Gerfrid G.W. Müller and Daniel Schwemer

13 Hethitologie-Portal Mainz (HPM). A Digital Infrastructure for Hittitology and Related Fields in Ancient Near Eastern Studies

Abstract: The Hethitologie-Portal Mainz is a not-for-profit, open-access digital infrastructure for Hittitology and related fields of research in Ancient Near Eastern studies. HPM, which was first established in 2001, gives access to an array of interconnected research documents, including critical editions of Hittite cuneiform texts, catalogues, bibliographies, onomastic databases as well as media archives with digital photos, drawings, and 3D models. The HPM community has been constantly growing over the past years and currently comprises more than fifty creators of contents and approximately 3,000 individual human users. User statistics and feedback by peers show that HPM has become essential to Hittitological research. Its digital strategy favours open-source, widespread software and standardized, well-documented data formats in order to ensure long-term sustainability. The absence of low-level, permanent funding opportunities for digital infrastructures in the Humanities in Germany is one of the challenges faced by HPM.

Keywords: cuneiform, Hittite, edition, bibliography, digital humanities

13.1 Remit and Unique Proposition

The Hethitologie-Portal Mainz1 (Figure 13.1) is one of the main digital infrastructures in Ancient Near Eastern studies. HPM’s specific remit is the study of the cultures of ancient Anatolia (Turkey), in particular those of the Late Bronze Age (second half of the second millennium BCE). The kingdom of the Hittites (17th–13th cent. BCE) plays a prominent role in this period of ancient Near Eastern history. Thus the epigraphic finds from Hittite sites, not least the continuously growing body of cuneiform tablets and fragments (current count approximately 33,000, including unpublished texts) are at the centre of HPM (cf. generally Schwemer, 2017).

1 HPM [http://hethiter.net].
The move of cuneiform studies as a whole into the online digital age began in the late 1990s in a number of different initiatives, most of which focused on one sub-corpus of cuneiform texts defined by parameters such as provenance, date, language, or text genre; usually these restricted bodies of cuneiform texts also correspond to sub-disciplines of Ancient Near Eastern studies. Hence HPM, the digital infrastructure of Hittitology, forms part of a varied and international landscape of online corpora of cuneiform texts.²

Within the range of digital cuneiform online databases, HPM is unique not only with regard to its focus on Hittitology. It also stands out due to its combination of complex philological text editions with related databases including catalogues of epigraphic objects, bibliographies, onomastic indexes, and gazetteers as well as media databases of drawings, photos, and 3D models.

² Most importantly: Ebla Digital Archives (EbDA: Eblaite, Early Dynastic III period; [http://ebda.cnr.it/]); Database of Neo-Sumerian Texts (BDTNS: Sumerian, archival texts, Ur III-period; [http://sefarad.filol.csic.es]); The Electronic Text Corpus of Sumerian Literature (ETCSL: Sumerian, literary texts; [http://etcsl.orinst.ox.ac.uk]); Archives babyloniennes XXe–XVIIe siècles av. J.-C. (ARCHIBAB: Akkadian, archival texts, Old Babylonian period; [http://www.archibab.fr]); The Neo-Babylonian Cuneiform Corpus (Nabucco: Akkadian, archival texts, first-millennium Babylonia; [http://nabucco.arts.kuleuven.be]); Sources of Early Akkadian Literature (SEAL: Akkadian, literary texts, 3rd and 2nd millennium BCE; [http://www.seal.uni-leipzig.de]). In contrast to these sites, the Open Richly Annotated Cuneiform Corpus (ORACC; [http://oracc.museum.upenn.edu]) is not restricted by language, provenance or text genre, but hosts a range of independent ‘projects’ with various editions of Sumerian and Akkadian texts. A complementary tool to these period- or genre-specific corpora is the database of the Cuneiform Digital Library Initiative (CDLI: [https://cdli.ucla.edu]), which aims to provide a complete catalogue of cuneiform tablets and fragments. The CDLI site offers a wide variety of digital tools and materials for Assyriology and has become an essential platform for the publication of photos of cuneiform texts by museums around the world.
13.2 **Objectives: Innovation, Collaboration, Acceleration**

For all its dynamic development, the basic objectives of HPM have not changed since its first inception in 2001. They flow from HPM’s design as a digital infrastructure serving the field of Hittitology as a sub-discipline of cuneiform studies.

