

# Textdatenbank und Wörterbuch des Klassischen Maya

Arbeitsstelle der Nordrhein-Westfälischen Akademie der Wissenschaften und der Künste  
an der Rheinischen Friedrich-Wilhelms-Universität Bonn



ISSN 2366-5556

PROJECT REPORT 5 | DE



Published 13 Mar 2018

DOI: 10.20376/IDIOM-23665556.18.pr005.de

## Arbeitsbericht 2017

Nikolai Grube<sup>1</sup>, Christian Prager<sup>1</sup>, Katja Diederichs<sup>1</sup>, Sven Gronemeyer<sup>1,2</sup>, Antje Grothe<sup>1</sup>, Céline Tamignaux<sup>1</sup>, Elisabeth Wagner<sup>1</sup>, Maximilian Brodhun<sup>3</sup> & Franziska Diehr<sup>3</sup>

<sup>1</sup> Rheinische Friedrich-Wilhelms-Universität, Bonn

<sup>2</sup> La Trobe University, Melbourne

<sup>3</sup> Niedersächsische Staats- und Universitätsbibliothek, Göttingen

### Digitaler Zeichenkatalog

Wir entwickelten einen digitalen Zeichenkatalog für die Hieroglyphenschrift der Maya, der als Inventar aller Grapheme zur Identifikation der in den Textzeugen verwendeten Zeichen dient. Für die Arbeit mit diesem noch nicht vollständig entzifferten logo-syllabischen Schriftsystem stellt er ein unverzichtbares Werkzeug dar. Die Identifikation und Klassifikation der Zeichen ist eine Herausforderung, da sie etliche Schreibvarianten aufweisen und mehrere Funktionen annehmen können, z.B. als Wort- oder als Silbenzeichen. Über die Entzifferung der ca. 1000 Zeichen herrscht ein reger Forschungsdiskurs, aus dem verschiedene Hypothesen zur sprachlichen Lesung einzelner Zeichen hervorgehen.

Die Anforderungen an die Analyse dieses komplexen Schriftsystems sowie die Integration der sich ständig verändernden Forschungsergebnisse in den Zeichenkatalog erfordern eine flexible Datenmodellierung. Diese muss auf etwaige Änderungen reagieren können und die dokumentierten Informationen müssen nachvollziehbar sowie überprüfbar sein. Wir wählten einen ontologisch-basierten Modellierungsansatz auf Grundlage von CIDOC CRM<sup>1</sup> und GOLD<sup>2</sup>. Um die semantischen Relationen zwischen den Entitäten optimal zu repräsentieren, wurde das Datenmodell in RDF realisiert. Durch die Nutzung von Graphentechnologie sind semantische Abfragen bei der Nutzung des Zeichenkatalogs möglich.

Eine besondere Herausforderung bei der Entzifferung und Analyse der Mayaschrift liegt in der Einbeziehung der vorliegenden Lesungshypothesen. Da für ein Zeichen mehrere, zunächst als gleichwertig zu betrachtende Lesungen vorliegen, wollen wir diese nicht nur dokumentieren, sondern auch qualitativ bewerten. Basierend auf der Forschungsliteratur zur Mayaschrift entwickelten wir Kriterien-Sets, die sich u.a. am sprachlichen Nutzungskontext (z.B. korrekte Wortart, plausibler Text-

---

<sup>1</sup> CIDOC Conceptual Reference Model <http://www.cidoc-crm.org/>.

<sup>2</sup> General Ontology for Linguistic Description <http://linguistics-ontology.org/>.

Bild-Bezug etc.) orientieren (vgl. Kelley 1976). Pro Set sind die Kriterien mittels Aussagenlogik so miteinander verbunden, dass sich je nach Kombination eine entsprechende Konfidenzstufe ergibt.

Ziel ist es, die bisher bekannten Texte und deren Textträger in einem maschinenlesbaren Korpus zu erschließen und darauf basierend ein Wörterbuch zu erstellen, das den gesamten Sprachschatz und dessen Verwendung der Schrift abbildet. Der digitale Zeichenkatalog agiert als Hilfsmittel zur Erstellung des Textkorpus, welches in TEI/XML kodiert wird. Der Text besteht jedoch nicht aus phonemisch-transliterierten Werten, sondern jedes erfasste Zeichen wird mit einer Referenz auf die URI der entsprechenden Instanz im Zeichenkatalog kodiert. Einmal erstellt, bleibt das Korpus eine stabile, sich nicht verändernde Datengrundlage. Alle Erkenntnisse über Lesung der Schrift und Grammatik der Sprache werden außerhalb des Korpus erfasst.

Die ermittelte Konfidenzstufe einer Lesungshypothese gibt bei der linguistischen Analyse der Texte Hilfestellung bei der Vorauswahl der zu untersuchenden Lesungen. Die Prüfung einer Lesung im Korpus kann neue oder weitere Kriterien liefern, so dass ihre Konfidenz steigt oder sinkt. Die ermittelten Kriterien können problemlos im Zeichenkatalog nachgetragen werden und dementsprechend passt sich die Konfidenzstufe der Lesungshypothese an. Damit tragen wir insbesondere der Anforderung Rechnung, Entzifferungen, die erst durch linguistische Analyse und Untersuchung des Textkorpus entstehen können, nachvollziehbar zu dokumentieren.

Ziel der graph-basierten Modellierung des Zeichenkatalogs in Kombination mit Textkorpus und linguistischer Analyse ist es, gesicherte Aussagen über die Lesung von Zeichen der Mayaschrift zu treffen und optimalerweise auch neue Entzifferungen für noch nicht entschlüsselte Zeichen vorlegen zu können (vgl. Diehr et al. 2017).

Die Inventarisierung der Zeichen mittels des digitalen Katalogs befindet sich derzeit in Bearbeitung und wird voraussichtlich Mitte des Jahres 2018 abgeschlossen werden. Neben der Dateneingabe und der Erstellung einer Konkordanz mit anderen Zeichenkatalogen arbeiten wir derzeit auch an der Erstellung von Vektorzeichnungen der Graphen und ihren Varianten. Zeichen, die bisher noch nicht in Thompsons Zeichenkatalog berücksichtigt sind, werden neu klassifiziert und in der Zeichen- und Graphendatenbank eingegeben. Wir rechnen mit rund 1000 verschiedenen Zeichen und rund 3000 Graphen. Der Katalog wird dann auf unserem Projektportal veröffentlicht und die RDF-Daten über einen SPARQL-Endpoint zugänglich gemacht. Weiterhin werden die Daten im TextGrid Repository veröffentlicht, wo sie mittels einer OAI-PMH Schnittstelle abrufbar sind. Die Dokumentation des digitalen TWKM-Zeichenkatalogs ist unter <http://idiom-projekt.de/catalogue> erreichbar.

Derzeit arbeiten wir an einem umfangreichen Aufsatz mit dem Arbeitstitel "Standards for Maya Epigraphic Analysis I: Principles of Maya Graphemics", der im Laufe des Jahres auf der Projektwebseite veröffentlicht werden soll. Dieser wird eine ausführliche Besprechung der Graphvariantenbildung, der Konfidenzstufen und -kriterien sowie weiterer, hier nur ansatzweise betrachteten, Forschungsergebnisse enthalten. Damit wird dieser Aufsatz auch insbesondere für die Fachcommunity der Mayaschriftforschung von Interesse sein.

## TEI-Parser

Die Transkription der Texte ist eine sehr komplexe Aufgabe, bei der viele Schritte beachtet werden müssen. Innerhalb der TEI-Datei werden alle Referenzen zu den Zeichen nicht mit ihrem konkreten Lautwert oder ihrer Katalognummer aus dem Zeichenkatalog angegeben, sondern durch die Referenz auf eine TextGrid-URI. Daraus ergibt sich der Umstand, dass bei der Transkription stets die TextGrid-URI zu einem Zeichen herausgesucht werden muss, um diese im TEI-Dokument anzugeben. Um diesen

Aufwand gering zu halten wurde mit der Entwicklung eines TEI-Parsers begonnen. Dieser generiert aus einem Transkriptions-Code die entsprechende TEI-Struktur gemäß des entwickelten Schemas.

Der Parser wurde als Plug-In für das TextGrid-Lab konzipiert. Dadurch besteht stets der Schreibzugriff auf den internen Speicher des Projekts.

Durch die Nutzung des Parsers wird die Produktivität der Erzeugung der Transkriptionen erheblich erhöht. Einige Elemente des Schemas müssen jedoch manuell eingefügt werden. Ein solches Element ist zum Beispiel die Auszeichnung einer Beschädigung auf dem Textträger.

