Table 2: Condensed standardised data dictionary (a full version of this table can be found in Github.

Variable number	Variable Name	Variable Label	Variable Data Type	Allowed Values	Allowed Value Description
1	Data-collectors-name	Name of data collector	string	Carla Lancelotti, Celine Kerfant, Emma Karoune, Javier Ruiz-Pérez, Juan José García-Granero, Marco Madella	Names of all of the data collectors from this project
2	Journal-name	Name of the Journal the article is published in	string	A journal name	Any journal name that appeared in selected articles
3	Full-Reference-APA	Full reference for the article in APA style	string	APA journal citation	Full citation reference in APA format
4	Year	Year the article is first published online - not a full date	integer	2016, 2017, 2018, 2019, 2020	Five year range used for selecting articles in this study
5	Period/Date	Archaeological or palaoecological studies named period or date that appears in the article or modern for studies that only use modern plant material	string	A specific period or date, or modern	Open text box for period or date found in the text of the article. If a study conducted on modern plants then response should be modern
6	Type-of-study	Main study focus of the article	string	Archaeological, Geochemical, Ecology, Methodological - modern reference, Methodological - morphometric, Methodological - ethnographic, Methodological - extraction method, Methodological - isotopes, Methodological - taphonomy, Methodological - radiocarbon, Palaeoecology, Palaeontology,	Categories of types of studies that use phytolith analysis

				Plant Physiology	
7	Geographic-location	Geographic location of study	string	A country/countries or region name	Country or region that was within the text of the article - standardised using geonames
8	oa_status	Open access status of the article	string	Green, Gold, Hybrid, Bronze, Diamond, Closed	unpaywall categories used but with addition of diamond
9	Journal-type	Type of journal	string	Open, Hybrid, Closed	
10	What-article-repository	Repository used for green open access articles	string	Arxiv, Bioarxiv, EarthArXiv, EcoEvoRxiv, F1000 research, Figshare, OSF Preprints, Palaeorxiv, PeerJ Preprints, Preprints.org, University or Institution repository, WikiJournal Preprints, Zenodo, None-gold, None-diamond, None-not-oa	None-gold = gold open access and also includes a few bronze oa articles, None-not-oa = not open access
11	Is it in a signed up repository?	Whether the article was placed on a repository such as ResearchGate or Academia, rather than a trusted research repository or pre-print server	string	Research gate Academia	
12	Processing-method-communication	Extent of detail the method written in the article or attached to the article	string	Summary of method, Full protocol in the text/supplementary, Full protocol in repository, One reference to specific protocol, One reference generic ref/Multiple reference,	

				No method reported	
13	Counting-method-is-replicable	Can the assessor replicate exactly what was counted on the phytolith slide?	string	Yes, No, Not applicable	
14	Instrument-used-described	Is there a description of the instrument (microscope) used for phytolith counting or other analysis?	boolean	Yes, No	
15	Nomenclature-used	Was a name phytolith nomenclature used?	string	ICPN 1.0, ICPN 2.0, MU Phytolith Classification System, Bertoli 2013, Zucol & Brea 2005, Not specified/not clear	
16	Have they used it fully?	Have they used the standardised nomenclature fully? ICPN1.0 or ICPN 2.0	boolean	Yes, No, Not used	Yes = fully used, No = not fully used, Not used = not used either ICPN 1.0 or ICPN 2.0
17	Data-location	Where is the data located in the article, supplementary material or repository?	string	No data, Data in text of article, Data in supplementary, Data in repository with DOI, Data in repository without DOI,	
18	What-data-repository	What data repository is the data located in?	string	ADS, Code Ocean, Dataverse, Dryad, Zenodo, Figshare, GitHub, Mendeley, Neotoma, Open Context, Open Science Framework, Pangaea, tDAR, None	

19	Data-type	What type of data is in the article?	string	i) Raw counts, ii) Raw counts + processing weights, iii) Presence/absence, iv) Relative presence/percentages, v) Mean or range, vi) Summary data - plant groupings, vii) Unclear what is presented/provided, viii) No data provided in a table - only graph, ix) No data provided	
20	Data-format	What type of data file format?	string	No data, In text of article, .pdf, .doc, .xslx, .csv	
21	Data-availability -statement	How is the data made accessible in the data availability statement?	string	None, Data on request, All data included, Dataset link to repository with DOI, Data paper with DOI	
22	Data-license	What license is on the data, if the data is in a repository	string	Data not available = not in repository, No license, CC-0, CC BY	
23	Pictures or photos included for identifications	Are all, significant morphotypes or no photos included in the article	string	None, Only significant ones, All	
24	Statistical-softw are-used	Names of the statistical software packages named in the articles	string	Any software package named in the article used for statistical analysis	
25	Other-comment s	Open text box for additional comment from data collectors	string	N/A	N/A