HPM provides sustainable online access to primary sources in the form of critical editions of texts, transliterations of individual cuneiform manuscripts, and representations of the archaeological objects on which these texts are inscribed, most commonly (fragments of) clay tablets; these representations include digital images (technical drawings and photos) as well as 3D models. HPM’s goal is to present the sources in a form that is compliant with the academic standards of Hittitology. The projects associated with HPM also take on an active role in the further development of these academic standards.

In addition to the presentation of sources, HPM provides sustainable online access to research documents of various types, especially catalogue databases, bibliographies, and onomastic indexes. These tools and materials include legacy data collections whose accessibility is preserved by hosting them on HPM.

HPM strives to develop dynamic, digital interconnections between the primary sources, research documents and data collections that form part of the infrastructure. From a current user’s perspective, the transition between various components hosted on HPM is already fluid. Frequently, Hittitologists are able to move seamlessly between editions, catalogues, bibliographies, and media databases.

Today, a considerable number of Hittitologists, including junior and postdoctoral researchers, present their data collections on HPM. Thus they reduce the burden of routine tasks for the individual researcher and help to avoid duplicating efforts. The amount of collective research time saved by the digital publication of S. Košak’s *Konkordanz der hethitischen Keilschrifttafeln*³ and D. Groddek’s *Groddeks Liste*⁴ is immeasurable; these are only two game changers that have fundamentally transformed Hittitology’s working methods. By functioning in this way as a digital platform for individual scholars and research projects of any size, HPM aims to foster collaboration and accelerate research procedures in Hittitology.

The digital medium, which enables authors and creators to update their research data, is especially suitable for any kind of catalogue or data collection requiring growth and modification as knowledge advances and the available sources constantly increase (a typical feature of cuneiform studies). This potential for openness of the digital medium, in contrast to print, also offers an appropriate framework for encouraging the publication of less definitive research efforts and collections of

---

³ [http://hethiter.net/hetkonk]
⁴ [http://hethiter.net/grodlist]
raw data. As a matter of principle, the competence to make editorial changes in any document on HPM stays with its creator(s) if they do not decide otherwise.

Last but not least, HPM has also become a space for developing innovative digital research methods and strategies in Hittitology; e.g., the metrological analysis of digitized cuneiform tablets (3D models) as a revolutionary method of palaeography and script classification\(^5\), or the development of a fully automated digital annotation of transliterated Hittite texts with lexical and morphological metadata\(^6\).

### 13.3 History and Status Quo 2017

HPM was conceived and created by G. Wilhelm in cooperation with G.G.W. Müller from 2001 onwards in a collaboration between the Academy of Sciences and Literature, Mainz (Academy Programme project *Hethitische Forschungen*, 1961–2015; Wilhelm, 2008; 2015), and Ancient Near Eastern studies at Würzburg University. Initial funding was provided by the Deutsche Forschungsgemeinschaft in 2001–2007 within the framework of the project *Informationsinfrastruktur für digitale Publikation keilschriftlicher Staatsverträge der Hethiter und für darauf bezogene netzbasierte Forschungskooperation* (Würzburg University; Wilhelm, 2013; Müller & Wilhelm, 2015).

In the years 2008–2015, HPM was continuously expanded within the framework of the project *Hethitische Forschungen*, directed by Wilhelm at the Mainz Academy. In that period of time, the most significant extension of the text editions presented on HPM was realized by the following research projects, all funded by the Deutsche Forschungsgemeinschaft: *Digitale Publikation hethitischer Texte: Die Beschwörungsrituale der Hethiter (CTH 390–500)* (Mainz University, 2010–2017; director: D. Prechel); *Hethitische mythologische Texte* (Marburg University, 2005–2008; director: E. Rieken); *Sprachlich-philologische Bearbeitung und digitale Edition der Hymnen und Gebete in hethitischer Sprache (CTH 371–389)* (Marburg University, 2011–2014; director: E. Rieken).