Der Parser übernimmt die Angabe der Transkription in: *12st.[\*5009st:\*128st:679st]* und generiert daraus folgenden Code:

```
<ab xml:id="pB1" type="glyph-block">
  <!-- 12st.[*5009st:*128st:679st] -->
  <g xml:id="pB1G1" ref="" n="12st" rend="left_beside" corresp="pB1S1"/>
  <damage agent="fracture" degree="0.75" quantity="3" unit="g">
    <supplied reason="damage" evidence="external" precision="high" ana="#pB1_note1">
      <note xml:id="pB1_note1" resp="SG">The same exonym also appears on DPL: P. 19,
P3 and DPL St. 15, J1, although in different glyphic renditions.
      <rs ref="textgrid:1234aa">Dos Pilas, Panel 19</rs>
      <rs ref="textgrid:5678bb">Dos Pilas, Stela 15</rs>
    </note>
    <seg xml:id="pB1S1" type="glyph-group" rend="right_beside" corresp="#pB1G1">
      <g xml:id="pB1G2" ref="" n="5009st" rend="above" corresp="#pB1G3"/>
      <g xml:id="pB1G3" ref="" n="128st" rend="above" corresp="#pB1G4"/>
      <g xml:id="pB1G4" ref="" n="679st" rend="beneath" corresp="#pB1G3"/>
    </seg>
  </supplied>
</damage>
</ab>
```

Bei dem Parser handelt es sich nicht um einen Editor, daher können Dateien durch diesen lediglich erstellt werden. Ein Öffnen und nachträgliches Bearbeiten ist daher nicht möglich. In zukünftigen Arbeiten könnte der Parser durch größeren Funktionsumfang zu einem Editor ausgebaut werden, um Dateien nicht nur erstellen, sondern auch in der Erstellungsansicht bearbeiten zu können.

## Entwicklung des TEI-Schemas

Die Entwicklung des TEI-Schemas besteht aus mehreren Arbeitspaketen, die einzelne, teils miteinander verzahnte Aspekte der Textauszeichnung behandeln. Hierbei wurde darauf geachtet, zuerst diejenigen Strukturen des TEI/XML zu erstellen, die es ermöglichen aus dem Zeichenkatalog heraus einen Text in seiner intendierten Lesefolge maschinenlesbar zu machen. Während diese Essentialia am Material getestet wurden und Korrekturschleifen in der Schemadefinition vorgenommen wurden, konnten parallel andere Arbeitspakete in Angriff genommen werden, wie etwa der sogenannte *Header* mit den Metadaten der Datei, die Auszeichnung beschädigter Textstellen oder gestalterische Aspekte der Textausführung.

### AP 01: Transkription und semantische Textstruktur

Im ersten Arbeitspaket behandelten wir die semantische Struktur der Texte. Ziel war es, den Text in seiner logischen Abfolge, also der Lesereihenfolge, in TEI/XML wiederzugeben. Parallel dazu steht die topografische Textanordnung, die beschreibt wo der Text auf dem Textträger verortet ist (s.u. AP 02).

Im AP 01 wurden folgende Fragestellungen beantwortet und Ziele erörtert: Wie können die Hieroglyphen in XML kodiert werden? Wie ist die Lesefolge in einem Textfeld und wie sind Graphvarianten im Hieroglyphenblock angeordnet? Deliverables des AP 01 waren die Erzeugung erster Beispieltexthe (von einfach bis komplex), so dass alle Textstruktur-Phänomene der Inschriften des Klassischen Maya berücksichtigt wurden. Die aus dem AP 01 hervorgegangene Strukturauszeichnung ist für alle Texte anwendbar.

### AP 02: Topographische Textanordnung

Im zweiten Arbeitspaket behandelten wir die topographische Anordnung der Texte. Topographisch bedeutet hierbei, dass es um die Verortung des Textes auf seinem Träger geht. Parallel dazu steht die semantische Textstruktur, die aussagt wie der Text gelesen wird und aus welcher logischen Abfolge er besteht. Siehe dazu AP 01.

Im AP 02 wurden folgende Anforderungen berücksichtigt: Wo sind Textfelder und Abbildungen auf dem Textträger verortet? Ergebnisse des AP 02 war es, einfache bis komplexe Beispieltexthe in der TEI/XML Struktur abzubilden, so dass alle Phänomene der Textanordnung in Inschriften des Klassischen Maya berücksichtigt werden. Diese aus AP 02 resultierende topographische Textanordnung ist grundsätzlich für alle Texte anwendbar, benötigt aber spezielle Erweiterungen für die Kodizes.

### AP 03: Philologische und textkritische Auszeichnung

Mit dem dritten Arbeitspaket behandelten wir den editorischen Umgang mit unlesbaren, vagen oder rekonstruierten Textstellen. Editorisch bedeutet dabei, dass es um eine textkritische Betrachtung geht, die bei der späteren epigraphischen und linguistischen Analyse berücksichtigt werden kann. Dieses Paket spezifiziert damit die Auszeichnung aus AP 01.

Das AP 03 regelt den Umgang mit folgenden Fragestellungen: Wie können durch unklare ursprüngliche Schreibungen oder durch spätere Restaurierungen unklar gewordene, durch physische, chemische oder biologische Einwirkung beschädigte oder zerstörte Textpassagen qualifiziert und in ihrem Ausmaß quantifiziert werden? Ziel des AP 03 war es, Beispieltexthe so auszuzeichnen, dass sich klare editorische Richtlinien für den Umgang mit vagem Material ergeben, aus denen durch Kommentierung im XML und Verweise in die Zotero-Literaturdatenbank ein *apparatus criticus* entsteht.

### AP 04: TEI-Header und Bearbeitungsmetadaten

Im vierten Arbeitspaket behandelten wir die deklarativen Eigenschaften des gesamten TEI-Dokuments. Deklarativ bedeutet dabei, dass es um die beschreibenden Metadaten für die Edition des Textes geht. Gemäß den Projektsprachen erfolgen diese Informationen in Deutsch, Englisch und Spanisch.

Das AP 04 behandelt dabei folgende Aspekte: statische Rahmeninformationen wie die Akademie und die Akademienunion als Projektträger, die Nennung der Kooperationspartner, eine Kurzbeschreibung des Projekts, die Nennung der Mitarbeiter, der Herausgeberschaft und die Lizenzierung der Edition als CC-BY-4.0. Als dynamische Informationen werden mit Zugriff auf die Textträgerdatenbank in TextGrid Informationen zum Artefakt angegeben (z.B. dessen Name), wer den Text mit den Graphnummern des Zeichenkatalogs transliteriert hat (siehe AP 01) oder wer das weitere Markup (siehe AP 02, 03, 05) erstellt hat. Weiter wird eine Revisionshistorie der Edition im Header angelegt.

## AP 05: Text- und Textträgergestaltung

Im fünften Arbeitspaket behandelten wir die gestalterische Ausführung der Texte. Das bedeutet, dass wir hier das Design der Texte (weniger das Layout, siehe AP 02) und die Typographie auszeichnen. Parallel dazu wird auch das Anordnungsverhältnis von Text zu bildlicher Darstellung berücksichtigt.

Das AP 05 behandelt dabei folgende Fragestellungen: welche Kriterien für die Textgestaltung können für Forschungsfragen relevant sein und in welchem Verhältnis stehen Design und Semantik in einer Beziehung? Hierfür werden mit TEI/XML Eigenschaften wie die Form eines Textfeldes, die Reliefierung, Rahmung, Kolorierung oder die Schriftgröße erfasst und weiterhin auch individuelle Schreiber oder -schulen ausgezeichnet werden. Ferner regelt das AP 05 auch den Umgang mit Text-Bild-Relationen, etwa bei Tags oder direkter Rede.

Mit Abschluss der Arbeitspakete steht ein umfassendes Schema zur Auszeichnung der Maya-Hieroglyphentexte zur Verfügung, das ggf. auch von anderen Projekten, die sich mit komplexen Schriftsystemen befassen, nachgenutzt werden kann. Aufgrund von Erfahrungswerten aus Editionsprojekten werden sich noch weitere Anforderungen bzw. Erweiterungen an die bisherigen Spezifikationen während der produktiven Phase (Beginn: Anfang 2018) ergeben. Diese Anpassungen in das TEI-Schema zu integrieren wird eine wiederkehrende Aufgabe während der kommenden zwei Jahre sein. Neben der technischen Dokumentation des Metadatenschemas werden wir auch unsere editorischen Richtlinien veröffentlichen, die unsere Praxis der Textkodierung anhand von Beispielen erläutern und darüber hinaus auch Handlungsanweisungen für Auszeichnungsstrategien erteilen.

## Annotationstool ALMAH (Annotator for the Linguistic Analysis of Maya Hieroglyphs)

Für die semi-automatische morphologische Annotation zur Nutzung im Forschungsprojekt "Textdatenbank- und Wörterbuch des Klassischen Maya" werden Anpassungen, Modifikationen und Funktionsergänzungen der linguistischen Annotationssoftware "GeTa" vorgenommen.

