UPPSALA UNIVERSITET

TextWorlds: Digital Perspectives on the Composition and Distribution of Writing from the Ancient World

Thursday 16 June 09.00–17.00 CEST (UTC+2) Cassiopeja Meeting Room, Centre for Integrated Research on Culture and Society (CIRCUS), the Old Observatory, Kyrkogårdgatan 8 A, 75210 Uppsala, Sweden. Please register to attend remotely or in-person.

Programme

First session

09.00-09.15 ***Welcome and introduction*** (Rune Rattenborg)
09.15-09.45 ***All the Things They Wrote on Clay: Exploring the Quantitative Dimensions of the Cuneiform Corpus c. 3,400 BCE–100 CE*** (Rune Rattenborg, SerainaNett, and Émilie Pagé-Perron)
09.45-10.15 ***DAMOS–Database of Mycenaean at Oslo*** (Federico Aurora)

10.15-10.30 Break

Second session

10.30-11.00 ***A Survey of Andean Knotted String Records in the Open Khipu Repository*** (Manuel Medrano)
11.00-11.30 ***The Materiality of Classic Maya Textual Corpora*** (Christian Prager)
11.30-12.00 ***The Australian Message Stick Database: Representing a Variable System with Partial Sources*** (Piers Kelly)

12.00-13.30 Lunch

Third session

13.30-14.00 ***Latin inscriptions from the Epigraphic Database Heidelberg (EDH): one piece of a puzzle*** (Petra Heřmánková)
14.00-14.30 ***The Palaeohispanic Corpus of Inscriptions, and the Hesperia Database*** (Eduardo Orduña)
14.30-15.00 ***From an Expanded Database of Papyrological Metadata to an Interdisciplinary Portal for the Study of Ancient World Texts, the Case of Trismegistos*** (Tom Gheldof)

15.00-15.30 Break

Fourth session

15.30-16.00 ***The Scandinavian Runic Corpus and Chronological Categorization*** (Hanna Åkerström)
16.00-16.30 ***Studying the Textual Corpus of Ancient Arabia in a Comparative Perspective: Quantitative and Qualitative aspects*** (Irene Rossi)
16.30-17.00 Respondent (Anna Foka)

Abstracts

First session

All the Things They Wrote on Clay: Exploring the Quantitative Dimensions of the Cuneiform Corpus c. 3,400 BCE–100 CE

Rune Rattenborg (Uppsala University), Seraina Nett (Uppsala University), and Émilie Pagé-Perron (Oxford University)

The number of inscriptions written in cuneiform, a script in widespread use across the Middle East over more than three millennia of recorded human history, is immense by any standard. Current estimates hover somewhere between 500,000 and 1,000,000 objects housed in museums, collections, and storerooms on six continents, the product of a long, complex, and contentious history of archaeological and philological research in Iraq, Syria, Turkey, Iran, and adjoining countries. In this paper, we explore primary aspects of the composition and distribution of this enormous body of written source material through the augmentation of data from the core digital catalogue of cuneiform inscriptions, namely the Cuneiform Digital Library Initiative, with updated spatial indices provided by Geomapping Landscapes of Writing. Applying a birds-eye perspective on the materiality of writing in the ancient world, we trace broader patterns and trends in the waxing and waning of materials, artifact types, and genres over the long term. Using basic methods of spatial analysis, we suggest that cuneiform offers unique insights on the prevalence of writing in early history. As such, processes of formation and discovery characteristic of this corpus may rewardingly inform perspectives on other bodies of written material from the ancient world.

DAMOS - Database of Mycenaean at Oslo

Federico Aurora (University of Oslo)

The Mycenaean Greek corpus is comprised of ca. 6000 administrative texts found in the Mycenaean palaces of Crete and mainland Greece and dated between ca. 1350 and 1150 BCE. The texts are written in a syllabary which we call Linear B and constitute the oldest attestation of an Ancient Greek dialect. The documents (clay tablets, sealings and jars) were not meant to be preserved for more than a relatively short administrative period, but were accidentally baked in the fires that took place at different times in the Mycenaean palaces and have thus been preserved to us as the sole testimony of Mycenaean literacy.

DAMOS (Database of Mycenaean at Oslo) is a digital repository of transliterations of all published texts (only ca. 100 recently found documents still await publication) and related metadata. Its online interface allows users to browse and search the texts by metadata, content, and a number of epigraphic features. Finally, each document post is completed by external links,

which make DAMOS also an entry point to other related online resources, as those providing pictures of the documents or related artifacts (e.g. the seals which were used to produce seal impressions on the sealings).