Since 2016, HPM has been an essential component of the digital strategy and publication plan of the Academy Programme project *Das Corpus der hethitischen Festrituale: staatliche Verwaltung des Kultwesens im spätbronzezeitlichen Anatolien*\(^7\) at the Mainz Academy (2016–2036; directors: E. Rieken and D. Schwemer). HPM as such, however, has no permanent funding arrangement.

Internal user statistics show that the various components of HPM have become an essential everyday tool of Hittitological research. HPM has approximately 3,000

---

\(^5\) [http://www.cuneiform.de].

\(^6\) A prototype of this tool is currently tested within the project HFR – *Das Corpus der hethitischen Festrituale*, and will be fully operational in 2019.

\(^7\) [http://www.adwmainz.de/projekte/corpus-der-hethitischen-festrituale].
individual human users per annum. About half of HPM’s users are based in Germany, ca. 1,000 users are based outside Germany but still within Europe and ca. 500 are outside Europe; a distribution corresponding to the locations of Hittitology at universities worldwide. The sites of HPM have approximately 5.5 million accesses by individual human users per annum. This high number confirms that today “studying Hittite is unthinkable without the Portal” (de Roos, 2007, p. 187).

13.4 Organization: A Network of Researchers and Projects

HPM has been recognized by the Deutsche Forschungsgemeinschaft as a research infrastructure. It is a not-for-profit, open-access host and platform. All its content is openly accessible upon publication (“gold open access”). HPM also serves as a gateway to other websites offering research content in Hittitology and related fields; however, the maintenance of (ever changing) external links poses challenges.

HPM is led by a steering committee that, in addition to the two present authors, currently comprises three further specialists in Hittitology (Prechel; Rieken; Wilhelm). The steering committee is tasked with the strategic and technical development of the HPM site. Most importantly, the members of the steering committee liaise with colleagues who create and present content on the HPM site. The committee is assisted in its work by an international scientific advisory board whose members are senior academics and leaders of Hittitological research in their country.

HPM is a platform that offers research projects and individual researchers a sustainable space for presenting and interconnecting digital content such as text editions, media, and data collections. At present, more than fifty colleagues are creating content on HPM either as individuals or as project researchers, with contributions ranging from text editions to bibliographies and geographical databases. The long-term Academy-programme projects Hethitische Forschungen (up to 2015) and Corpus der hethitischen Festrituale (2016–2036) use HPM as a digital publication platform. Due to their extensive funding periods and the role of their researchers on HPM’s steering committee, these two projects have been essential for the maintenance and further development of HPM; this situation will not change for the foreseeable future.

As a platform and portal, HPM considers it very important that the creators of contents and their sources of funding are clearly identifiable and visible on the website, and that clear information is provided on how to refer to and quote from the research documents it hosts. It has been discussed for some years in the HPM community to what extent and in which form older, outdated versions of research documents should be archived and kept available to all users. Indications are that HPM will now move to a public archive solution that provides access to previous

---

8 See [http://risources.dfg.de/detail/RI_00500_de.html].
versions of some components (e.g., text editions and the Konkordanz) if significant changes have occurred. The decision on how their content is presented will, however, always stay with the relevant creators.

Institutionally, HPM is located at the Mainz Academy of Sciences and Literature (Department Hethitologie-Archiv) and also has office space at the Ancient Near Eastern studies section of the Department for Ancient Cultures at Würzburg University. For data storage and retrieval, it uses local servers at the Academy as well as servers of the computing centres of Würzburg University and Mainz University.