Die von der Computerlinguistin Dr. Cristina Vertan (Hamburg) konzipierte und programmierte Java-Client-Anwendung "GeTa" (Ge'ez Tagging) für die Annotation von in nicht-alphabetischen Schriftsystemen (Silben- und Wort-Silbenschriftsysteme) verfassten Texten wird im Rahmen eines Werkvertrags zwischen der Entwicklerin Frau Dr. Vertan (Auftragnehmerin) und unserem Forschungsprojekt als Auftraggeber für die semi-automatische Transliteration, morphologische und semantische Annotation des Klassischen Maya angepasst, modifiziert und mit zusätzlichen Annotationsfunktionen ergänzt, was in fachwissenschaftlicher Abstimmung mit dem Auftraggeber erfolgt. Gleichzeitig wird eine Verlinkung mit dem parallel entwickelten Lexikon realisiert. Der Werkvertrag definiert dabei sechs Arbeitspakete.

## Computergestützte Annotationstools

Unser Projekt greift auf korpuslinguistische Methoden zurück, um sprachwissenschaftliche Fragestellungen auf der Basis von digitalen Textsammlungen zu bearbeiten. Dazu werden Sprachdaten des Klassischen Maya digital aufbereitet und mit Grundannotationen wie etwa Part-of-Speech-Labeln versehen. Für das Wörterbuch werden große Mengen an Sprachdaten benötigt, die nicht nur manuell annotiert werden sollen, sondern größtenteils semi-automatisch bearbeitet werden müssen. Annotationstools sind computergestützte Werkzeuge, mit deren Hilfe Forscher Texte oder ihre Bestandteile mit individuellen Informationen anreichern können, etwa mit Notizen, Sach- oder Textkommentaren, oder Inhalts- und Semantikanalysen. Linguistische Annotationstools erlauben

dabei die computergestützte Transliteration, Transkription und Annotation von Texten und ermöglichen das Erstellen und Auswerten von Korpora gesprochener sowie verschrifteter Sprachen.

Bislang existieren keine computergestützten Verfahren und Werkzeuge für die Dokumentation der Hieroglyphenschrift des Klassischen Maya bzw. der Mayasprachen. Des Weiteren müssen wir unser Lexikon und Korpus gleichzeitig entwickeln, da bislang nur Wortlisten des Klassischen Maya vorliegen. Bisherige Annotationstools erfüllen diese Bedingungen nicht und könnten für diese Aufgaben grundsätzlich nicht eingesetzt werden, da sie nur die Annotation von Texten erlauben, die in alphabetischen Schriftsystemen verfasst wurden. Für das Klassische Maya gibt es bislang mit Ausnahme von GeTa keine andere Software, die nachnutzbar wäre, um den Lexembestand computergestützt zu erschließen. Dabei müssen verschiedene Ebenen der Transliteration und Transkription semi-automatisch mit Part-of-Speech- und Morphologie-Tags annotiert werden können. Zudem müssen manuell Korrekturen während des Annotationsprozesses eintragbar sein und Textbereiche, die bislang nur unsicher oder überhaupt nicht entziffert sind, markiert werden können. Der semi-automatische Charakter eines derartigen Annotationstools ist notwendig um mit dem wachsenden Materialkorpus den Annotationsvorgang der Hieroglyphentexte zu beschleunigen. Hierbei ist auch die Korrekturmöglichkeit zu betonen, die dem Fachwissenschaftler erlaubt, automatisch generierte Annotationen zu prüfen und ggf. manuell zu verändern.

### Anpassung des Annotationstools GeTa für die Mayaschrift

Für die Nachnutzung in unserem Projekt eignet sich die für das Ge'ez entwickelte Datenstruktur des GeTa-Tools in mehrerer Hinsicht. GeTa erlaubt die automatische Transliteration der Primärquelle(n) und die Annotation erfolgt semi-automatisch und selbstlernend. Diese Eigenschaft beschleunigt die Annotation und garantiert ihre Konsistenz. GeTa erlaubt die Annotation auf Lexem- und Morphemebene, eine weitere wichtige Voraussetzung für die Lexikographie des Klassischen Maya. In Schriftsystemen mit syllabischen Komponenten sind Morphemgrenzen in der Schrift mitunter nicht darstellbar, so dass ihre Laute für die morphologische Segmentierung einzeln tokenisiert werden müssen.

In der existierenden Version von GeTa sind Originaltext, Transliteration und Transkription synchronisiert, d.h. auf diese Weise bleiben alle Ebenen einschließlich weiterer Annotationen im Fall von Änderungen, Korrekturen oder Hinzufügungen stets synchron. Korrekturen während des Annotationsprozesses können dadurch synchron je Text oder global über das Textkorpus vorgenommen werden. Die Einbettung von Varianten zur Analyse alternativer Interpretationen auf jeder Ebene wird neu implementiert. In der aktuellen Version von GeTa für das Altäthiopische können textstrukturelle Annotationen (z.B. Sätze, Sektionen), morphologische (Part-of-Speech), inhaltliche (Named Entities) und editorische (Seitenumbrüche, Korrekturen etc.) gesetzt werden. Die Daten werden im JSON-Format gespeichert und erlauben den Export in TEI/XML, die Annotationen können durchsucht, visualisiert und statistisch ausgewertet werden.

Die Funktion, Lexikon und Korpus in gegenseitiger Synchronizität zu halten, ist für die Wörterbucherstellung des Klassischen Maya fundamental. Neu wird die gleichzeitige Entwicklung des Lexikons eingebaut. Nachdem GeTa bereits für die Annotation altsabäischer, hebräischer und jiddischer Texte angepasst wurde, werden aufgrund der beschriebenen Grundfunktionen jetzt Anpassungen für die Analyse der Mayahieroglyphentexte durchgeführt.

Architektur, Datenmodell, Funktionen und Module von GeTa müssen an die vorhandene informationstechnische Umgebung unseres Projekts angepasst werden, insbesondere für eine Schnittstelle mit TextGrid. Diese Modifikationen für ein Annotationstool für das Klassische Maya

werden unter dem Namen ALMAH (Annotator for the Linguistic Analysis of Maya Hieroglyphs) geführt, abgeleitet vom yukatekischen *a'Imah*, "decir, mandar, ordenar (sagen, schicken, ordnen)" (Barrera Vásquez 1980: 13). ALMAH muss Zugriff auf die Benutzerkonten in TextGrid haben, damit eindeutig während der Arbeit mit Texten Daten im TextGrid Lab gelesen, geöffnet, gesperrt, gespeichert und gelöscht werden können. Dazu muss über den Fuseki-Triplestore auf die RDF-Objekte zugegriffen werden, damit diese mittels ihrer URI in der ALMAH-Datei verlinkt werden können. Weiter muss eine Kommunikation mit einer Exist-DB eingerichtet werden, weil die numerisch transliterierten Texte im TEI/XML-Format in TextGrid vorliegen. Weiter müssen neben den Verweisen zu TextGrid auch Verlinkungen zur Bibliographiedatenbank Zotero ermöglicht werden, um Annotationen mit Quellenangaben belegen zu können.

Aufgrund einer unterschiedlicher Grammatik müssen sämtliche morphologischen Auszeichnungsmöglichkeiten neu definiert und implementiert werden, um linguistische Analysen durchführen zu können. Hier müssen die unterschiedlichen Annotationstypen, -ebenen und -kategorien definiert, vorbereitet, implementiert und letztlich auf ihre Funktionalität getestet werden. Für das Wörterbuchprojekt benötigen wir eine Reihe von Annotationsmöglichkeiten, die neu konzipiert und in die bisherige GeTa Architektur implementiert werden müssen: Neben einer allgemeinen Kommentierungsfunktion der Texte, benötigen wir Annotationsebenen für die schrittweise Überführung von Schrift in Sprache, die grammatologisch und linguistisch transparent erfolgen soll. Dafür sind acht Analyseebenen plus die Übersetzung vorgesehen: 1) Transliterationen: alphanumerisch, numerisch, graphemisch, phonemisch; 2) Transkriptionen: morphologisch segmentiert, morphophonemisch angepasst, morphosyntaktisch glossiert, und final konsolidiert. Weiter wird es Annotationen für die Syntax oder eine semantisch-funktionale Analyse geben, etwa für Metaphern und Metonyme, kalendarische Inhalte oder onomastische Inhalte und Etymologien. Jede Ebene muss innerhalb von ALMAH aufeinander referenzierbar sein, aber auch Verweise mittels URIs nach TextGrid müssen hier ermöglicht werden. Für jede Annotationsebene müssen fachwissenschaftliche Begriffslisten mit festen Werten vorbereitet und implementiert werden. Für die morphosyntaktische Analyse sind Glossierungslisten zu implementieren.

