https://doi.org/10.2352/issn.2168-3204.2018.1.0.26

A Complex Database for Documentation of Cuneiform Tablet Collection Enabling Cross-Domain Queries

Jaroslav Valach, Institute of Theoretical and Applied Mechanics of AS CR; Prague; Czech Republic
Petra Štefcová, National Museum; Prague; Czech Republic
Petr Zemánek, Faculty of Arts, Charles University; Prague; Czech Republic

Abstract

The paper introduces multi-domain database for documentation of Prague's collection of cuneiform tablets. The complexity of documentation of the individual tablets is the most important innovation the database represents. It allows scholars to study the tablets in previously unachievable complexity of relations and context. Open formats and strict quality of data standards support reasonable hope of avoiding 'digital obsolescence' frequently observed in digitization projects.

Introduction

The same object can be of interest for several scientific fields. These fields apply distinct methods to study the object and their results are only seldom available to other experts from different science domains. This is an unfortunate situation as a combination of knowledge could yield new insights into the studied objects.

Bearing this in mind, one of the most important goal of the project "Analysis, description and archivation of aggregate information on properties of cultural heritage artifacts and usage of such data in restoration, conservation and research", supported by the Ministry of culture of the Czech Republic is to overcome this division based on different scientific fields and to combine several sources of data into one unified structure.

Prague's cuneiform tablet collection seems ideally suited for such a demonstration as it contains a limited number (about four hundred) of well-defined objects. The collection consists of cuneiform tablets from the Old Assyrian period excavated after WWI in Kültepe (Turkey) by Bedřich Hrozný and is almost entirely homogeneous in the sense that it comes from one location and rather narrow chronological layer (ca 19th century BCE). From the point of view of the content, majority of tablets represent correspondence among the members of the ancient Assyrian society, connected with the Old Assyrian trade network [1-2].

There are other databases of cuneiform tablets (e.g. see [3]), concentrating mostly on their textual and historical contents. There are also papers devoted to study of atomic composition of tablets [4], but truly multidisciplinary approach attempted in the presented project is a novel way, how to bring together results from several scientific disciplines to get synergies from the concerted effort.

Existence of this database intends to concentrate a substantial body of knowledge in one place to allow for yet unconceived scientific inquiry. The idea is to rationally use characterization techniques and enable researchers to formulate and verify ideas combining disciplines.

Database description

The cuneiform collection is represented by a content rich database combining knowledge from humanities and science domains as well as information important for curators.

The database record for every tablet is described in detail in Figure 1 and is composed of the following parts:

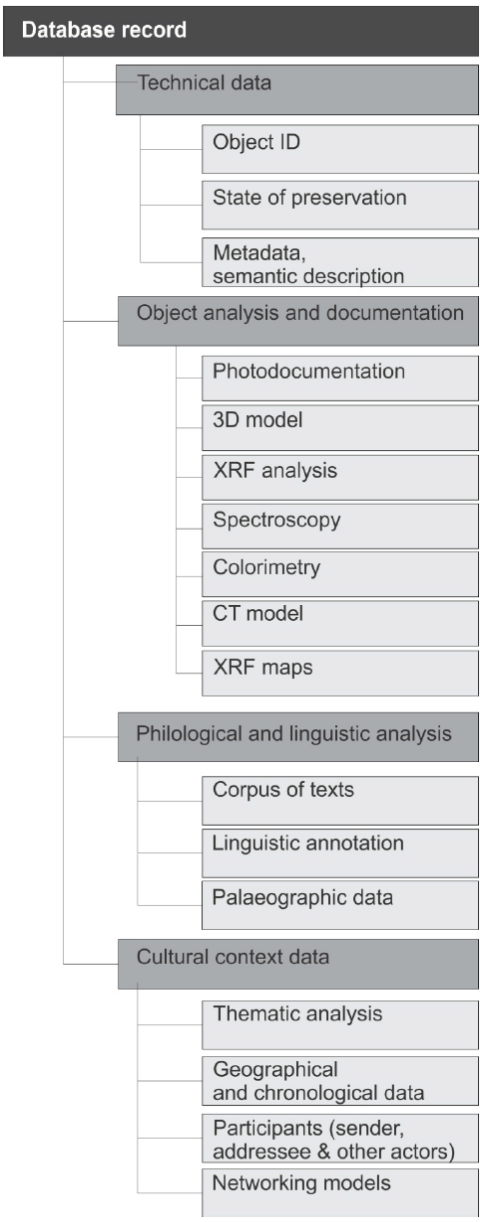

Figure 1. Structure of the records in the multi-domain database of Prague's collection of cuneiform tablets [4].

