

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 4 | DE



Published 11 Apr 2017

DOI: 10.20376/IDIOM-23665556.17.pr004.de

Meilenstein-Bericht 2014-2016

Nikolai Grube¹, Christian Prager¹, Katja Diederichs¹, Sven Gronemeyer^{1,2}, Elisabeth Wagner¹, Maximilian Brodhun³ & Franziska Diehr³

¹ Rheinische Friedrich-Wilhelms-Universität, Bonn

² La Trobe University, Melbourne

³ Niedersächsische Staats- und Universitätsbibliothek, Göttingen

I. Beschreibung und Zielsetzung des Vorhabens

Die aus Wort- und Silbenzeichen bestehende Hieroglyphenschrift der Klassischen Maya gehört zu den bedeutendsten Schrifttraditionen der antiken Welt. Schrift als graphische Realisierung von Sprache medialisiert menschliches Denken, Kommunikation und kulturelles Wissen in Form von Texten. Ihre Erschließung ermöglicht es, Ideen, Werte, Vorstellungen und Überzeugungen zu rekonstruieren und damit in die Lebenswelt und das Gedächtnis vergangener Gesellschaften hineinzublicken. Voraussetzung dafür ist, dass das Schriftsystem und die zugrunde liegende Sprache bekannt sind und der kulturelle Code entschlüsselt ist. Für das Klassische Maya ist der Durchbruch bei der Entzifferung gelungen, aber trotz großer Fortschritte in den vergangenen Jahrzehnten bleiben bis heute etwa 40% der über 800 Zeichen unlesbar. Bisher wurde das Textmaterial noch nicht in seiner Gesamtheit epigraphisch erschlossen und Vorkommensnachweise von Zeichen fehlen. Selbst dort, wo die Schriftzeichen lesbar sind, entziehen sich Texte unserem Verständnis, weil das Klassische Maya nicht überliefert ist, sondern nur aus dem Vergleich der 30 heute noch gesprochenen Mayasprachen rekonstruiert werden kann. Vieles vom kulturellen Vokabular der vorspanischen Zeit ist als Folge der europäischen Kolonisierung verloren gegangen. Die vollständige Dokumentation und Entzifferung der rund 10.000 Hieroglyphentexte, die Rekonstruktion der in ihr enthaltenen Sprache und deren Dokumentation in einem Wörterbuch sind daher notwendige Voraussetzungen für ein tieferes Verständnis der Mayakultur, der Geschichte ihrer Königsdynastien und ihrer Religion.

Gegenstand des Projektes ist ein unvollständig entziffertes, komplexes Schriftsystem und unser Ziel ist es, dieses weitgehend mit Hilfe digitaler Werkzeuge zu entschlüsseln und die zugrunde liegende Sprache in einem Wörterbuch zu beschreiben. Hierzu werden die Hieroglyphentexte maschinenlesbar gemacht, in einer Textdatenbank mit Analysen und Kommentaren gespeichert und das Klassische Maya in einem webbasierten Wörterbuch mit Originalschreibung abgebildet, damit die Nutzer diese mit der Analyse vergleichen können. Bis heute existiert kein Projekt auf dem Gebiet der digitalen Schriftforschung, das vergleichbare Voraussetzungen, Zielsetzungen und Anforderungen aufweist und

als Modell für die Konzipierung und Entwicklung unserer Datenbank dienen könnte. Forschungsprojekte zu griechischen¹, lateinischen² oder altägyptischen³ Sprachdenkmälern stehen bei der Realisierung ihrer Datenbanken nicht vor der Herausforderung, dass die Schrift und die ihr zugrunde liegende Sprache gar nicht oder nur teilweise entziffert sind. Unser Ziel ist es, mit den zu entwickelnden digitalen Tools Zeichenlisten zu erstellen und neu klassifizierte Zeichen darin zu erfassen, die Texte maschinenlesbar zu machen, Lesungen zu gewinnen und den Sprachschatz in Originalschreibung zu dokumentieren. Dieser innovative Charakter unseres Projekts erfordert eine flexible Prozessgestaltung bei der Entwicklung der digitalen Infrastruktur, beim Einsatz der finanziellen Mittel, sowie methodische Grundlagenarbeit zur Entwicklung nachhaltig nutzbarer Arbeitskonzepte und wirkt sich dadurch auf die Projektarbeit, den Zeitplan sowie die Personalplanung in den ersten beiden Arbeitspaketen aus.