Im Bereich der Programmarchitektur und internen Funktionalität sind Änderungen und Erweiterungen nötig: GeTa muss so erweitert und ergänzt werden, dass es alternative Transliterationen, Transkriptionen und Analysen in allen Annotationsebenen zulässt. Da die Mayaschrift noch nicht vollständig entziffert ist, müssen Epigraphiker die Möglichkeit haben, mit alternativen Lesungen und Deutungen zu arbeiten und diese im weiteren Analyseprozess zu entwickeln. Hier sind grundlegende Anpassungen im Datenmodell notwendig, die es aber erlauben, sich nicht auf eine hypothetische Lesung festlegen zu müssen. Zudem ist auch ein Interface zum Lexikon des Klassischen Maya zu modellieren und implementieren. Im Zuge der Analyse und Annotation der Hieroglyphentexte entsteht im Hintergrund das Lexikon des Klassischen Maya, das es ermöglicht, die automatisch generierte Lemma-Liste mit fachwissenschaftlichen Kommentaren und Analysen zu ergänzen. Diese Funktion ist ebenfalls neu zu modellieren und implementieren. Diese oben genannten zahlreichen Erweiterungen und Funktionsanpassungen von GeTa für unser Projekt machen eine Anpassung der Benutzeroberfläche notwendig, indem die einzelnen Eingabemasken einschließlich Suchfunktionen für die unterschiedlichen Annotationstypen dem fachwissenschaftlichen Workflow entsprechend konzipiert, implementiert und visualisiert werden. Diese werden gemäß Vorgaben und Anforderungen der Bonner Fachwissenschaftler von Dr. Cristina Vertan geplant, umgesetzt und implementiert.

## 3D-Dokumentation

Das Projekt führte 2017 drei Kampagnen zur Dokumentation von Textträgern mit seinem 3D-Streifenlichtscanner durch.

Im Februar wurde an fünf Tagen das 2016 begonnene Scanvorhaben am *Museum of Archaeology and Anthropology* an der *University of Cambridge* fortgeführt und beendet. Dabei wurde das letzte fehlende Fragment des Gipsabgusses von Quirigua Stele E in der Ausstellungshalle aufgenommen, was im vorangegangenen Jahr aufgrund eines fehlerhaft kalibrierten Sensors nicht abgeschlossen werden konnte. Weiterhin konnten aus dem Magazin diverse Kleinartefakte, darunter beschriftete Keramiken und Knochen, sowohl mit dem 3D-Scanner als auch photographisch dokumentiert werden.

Zwischen Oktober 2016 und April 2017 fand im *Historischen Museum der Pfalz* in Speyer die Ausstellung "Maya: Das Rätsel der Königsstädte" statt. In Zusammenarbeit mit den Kuratoren der Ausstellung, der guatemaltekischen Denkmalbehörde *Instituto de Antropología e Historia* (IDAEH) und der Stiftung *Fundación Ruta Maya* als Leihgeber der Exponate konnten an insgesamt sechs Terminen 12 monumentale Steinskulpturen in 3D erfasst werden. Unser Projekt wurde dabei auch an ausgewählten Terminen mit zusätzlichen 3D-Scannern unterstützt, einmal direkt durch die Herstellerfirma *Aicon 3D Systems*, andere Male durch Dr. Hubert Mara vom *Forensic Computational Geometry Laboratory, Interdisciplinary Center for Scientific Computing* an der Universität Heidelberg. Das Scannen in der Ausstellung wurde öffentlichkeitswirksam durchgeführt (siehe unter *Außenwirkung des Projekts*).

Im Juni konnten an fünf Tagen im Magazin des *British Museum* in London ausgewählte Gipsabgüsse, die in den 1880er Jahren von Alfred Maudslay angefertigt wurden, mittels 3D-Scanning dokumentiert werden.

Im September konnte das Projekt weiterhin an einem Workshop des 3D-Labors am Institut für Mathematik der TU Berlin teilnehmen, wo verschiedene 3D-Digitalisierungs- und Visualisierungstechnologien vorgestellt und diskutiert wurden.

Für zukünftige Scanvorhaben haben dieses Jahr weitere Museen mit Inschriftenartefakten für ihre Sammlungen ihre Zusage gegeben (siehe unter *Kooperationen*).

Die 2016 durchgeführten Scanprojekte wurden dieses Jahr zusammen mit allen in diesem Jahr durchgeführten Arbeiten einem intensiven Postprocessing unterzogen, das im Wesentlichen aus folgenden Schritten besteht: 1) Bereinigung der Rohdaten, 2) Mergen der Rohdaten zu einem einzelnen Mesh, 3) Lochfüllung von nicht erfassten Oberflächenpartien und 4) Retexturierung. Das dabei entstehende hochqualitative Mesh dient dann als Grundlage für weitere Visualisierungen. Einzelne Monumente bestehen teilweise aus mehreren Einzelteilen, deren Meshes dann wiederum virtuell zusammengefügt werden müssen. Folgende Objekte konnten dabei finalisiert und teilweise auf unserer *Sketchfab*-Seite veröffentlicht werden:

1. Cancuen, Panel 1 (Original, Speyer)
2. Dos Pilas, Panel 19 (Original, Speyer)
3. La Amelia, Stela 1 (Original, Speyer, 2 Einzelteile)
4. Lacanha, Panel (Original, Speyer)
5. La Corona, Panel 1 (Original, Speyer, 2 Einzelteile)
6. La Pasadita, Panel 1 (Original, Speyer)
7. Machaquila, Stela 3 (Original, Speyer)
8. Machaquila, Stela 8 (Abklatsch, London, 1 Einzelteil)
9. Palenque, Hieroglyphic Stairway (Kopie, London, 14 Einzelteile)
10. Palenque, Hieroglyphic Stairway (Abklatsch, London, 1 Einzelteil)

11. Palenque, Tablet of the Foliated Cross, centre slab (Kopie, Bonn)
12. Palenque, Temple of the Inscriptions, east panel (Kopie, London, 2 Einzelteile)
13. Provenance Not Known, Altar (Original, Speyer)
14. Provenance Not Known, Bone 267 (Original, Cambridge)
15. Provenance Not Known, Jade 618 (Original, Cambridge)
16. Provenance Not Known, Panel (Original, Speyer)
17. Provenance Not Known, Vessel 238 (Original, Cambridge, 3 Einzelteile)
18. Provenance Not Known, Vessel 238XHI (Original, Cambridge)
19. Provenance Not Known, Vessel 257 (Original, Cambridge)
20. Provenance Not Known, Vessel 273 (Original, Cambridge)
21. Quirigua, Altar L (Original, Speyer & Abguss, Cambridge)
22. Quirigua, Altar Q (Kopie, Cambridge)
23. Quirigua, Stela A (Kopie, Cambridge, 18 Abgüsse)
24. Quirigua, Stela C (Kopie, Cambridge, 4 Abgüsse)
25. Quirigua, Stela D (Kopie, Cambridge, 6 Abgüsse)
26. Quirigua, Stela E (Kopie, Cambridge, 14 Abgüsse)
27. Quirigua, Stela F (Kopie, Cambridge, 4 Abgüsse)
28. Quirigua, Stela J (Kopie, Cambridge, 5 Abgüsse)
29. Quirigua, Stela K (Kopie, Cambridge, 2 Abgüsse)
30. Quirigua, Zoomorph B (Kopie, Cambridge, 1 Abguss)
31. Seibal, Stela 3 (Original, Speyer)
32. Tikal, Lintel 3, Tempel I (Original, Basel)
33. Tikal, Lintel 2, Tempel IV (Original, Basel, 7 Einzelteile)
34. Tikal, Lintel 3, Tempel IV (Original, Basel)
35. Yaxchilan, Lintel 2 (Kopie, Basel)
36. Yaxchilan, Lintel 16 (Kopie, Cambridge)
37. Yaxchilan, Lintel 17 (Kopie, Cambridge)
38. Yaxchilan, Lintel 24 (Kopie, Cambridge)
39. Yaxchilan, Lintel 30 (Kopie, Cambridge)
40. Yaxchilan, Lintel 42 (Kopie, Cambridge)
41. Yaxchilan, Lintel 43 (Kopie, Cambridge)
42. Yaxchilan, Lintel 45 (Kopie, Cambridge)
43. Yaxchilan, Stela 5 (Kopie, Cambridge)
44. Yaxchilan, Stela 7 (Kopie, Cambridge, 2 Abgüsse)
45. Zapote Bobal, Panel (Original, Speyer)

## Dateneingabe Metadaten

2017 konzentrierten wir uns auf die Dateneingabe im Bereich Artefakte, epigraphische Akteure und Aktivitäten. Die in den 20 Bänden des *Corpus of Maya Hieroglyphic Inscriptions Project* (Harvard University) veröffentlichten Artefakte aus über 20 Fundstätten wurden mit Unterstützung studentischer Hilfskräfte bis auf einen Band vollständig mit Hilfe der Eingabemaske in TextGrid erfasst (über 700 Objekte). Im November 2017 beinhaltete unsere Datenbank über 3000 Objekte, bestehend aus 1442 Objekten mit Text- und Bildträgern (artefacts), 1337 Objekte mit Angaben zu Orten (locations), 523 Objekte mit aus der Literatur recherchierten epigraphische Aktivitäten (epigraphic activities), 441 Objekte zu epigraphischen und 28 nicht-epigraphische Akteuren (epigraphic and non-epigraphic actors), 436 Objekte mit Entdeckungsereignissen von Artefakten (discovery events) und 103 Objekte zu Herstellungsereignissen (production events). Der digitale Zeichenkatalog umfasst derzeit 190 Objekte mit Zeichen und 71 mit Graphen.