Digital models of tablets

The first part of database record contains digital 3D models of the tablets. In the presented project, two approaches to obtaining 3D model of the cuneiform tablets are employed based on optical methods and X-ray computed tomography.

True model of tablet based on X-ray micro-computed tomography (uCT) technique documents not only its surface, but also its inner structure. This allows investigation of tablet's preparation process, its inner structure, layering (placing layer of extra fine clay on the surface of the core made of not so refined material).

The drawback of CT data is that CT models do not contain any surface color information; therefore, to obtain full digital copy, optical technique has to be implemented. In order to get full 3D shape, automatized method based on photogrammetric principles - Structure from motion is implemented (for details see [5-6]). In order to record the finest surface details, method called photometric stereo [7-8] is used in own-build laboratory device.

Physical characteristics of tablet

The second part of the database record contains metadata and physical properties of tablet as acquired by spectrometry, colorimetric data and X-ray fluorescence. This type of information enables to determine chemical composition of the tablet and to determine the provenience of the clay of the tablet.

Linguistic, philological and historical description

The third part of the record unites linguistic, philological and historical description of the tablets. In comparison with other types of pottery, cuneiform tablets contain textual information that is a direct reflection of their use by those days society. The textual level is analyzed by the current corpus linguistics methods, including the linguistic annotation, paleographical analysis (variation of signs) as well as layers for further analysis and discussion. Further, data on the content of the texts is provided. This includes network analysis of the correspondence participants, containing several other layers, such as the geographical one, logistic one, historical one, etc.

All of these layers are accessible across the system. E.g., it is important to connect the graphical information (images, 3D models) with the textual level in order to analyze the damaged parts of the tablets and other types of the interpretation of the textual content. Software tools for virtual manipulation with tablet as well as rendering their visual appearance in user's chosen illumination is essential in addition to observing tablet model in virtual reality, additional data would be accessible. (Importance of capturing tablets in 3D is demonstrated on Figure 2).

Another example is that grouping of data on clay composition and paleographic analysis adds the probability of the correct identification of the scribe; this connection can be further supported by grouping of people using the services of the actual scribe. All of these combinations increase the probability of correct conclusions.

Implementation of the system

The data is stored in a distributed system on several servers, but to the user it appears as one location [9]. All the components responsible for different parts of the record are interlinked and the type of data structure determines the application suitable to manipulate the part of data.

The system division follows partners´ specialization: partner responsible for particular field and data of database, is also responsible for their maintenance and at the same time, it mirrors data of other two partners. Database would be accessible by users on-line and only requirement would be internet browser.

Use of XML format, rich metadata should assure the database is future-proof and contains computer readable and understandable information. Open format increases chances to overcome otherwise common fate of digitalized 3D model collection - becoming obsolete. Data are not only for researchers, but also mentioned for later re-use.

Motivation of database

The database enables researchers to pose new type of queries combining approach of humanities and science, combining meaning in text and physical methods for materiality in objects.

The aim of the database is the preservation of the cuneiform tablet collection and at the same time making it publicly available; it is supposed to provide an open access to it. Another important feature is its potential connectivity to other databases of related collections, which should open space for the creation of a wider image of the past reflected by the collection and a verification of conclusions reached on one collection.

Figure 2. Comparison of appearance of the same cuneiform tablet on photography (left) and with artificial rendered shading on 3D surface model (right).

Curators of exhibitions face a hard choice to select objects from their collection to be displayed, leaving in depositories usually more than 90% of collected artifacts. In addition, the similar uneasy choice applies determining what is to be shown as information alongside object. At this situation augmented reality tools can be of great help as it allows all known and in database stored data be made accessible to researcher or interested visitor.

Discussion of preliminary results

It is well known that computed tomography is instrumentation intensive technique carried out usually only in few highly specialized laboratories. It is also correct to say that curators prefer elimination of moving their collection outside promises of their institution. In order to reconcile these two conflicting requirements, it was decided to design and build transportable CT device to be deployed directly in collection The device would be able not only to capture full 3D structure of the object, but also it would be capable of creating surface maps based on x-rays fluorescence (simplified drawing of the device is presented in Figure 3).

The unique design and functionality of the device was submitted for patent protection [10].

