for regional syntheses. As a consequence, standardisation issues may be resolved meaning the greater likelihood of FAIR phytolith data in the future.

Guidelines will be created for authors, peer reviewers and journal editors. This will have greater impact to change data sharing practice through the publication route - researchers are more likely to supply data in line with the FAIR principles if enforced.

For EOSC and the wider scientific community, this project will be a challenging case study of FAIRification. Most FAIR assessments already have data deposited in open repositories but we are starting with the assumption that this is not common. Therefore, the FAIR assessment tool created could be a template for other disciplines at a similar stage of FAIRification.

We will be recording the time, and therefore cost, to convert poorly formatted and located datasets into csv files. Reporting this will impact wider use of good data sharing practices and potentially better enforcement of these practices by funding bodies due to economic considerations.

This project will diversify the adoption of FAIR data by influencing closely related fields such as Archaeology and Palaeoecology to embark along the same path.

The development of resources to improve the accessibility of novice open scientists, such as upskilling in Github, collaborative working and FAIRification of data, will make a large contribution to EOSC. These resources can be used to help train researchers from other fields and make FAIR data more mainstream.

**Scientific domain:**

Bioinformatics

**Budget: - there is a spreadsheet for this. I will find out the salary costs from Historic England.**

I have put this info into the spreadsheet:

PI (3 days per week for 1 year) - establishing workflow in Github and webpage, work on European datasets, contacting researchers of existing European datasets, developing criteria for data minimum information and best practice FAIR phytolith data, leading on training material for communicating findings to phytolith researchers, leading on research paper of FAIR phytolith criteria, collaborating on data paper for project datasets.

Post-doctoral Research Assistant (3 days per week for 1 year) - work on South American datasets, contacting researchers of existing South American datasets, developing criteria for data minimum information and best practice FAIR phytolith data, contributing to developing training material for communicating findings to phytolith researchers, contributing to research paper of FAIR phytolith criteria, leading on data paper for project datasets.

**Relevant publications (max 10):**

Ball, T., Davis, A., Evett, RR., Ladwig, JL., Tromp, M., Out, WA., and Portillo, M. 2016. Morphometric analysis of phytoliths: recommendations towards standardization from the International Committee for Phytolith Morphometrics. *Journal of Archaeological Science*, 68:106-111. DOI: [dx.doi.org/10.1016/j.jas.2015.03.023](https://doi.org/10.1016/j.jas.2015.03.023)

David, R., Mabile, L., Specht, A., Stryeck, S., Thomsen, M., Yahia, M., Jonquet, C., Dollé, L., Jacob, D., Bailo, D., Bravo, E., Gachet, S., Gunderman, H., Hollebecq, J.-E., Ioannidis, V., Le Bras, Y., Lerigoleur, E., Cambon-Thomsen, A. and Alliance – SHARing Reward and Credit (SHARC) Interest Group, T.R.D., 2020. FAIRness Literacy: The Achilles' Heel of Applying FAIR Principles. *Data Science Journal*, 19(1), p.32. DOI: <http://doi.org/10.5334/dsj-2020-032>

Garcia, L., Batut, B., Burke, M.L., Kuzak, M., Psomopoulos, F., Arcila, R., et al. 2020. Ten simple rules for making training materials FAIR. *PLoS Computational Biology*, 16 (5): e1007854. DOI: <https://doi.org/10.1371/journal.pcbi.1007854>

Hodson, M.J., Song, Z., Ball, T.B., Elbaum, R. and Stuyt, E. 2020. Editorial: Frontiers in Phytolith Research. *Frontiers in Plant Science*, 11(454): 5-7. DOI: 10.3389/fpls.2020.00454

International Committee for Phytolith Taxonomy (ICPT), International Code for Phytolith Nomenclature (ICPN) 2.0, *Annals of Botany*, Volume 124, Issue 2, 24 July 2019, Pages 189–199, <https://doi-org.eres.qnl.qa/10.1093/aob/mcz064>

Karoune, E. 2020. Data from “Assessing Open Science Practices in Phytolith Research”. *Journal of Open Archaeology Data*, 8(1), p.6. DOI: <http://doi.org/10.5334/joad.67>

Madella, M. and Lancelotti, C. 2012. Taphonomy and phytoliths: a user manual. *Quaternary International*, 275: 76-83. DOI: 10.1016/j.quaint.2011.09.008

Madella, M., Alexandre, A., and Ball, T. 2005. International code for phytolith nomenclature 1.0. *Annals of Botany*, 96: 253-260. DOI: <https://dx.doi.org/10.1093/aob/mci172>

Saunders B, Sim J, Kingstone T, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. *Qual Quant*. 2018;52(4):1893-1907. doi:10.1007/s11135-017-0574-8

Zurro, D., Garcia-Granero, J.J., Lancelotti, C. and Madella, M. 2016. Directions in current and future phytolith research. *Journal of Archaeological Science*, 68: 112-117. DOI:10.1016/j.jas.2015.11.014

### **Duration:**

One year for the FAIR assessment of existing data, developing minimum and best practice FAIR criteria for phytolith data and developing a training package for researchers for the implementation of criteria. All reporting will be within the given timescale for the funding.

### **Sustainability plan (2000 characters):**

This project embraces an open science approach and therefore all research outputs will be made as sustainable as possible. They will be stored in an open Github repository and archived on Zenodo for long-term storage. Please see the Data

Management Plan attached to this application.

[https://docs.google.com/document/d/1H-OdaqPDe84RrpIDbif2jgekXFXFXLQOhu\\_hPiQVHU/edit?usp=sharing](https://docs.google.com/document/d/1H-OdaqPDe84RrpIDbif2jgekXFXFXLQOhu_hPiQVHU/edit?usp=sharing)

This is only the first phase of the FAIRification process for phytolith data. Therefore, the issues arising from this research and the guidelines developed will be used to move forward with FAIRifying data. It is planned to make several funding applications during 2021 to apply for funding to continue this process such as the Future Leaders Fellowship (UKRI) by Emma Karoune and a large research grant to the Arts and Humanities Research Council (UK) or within Horizon Europe. These projects will be used to apply the newly developed guidelines in building an online open data repository for phytolith data and start research to demonstrate how FAIR phytolith data can be created and used for reproducible studies to strengthen existing and new methodological approaches.

The project is supported by the International Phytolith Society (IPS) and the working group will become an official standing committee of the IPS - the International Committee on Open Phytolith Science (ICOPS). This will ensure that we have access to a large number of researchers on an ongoing basis and we can report findings at the society conferences to enable the sustainability of the project in the phytolith community.