Gerade die Schwerpunkte digitale Epigraphik, Entwicklung von Datenbanken, sowie nachhaltige und interoperable Speicherung von Forschungsdaten unterstreichen die enorme Bedeutung der digitalen Geisteswissenschaften in unserem Projekt. Die Arbeitsstelle kooperiert deshalb mit der Niedersächsischen Staats- und Universitätsbibliothek Göttingen (SUB), welche die informationswissenschaftlichen und -technologischen Kompetenzen im Projekt vertritt. Das Projekt ist interdisziplinär ausgerichtet und arbeitet am Schnittpunkt zwischen angewandter Informatik und den Kulturwissenschaften. Das Projekt bildet nicht nur eine Nahtstelle beider Disziplinen im Bereich der digitalen Epigraphik, vielmehr erwerben unsere fach- und informationswissenschaftlichen Mitarbeiter durch ihre intensive Zusammenarbeit fächerübergreifende Qualifikationen, die sie nicht nur im Forschungsprojekt, sondern auch in der Lehre einbringen. Zu den gemeinsamen Arbeiten im Projekt gehören die Dokumentation der Forschung zur Mayakultur in einer Literaturdatenbank, das Sammeln, Archivieren und Systematisieren von Hieroglyphentexten, die Entwicklung von Metadatenmodellen für die Beschreibung und Annotation der Texte, deren inschriftenkundliche und sprachwissenschaftliche Analyse, sowie die Anlage dieser Forschungsdaten in einer Text- und Textträgerdatenbank, aus der sich das Wörterbuch speisen wird. Darin finden sich nicht nur die Originalschreibungen, sondern auch Varietäten und Entwicklungen von Lexikon, Grammatik und Schrift werden abgebildet und mit Literaturangaben referenziert.

Um diese Aufgaben zu bewältigen und die große Datenmenge optimal zu verarbeiten, werden Methoden der digitalen Geisteswissenschaften eingesetzt um mit Hilfe von *Text Mining*-Verfahren Textmuster zu finden; Zeichen, Zeichenfolgen oder Passagen herauszuschälen, Entzifferungen in ihrem jeweiligen Verwendungskontext zu prüfen, neue Lautwerte zu bestimmen und den Wortschatz des Klassischen Maya korpuslinguistisch in seinen zeitlichen, räumlichen und gesellschaftlichen Dimensionen zu erfassen. Die digitale Arbeitsumgebung wird in der Virtuellen Forschungsumgebung (engl.: Virtual Research Environment, kurz: VRE) TextGrid realisiert und unsere Anpassungen sowie epigraphischen Tools laufen dort unter der Bezeichnung *Interdisciplinary Database of Classic Mayan (IDIOM)*. Die Forschungsdaten werden langfristig im TextGrid-Repository angelegt; sie werden vollständig über ein Webportal frei verfügbar gemacht. In Kooperation mit der Universitäts- und Landesbibliothek Bonn (ULB) werden ausgewählte Inhalte aus dem TextGrid-Repository in einem virtuellen Inschriftenarchiv in den Digitalen Sammlungen der ULB publiziert, wie zum Beispiel Abbildungen der Inschriften, kurze Informationen dazu, sowie Analysen und Übersetzungen.

