

Mapping the Dawn of History: A Mixed Methods Approach to the Distribution and Composition of the Cuneiform Corpus (c. 3,400 BCE - 100 CE)

[Rune Rattenborg](#),¹  [Carolin Johansson](#),¹  [Seraina Nett](#),¹ [Gustav Ryberg Smidt](#)² 

¹ Department of Linguistics and Philology, Uppsala University, Sweden

² Department of Languages and Cultures, Ghent University, Belgium

<https://www.lingfil.uu.se/research/assyriology/glow/>

This paper reviews the aim, structure and initial outcomes of a three-year research project with the aim to provide an updated survey of the cuneiform corpus, the material product of a script in widespread use across the Middle East from c. 3,400 BCE until 100 CE. Counting a conservative 500,000 inscriptions known from public and private collections across the globe, the corpus ranks among the largest discrete bodies of textual source material from the ancient world. Moreover, the specific circumstances and processes of formation and discovery of this corpus mean that the relative prevalence of textual genres represented in this corpus is markedly different from most other textual corpora from the ancient world. As such, this corpus offers unique opportunities for the study of the use and spread of writing across early human societies.

Studies on text production and consumption in a variety of Old World historical periods, such as the European Middle Ages and Dynastic China (Buringh and van Zanden 2009; Xu 2013) have demonstrated that long-term perspectives on the material signature of writing may yield profound insights on the changing uses of text in human societies through time and space. Data-driven, large-scale investigation of social dynamics across regions and continents has by now become commonplace in archaeology, tapping into a rapidly growing array of standardised and open access datasets on virtually any aspect of human material culture. While such studies have proven extremely capable in bridging previously insulated fields of regional or temporal specialisation, similar perspectives on ancient writing as an integral element of material culture have remained largely absent in comparative and diachronic analyses, even if the tools requisite for such perspectives are becoming increasingly available (see for an innovative example Heřmánková et al. 2021).

Within the field of cuneiform studies, the last two decades have seen a steady growth in the number and extent of digital text catalogues, creating an immensely rich, if highly heterogenous, body of digitised information on primary texts (Charpin 2014). Even so, significant transects of the corpus remain available only from print publications and grey papers. Placed at the intersection of philology, archaeology, and digital and spatial

humanities, our project works to integrate and augment metadata from existing digital catalogues into a coherent and comprehensive index of the entire corpus, through collaboration with existing repositories and through the digitisation of analogue records. A central structuring element of this approach has been the development of a comprehensive open access gazetteer of archaeological finds of cuneiform texts (Rattenborg et al. 2021). Through the development of formal and transparent criteria for dealing with fuzzy geographies and imperfect geographical data, this resource allows for easy visualisation and querying of an immense number of primary records and underpins analyses of more intricate patterns of corpus formation, deposition, and discovery through the examination of a variety of metadata variables.

Here, we offer first an initial review of basic spatial statistics of the entire corpus as an archaeological artefact, thereby gauging the relative historical authority of the corpus in addressing more complex research questions. We then proceed to present a set of case studies that will focus on the changing roles of specific materials as media of writing during the extremely long history of the script, the waxing and waning of specific genres of text over time and space, and prospective diatopical patterning in assemblage composition. These examples aim to underscore the potential insights that can be gained on ancient writing when approached from a spatial, macro-historical dimension. Given its immense size and particular material characteristics, the cuneiform corpus is uniquely suited for such an approach.

Buringh, Eltjo, and Jan Luiten van Zanden. '[Charting the "Rise of the West": Manuscripts and Printed Books in Europe, A Long-Term Perspective from the Sixth through Eighteenth Centuries](#)'. *The Journal of Economic History* 69 (2009): 409–45.

Charpin, Dominique. '[Ressources Assyriologiques Sur Internet](#)'. *Bibliotheca Orientalis* 71, no. 3–4 (2014): 331–57.

Heřmánková, Petra, Vojtěch Kaše, and Adéla Sobotková. '[Inscriptions as Data: Digital Epigraphy in Macro-Historical Perspective](#)'. *Journal of Digital History* 1 (2021).

Rattenborg, Rune, Carolin Johansson, Seraina Nett, Gustav Ryberg Smidt, and Jakob Andersson. '[An Open Access Index for the Geographical Distribution of the Cuneiform Corpus](#)'. *Cuneiform Digital Library Journal* 2021, no. 1 (2021): 1–12.

Xu, Ting. '[Knowledge Formation and the Great Divergence between China and Europe: Manuscripts and Printed Books, ca. 581–1840](#)'. *Journal of Comparative Asian Development* 12, no. 2 (1 August 2013): 245–84.