Second session

A Survey of Andean Knotted String Records in the Open Khipu Repository

Manuel Medrano (University of St Andrews)

The Open Khipu Repository (OKR) is the largest existing compilation of data from Andean khipus — the undeciphered knotted string devices used for recording information in the Andes for at least 1,000 years (c. 950–1950 CE). The OKR's relational database, which contains cord-by-cord observations produced since the early 1970s by various researchers, has already informed some 15 years' worth of data-driven hypotheses regarding the nature of recordkeeping during the Inka Empire (c. 1400–1532 CE), the most prolific period of khipu use. However, to what extent is the OKR representative of a singular (Inka) khipu tradition? Where do its records come from? How has the history of khipu cataloging left its mark on the composition of the data set? In this talk, I survey the 629 Inka-style khipus in the OKR, with particular emphasis on provenance, age, material, and morphology. Though not exhaustive, these variables offer a window into the heterogeneity of the corpus, suggesting a greater diversity of khipu traditions than is often acknowledged. Interpretive nuance is discussed throughout within the broader aims of the OKR Advisory Board — a newly formed consultative body overseeing the OKR's administration.

The Materiality of Classic Maya Textual Corpora

Christian Prager (Text Database and Dictionary of Classic Mayan, University of Bonn)

This paper addresses the semi-deciphered written language of the Classic Maya, whose cultural area extended over territories of the present-day nation states of Mexico, Guatemala, Belize and Honduras. Maya hieroglyphic writing was used between between 300 BC and AD 1500. As a visual language, Classic Mayan survived in more than ten thousand texts. Most sources exhibit biographical information on political elites and provide written evidence for political relations between the more than sixty ruling dynasties. The inscription's focus lies on religious and political events that marked elite daily life. Maya kings made their public claim to power through writing and iconography. In this context, written and pictorial records, especially those on stone, wood, ceramics, bone and fig-bark paper, not only served as vehicles for cultural memory at the time, but today form the most important material basis for reconstructing elite history and culture. The materiality, function, and morphology of the inscribed artifacts are semiotically charged and semantically closely related to specific text genres, textual content, and pictorial information. For example, publicly accessible staircases were used to record war narratives in words and images, clay vessels were painted with ownership marks and provided with mythologies, or books made of bark paper were filled mainly with religious content. The close connection between genre and material will be systematically demonstrated in this paper.

The Australian Message Stick Database: Representing a Variable System with Partial Sources

Piers Kelly (University of New England)

Message sticks are marked objects, typically made of wood, used in Indigenous Australia for facilitating important long-distance communications. Between the 1880s and the 1910s, settlers and scholars took great interest in message sticks and this was reflected in efforts to document, collect and conserve them in museums worldwide. However, these collectors often failed to grasp the relationship between individual symbols and their meanings, the complex pragmatics

of the system, and its variable application across the continent. My presentation introduces the Australian Message Stick Database, a dataset of 1322 items compiled from museum registers, auction catalogues, photographs, sketches, and written descriptions. As such it aggregates diverse metadata standards from different institutions and record types in an effort to capture as much information as possible about individual objects. Making this data meaningful has required the cautious interpretation of sources that are 'partial' in both senses of the term: incomplete as well as distorted by colonial biases. Solutions to this dilemma have involved the generation of flexible standards, a mechanism for sources to be annotated without modification, and a hierarchy of descriptive adequacy based on the degree to which relevant Indigenous specialists were consulted about the meaning of the object.

Third session

Latin inscriptions from the Epigraphic Database Heidelberg (EDH): one piece of a puzzle

Petra Heřmánková (Aarhus University)

The Epigraphic Database Heidelberg (EDH), with almost 82,000 records, represents one of the best digital sources for studying Latin inscriptions over several centuries on an Empire-wide scale. These inscriptions were systematically collected for over 35 years and meticulously published and curated online, providing the best and one of the largest digital corpora of Latin Provincial epigraphy. However, the recent cuts in funding made the source vulnerable and unfortunately, its future development halted in 2021.

I will be discussing the dataset itself from a user perspective, describing a snapshot of data we made in 2020 for a longitudinal study of epigraphic trends for the *Social Dynamics in the Ancient Mediterranean* (SDAM) Project at Aarhus University. After a basic introduction of the contents of the data, aligned with the TextWorlds classification system, I will discuss the representativeness of the data with all its biases and shortcomings and the challenges any comparative and large-scale study has to inevitable face. To read more about our experience working with EDH, see Heřmánková *et al.* 2021 in the Journal of Digital History 1.
The presentation will be prerecorded.

