PaleOrdia: Semantically describing (cuneiform) paleography using paleographic linked open data

Timo Homburg
timo.homburg@hs-mainz.de
Hochschule Mainz, University Of Applied Sciences
funded by DFG grant 424957759: "Die digitale Edition der Keilschrifttexte aus Haft Tappeh"

Problem

Modeling paleography of scripts in linked open data has not necessarily been tackled in the LOD community. For many scripts, the shape of the individual sign does not necessarily matter for tasks such as natural language processing or would provide too little value to be relevant for classifications. For scripts where paleography matters for classification, paleographic relations have often not been formalized or being put into training data. Therefore, the synnergies that may lead to better machine learning classifications or the insights which could be gained by modelling paleographic data for scripts, also independent of languages are yet to be discovered by the linked open data community.

Related Work

First ideas to capture paleography with LOD in the digital humanities were brought forward by [1] to capture sign encodings [2] which are used to describe cuneiform signs which was later formalized and presented in 2022 [3] as a possible extension to the Ontolex-Lemon model [4].

References

[1] Timo Homburg. Towards paleographic linked open data (plod): A general vocabulary to describe paleographic features. In *DH*, 2020.

[2] Timo Homburg. Paleocodage—enhancing machine-readable cuneiform descriptions using a machine-readable paleographic encoding. *Digital Scholarship in the Humanities*, 36(Supplement_2):ii127–ii154, 2021.

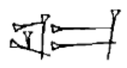
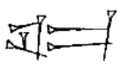
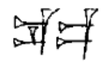
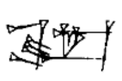
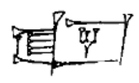
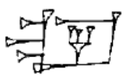
[3] Timo Homburg and Thierry Declerck. Towards the integration of cuneiform in the ontolex-lemon framework, 2022.

[4] John P McCrae, Julia Bosque-Gil, Jorge Gracia, Paul Buitelaar, and Philipp Cimiano. The ontolex-lemon model: development and applications. In *Proceedings of eLex 2017 conference*, pages 19–21, 2017.

[5] Finn Årup Nielsen. Ordia: A web application for wikidata lexemes. In *The Semantic Web: ESWC 2019 Satellite Events: ESWC 2019 Satellite Events, Portorož, Slovenia, June 2–6, 2019, Revised Selected Papers 16*, pages 141–146. Springer, 2019.