Textdatenbank und Wörterbuch dienen nicht nur zur Dokumentation des Klassischen Maya, sondern sind unsere Werkzeuge für die Entzifferung der Mayaschrift: neue Lautwerte oder Lesungshypothesen für Wort- und Silbenzeichen werden erschlossen und laufend im digitalen Zeichenkatalog erfasst. Sie

¹ <http://stephanus.tlg.uci.edu/>

² <http://www.thesaurus.badw.de/>

³ <http://aaew.bbaw.de/tla/>

können direkt am Textmaterial auf ihre Plausibilität überprüft werden. Datenbank und Wörterbuch berücksichtigen in ihrer Architektur und Realisierung immer den aktuellen Stand der Schriftentzifferung und bilden damit die Dynamik der Mayaschrift-Forschung ab. Weil die Schrift nur teilweise entziffert ist und viele Texte unvollständig lesbar sind, ist es notwendig, sämtliche verfügbaren Informationen zum Schrifträger zu dokumentieren, *denn die Bedeutung von Wörtern ist ihre Verwendung*. Aus diesem Grund erfassen wir Informationen über den Schrifträger und seinen Kontext (so genannte nicht-textuelle Informationen) und verbinden diese mit der Analyse von Schrift und Sprache. Texte der Maya beziehen sich häufig auf den Textträger, seinen Standort und Auftraggeber, weshalb die Materialität, das kulturelle und gesellschaftliche Umfeld von Textträgern wichtige Informationen sind. Wir berücksichtigen bei der Konzipierung und Umsetzung der VRE diese komplexen Informationszusammenhänge und schenken daher ihrer Entwicklung und Anpassung an die Bedürfnisse unseres Projekts erhöhte Aufmerksamkeit.

2. Projektphase I (2014 - 2016) - Dokumentation und Erschließung von Textträgern

Die Dokumentation und Erschließung von Textträgern bilden die Grundlage der Textanalyse, die zentraler Bestandteil der Arbeitspakete von 2017 bis 2028 ist. Aus diesem Grund begann die Projektarbeit mit der Konzeption und der darauf aufbauenden Umsetzung der informationstechnologischen Infrastruktur. Um das Wörterbuch des Klassischen Maya zu erstellen und die Bedeutung von Wörtern zu analysieren, dürfen die Texte nicht losgelöst von ihrem Trägerobjekt und dessen räumlich-zeitlichen Kontext betrachtet werden: nicht-textuelle Informationen, also Metadaten über den Textträger selbst, seinen Standort, Nachbarmonumente und Begleitfunde, Auftraggeber und den historischen Rahmen sind für die Entzifferung und Interpretation der Inschriften enorm wichtig. Daher ist die sorgfältige Dokumentation dieser Metadaten in der Datenbank Voraussetzung für die Entzifferung und das Verständnis der Mayatexte.

Neben der Textdatenbank selbst ist die Möglichkeit, Beziehungen zwischen mehreren Texten und Textträgern herzustellen und abzufragen, ein wesentlicher Bestandteil der Datenbank, nicht nur, weil sich ein Text über mehrere Textträger erstrecken kann. Erst mit der Erfassung granularer Metadaten wird dies ermöglicht und Relationen zwischen Personen, Ereignissen und Orten können dadurch sichtbar gemacht werden. Die Datenbank enthält aber nicht nur Beschreibungen über die Textträger und den Inhalt der Texte, sondern mit Hilfe der Literaturdatenbank gewinnt der Nutzer einen Überblick, welche Autorinnen und Autoren sich mit einem Monument befasst oder es publiziert haben, eine Textpassage diskutiert oder erstmals eine bis heute gültige sprachliche Lesung einer Hieroglyphe bzw. eines Zeichens der Öffentlichkeit präsentiert haben. Bei der Projektarbeit ergibt sich durch die engmaschige Datenvernetzung somit ein Mehrwert in der Nutzung der Text- und Objektdatenbank, die unmittelbar das Quellenmaterial für das eigentliche Projektziel zur Verfügung stellt: das Wörterbuch des Klassischen Maya.

