Journal of open archaeology data

Neolithic Settlements in Central Europe: Data from the Project 'Lifestyle as an Unintentional Identity in the Neolithic'

DATA PAPER

PETR PAJDLA
FRANTIŠEK TRAMPOTA
*Author affiliations can be found in the back matter of this article

ubiquity press

ABSTRACT

This dataset comprises the core spatial and temporal structure of the 'Lifestyle as an Unintentional Identity in the Neolithic' project. The data consist of spatial and chronological information on 2,154 Neolithic settlement sites from c. 4900 to 3300 BCE in two separate regions, the eastern part of Bohemia (Czech Republic) and the Morava River basin (Czech Republic, parts of Austria and Slovakia). To gather as much data as possible, all available sources including published works, find and excavation reports in archives, museum collections and unpublished records and both existing and legacy databases were examined. The dataset is deposited at the Zenodo data repository with an open access license. There is a reuse potential in aggregating the data with data from different regions and/or time periods for various spatial and temporal analyses.

CORRESPONDING AUTHOR:
Petr Pajdla
Department of Archaeology and Museology, Masaryk University, Brno, CZ
petr.pajdla@protonmail.com

KEYWORDS:
Neolithic; Central Europe; settlement analysis; landscape; lithic raw materials; relative chronology

TO CITE THIS ARTICLE:
Pajdla P, Trampota F. 2021 Neolithic Settlements in Central Europe: Data from the Project 'Lifestyle as an Unintentional Identity in the Neolithic'. *Journal of Open Archaeology Data*, 9: 13, pp. 1–6. DOI: *https://doi.org/10.5334/joad.88*

Another constraint is that some sites are repeatedly mentioned under different names and, on the other hand, different sites can be recorded under one local name covering a large spatial extent. Although this was controlled for in the dataset, it must be acknowledged.

(3) DATASET DESCRIPTION

OBJECT NAME

- **sites.csv** – is a main list of sites with unique identifiers in the ID field; an ID starting with B means the site is from the eastern part of the Bohemia section of the study area, while in the case of an ID starting with M the site is from the Morava River basin area. The *orig_id* field contains an identifier by which the site is referenced in cited works and the field *site* contains the site name;
- **pot_traditions.csv** – contains site IDs, field *chrono* giving the general pottery tradition and field *period* listing the occurrence of the site in one of the nine time slices;
- **pot_groups.csv** – has the same fields as the previous file, with a difference in the *chrono* field containing information on detailed pottery groups;
- **references.csv** – a list of references, where possible, the excavation reports are linked to their source in the Digital Archive of the Archaeological Map of the Czech Republic (*https://digiarchiv.aiscr.cz/*). Column *ref_id* is linked through file *references_sites.csv* to the database of sites in the *sites.csv* file;
- **references_sites.csv** – connects files *references.csv* and *sites.csv*.

Geodata (in S-JTSK/Krovak East North coordinate reference system):

- **site_locations.gml** and *site_locations.xsd* – settlement sites locations. The ID field gives a unique identifier for each site. Column *accuracy* indicates how accurately the site location is defined; value **1** means an accurate location (instrumentally measured); value **2** is precision in hundreds of metres, i.e., the site location is known by the local name, street name, etc.; and value **3** means the location is not very accurate, in a roughly 1-km range. The field *surface* is **TRUE** if the site is only defined based on a surface survey, and field *altitude* gives altitude in metres;
- **study_area.gml** and *study_area.xsd* – polygon giving the borders of the initial study area where data was collected. For any analysis, spatial extent given in the *regions.gml* file should be used;
- **regions.gml** and *regions.xsd* – polygons defining the extent of studied regions (i.e., the eastern part of Bohemia and the Morava River basin). The polygons were created by buffering the site locations in 10-km range and the resulting polygons were cropped by the polygon of the study area (*study_area.gml* file);
- **raw_material_sources.gml** and *raw_material_sources.xsd* – locations of raw material sources as points or lines. Points are based on places where prehistoric procurement activities are known or outcrops of the given raw materials are present. Lines give the border of the raw material occurrence in the case of erratic flint or river courses in which the raw materials can be procured. The *rm* column gives an abbreviated name of the raw material and the *type* field is either **l** for chipped stone tools or **p** for polished stone tools.

Vocabularies:

- **voc_periods.csv** – contains period labels;
- **voc_pot_traditions.csv** – contains pottery traditions labels, where possible, field *periodo_link* maps the period to AMCR Periods Vocabulary at Periodo (*http://n2t.net/ark:/99152/p0wctqt*);
- **voc_pot_groups.csv** – contains pottery groups labels, same fields as previous file;
- **voc_pot_groups_facets.csv** – general labels for pottery groups;
- **voc_raw_materials.csv** – list of raw material abbreviations in the *rm* column of *raw_material_sources.gml* file with full names derived from [3].

DATA TYPE

Secondary and processed data collected largely from published studies (for a detailed list of citations, see file *references.csv*) and excavation reports.

FORMAT NAMES AND VERSIONS

CSV, GML

Flat tables are represented as comma-separated values (CSV), spatial information is in Geographic Markup Language (GML) file format, i.e., a .gml file and an associated schema definition file (.xsd).

CREATION DATES

The dataset was created in several phases between 2010 and 2020. At first, data in the Thaya River basin, a segment of the Morava River basin, were collected as part of the PhD thesis of the second author [8]. Later, in the course of the 'Lifestyle as an Unintentional Identity in the Neolithic' project (2019 to 2021), the geographical scope was expanded to the whole Morava River basin and eastern part of Bohemia.

DATASET CREATORS

František Trampota (Department of Archaeology and Museology, Masaryk University) collected most of the data.