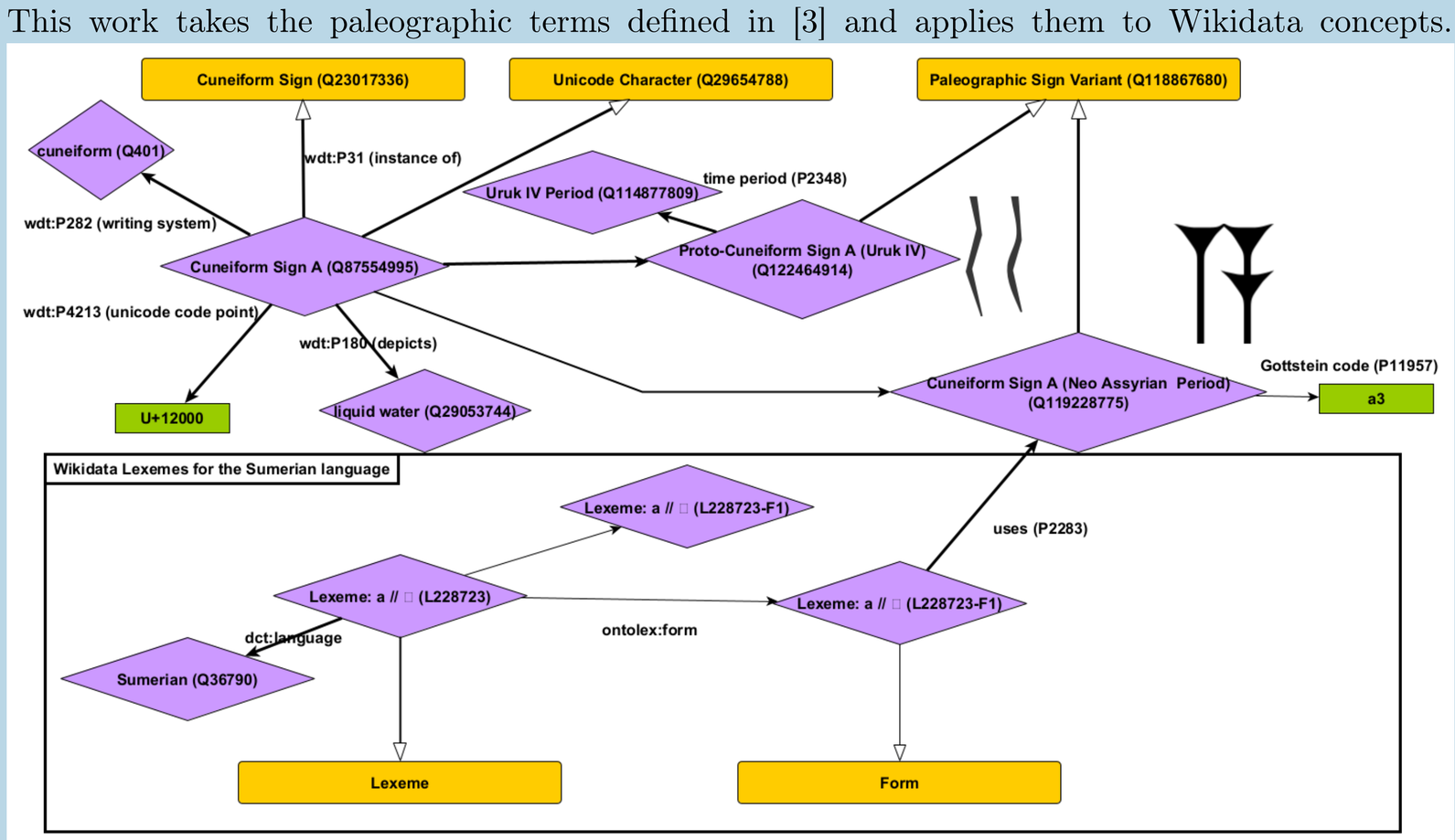
Cuneiform Paleography Model: Motivation

Paleographic changes, here in the cuneifrom script, occur over the period of the scripts' existence. Etymological relations can be inferred and modelled in RDF. Individual sign variants can be described using character encodings.

Time Period Designation Point in time	Late Uruk ca. 3100	Djemdet Nasr ca. 3000	Early Dynastic III ca. 2400	Ur III ca. 2000	Old Assyrian ca. 1900	Old Babylonian ca. 1700	Middle Assyrian ca. 1200	Neo-Babylonian ca. 600	meaning
Glyph (U+12295)									SAG "head"
Gottstein	N/A	N/A	a3b2c1d1	a3b2c1d1		a3b4	a3b3c1	a3b4	
Glyph (U+120FB)									NINDA "ration"
PaleoCode	N/A	N/A	:sa~a-:c~;<F	a-:sa-sa~~~;f	sa-sa;w-sa	sa-:sa;w-sa	sa-sa;w-sa	f-;>c–:sb-a	
Gottstein	N/A	N/A	a2c1d1	a3d1	a3c1	a3c1	a3c1	a1c2d1	
Glyph (U+12165)									GU_7 "to eat"
Gottstein		N/A	a5b2c6d2	a5b2c4d2		a6b7c2d1	a5b5c1	a5b4c2	

Cuneiform Paleography LOD Model in Wikidata

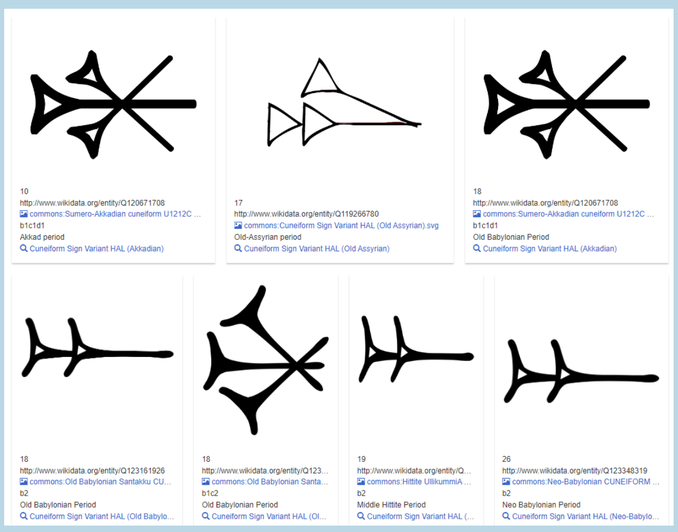
This work takes the paleographic terms defined in [3] and applies them to Wikidata concepts.