2.I. Aufbau der technischen Umgebung zur Erfassung der Textträger

Um eine Erfassungsumgebung zu schaffen, die den fachwissenschaftlichen Anforderungen der Bonner Arbeitsstelle gerecht wird, ist ein enger Austausch der im Projekt vertretenen Wissenschaftsdisziplinen notwendig. Die gegenseitige Vermittlung von Wissen (Epigraphik, Linguistik, Altamerikanistik auf der einen - Datenmodellierung, Vokabularentwicklung und Informatik auf der anderen Seite), die kollaborative Definition und Spezifizierung von Anforderungen und das Erarbeiten informationstechnologischer Möglichkeiten sind Prozesse, die den Informations- und Wissenstransfer im Projekt befruchten und vor allem maßgeblich dazu beitragen, das gewünschte Ergebnis zu erzielen.

2.1.1. Fachwissenschaftliche Anforderungsanalyse und Entwicklung des Metadatenschemas

Anhand eines gemeinsam spezifizierten und fachlich diskutierten Anforderungskatalogs modellierte die Göttinger Arbeitsstelle das Metadatenschema zur Erfassung der Textträger und des historischen Kontexts. Die Entwicklung gestaltete sich als iterativer Prozess, durch den die fachlichen Anforderungen nach und nach definiert werden konnten. Um eine hohe Qualität der eigenen Konzepte zu garantieren, wurde das Schema auf Basis internationaler und etablierter Standards aufgebaut. Die Nachnutzung bereits bestehender und fachlich anerkannter Terme trägt darüber hinaus zu einer hohen Interoperabilität mit anderen Datenbeständen und Informationssystemen bei. Das ausgestaltete Schema hat eine ontologisch-vernetzte Struktur, die komplexe Beziehungen und Zusammenhänge abbildet.⁴

2.1.2. Entwicklung kontrollierter Vokabulare

Für die fachwissenschaftliche Erschließung der Textträger wurden insgesamt 10 mehrsprachige Thesauri in interdisziplinärer Zusammenarbeit entwickelt. Die Aufgabe der Bonner Arbeitsstelle bestand in Recherche, Auswahl und Definition fachlich korrekter Begriffe, während in Göttingen die Vokabularentwicklung durch die Anwendung dokumentationswissenschaftlicher Methoden unterstützt wurde.

Zur Auswahl geeigneter Terme wurden alle bisher in der Literatur verwendeten Fachbegriffe auf ihre Plausibilität, Vergleichbarkeit und ihre Nutzbarkeit hin überprüft. Die so entstandene Begriffssammlung wurde nach terminologischen Prinzipien geordnet und zur maschinenlesbaren Repräsentation und Integration in das Metadatenschema im Format SKOS (Simple Knowledge Organisation System)⁵ modelliert. Dadurch konnte gleichzeitig ein Mapping der Begriffe zu den Normdaten des Getty-Thesaurus⁶ realisiert werden, das die Referenzierung der nachgenutzten Konzepte ermöglicht.

Die Entwicklung der Vokabulare hat nicht nur für die Arbeit im projekteigenen Kontext, sondern auch darüber hinaus für den gesamten Fachbereich eine hohe Relevanz. Bisher existierten in der Mayaschriftforschung zahlreiche Begriffe, Vokabularien und Beschreibungsschemata, was zu einer großen Bandbreite disparat dokumentierter Schriftträger führte. Die verwendeten Begrifflichkeiten zeigen teilweise wenig Übereinstimmung und sind oft unvollständig, fehlerhaft, unpräzise oder stark vereinfachend.

Mit der Entwicklung der Vokabulare leistet das Projekt einen maßgeblichen Beitrag zur Standardisierung im Fachbereich der Mayaschriftforschung, da es die bereits in der Fachwelt etablierten Begriffe nachnutzt, diese erstmals eindeutig definiert und sie darüber hinaus in einen terminologischen Zusammenhang stellt.