13.5 Digital Components and Concepts

13.5.1 Components of HPM

The heart of HPM is the Konkordanz der hethitischen Keilschrifttexte created by Košak. It lists all known Hittite fragments and tablets with their place of discovery, date, joins, and bibliographical references. In addition to its own search interface, the Konkordanz can be accessed content-wise through CTH, the Catalogue of the Texts of the Hittites (the digital continuation of Laroche, 1971 and supplements). From the Konkordanz one can also reach the Joinskizzen, which show the placement of the fragments within a fragmented clay tablet, as well as the text editions on HPM.

The reconstruction and online publication of the Hittite texts is one of the main objectives of HPM. In addition to Hittite texts (see section 3 for the most extensive text groups currently available), HPM also includes Old Assyrian texts from Anatolia by K. Hecker and, in the future, documents from the Hurrian cultural area (Nuzi).

For the exploration of research literature, there are several bibliographies that have a different scope than the manuscript-based Konkordanz and the passage-based Groddeks Liste. The comprehensive Hethitische Bibliographie was started by J. Součková (Prague), Müller and Wilhelm for older literature, and has now been maintained for many years by M. Marazzi (Naples) in cooperation with several other colleagues. It is complemented by the Systematische Bibliographie (mainly supervised by Součková) and a bibliography of Hittite lexemes (led by Marazzi and N. Bolatti Guzzo in collaboration with other colleagues).

Various onomastic databases for personal names, place names and divine names are, or will, soon be available and can be used for indexing the text editions and other data sets on HPM.

The media archive contains approximately 70,000 photos of Hittite texts, as well as a photo collection of Alalakh texts and photos of Old Assyrian texts. Since 2017, HPM gives access to more than 2,000 3D scans of cuneiform tablets. Viewing of the 3D models is enabled by the programme Cuneiform WebGLViewer (Figure 13.2),

Unauthenticated
Download Date | 5/25/19 9:25 PM
which was created by D. Fisseler for HPM, and allows an exact examination and measurement of the surface.⁹

Figure 13.2: The WebGLViewer of HPM allows the collation of cuneiform tablets in the web browser and provides several tools for measuring and enhancement

Finally, HPM offers various services, including downloadable fonts and e-books (Studien zu den Boğazköy-Texten; HPM – Materialien) as well as some general information on Hittite history and culture.

13.5.2 Open Standards and Widespread Open-Source Software

When, in 2000, Wilhelm and Müller first discussed the development of an information infrastructure for the digital publication of Hittite texts, it was clear from the start that only international standards with the widest possible dissemination could be used for an enterprise of this type.

From the beginning, preference was given to open-source software. As with standards, wide dissemination and large user numbers were important criteria for the choice of software solutions in order to safeguard stability, sustained compatibility and continuous further development. Too much work and effort had been spent on the transitions from Apple II to Atari to DOS and Windows, from 7-bit to 8-bit data format, from the inadequate ASCII text editor in the DOS environment (via various

⁹ [http://www.cuneiform.de/fortschritte/webviewer.html].
adaptable programmes such as Signum and Word Perfect) to Word for Windows, which still did not guarantee compatibility and a smooth file exchange. For a project like HPM, however, it was crucial to store and provide its data in a documented, open-source format and thus be independent not only of proprietary software, but also of specific computer platforms.

After twenty years of such struggles, the dynamics of the World Wide Web created an adaptation and standardization pressure that paved the way for creating platforms like HPM. When HPM went online, most of the transliteration letters for ancient Near Eastern languages were already available and could be displayed with some universal fonts (e.g., Arial Unicode, Code2000), which had to be installed in the operating system. For others (e.g., half brackets) similar characters had to be used and adapted in the Semiramis Unicode font that was developed for HPM. In 2008, in the course of an extension of Unicode, the half brackets and some ancient Egyptian transcription characters were given their own code point, prompting an adaptation in HPM. Today the encoding of the characters can be regarded as stable in the long term.

Based on Linux Libertine and Linux Biolinum, HPM created its own new set of fonts, which will be used for publications in print and online. Semiramis Unicode exists in a version 3, but is deprecated and will no longer be updated. In addition, S. Vanséveren provided a Unicode font for Hittite cuneiform and G. Anders for Luwian hieroglyphs.