## Vokabularentwicklung

Die nach der 2015 erfolgten Konzipierung im Vorjahr begonnene Entwicklung kontrollierter Vokabulare zur Unterstützung der Eingabe von Textträgerdaten sowie zur ikonographischen Beschreibung der Graphe im Zeichenkatalog wurde fortgeführt. Die Eingabe erfolgte mittels des Vokabularentwicklungstools SKOtch im SKOS (Simple Knowledge Organisation System)-Format. Derzeit (Stand 20. Dezember) liegen nun in 11 Vokabularen (Activity, Actor Appellation, Artefact Orientation, Position and Assembling, Artefact Shape, Artefact Type, Condition, Group Type and Actor Role, Material, Technique and Style, Graph Icon) und insgesamt 1770 Konzepte (concepts) mit jeweils präferierten (preferred terms) und alternativen Begriffen (alternative terms) vor, zusammen mit ihrer jeweiligen Definition sowie Quellennachweisen. Bei der Erstellung der Vokabulare wurde soweit wie möglich auf Normdaten (Getty Art & Architecture Thesaurus, Getty Thesaurus of Geographic Names) zurückgegriffen und die entsprechenden Einträge mit den jeweiligen Normdatensätzen verlinkt. Da jedoch den mesoamerikanischen Kulturraum und insbesondere die Mayakultur betreffende Begriffe in den erwähnten Normdatenbeständen bisher sehr unterrepräsentiert sind, mussten zudem Begriffe und Definitionen in der Fachliteratur recherchiert und übernommen als auch eigene Definitionen formuliert werden. Der so aufgebaute umfangreiche Thesaurus einschließlich der Definitionen liegt jeweils in den Projektsprachen Deutsch, Englisch und Spanisch vor, enthält jedoch zusätzlich auch eine Reihe Begriffe (terms) des Klassischen Maya. Die erstellten Vokabulare können bei Bedarf jederzeit ergänzt werden; wobei eine erste Version im kommenden Jahr veröffentlicht wird.

## Bibliographie (Zotero)

Seit November 2016 wurden über 5000 neue Einträge in die Literaturdatenbank Zotero eingearbeitet bzw. bestehende Einträge korrigiert bzw. mit Internetlinks für den direkten Download ergänzt. Im Dezember 2017 umfasst die Literaturdatenbank 21.241 Einträge - sie ist unter anderem mit TextGrid verknüpft und wird dort verwendet, um Einträge mit Literaturangaben zu belegen. In Zusammenarbeit mit dem Projektmanagementbüro Beuse, Köln, wird derzeit eine Suchmaske für die Webseite [www.mayawoerterbuch.de](http://www.mayawoerterbuch.de) vorbereitet, die es den Nutzern erlaubt, Einträge zu suchen und herunterzuladen. Darüber hinaus ist die Literaturdatenbank mit der Bilddatenbank ConedaKOR verknüpft um die Einträge dort ebenfalls mit Literaturangaben zu belegen. Die Arbeit an der Literaturdatenbank wurde in diesem Jahr von Praktikanten unterstützt. Neben der Eingabe von Literaturangaben über Printmedien wurde in diesem Jahr vermehrt die Eingabe von Onlinepublikationen vorgenommen.

## Bilddatenbank

Das fotografische Archiv des Projekts besteht aus den Sammlungen mehrerer Forscher, die teilweise auf mehrere Dekaden an Forschungsreisen im Mayagebiet zurückblicken können. Daher sind diese Fotos eine unschätzbare Quelle für das Studium von Inschriften, Kunst und Architektur der klassischen Maya. Zu jedem Bild sind einige grundsätzliche Informationen beigefügt. Jedoch sind diese Beschreibungen meist sehr unspezifisch gehalten. Es stellte sich daher sehr bald heraus, dass eine systematische Revision des Materials notwendig ist, um eine präzise Identifikation der abgebildeten Objekte zu erhalten.

Hier wurde entschieden, mit der inhaltlichen Aufarbeitung des Schwarzweißfotoarchivs von Karl Herbert Mayer zu beginnen. Es besteht aus etwa 20.000 Bildern aus den Jahren 1974 bis 2006. Während seiner zahlreichen Reisen legte Karl Herbert Mayer seinen Fokus auf skulptierte Monumente

mit ikonographischem oder epigraphischem Inhalt. Seine Dokumentation legte den Grundstein für eine Schriftenreihe zu Monumenten unbekannter Herkunft (Mayer 1978, 1980, 1984, 1987, 1989, 1991, 1995) und zahlreichen Artikeln in Fachzeitschriften.

Den Einstieg bildeten Monumente, die Karl Herbert Mayer selber publiziert hat und für die ausführliche Beschreibungen vorlagen. In einem zweiten Schritt wurden abgebildete Monumente aus Museen aufgearbeitet (z.B. das *Museo Nacional de Antropología* in Mexiko-Stadt, das *Museo Nacional de Arqueología y Etnología* in Guatemala-Stadt oder das *Museo Regional de Antropología Palacio Cantón* in Mérida). Weitere bisher erschlossene Inhalte zu Bildern decken archäologische Fundstätten in den mexikanischen Bundesstaaten Campeche und Yucatan ab, die Monumente im Magazin des Nationalparks Tikal in Guatemala, sowie besonders Fundstätten mit Wandmalereien und Graffiti. Das gesamte Bild- und Dokumentenarchiv wurde im November 2017 von Christian Prager und Antje Grothe in Graz gesichtet, in Kisten verpackt und nach Bonn transportiert, wo es derzeit für die Digitalisierung durch studentische Hilfskräfte im kommenden Jahr vorbereitet wird.

Gegenwärtig sind ungefähr 60% der Schwarzweißfotos mit Metadaten zum abgebildeten Objekt (Beschreibung, Aufbewahrungsort zur Zeit der Aufnahme, Eigentümer, Sammlung, archäologischer Fundort, Katalognummer, Publikation etc.) versehen.

Neben dem Archiv Karl-Herbert Mayers werden weitere Fotoarchive sukzessive aufbereitet und in eine Online-Bilddatenbank überführt werden.

## ConedaKOR

Die Bildverwaltungssoftware ConedaKOR dient der Verwaltung und Präsentation unserer Online-Bilddatenbank, deren Sammlung somit per Webzugang nutzbar sowie öffentlich voll zugänglich gemacht wird. Das browserbasierte Frontend sowie die Metadaten zu den Inhalten der Datenbank sind in englischer Sprache verfasst.

In der Datenbank werden alle digitalisierten und inhaltlich erschlossenen Fotoarchive miteinander kontextualisiert präsentiert. Damit an den abgebildeten Objekten und den zugehörigen Eigenschaften der verschiedenen Archive zielgerichtete Suchen durchgeführt werden können, müssen die heterogenen Informationen der unterschiedlichen Sammlungen zu jedem abgebildeten Objekt in einem kohärenten Metadatenchema abgebildet werden.

Das Schema umfasst übergreifend beschreibbare Eigenschaften der abgebildeten Inhalte wie den Aufbewahrungsort zur Zeit der Aufnahme, Eigentümer, Sammlung, archäologischer Fundort, Artefakttyp, Beschreibung, Katalognummer, Publikation, GND-ID etc.

Die Eigenschaften werden bestimmten Entitäten zugeordnet und mittels möglicher Relationen verbindbar gemacht. Hier werden folgende Entitäten abbildbar:

1. "Medium", die digitalisierte Abbildung (in Form einer Tiff-Datei).
2. "Archaeological Site", die Fundstätte des abgebildeten Objekts.
3. "Artefact", ein abgebildetes Artefakt.
4. "Holder", der Ort oder die Institution oder Person, zu der das abgebildete Objekt zur Zeit der Abbildung zugehörig war.
5. "Collection" bezieht sich auf einen Holder (Museum, Privatperson etc.). Collection ist immer in Verbindung mit einer Holder-Entität vorzufinden.
6. "Place", der Ort des abgebildeten Objekts.
7. "Person" bezieht sich auf die Person, die Urheber des Fotos ist, welches sich als Medium in der Datenbank befindet.

Schließlich sind zurzeit die folgenden bidirektionalen Relationen abbildbar:

1. "Artefact is held by holder - Holder holds Artefact"
2. "Person visited Place - Place was visited by Person"
3. "Place is located in Place - Place locates Place"
4. "Holder is located in Place - Place locates Holder"
5. "Archaeological Site is in Place - Place locates Archaeological Site"
6. "Archaeological Site is origin of Artefact - Artefact originates from Archaeological Site"
7. "Medium depicts Artefact - Artefact is depicted by Medium"
8. "Person created Medium - Medium is created by Person"
9. "Artefact is related to Artefact - Artefact is related to Artefact"
10. "Person is depicted by Medium - Medium depicts Person".