Die Publikation der Vokabulare erfolgt sowohl in maschinen- als auch menschenlesbarer Form über die Website des Projekts und ist für Ende 2016 geplant.

⁴ Die vollständige Dokumentation ist erreichbar unter <http://idiom-projekt.de/idiommask/schema.html>

⁵ <https://www.w3.org/TR/2009/REC-skos-reference-20090818/>

⁶ <http://www.getty.edu/research/tools/vocabularies/>

2.1.3. Entwicklung und Konzeption der technischen Infrastruktur

Im Projekt werden Daten verschiedener Art erstellt und gespeichert. Bilddaten, Metadaten- und Textanalysedateien müssen zusammenhängend in einer Infrastruktur verwaltet werden. Dies betrifft die Erstellung, Speicherung, Bearbeitung und die Zugriffsregelung. Für diese Aufgaben wird die VRE TextGrid benutzt. Das Frontend TextGrid-Laboratory (TG-Lab) ermöglicht die Erstellung und Bearbeitung von Dateien und ein feingranulares Rechtemanagement. Durch das Backend besteht Zugriff auf ein Repositorium (TG-Rep), das die Daten in einer sicheren Umgebung speichert. Diese wird von der GWDG bereitgestellt und gewährleistet die Datenspeicherung durch Methoden der Langzeitarchivierung.

Aufgrund ihrer komplex vernetzten Struktur benötigen die Dateien mit den Metadaten der Textträger eine adäquate Speicherung. Das bietet das Format RDF (Resource Description Framework)⁷ und die Speicherung in einer Graphdatenbank in Form eines Triplestores.

Zur benutzerfreundlichen Erfassung der Metadaten wird eine Eingabemaske verwendet. Diese basiert auf HTML und JavaScript und stellt dem Anwender mehrere Eingabehilfen zur Verfügung. Als Plug-In für das TG-Lab realisiert, kann die Eingabemaske installiert und direkt aus dem TG-Lab verwendet werden. Suchen in internen und externen Datenbanken (um Objektrelationen herzustellen), Validierung der Eingabefelder, automatisierte Umrechnung von Datumsformaten sind Beispiele für diese unterstützenden Funktionen. Eine weitere wichtige Eingabeunterstützung ist die Navigation in der Vokabularhierarchie. Somit muss der Anwender nicht erst nach konkreten Schlagworten suchen. Auf dem Projekt-Server werden die Triplestores für die Speicherung der Metadaten und der Vokabulare, die Eingabemaske und die Projekt-Webseite zur Verfügung gestellt.

2.2. Fachwissenschaftliche Tiefenerschließung von Text- und Bildträgern und des historischen Kontexts

Die Inschriften- und Bildträger werden von der Bonner Arbeitsstelle seit November 2015 mit Hilfe des TG-Lab und der Eingabemaske dokumentiert. Dafür wird von den wissenschaftlichen Mitarbeitern (mit Unterstützung der WHKs und SHKs) eine systematische Recherche aller verfügbaren Publikationen, Datenbanken und Webseiten von Sammlungen, Museen und anderen Forschungseinrichtungen durchgeführt. Diese Daten bilden die Grundlage für die Erschließungsarbeit, die sich grob in drei Bereiche teilen lässt:

1. Dokumentation der Artefakte (deren Titel, Identifikatoren, Objektart und -form, Abmessungen, Erhaltungszustand, Herstellungs-, Fund-, Erwerbungs- und Aufbewahrungskontext);
2. Archäologisch relevante Orte (mit archäologischen und geographischen Koordinaten, normierten und alternativen Ortsnamen) und deren Einordnung in eine komplexe Ortshierarchie (von der Fundstelle, deren Verortung innerhalb des architektonischen Kontexts in einem Bauwerk, dessen Lage in einer Gebäudegruppe innerhalb der Fundstätte, sowie deren Lage innerhalb der Hierarchie moderner politisch-administrativer Gebietskörperschaften);
3. Einbettung in den historischen Kontext: Ereignisse und Personen (z.B. Kriege, Weihe von Monumenten, Inthronisationen, Herrscherbiographien, soziopolitische Verhältnisse).