Above the level of single characters, every document has to exist in an intelligible and well-documented format: as a kind of XML to describe textual data, as TIFF for photographs, SVG for drawings, or PLY for 3D data. For each of these data formats plenty of software is available; they are well documented and will permit the creation of new software in case this should be required. At HPM, this does not necessarily imply that all data will be made available to external users in these formats, mainly due to copyright issues.

13.5.3 Continuity Online: Development and Experiences

The functionality and presentation of web pages with HTML has remained largely stable. The diversity of displays and media in the web has led to a growing awareness of the separation of content and form through multiple media. HPM is currently undergoing a revision in order to remove display-oriented elements deprecated by HTML 5.

The development of CSS was less consistent and the implementation of standards in web browsers sluggish. A careful use of HTML and CSS has low maintenance requirements and ensures sustainability. The implementation of the new media features of HTML 5 will further increase HPM’s longevity.

Most websites on HPM also contain a dynamic component, e.g., the output of a database. The choice of technology for dynamic websites must be based on long-term
functionality. The programme code should be designed in a simple and modular way with little nesting in order to facilitate the acquisition of maintenance competence. In the humanities, but especially in “small disciplines” like Hittitology, highly qualified maintenance specialists may not be available or affordable at all times.

The uniformity and popularity of Content Management Systems (CMS) on the Internet, and the complexity of their numerous features, produce security vulnerabilities that are a popular gateway for hacker attacks. The use of a CMS therefore requires permanent and professional system management. The effort required for the integration of additional functions into a CMS causes further costs. The dependency on one CMS also limits flexibility and may impede a move from one host institution to another. For all these reasons, HPM does not use a CMS. The strength of a CMS as a multi-user system that ensures front-end homogeneity is less significant for HPM, where the visibility of the individual researchers and projects would even be undermined by a corporate approach to web design.

HPM uses PHP in conjunction with an Apache server and MySQL database. This is the standard configuration on university servers and thus ensures efficient hardware provisioning and data backup. With the introduction of PHP 7, revisions have now become inevitable. The outdated POSIX engine for regular expressions will be replaced by PCRE; also the deprecated connection to the MySQL database has to be replaced. These changes produce some uncertainty with regard to long-term maintenance.

Another instability concerns the presentation of media. The first photo viewer of HPM was written in Java. It had to be installed in web browsers, causing countless help requests from the HPM user community; also the performance of Java (speed) was unsatisfactory. The Flash viewer as its successor does a good job, but is notorious for its security flaws and will now be replaced by an HTML 5 viewer based on the HTML canvas implemented in the browsers. This promises long-term stability. JavaScript has also made great strides in the standardization process, so that a careful use of JavaScript can be considered sustainable. JavaScript is now so powerful that even 3D objects can be displayed in HTML 5 with its libraries.

### 13.5.4 Tools for Scholars, not Scholars for Tools

One of HPM’s basic goals is to collaborate with scholars without requiring them to leave their accustomed digital workflows and work environment. HPM offers simple avenues for accessing and processing data and manuscripts, which may have been created for a wide variety of purposes. In preparing materials for publication on HPM (manual), routine work should be avoided, and the revision of existing data kept to a minimum. For example, the manuscript of *Groddeks Liste* consists of several thousand pages arranged consistently by the author. By following a few rules, this document can be converted into a searchable online database, and an XML version is created automatically.
In other cases with more complicated structures, meaningfully named styles are added into the manuscript. In OpenOffice/LibreOffice this can be done in Fill Format Mode comparable to using a text marker on paper (Figure 13.3). The underlying XML structure can then be evaluated automatically, and the document can be converted into any XML format. In principle, the same method is used for creating the text editions on HPM. In the early years, HPM had developed a proprietary XML editor. For various applications, however, OpenOffice, which introduced XML into the Office sector, proved to be the best solution. It allows the quick definition of tags, records revisions and permits the use of foot- and endnotes, all as XML that can be processed and read in any text programme (Figure 13.4).

