

Archaeology Data Infrastructures

Data reuse potentials and limitations to modelling settlement systems (...)

Petr Pajdla

2022-11-12

Table of contents

Preface	3
Notes on writing in the open	3
Introduction	5
1 Theory and Method	6
1.1 Archaeology as theory- and/or data-driven science	6
1.2 Theorizing data	6
1.3 Methodological Approaches	6
2 Data	7
3 Discussion	8
4 Summary	9
References	10

Preface

⚠ This is a website for the **work-in-progress** PhD thesis of mine. It is **not** intended to be read by anyone except me (*and maybe few other people*) yet. If you do flick through it anyway, consider yourself warned. It might be messy at some places and will definitely undergo serious rewriting.

i This work can be read online at <https://petrpajdla.github.io/dataInfrastructures/>. The source repository is on GitHub at <https://github.com/petrpajdla/dataInfrastructures/>.

This document is created in an open-source [Quarto](#) scientific and technical publishing system. You might be asking why is it published and written like this even if it is not intended for any audiences except myself yet. I have no answer to this. One evening I simply decided to give *Quarto* publishing a try and set this whole thing up in less than an hour or so.

Notes on writing in the open

This note is written mostly for a future me, in case I need to set up the working environment again on a different machine and to serve as a memo if I forget how to continue.

As of November 2022, this is written on [Archlabs](#) *GNU/Linux* machine, mostly in [Visual Studio Code](#) editor and sometimes in [RStudio](#). Changes are tracked with *Git* and a remote repository is on *GitHub* (see the note above), same as the rendered website. The rendered version of the manuscript is in the branch **gh-pages**. See a guide on how to set this up [here](#). The online version is published with this command:

Terminal

```
quarto publish gh-pages
```

In my point of view, there are numerous advantages to writing in this manner. A non-exhaustive list is below.

- **Plain text**

Writing in plain text enhanced with a simple *Markdown* syntax and some *Quarto* elements is great because from one source document, a .pdf, .html and a .docx documents can be rendered using [pandoc](#).

- **Version control**

Tracking changes using *git* is easily implemented when writing in plain text and keeping track of any changes in the manuscript is crucial for any later revisions etc.

- **Simple citation management**

Bibliography is organized using [Zotero](#) with [Better BibTeX](#) extension which is used to export (and keep updated) necessary collections in a parent folder of the manuscript. My *Zotero* library is [here](#).

Introduction

This is a book created from markdown and executable code.

See Zhang, Zhao, and Ventrella (2018) for additional discussion of literate programming.

1 Theory and Method

1.1 Archaeology as theory- and/or data-driven science

(based on TAG Brno 2021 talk)

1.2 Theorizing data

Defining archaeological data, micro- to macro-scales;

1.3 Methodological Approaches

Review of current approaches: Spatial and/or Landscape archaeology, Macroarchaeology, Big data archaeology etc.

2 Data

Sources of (archaeology) data in the Czech Republic, an overview:

Data models, datafication of past reality, simple vs complex data models; Assessing findability, accessibility, interoperability, and reusability (FAIR) principles; Cultural heritage management data vs research data domains; Archaeological information system of the Czech Republic (AIS CR) as the main data infrastructure.

3 Discussion

4 Summary

In summary, this book has no content whatsoever.

References

- Zhang, Shaozeng, Bo Zhao, and Jennifer Ventrella. 2018. “Towards an Archaeological-Ethnographic Approach to Big Data: Rethinking Data Veracity.” *Ethnographic Praxis in Industry Conference Proceedings* 2018 (1): 62–85. <https://doi.org/10.1111/1559-8918.2018.01197>.