

Specifications table (Detailed overview of the study design)

Subject	Phytolith research
Specific subject areas	Archaeology Ecology Geochemistry Methodology (related to phytoliths) Palaeoecology Palaeontology Plant physiology
Specific subject glossary	
FAIR	FAIR stands for Findable, Accessible, Interoperable and Reusable. These principles aim to make data available whenever it is possible.
Metadata	Data which describes and gives context to other data.
Morphometry	Phytolith analysis based on measurements of size and shape.
Multiplicity	Multiplicity refers to a plant that can produce a variety of phytolith morphotypes.
Phytolith morphotype	A type of phytolith whose shape is diagnostic to various taxonomic levels. The pictures of those phytoliths are the most published.
Phytolith nomenclature	A system of conventions and rules to name the phytolith morphotype and the phytolith. Two phytolith nomenclatures exist ICPN 1.0 (Madella <i>et al.</i> 2005) and ICPN 2.0 (Neumann <i>et al.</i> 2019)
README file	A README file is a text file that aims at describing a project in terms of methodologies and key information about data collection and analysis. It can be used to provide a well-described metadata
Redundancy	Redundancy refers to the same phytolith morphotype

	which is commonly produced by several plants.
Dataset description	This dataset contains information collected from 100 phytolith papers - 50 selected from European studies and 50 selected from South American studies, over a 5 year publication time period (2016-2020).
Data format	Primary and secondary data
Type of data	Qualitative data
Full description of the repository files	<ul style="list-style-type: none"> • Data-dictionary-FAIR-assessment-final.csv • FAIR assessment data categories table • Observations made during data assessment table • Saturation and representativity of data collected.pdf • Specifications table (Detailed overview of the study design).pdf • README files <ul style="list-style-type: none"> ■ README file for European dataset ■ README file for South American dataset • Data collection forms <ul style="list-style-type: none"> ■ FAIR Phytoliths Data Assessment Form – Europe ■ FAIR Phytoliths Data Assessment – South America • <i>Roadoi</i> R package documentation <ul style="list-style-type: none"> ■ <i>roadoi</i> R package_Europe.csv ■ <i>roadoi</i> R package_South_America.csv ■ <i>roadoi</i> R package_Europe_code.md ■ <i>roadoi</i> R package_South_America_code.md <p>FAIR assessment trial</p> <p>This folder contains information about the trial we conducted prior to starting our main data collection phase.</p> <ul style="list-style-type: none"> • Data collection methodology development document • Data collection form for trial • Table summarising how the FAIR principles were linked to a practical set of questions shaped specifically for phytolith research.pdf

	<p>data-raw</p> <p>This folder contains the final full raw dataset as well as the two separate datasets for Europe and South America.</p> <ul style="list-style-type: none"> • Contains https://doi.org/10.5281/zenodo.7215572/data_raw • Also contains the two separate raw datasets for Europe and South America as csv files. <p>data-search</p> <p>This folder contains the files of the results of the searches we conducted to produce lists of articles.</p> <ul style="list-style-type: none"> • contains csv files of the searches conducted using publish or perish to get the list of relevant articles for Europe and South America articles. <p>paper-tables</p> <p>This folder contains the tables in the article and as supplementary files.</p> <ul style="list-style-type: none"> • Table-1-Specifications-table.pdf • Supplementary-Table-1-Data-dictionary-FAIR-assessment.pdf
<p>How data were acquired</p>	<p>Data were collected by six assessors in total (Emma Karoune, Carla Lancelotti, Juan José García-Granero, Javier Ruiz-Pérez, Marco Madella, Céline Kerfant). A trial of FAIR variables was conducted on five articles to check the criteria relevance by five assessors while the main data collection was carried out by four assessors.</p> <p>Access to the literature was carried out using <i>Publish or Perish</i> software (<i>Harzing, A.W. (2007) Publish or Perish, available</i>) to conduct two searches using the decided criteria ('phytoliths' and 'Europe' or 'South America') to produce a static list of papers. The first 100 articles selected from these lists using the inclusion criteria ('primary data', 'phytoliths' and 'Europe' or 'South America') were assessed on the basis of the availability and format of data and metadata presented. The R package <i>roadoi</i> (version 0.7.2) was used to transparently assess the</p>

	accessibility of the articles. The languages of most of the studies are English and Spanish.
Open science tools used	<i>Publish or Perish</i> software (Harzing, A.W. (2007) <i>Publish or Perish</i> , available) - https://harzing.com/resources/publish-or-perish roadoi R package (version 0.7.2) - https://github.com/ropensci/roadoi
Data availability	Zenodo - https://doi.org/10.5281/zenodo.7215572 GitHub repository - https://github.com/open-phytoliths/FAIR-assessment-data-paper-documentation Data is available under the Creative Commons Attribution 4.0 International licence