PaleOrdia Application on the example of cuneiform

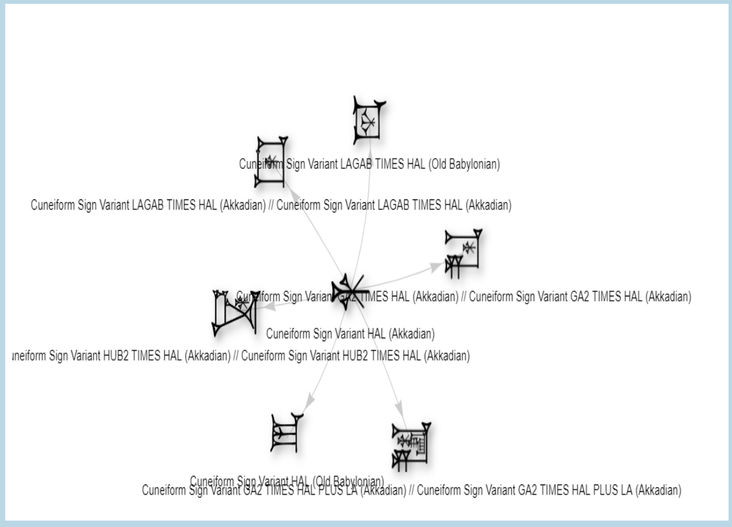
PaleOrdia (`https://situx.github.io/paleordia/`) is a fork of the tool Ordia[5], which allows to browse Wikidata Lexemes and paleography. PaleOrdia is a static web application on Github. It executes SPARQL queries in JavaScript to generate views on paleographic data within Wikidata.

Search



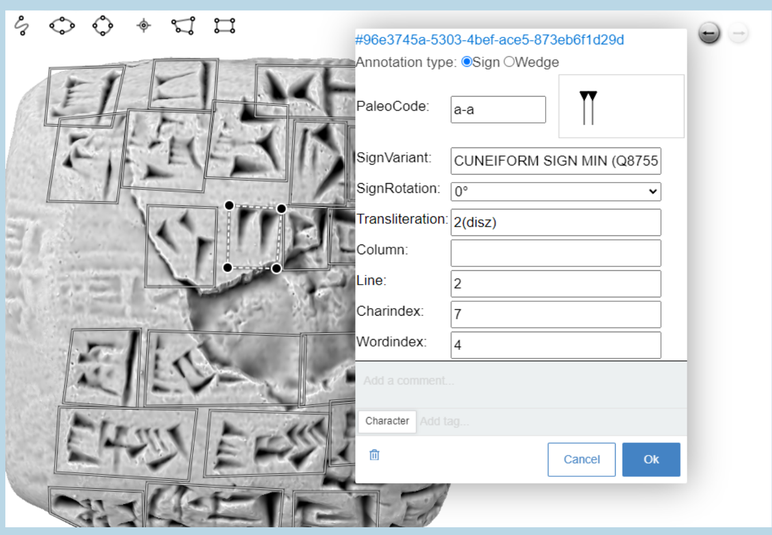
- Search for etymology, sign variant per time period and reference work
- Search for sign variants per time period
- Search for sign variants per reference work
- Search for sign variants used in Lexemes

Exploration



- Explore compound signs, i.e. signs in which a sign occurs
- Find signs without Unicode codepoints
- Find signs with disputed readings

Annotation



- Use the Wikidata paleographic LOD cloud in annotation applications
- Cuneiform Annotator allows to search for Wikidata paleographic signs
- Saves result as Web Annotations