⁷ <https://www.w3.org/TR/rdf-schema/>

Zusätzlich wird die Erfassung durch die Verwendung kontrollierter Vokabulare und den Verweis auf Normdaten⁸ unterstützt. Weiterhin können die Informationen mit Quellenangaben versehen werden, die sich aus der vom Projekt gepflegten Literaturdatenbank in Zotero speisen.

Die ontologisch-vernetzte Datenstruktur ermöglicht es, komplexe Fragen an das Material zu stellen, so z.B.: Welche Schriftträger befinden sich auf der Großen Plaza von Tikal? Wann wurden diese eingeweiht, wer hat sie in Auftrag gegeben und was sind die dort erwähnten historischen Ereignisse z.B. im Vergleich mit den Altären (Text-Objekt-Relation)?

Inhaltlich wurden die Inschriften der Fundstätte Tikal, Guatemala, und damit verbundene Personen und Ereignisse erschlossen. Tikal ist eine der größten und bedeutendsten Städte der klassischen Mayakultur. Im Gegensatz zum im Antrag genannten Bundesstaat Campeche, Mexiko, wurde Tikal als erster Use Case für die Tiefenerschließung herangezogen, da sich die Fundstätte aufgrund ihres guten Dokumentationsstands in der Forschungsliteratur optimal eignet, um die fachwissenschaftlichen Anforderungen an die informationstechnologische Umgebung (u.a. Metadatenchema, Eingabemaske, Abfrage der Daten) zu überprüfen. Im Gegensatz zu Tikal verteilen sich die Textträger aus Campeche auf eine Vielzahl von Fundstätten und sind teilweise nur oberflächlich in der Forschungsliteratur beschrieben.

Darüber hinaus wurde bereits mit der Erfassung weiterer Artefakte aus anderen Inschriftenorten begonnen. Aktuell⁹ umfasst die Datenbank in unterschiedlichem Erschließungsgrad die folgende Anzahl an Objekten: 597 Artefakte in Verbindung mit 76 Fund-, 88 Herstellungs-, 2 Erwerbungs-, und 56 Aufbewahrungsereignissen sowie 439 Orte. Zum forschungs- und objektgeschichtlichen Kontext der Artefakte wurden weiterhin jeweils 25 Personen und Gruppen (u.a. Forscher, Kuratoren und Museen) und zur Einbettung in den historischen Kontext 302 Personen (zumeist Herrscher) und 91 epigraphisch attestierte Ereignisse erfasst. Damit sind insgesamt 1706 Datenbankobjekte erfasst.

2.3. Dokumentation und Digitalisierung von Forschungsmaterial

Schrift und ihre Schriftträger bilden als Forschungsgegenstand eine Einheit. Dadurch wird auch die Forschung am Schriftträger selbst notwendig. Durch geographische Verteilung und Immobilität ist eine Digitalisierung in hoher Qualität notwendig, um effektiv und transparent forschen zu können. In diesem Kontext wird ein Datenaustausch und damit die kollaborative Forschung an verschiedenen Standorten ermöglicht.

Für die epigraphische Analyse in einer VRE ist es daher unabdingbar, dass die Einheit aus Textträger und Text auch digital vorliegt und damit gegebenenfalls digitalisiert werden muss. Ebenso soll auch unpublizierte Forschung erfasst werden. Zu vielen Archivalien finden sich Beschreibungen oder epigraphische Notizen, gewissermaßen 'analoge Metadaten', die nachnutzbar sind. Auch diese müssen digitalisiert werden und können somit auch erstmals der Forschung zugänglich gemacht werden. Weiterhin erstellt das Projekt eigenes digitales Forschungsmaterial.