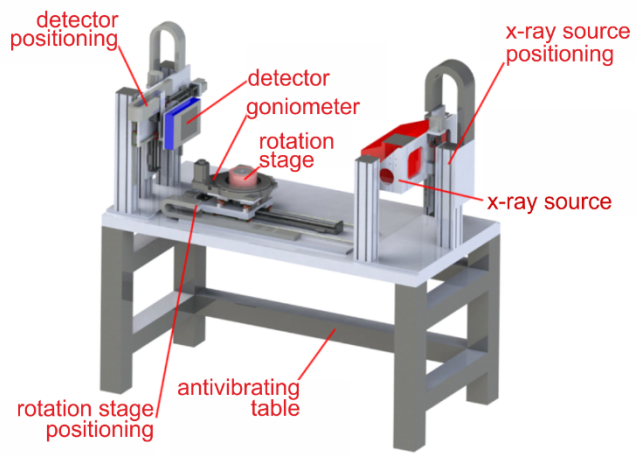

Figure 3: Schematic drawing of transportable device for CT and x-rays fluorescence mapping (authors: T. Fíla and D. Vavřík [5])

As mentioned above, the database of tablets is intended to be freely accessible. If virtual tablet is to be studied by a researcher, it is important, that it includes 3D models, because the nature of reading cuneiforms is to expose the text to lateral illumination enhancing shadows in grooves of characters. The lateral represents a necessary condition for true legibility of the cuneiform tablet. Therefore incorporating controlled virtual illumination is prerequisite for database usability.

Figure 4 represents 3D digitized 3D surface acquired by device utilizing photometric stereo method. The surface of the cuneiform tablet is shown with clearly visible fingerprints left by scribe four thousand years ago. In addition, also signs of damage can be identified in 3D model/data, thus this technique can also be used for monitoring conditions of objects' preservation.

Preliminary XRF scans show variation of atomic composition among tablets. Grouping tablets according to the composition can yield a new insight how to find connections between texts outside textual analysis. Composition outliers can indicate tablet deserving attention because of its unique properties.

Conclusions and future work

The digitalization of collection brings a great promise, as it represent an opportunity to make objects public and protect them at the same time.

Ideal digitization of collections involves a degree of precision that allows future researchers to find answers in the created digital models that were not known and asked at the time of digitization. Virtual copy, replica is similar/faithful to authentic object only in some predefined aspects. Thus, once the query about object cannot be answered studying the digital replica, it is necessary again to return back to the original authentic object. Keeping this limitation in mind, data on every cuneiform tablet are prepared as rich as possible providing for unknown queries future researchers can pose concerning the tablet and its content.3D models make available the virtual replicas not only suitable for virtual manipulation with the tablet, but they also provide a suitable tool for curators to inspect objects for any loss of material due to degradation or improper handling. Thus, the data is usable to improve collection care.

First, subsequent scanning can indicate changes undergoing in the tablet. Atomic composition can also play a role as a "signature" of its kind, allowing indicating counterfeit objects.

In the course of final project's year (2020), finalized database of Prague's cuneiform collection is to be presented in dedicated exhibition combining authentic physical objects with virtual and augmented reality – visitors would be able to see additional 'layers' of information using various types of special tools (e.g., VR googles).

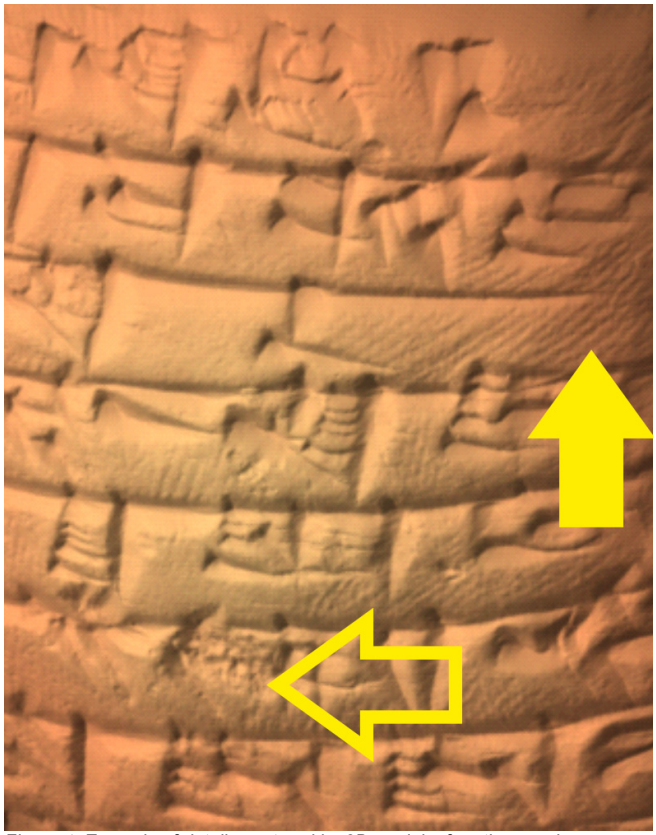

Figure 4. Example of details captured by 3D model – four thousand year old fingerprint of scribe (vertical arrow) and surface damage of the tablet (horizontal arrow).