Figure 13.3: Hittite text with mark-up in LibreOffice

Figure 13.4: An example for the reuse of an older manuscript as database by tagging it with styles
To facilitate and speed up work, Müller developed the programme Simtex for the
digitization of larger bodies of texts. This simple input method generates XML from
text files automatically, which is then further processed as an OpenDocument.

13.5.5 Connecting Data

HPM makes data collections and tools for research available online. A particular
advantage of the digital medium is the possibility of linking data sets, which often
results in new findings. As far as possible, information should be collected only once
and then be linked to other information, whether it is word forms, datings, locations,
place names, persons, or inscribed objects. This does not exclude competing projects,
but the goal is to preserve existing data collections and develop them further in a
continuous, joint effort.

The data must be arranged in such a way that any correction has to be made in
only one place, and is then available at all points of reference automatically. This
seems trivial for an Internet project, but does indeed pose a challenge due to the
heterogeneity of the individual projects that collaborate with HPM.

The principle of collecting any given information in only one place has further
important implications, not least with regard to information that has to be retrieved
multiple times. Thus a text edition should be limited to the text and not contain a
lexical, morphological and syntactic annotation. An integrated annotation not only
requires the repeated annotation of the same word form; it also impedes future
changes. If the analytic annotation is separate from and, at the same time, linked
to the text editions, the work process is more efficient, and it is easier to preserve
consistency over time.¹⁰

An example of the efficient arrangement of information is the Joinskizzen
component, which provides a documentation of the reconstruction of fragmented
cuneiform tablets in drawings. As a stable reference, the join sketches include the
inventory number of the individual fragments in the drawing. But the citation of
fragments in the literature usually is by publication number. The join sketches were
originally drawn in ink and published in print. Later they were digitized for the
online version. The drawings were transferred to OpenOffice Draw, which uses a
vector graphics format. This is automatically processed and the individual parts can
be addressed. When the user is loading a join sketch, a link to the relevant photos
is automatically generated. The publication number is automatically queried in the
Konkordanz and appears on mouse over (Figure 13.5).

¹⁰ The fact that in the Hurrian language no fewer than five new nominal cases have been discovered
over the last thirty years may serve as a warning against hard-coding every analysis in the text edition.
Figure 13.5: On mouse-over the SVG join sketch displays the publication number of the fragment and a link to the relevant photos.

In the future, HPM intends to extend the offer of such automatic links. Sometimes there are, however, compatibility issues: for example the *Konkordanz* and *Groddeks Liste* do not always follow the same bibliographical standards for referencing cuneiform manuscripts by publication.

### 13.6 Outlook: Expansion, Connectivity, Sustainability

All components of HPM – catalogues, bibliographies, onomastica, text editions, media archives, and services – are engaged in a continuous process of further development and expansion. The single most important new component currently under preparation is the *Thesaurus Linguarum Hethaeorum digitalis* (TLHdig), a tool that is not conceived as a dictionary, but as a basic, searchable database of all cuneiform manuscripts from Hittite tablet collections in transliteration. In this context, HPM is planning to develop a “creator interface” that will allow users to actively contribute newly discovered texts to the growing TLHdig. This will add a new dimension to HPM as a truly collaborative digital infrastructure.
External connectivity is an important challenge for the future, especially when more Hittitological research publications will become available in a digital format. This particularly concerns the Hittitological philological dictionaries, which currently have only a limited or no digital presence, and the various databases of excavations in the Anatolian cultural area, only some of which are accessible online. Also the creation of dynamic interconnections (rather than static links) between the various digital corpora of cuneiform texts (see fn. 2) is an important task for the future.

Expansion and connectivity must be underpinned by sustainability in order to ensure the long-term availability of HPM. The structure of the existing web pages must be further developed to allow easy maintenance. Not only the data themselves should be present in documents (XML) that are self-explanatory, but also the programme logic should be stored in forms that allow automatic reconfiguration. This is likely to involve an increased use of XML technologies (with XSL and XML databases), but such changes should be approached with the necessary caution and employ only common, wide-spread and proven technologies.

**Bibliography**