In ConedaKOR werden alle Inhalte technisch innerhalb eines Graphen organisiert, der Entitäten mit ihren Eigenschaften als Knoten umsetzt und deren Beziehung zueinander über Kanten verbindet. Neben einer gezielten Suche wird sozusagen per Klick ein Stöbern von Entität zu Entität ermöglicht, wodurch vorhandene Verknüpfungen entlang der Relationen entdeckt werden können. Diese Netzwerkstruktur erlaubt es somit komplexere Zusammenhänge zwischen den Objekten und Eigenschaften der verschiedenen Sammlungen in der Datenbank aufzubewahren, abzubilden, und recherchierbar zu machen.

## Internationalisierung und Nachwuchsförderung

### Internationaler Fellow

Für den Ausbau der internationalen Beziehungen finanzierte die Nordrhein-Westfälische Akademie der Wissenschaften und der Künste den Forschungsaufenthalt von Dr. Albert Davletshin vom Zentrum für Komparative Linguistik an der *Russischen Staatlichen Geisteswissenschaftlichen Universität* in Moskau. Von März bis Dezember forschte Dr. Davletshin an der Bonner Arbeitsstelle zum Thema *Historische Phonologie als Grundlage für die Entzifferung des Klassischen Maya*. Mit diesem Forschungsvorhaben leistet Dr. Davletshin essentielle Grundlagenarbeit für die Sprachbeschreibung und damit für das Wörterbuch des Klassischen Maya. Als komparativer Linguist mit Expertise für amerindische Sprachen mit speziellem Fokus auf Mayasprachen ergänzte er als Linguist die Bonner Schriftforscher bei ihrer Beschreibung und Entzifferung des Klassischen Maya. Dr. Davletshin füllte mit seinem Projekt über die historische Phonologie des Klassischen Maya eine Forschungslücke, die wir im Antrag für dieses Forschungsprojekt definiert haben: die phonologische Struktur des Klassischen Maya, die bislang nur punktuell erforscht wurde und bis heute ein Desideratum bei der Lexikographie des Klassischen Maya bildet.

Dr. Davletshin stellt fest, dass die Mayaschrift logosyllabisch sei und unter der Annahme entziffert wurde, dass die ihr zugrunde liegende Sprache Maya ist. Die Mayasprachfamilie bestehe allerdings aus über 30 verschiedenen Mayasprachen und die Unterschiede zwischen ihnen überwiegen im Vergleich zu ihren Gemeinsamkeiten, vergleichbar mit den grammatischen, syntaktischen und lexikographischen Unterschieden zwischen Spanisch und Französisch. In den vergangenen 40 Jahren der Mayaschriftentzifferung zeigte sich, dass die Sprache der Mayahieroglyphen zum Sprachzweig des Ch'olan gehört - keine der ihr zugehörigen heutigen Mayasprachen entspreche allerdings dem "hieroglyphenschriftlichen Maya" noch stammen diese *in strictu sensu* direkt vom Klassischen Maya ab. Bemerkenswert sei, so Davletshin, dass das hieroglyphenschriftliche Maya phonologisch

konservativ sei und systematisch zwischen kurzen, langen und glottalisierten Vokalen und zwischen velaren und laryngalen Frikativen unterscheidet. Diese phonologischen Merkmale des Klassischen Maya seien heute in den meisten Mayasprachen verloren und in keiner der Ch'olan-Sprachen hätten sich diese Merkmale erhalten. Unser Verständnis der Mayatexte und die Entzifferung der individuellen Zeichen basieren auf dem Prinzip der Etymologie, d.h. dem Vergleich mit lexikalischen Wörterbucheinträgen in kolonialzeitlichen sowie modernen Mayawörterbüchern. Diese Methode beruhe auf regulären Lautkorrespondenzen und auf Annahmen über die Entwicklung von Lauten in den verschiedenen Mayasprachen, die auf das Proto-Maya zurückgingen. Regelmäßige Lautkorrespondenzen und die Entwicklung des Lautsystems in den verschiedenen Mayasprachen bildeten nicht nur die Grundlage für die Entzifferung der Mayahieroglyphen, da sie lexikalische Vergleiche mit anderen Mayasprachen ermöglichten, sondern sie erlaubten den Forschern auch unmögliche Entzifferungen und Interpretationen auszuschließen.

Dr. Davletshins Forschungsprojekt in der Arbeitsstelle "Textdatenbank und Wörterbuch des Klassischen Maya" hat sich daher mit dieser wichtigen Methode befasst und am Beispiel von komplexen Vokalen und seltenen Konsonanten das System von Lautkorrespondenzen zwischen den verschiedenen Mayasprachen in Bezug auf das Klassische Maya erforscht. Diese Erkenntnisse unterstützen nicht nur die Bonner Forscher bei der Entzifferung der Texte, sondern sie erlauben Einblicke in die phonologische Struktur einer toten Schriftsprache. Die Ergebnisse dieses Projekts werden in zwei Artikeln veröffentlicht werden, welche die genannten Lautgesetze systematisch untersuchen und beschreiben.

### Praktikum im Projekt

Das Angebot an gemäß der Studienordnung an der Abteilung für Altamerikanistik erforderlichen Praktika für BA- und MA-Studenten wurde in diesem Jahr um Praktikumsplätze in unserem Projekt erweitert und es wird bereits von mehreren Studenten wahrgenommen. Die Dauer der Praktika umfasst 180 Stunden (BA) bzw. 150 Stunden (MA), wobei die Arbeitszeit studienbegleitend in 10 Stunden pro Woche aufgeteilt ist, die Praktikumsdauer also 18 bzw. 15 Wochen umfasst. Während des Praktikums sind sechs Einsatzbereiche zu absolvieren, wobei Kenntnisse in Epigraphik, Sprachen, wissenschaftlichem Arbeiten, und Recherche, als auch grundlegende IT-Kenntnisse sowie Erfahrung in digitaler Bildbearbeitung erforderlich sind. Die Einsatzbereiche bieten den Praktikanten einen umfassenden Einblick in Arbeitsabläufe und Methoden eines geisteswissenschaftlichen Langzeitprojekts und sind wie folgt festgelegt:

**Bibliographie:** Literaturrecherchen in bestehenden Datenbanken und ausgewählten gedruckten Fachbibliographien und Dateneingabe haben zum Ziel, Recherchewerkzeuge und -quellen kennenzulernen und die gewonnenen bibliographischen Daten durch Einsatz einer bibliographischen Datenbank zu strukturieren, inklusive der Verlinkung der Datensätze mit bestehenden Datenbanken, Webseiten, etc. Zudem sollen anhand von Recherchen in Katalogen von Universitätsbibliotheken und gedruckter Literatur die Einträge bisheriger Fachbibliographien erweitert und ggf. korrigiert werden.

**Inschriftenarchiv:** Der zweite Arbeitsbereich umfasst die Arbeit an sowohl dem analogen (Archivkästen) als auch dem digitalen Inschriftenarchiv (TextGrid). Dabei werden aus ausgewählten Archiven (Inschriften aus Campeche, Justin Kerrs Keramikarchiv, Karl Herbert Mayer-Archiv von Objekten unbekannter Herkunft) Artefakte mit Inschriften recherchiert, dazu in TextGrid die jeweils erforderlichen Objekte (Artefact, Production, Location) angelegt und die entsprechenden Daten eingepflegt.

**Digitalisierung:** Im dritten Arbeitsbereich wird an die Digitalisierung von 2D- und 3D-Objekten und die dafür erforderlichen Arbeitstechniken, Geräte und Software herangeführt. Die 2D-Digitalisierung umfasst das Einscannen von Dias und Fotos und Postprocessing sowie vorbereitende Arbeiten zur späteren Einpflege der digitalisierten Objekte in die Bilddatenbank ConedaKOR (Auswahl, Recherche und Beschreibung der Objekte). Zur Digitalisierung von 3D-Objekten wird das Scanning von Artefakten mittels eines 3D-Streifenlichtscanners und der dazugehörigen Software Optocat ausgeübt, einschließlich des Postprocessings der 3D-Scans, welches das Bereinigen von Rohdaten, Mergen, und das Zusammenfügen virtueller 3D-Modelle von gescannten Monumenten umfasst. Abschließend wird mit 3D-Modellen bereits gescannter Abgüsse der Umgang mit 3D-Modellen als Quellen zur Inschriftenforschung geübt.

**Datenbanken:** In diesem Arbeitsbereich erfolgen nach Einweisung und Einarbeitung in die Bilddatenbank ConedaKOR und die virtuelle Forschungsumgebung TextGrid zunächst Recherchen zu den digitalisierten Objekten, die in ConedaKOR eingepflegt werden. Darauf erfolgt dann die Angleichung dieser recherchierten Inhalte bezüglich Aufnahmeort und Artefakttyp an die TextGrid-Konzepte PlaceType bzw. ArtefactType unter Verwendung der in TextGrid angelegten Vokabulare. Parallel dazu erfolgt die Eingabe der entsprechenden ConedaKOR-Inhalte in TextGrid, um sie in ConedaKOR anzugleichen und verlinkbar zu machen. Ggf. sind Pflege und Korrektur der Inhalte in ConedaKOR durchzuführen.