The Palaeohispanic Corpus of Inscriptions, and the Hesperia Database

Eduardo Orduña (Institut El Pont de Suert - Proyecto Hesperia)

The Palaeohispanic languages spoken in the Iberian peninsula before and during the Roman times, at least until the change of era, have left up to 2800 epigraphic records. Most of them, about 2300, belong to the Iberian language, spoken in the Mediterranean coastal area of Spain and South France, in two different semi-syllabic scripts, NE and SE. Only the first one is completely deciphered, but the language can not be understood. The same script was used to write Celtiberian, a Celtic language whose epigraphic remains are among the longest of the ancient Celtic texts. A script close to the SE Iberian was used to write an unclassified language recorded only in stone steles in south Portugal. The Latin alphabet was used to write the very scarce Lusitan inscriptions, an Indoeuropean but probably not Celtic language. All those indigenous inscriptions are recorded, and most of them already made public, in the Hesperia database. The distribution of the inscriptions, the materials used and their chronology show interesting patterns related to the colonial contact with Phoenicians, Greeks and Romans.

From an Expanded Database of Papyrological Metadata to an Interdisciplinary Portal for the Study of Ancient World Texts, the Case of Trismegistos
Tom Gheldof (KU Leuven)

Trismegistos [TM] was created in 2005 as a database, containing information on published papyrological documents from Graeco-Roman Egypt. Data such as the material, provenance, attested persons, and more metadata from these texts were collected, standardized, and published in an online portal with a search interface. The following years, TM started including epigraphic material, broadened its spatial and temporal scope and began experimenting with digital humanities tools and quantitative methods to facilitate research with (parts of) the dataset.

In total, TM aggregates (internal and external) metadata about almost 1 million ancient world texts, ranging globally from 800 BC–AD 800. The papyrological (sub)corpus is still the determining standard (e.g., for several conventions such as dates or genres), although the expansion to the broader ancient world has had implications for the representation of the TM dataset in general. In this presentation, not only the challenges of bringing together textual material from different fields and (re)sources will be discussed, but also the opportunities that creating and expanding a digital corpus of such (meta)data provides for both its creators and researchers.

Fourth session

The Scandinavian Runic Corpus and Chronological Categorization
Hanna Åkerström (Uppsala University)

The Scandinavian runic corpus consist of ca. 7000 inscriptions spanning over almost 2000 years, from the 2nd century to the end of the 19th century. The absolute majority of the inscriptions are found in Sweden, Denmark and Norway, but also in Nordic settlements located in areas such as Greenland, England, Russia and others.

The Scandinavian Runic-text Database and the upgraded version that is a result from the project Everlasting Runes, are both tools used to study the runic corpus. I will address three issues related to the chronological categorization of the inscriptions in the database. The first issue concerns how the presentation of the inscriptions are affected by the selected time-periods, as they highlight certain aspects of the material and downplay others. The second issue relates to how the dates are presented as absolute, even though they rarely are. The third concerns how the time-periods used for categorizing inscriptions from all over Scandinavia in some cases in fact suit some areas better than others.

Studying the Textual Corpus of Ancient Arabia in a Comparative Perspective: Quantitative and Qualitative aspects
Irene Rossi (CNR-ISPC, Milan)

The civilizations of Ancient Arabia left a written documentation consisting of tens of thousands of inscriptions and graffiti in different Semitic languages and alphabetic scripts. The distribution of those sources is intimately related to the settlement pattern and to the political histories of the communities that inhabited Arabia from the beginning of the 1st mill. BCE until the advent of Islam.

In this paper, quantitative and qualitative aspects of this corpus will be explored based on the results of the analyses carried out within the TextWorlds project on the dataset of DASI – Digital Archive for the Study of pre-Islamic Arabian Inscriptions, which provides the curated

digital edition of more than 8,500 inscriptions from Ancient Arabia, most of which are part of the Ancient South Arabian corpus.

Keeping the latter at the centre of our discussion, special attention will be paid – in a comparative perspective – to focus themes of the project, such as: epigraphic corpus's vs digital corpus's composition and formation, and their variation over time; degree of the corpus's representativity of the historical reality; the critical divide between epigraphy and manuscript studies; affinities or mismatches of the disciplinary field's conventions with respect to other epigraphic traditions, which impact on cross-disciplinary mappings.