⁸ GeoNames Geographical Database: <http://www.geonames.org/>, Getty Thesaurus of Geographic Names (TNG): <http://www.getty.edu/research/tools/vocabularies/tgn>, Gemeinsame Normdatei (GND) der Deutschen National Bibliothek: http://www.dnb.de/DE/Standardisierung/GND/gnd_node.html

⁹ Stand: 1. Juli 2016

Dokumentiert wird Material aus Archiven im In- und Ausland (u.a. Ibero-Amerikanisches Institut Berlin, Carnegie Institution of Washington via ARTSTOR). Bei Privatarchiven bilden die Bild- und Textarchive von Prof. Karl Herbert Mayer und Prof. em. Dr. Berthold Riese besonders bedeutende Bestände.

Parallel legt die Bonner Arbeitsstelle detaillierte Informationen in Arbeitslisten zusammen. Diese umfassen: Konkordanz¹⁰ sämtlicher Zeichenkataloge und -klassifikationen, Fundstätten, Museen und Sammlungen, Textträger, grammatikalische Morpheme¹¹, Lemmata, sowie ein Inschriftenarchiv, das sämtliche Inschriftenträger auf Papier mit Kommentaren zur Inschrift und Chronologie in Akten abgelegt umfasst.

2.3.I. Digitalisierung der Archivalien

Berthold Rieses Inschriftenarchiv besteht aus 135 Aktenordnern mit Photographien, Zeichnungen und epigraphischen Notizen und ist zum Berichtszeitpunkt mit etwas über 14.000 Dateien zu knapp einem Drittel digitalisiert. Von den ca. 40.000 Bildobjekten aus Karl Herbert Mayers Photoarchiv (Diapositive, Filmnegative, Abzüge) sind bereits über 20.000 digitalisiert. Des Weiteren ergänzen private Bestände der Projektmitarbeiter sowie Schenkungen von Kollegen (z.B. Dr. Daniel Graña-Behrens, Stephan Merk) das Forschungsmaterial. Die Bestände werden von den SHKs in der Arbeitsstelle Bonn mittels geeigneter Scanner digitalisiert und von einer WHK mit Metadaten versehen. Insgesamt erwarten wir über 70.000 Dokumente, die wir in den kommenden Jahren scannen und in das digitale Archiv integrieren werden.

2.3.2. Dokumentation und Digitalisierung von Inschriften

Die Dokumentation von Inschriften erfolgt sowohl mittels traditioneller Verfahren wie Photographie und Zeichnung als auch mit neueren Techniken wie Photogrammetrie und 3D-Streifenlichtscanning. Gegenüber bisherigen bildgebenden Verfahren und Vermessungsmethoden erweisen sich Dokumentation und Vermessung von Artefakten mittels 3D-Scanning als vorteilhaft für die epigraphische Forschung. Das 3D-Scanning ermöglicht eine umfassende Auswertung der zahlreichen erodierten und für das bloße Auge nicht mehr lesbaren Inschriften, die auch mittels Photographie und anschließender Bildbearbeitung nicht mehr lesbar zu machen sind. Durch virtuelle Manipulationen (z.B. Simulation unterschiedlicher Beleuchtungswinkel) lassen sich feine Strukturen leichter erkennen. Zudem wird die virtuelle Rekonstruktion fragmentierter Inschriftenträger ermöglicht und die mittels bisheriger Methoden gewonnenen Daten werden durch 3D-Scanning weiter ergänzt und präzisiert. Der Zeitaufwand für den einzelnen Dokumentations- und Messvorgang am zu dokumentierenden Objekt ist gegenüber den anderen genannten Verfahren deutlich geringer. Darüber hinaus trägt das 3D-Scanning zur originalgetreuen Archivierung und Sicherung bestehenden Kulturguts bei.