Des Weiteren sollen mittels Literaturoauswertung weitere nicht-textuelle Informationen zu Inschriftenträgern gesammelt werden und deren Einpflege in TextGrid dort bereits bestehende Datensätze vervollständigen.

**Öffentlichkeitsarbeit und Website:** Die Arbeiten des vorletzten Bereichs umfassen Pflege, Korrektur, Übersetzung und Ergänzung bestehender Inhalte auf der Webseite sowie in sozialen Medien.

**Epigraphik:** Im letzten Arbeitsbereich werden um epigraphische Kenntnisse zu vertiefen, unter Anwendung von Zeichenkatalogen die Zeichen ausgewählter Inschriften klassifiziert und numerische Transliterationen erstellt. Auch sollen Vorschläge für Zeichen, die in keinem Katalog erfasst sind, erörtert werden. Zudem wird das Erstellen von digitalen Zeichnungen mittels Graphiktablet und entsprechender Software vermittelt.

## Lehre

Im Rahmen der Lehre wurde das Projekt eingeladen, an der Unterrichtseinheit „3D scanning and imaging“ des Seminars *Sunoikisis Digital Cultural Heritage* zu partizipieren. *Sunoikisis* ist ein nationales Konsortium für klassizistische Programme, das am Center for Hellenic Studies der Harvard University entwickelt wurde. *SunoikisisDC* hat seinen Sitz am Alexander von Humboldt Chair of Digital Humanities an der Universität Leipzig. Die Einheit wurde gemeinsam von Dr. Sven Gronemeyer, Professor Dr. Graeme Earl (King's College London), Dr. Gabriel Bodard und Dr. Valeria Vitale (University of London) durchgeführt. Das Seminar fand virtuell über eine Videokonferenzplattform statt und wurde live auf YouTube übertragen und dort auch archiviert.<sup>3</sup>

---

<sup>3</sup> <https://youtu.be/NZK1yoXdAVY>.

## Vorträge, Präsentationen, Publikationen, Kooperationen, Außenwirkung des Projekts

### Eigene Tagungen

Vom 7. - 9. Dezember 2017 fand in Bonn die dritte Projektjahrestagung<sup>4</sup> mit dem wissenschaftlichen Beirat statt. Im Rahmen dieser Tagung konnten wir zwei weitere Mitglieder für unseren wissenschaftlichen Beirat gewinnen: Prof. Dr. Alexandre Tokovinine (Tuscaloosa) und Prof. Dr. Dmitri Beliaev (Moskau). Im Fokus der Tagung standen Präsentationen der Projektmitglieder, um über den aktuellen Arbeitsstand zu berichten. Am zweiten Tag organisierten wir einen Anwenderworkshop, um die anwesenden Kollegen in die Nutzung unserer Forschungsdatenbank IDIOM einzuweisen. Die Präsentation neuester Forschungsergebnisse auf dem Gebiet der Mayaepigraphie durch alle anwesenden Wissenschaftler am Samstag bildeten den Abschluss unserer Jahrestagung.

### Externe Tagungen

21. Januar 2017 **„XX. Mesoamerikanistik-Tagung“** Jahrestagung, Rautenstrauch-Joest Museum/Abteilung für Altamerikanistik, Rheinische Friedrich-Wilhelms-Universität Bonn, Köln/Bonn, 20. - 23. Januar 2017 *"Zeichen, Grapheme, Varianten und Metadaten: ein digitaler Zeichenkatalog des Maya"*

15. Februar 2017 **„DHd 2017: Digitale Nachhaltigkeit“** Jahrestagung, Universität Bern, Bern, 13. - 18. Februar 2017 *"Nachnutzungs- und Nachhaltigkeitsstrategien im Projekt Textdatenbank und Wörterbuch des Klassischen Maya"*

24. April 2017 **„Lateinamerikanische Perspektiven“**, Vortragsreihe in Kooperation mit der VHS Bonn, Abteilung für Altamerikanistik der Universität Bonn und dem Ibero-Club Bonn e. V., Bonn *"Mayahieroglyphen im digitalen Zeitalter Das Bonner Mayawörterbuch-Projekt"*

27. April 2017 **„Digital Humanities Kolloquium der Uni Köln SoSe 2017“**, Universität Köln, Köln *"Entzifferung der Mayahieroglyphen im digitalen Zeitalter: Herausforderungen, Methoden, Ergebnisse"*

30. August 2017 **„Digital Cultural Heritage (DCH2017)“** Jahrestagung, Staatsbibliothek zu Berlin, Berlin, 30. August - 1. September 2017 *"Cracking the Code: An Ontological Interlinked Working Environment for the Analysis of Classic Mayan Language and Script"*

29. September 2017 **„Modellierungsfragen in den Digital Humanities“** Workshop und Informatik Tagung: Chemnitz, 26. - 29. September 2017, TU Chemnitz *"Modellierung eines digitalen Zeichenkatalogs für die Hieroglyphen des Klassischen Maya - Ein neues Konzept zur Klassifikation von Schriftzeichen sowie der qualitativen Bewertung und Einstufung von Entzifferungshypothesen"*

7. Oktober 2017 **„Visible Words: Digital Epigraphy in a Global Perspective & Epidoc Editing Workshop“**, Center of Digital Epigraphy, Brown University, Providence, 5. - 7. Oktober 2017 *"Maya Hieroglyphic Writing and its Digital Exploration: The Bonn Classic Mayan Dictionary Project"*

19. Oktober 2017 **„Geisteswissenschaftliche Forschungsdaten - Methoden zur digitalen Erfassung, Aufbereitung und Präsentation“**, Workshop und Tagung, Akademie der Wissenschaften und der Literatur Mainz und der Union der deutschen Akademien der Wissenschaften, Mainz, 18. - 20.

<sup>4</sup> Programm: <http://mayawoerterbuch.de/third-annual-project-workshop/>.

Oktober 2017 *“Digitale Erforschung des Klassischen Maya im Bonner Mayawörterbuch-Projekt: eine Blaupause für die moderne Epigraphie?”*

10. November 2017 **„3D Imaging in Cultural Heritage“** Tagung, British Museum, London, 9. - 10. November 2017 *“Old Casts in a New Light: 19th Century Plaster Casts of Classic Maya Artefacts as Cultural Heritage and Efforts Towards their Digital Preservation and Analysis”*

1. Dezember 2017 **„Wissensorganisation ‘17 - Knowledge Organization for Digital Humanities“** Tagung, Freie Universität Berlin, 30. November – 1. Dezember 2017 *“Ein digitaler Zeichenkatalog als Organisationssystem für die noch nicht entzifferte Schrift der Klassischen Maya”*

## Kooperationen

Zusätzlich zu den bereits bestehenden Kooperationen konnten wir die nachfolgend genannten Personen und Institutionen für Kooperationen hinsichtlich der Ergänzung unserer Forschungsinfrastruktur mit für die linguistische Analyse und Lexikographie, die Archivierung von Bilddokumenten, sowie der Analyse von 3D-Daten essentiellen Tools gewinnen:

Dr. Cristina Vertan (*TraCES: From Translation to Creation: Changes in Ethiopic Style and Lexicon from Late Antiquity to the Middle Ages*, Universität Hamburg), passt das Annotationstool GeTa für die Annotation von Mayainschriften an (s. unter *Annotationstool ALMAH*).

Zum Aufbau eines Mayabildarchivs erfolgt die Zusammenarbeit mit *ConedaKOR* (Frankfurt mit *Max Weber Stiftung*, Bonn und *DAASI International GmbH*, Tübingen) zwecks Erstellung und Nutzung einer Open Source Datenbankinstanz.

In Zusammenarbeit mit Dr. Hubert Mara (*Forensic Computational Geometry Laboratory, Interdisciplinary Center for Scientific Computing*, Universität Heidelberg) ist die Entwicklung und Anwendung von Algorithmen für die automatisierte Erkennung von Mayaschriftzeichen und Textmustern auf der Basis von 3D Modellen sowie der Entwicklung von Metadatenstandards für Zitierung, Archivierung und Nachnutzung von 3D Modellen geplant. Bisher unterstützte er das Projekt mit topometrischen Analysen der Scandaten.

Die bereits zuvor aufgeführten 3D-Scanvorhaben im Jahr 2017 erfolgten im Rahmen von Kooperationen mit verschiedenen Institutionen in Deutschland, Großbritannien und Guatemala, die zum einen Zugang zu ihren Beständen an Artefakten mit Mayainschriften gewährt haben, und zum anderen uns hinsichtlich technischer Infrastruktur bei den Scanarbeiten unterstützt haben. Erstere umfassen das *Museum of Archaeology and Anthropology* (Cambridge), das *British Museum* (London), sowie das *Historische Museum der Pfalz* (Speyer) zusammen mit dem *Instituto de Antropología e Historia*, der nationalen Denkmalbehörde Guatemalas, und der *Fundación Ruta Maya*, einer privaten Kulturstiftung in Guatemala.