Priority is given to standards and open formats in order to avoid threat of "digital obsolescence", which is mainly caused by insufficiency of digitization standard from the viewpoint of future requirements. Another way to avoid the obsolescence is extending impact of the data by enlarging its audience. In this case sharing the data via integrators like EUROPEANA [11] is also planned as well as close adherence to its standards of metadata description.

Acknowledgement

The presented work is carried out within project "Analysis, description and archivation of aggregate information on properties of cultural heritage artifacts and usage of such data in restoration, conservation and research" is supported by the program of applied research and development of national and cultural identity (NAKI) of the Ministry of Culture of the Czech Republic – grant No. DG16P02M022.

References

[1] Larsen, M.T.: Ancient Kanesh: A Merchant Colony in Bronze Age Anatolia. Cambridge University Press, New York (2015)

[2] Barjamović, G.: The Geography of Trade. Assyrian Colonies in Anatolia c. 1975–1725 BC and the Study of Early Interregional Networks of Exchange. In: Dercksen, J.G. (ed.), Anatolia and the Jazira during the Old Assyrian Period, Old Assyrian Archives, Studies 3, PIHANS 111. Nederlands Instituut voor Het Nabije Oosten, Leiden, 87–100, (2008).

[3] Cuneiform Database Library Initiative (https://cdli.ucla.edu/, accessed 20/3/2017).

[4] Uchida, E. , Niikuma, D. and Watanabe, R.: Regional Differences in the Chemical Composition of Cuneiform Clay Tablets. Archaeological Discovery, 3, 179-207. doi: 10.4236/ad.2015.34015, (2015).

[5] Fonstad, M.A. et al.: Topographic structure from motion: a new development in photogrammetric measurement, Earth Surface Processes and Landforms, 38 (4), 421-430, (2013).

[6] Westoby, M. J. et al.: 'Structure-from-Motion' photogrammetry: A low-cost, effective tool for geoscience applications, Geomorphology, 179, 300-314, (2012).

[7] Valach, J. et al.: Digitizing 3D surfaces of museum objects using photometric stereo-device. In Bentkowska-Kafel, A.; Murphy, O. (ed.). COSCH e-Bulletin Proceedings of the colour and space in cultural heritage session at the Denkmäler 3D conference. From low-cost to high-tech. 3D-documentation in archaeology and monument preservation Vol. 1. Dortmund: The LWL Industrial Museum Zeche Zollern, ISSN 2409-9503, (2014).

[8] Valach J. et al.: 3D Digitization of Selected Collection Item Using Photometric Stereo, In: Digital Cultural Heritage (ed. M. Ioannides), LNCS 10605, ISBN 978-3-319-75825-1 (2018) (in print).

[9] Valach, J. et al.: Multidisciplinary information system of Assyrian cuneiform tablets enhancing new research possibilities via heterogeneous data in records, International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XLII-2/W5, 703-706, E-ISSN 2194-9034 (2017).

[10] Fíla T., Vavřík D.: Device for scanning of sculptures, Czech patent application PV 2018-60, Feb. 2018.

[11] Doerr, M. et al.: The EUROPEANA data model (EDM). In: Proceedings of the 76th IFLA General Conference and Assembly, Gothenburg, Sweden, p.9 (2010) (https://www.ifla.org/past-wlic/2010/149-doerr-en.pdf, accessed 20/3/2017).

Authors Biography

Jaroslav Valach has a background in materials science and physics and works in application of optical and physical methods in study of cultural heritage objects.

Petra Štefcová studied chemistry of corrosive processes at the Institute of chemical technology; she specializes in preventive conservation and collection care in the Czech National Museum. She has carried out many research project dedicated to preservation of collections.

Petr Zemánek studied the philology of Arabic and other Near Eastern languages (ancient and modern). His specialization is Arabic and Semitic corpus linguistics.