Desweiteren haben dieses Jahr bereits folgende Institutionen ihre Zusammenarbeit mit dem Projekt hinsichtlich des 3D-Scannings ihrer Inschriftenartefakte ab 2018 zugesagt: *Ethnologisches Museum* (Berlin), *Linden-Museum* (Stuttgart), *Rautenstrauch-Joest-Museum* (Köln), *Museum zu Allerheiligen* (Schaffhausen), *Museum Rietberg* (Zürich).

## Außenwirkung des Projekts

Neben der Präsentation auf Fachtagungen hat das Projekt eine vermehrte Öffentlichkeitswirkung in Printmedien und TV und Rundfunk entfalten können. Dies betrifft insbesondere die Arbeiten zum 3D-

Scanning in Speyer, die durch das Museum in einer Pressemitteilung angekündigt wurden. Dazu gab es insgesamt vier Artikel in lokalen und regionalen Tageszeitungen.

Weiterhin wurde unsere Arbeit im öffentlich-rechtlichen Rundfunk in zwei Formaten ausgiebig dargestellt. Im Rahmen des Magazins **“IQ - Wissenschaft und Technik”** von Radio Bayern 2 wurde das Projekt am 1.3. um 18:05 in einem 4:53-minütigen Beitrag insbesondere mit dem Fokus auf die Datenbank und das Wörterbuch vorgestellt. Für die Abendausgabe der rheinland-pfälzischen Regionalnachrichten **“SWR Aktuell”** vom 7.2. um 19:30 gab es einen 3:05-minütigen Fernsehbeitrag, der insbesondere auf die Dokumentation in der Ausstellung in Speyer fokussiert hat.

Um auch während der Öffnungszeiten des Museums dokumentieren zu können, wurde mit dem museumspädagogischen Dienst vereinbart, dass wir während der Arbeit Besuchern für Fragen zur Verfügung stehen, was jeweils durch Aushang angekündigt wurde. Dadurch konnte unsere Reichweite kollektiv für die Museumsbesucher der jeweiligen Tage erweitert werden, als auch individuell durch persönlichen Kontakt mit einem der drei Teammitglieder vor Ort.

## Publikationen

**Diehr, Franziska, Maximilian Brodhun, Sven Gronemeyer, Katja Diederichs, Christian Prager, Elisabeth Wagner & Nikolai Grube**

Modellierung eines digitalen Zeichenkatalogs für die Hieroglyphen des Klassischen Maya. In: *Informatik 2017, Lecture Notes in Informatics*, 275, herausgegeben von Maximilian Eibl & Martin Gaedke, pp. 1185-1196. Gesellschaft für Informatik, Bonn.

**Diehr, Franziska, Maximilian Brodhun, Sven Gronemeyer, Katja Diederichs, Christian Prager, Elisabeth Wagner & Nikolai Grube**

Ein digitaler Zeichenkatalog als Organisationssystem für die noch nicht entzifferte Schrift der Klassischen Maya. In: *Proceedings of Wissensorganisation 2017*, herausgegeben von Christian Wartena, Michael Franke-Meyer & Ernesto De Luca. (noch nicht veröffentlicht)

**Feldmann, Felix, Bartosz Bogacz, Christian Prager & Hubert Mara**

Histogram of Oriented Gradients for Maya Glyph Retrieval. In: *GCH 2017 Eurographics Workshop on Graphics and Cultural Heritage*, herausgegeben von Dieter Fellner, pp. 115-118. Eurographics Association, Goslar.

**Gronemeyer, Sven**

Die abgesagte Apokalypse: Der Blick der vorspanischen Maya auf das Ende 13. Bak'tun und das autochthone Konzept von Prophetie. In: *2012 – Die globalisierte Apokalypse aus lateinamerikanischer Perspektive, Interdisziplinäre Studien zu Lateinamerika*, 1, herausgegeben von Antje Gunsenheimer, Monika Wehrheim, Mechthild Albert & Karoline Noack, pp. 45-66. Bonn University Press & V&R unipress, Göttingen.

**Grube, Nikolai & Octavio Quetzalcoatl Esparza Olguín**

*Two Captives from Uxul*. Textdatenbank und Wörterbuch des Klassischen Maya, Research Note 6.

**Grube, Nikolai, Christian Prager, Katja Diederichs, Sven Gronemeyer, Elisabeth Wagner, Maximilian Brodhun & Franziska Diehr**

*Milestone Report 2014-2016*. Textdatenbank und Wörterbuch des Klassischen Maya, Project Report 4.

**Prager, Christian & Elisabeth Wagner**

Historical Implications of the Early Classic Hieroglyphic Text CPN 3033 on the Sculptured Step of Structure 10L-11-Sub-12 at Copan. Textdatenbank und Wörterbuch des Klassischen Maya, Research Note 7.

**Wagner, Elisabeth**

*Jun Yop Ixiim – Another Appellative for the Ancient Maya Maize God*. Textdatenbank und Wörterbuch des Klassischen Maya, Research Note 8.

## Website und Soziale Medien

Im Laufe des zweiten Quartals wurde ein Facelift der Website in Zusammenarbeit mit dem Projektmanagementbüro Beuse, Köln, entwickelt. Üblicherweise geht man bei Webauftritten von einer designerischen Lebensdauer von nicht mehr als drei Jahren aus, nachdem die Website offiziell im Dezember 2015 gelauncht wurde. Wesentlich problematischer waren technische Aspekte. Projektspezifische Anpassungen von Plugins und Templates waren seit Anfang 2017 nicht mehr mit dem neuen Standard PHP 7 und den neuesten Versionen von WordPress kompatibel, was zu Sicherheitslücken führte.

Auf Empfehlung wurde neu das Plugin "Essential Grid" eingeführt, das ein modernes, frei konfigurierbares Kacheldesign umsetzt. Anstatt durch das Projektmanagementbüro Designänderungen templateseitig umsetzen lassen zu müssen, können diese nun kostenneutral über die Redaktionsoberfläche vorgenommen werden, wodurch die Webseite weniger wartungsintensiv geworden ist. Da das Plugin mit WordPress weiterentwickelt wird, ist zukunftsicher eine Aufwärtskompatibilität gegeben. Das neue Plugin setzt weiter ein "responsive design" um, das sich dynamisch an die Bildschirmauflösungen mobiler Endgeräte anpasst und damit der zunehmenden Nutzung von Tablets und Smartphones zum Browsen gerecht wird. Der Facelift konnte im Juli ausgerollt werden.

In den Kanälen der sozialen Netzwerke hat das Projekt auch 2017 neue Publikationen, Vortragsankündigungen und 3D-Modelle verbreitet. Infolge der Scantätigkeiten in Großbritannien und vor allem in Speyer wurden hier Einblicke in unsere praktische Arbeit in Foto und Video geteilt, aber auch von internen Workshops zur Entwicklung der digitalen Infrastruktur. Für Facebook zeigen die internen Statistiken, dass ein Post üblicherweise zwischen 250 und 600 Personen erreicht (bei 225 Followern), in einem Fall erreichte ein Post über 11.600 Personen.

## Literaturverzeichnis

### **Barrera Vásquez, Alfredo**

1980 *Diccionario Maya Cordemex: Maya - Español, Español - Maya*. Ediciones Cordemex, Mérida.

### **Diehr, Franziska, Maximilian Brodhun, Sven Gronemeyer, Katja Diederichs, Christian Prager, Elisabeth Wagner & Nikolai Grube**

2017 Modellierung eines digitalen Zeichenkatalogs für die Hieroglyphen des Klassischen Maya. In: Eibl, M. & Gaedke, M. (Hrsg.), *INFORMATIK 2017*. Gesellschaft für Informatik, Bonn. (S. 1185-1196).

### **Kelley, David**

1962 Review of A Catalog of Maya Hieroglyphs, by J. Eric S. Thompson. *American Journal of Archaeology* 66, 436-438, 1962.

### **Mayer, Karl Herbert**

1978 *Maya Monuments: Sculptures of Unknown Provenance in Europe*. Acoma Books, Ramona, CA.

1980 *Maya Monuments: Sculptures of Unknown Provenance in the United States*. Acoma Books, Ramona, CA.

1984 *Maya Monuments: Sculptures of Unknown Provenance in Middle America*. Verlag von Flemming, Berlin.

1987 *Maya Monuments: Sculptures of Unknown Provenance, Supplement 1*. Verlag von Flemming, Berlin.

1989 *Maya Monuments: Sculptures of Unknown Provenance, Supplement 2*. Verlag von Flemming, Berlin.

1991 *Maya Monuments: Sculptures of Unknown Provenance, Supplement 3*. Verlag von Flemming, Berlin.

1995 *Maya Monuments: Sculptures of Unknown Provenance, Supplement 4*. Academic Publishers, Graz.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party Material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the Material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>