Das 3D-Scannen ist essentieller Bestandteil der im Projektantrag aufgeführten Dokumentationsreisen zu archäologischen Stätten, Museen und Archiven. Im Fokus steht dabei auch die Zusammenarbeit mit Museen und Sammlungen, aus deren Beständen Objekte mit Inschriften für die epigraphische Auswertung gescannt und als 3D-Objekte publiziert werden sollen. Die mittels Software bearbeiteten, zusammengesetzten und gerenderten 3D-Objekte werden nach Fertigstellung den Museen zur Verfügung gestellt. Im Projektjahr 2015 wurden bereits Maya-Artefakte aus der 'Reliefsammlung der großen Epochen' im Knauf-Museum (Iphofen) und die Holztürsturze aus Tikal im Museum der Kulturen

¹⁰ Diese Liste bildet die Grundlage für die Erstellung des neuen Zeichenkatalogs.

¹¹ Diese Liste wird als Leitfaden bei der Transliteration, Transkription und der morphologischen Analyse / Glossierung dienen.

(Basel) auf diese Weise dokumentiert. Der Aufwand an Dokumentationsreisen in 2014 betrug 3 Personentage, in 2015 waren es 68 PT, für 2016 sind 46 PT angesetzt.

2.3.3. Erstellung einer Bibliographie

Zur Erfassung der Literatur wird die freie und quelloffene Anwendung Zotero verwendet um unterschiedliche Online- und Offlinequellen zu sammeln, zu verwalten und zu zitieren. Diese Anwendung unterstützt die Bearbeitung und Edition von Literaturangaben und Literaturlisten und erlaubt kollaboratives Arbeiten an unterschiedlichen Standorten. Die Literaturdatenbank enthält mittlerweile 15.500 Einträge und wird stetig erweitert, wobei ein Gesamtvolumen von 70.000 Einträgen zu erwarten ist. Die Inhalte werden nicht nur ständig ergänzt, sondern auch thematisch getaggt und revidiert.

3. Präsentation und Veröffentlichungen

Das Projekt arbeitet daran, die zeitgemäße, offene und interdisziplinär kooperierende Wissenschaft zu fördern und soweit möglich einen allumfassend offenen Zugang gemäß einer "Open Science" zu ermöglichen, sowohl zu seinen wissenschaftlichen Publikationen ("Open Access"), einer freien Dokumentation seiner Methodik und Arbeitsprozesse ("Open Methodology"), als auch den Forschungsdaten ("Open Data") und der vom Projekt genutzten Software ("Open Source"). Ein ungehinderter Zugang zu den Forschungsinhalten des Projektes und ebenso die Gewährleistung deren produktiver Nachnutzung muss durch freie Lizenzen verwirklicht werden.

Alle im Rahmen des Projektes erstellten Daten werden daher unter der international geltenden Creative Commons Urheberrechtslizenz CC BY 4.0¹² "Open Access" online veröffentlicht werden.

Das Projekt versucht alle digitalisierten Bilder und Umzeichnungen von Maya-Monumenten aus verschiedenen Publikationen, deren Rechte zur weltweiten Veröffentlichung es erhalten oder erworben hat, im TextGrid-Rep und Portal mitsamt annotierten Metadaten der Öffentlichkeit unter Verweis auf Originalquellen und Urheberschaft zugänglich zu machen. Digitalisate und Publikationen, deren Rechte eingeschränkt sind, werden zumindest innerhalb des TextGrid-Lab dem registrierten Nutzerkreis an Fachwissenschaftlern zu Forschungszwecken zugänglich gemacht.

Innerhalb der Arbeitsprozesse des Projektes fließen die einzelnen "Open"-Strategien ineinander über oder bedingen einander. Hierbei nutzt das Projekt mehrere verschiedene Plattformen, auf denen die im Laufe des Projektes erstellten Methoden, Forschungsergebnisse, Daten und Metadaten geteilt und publiziert wurden und weiterhin zur Verfügung gestellt werden.

3.I. Website, Social Media und Portal

Die 2015 veröffentlichte Projekt-Webseite¹³, die auf Basis von WordPress aufgebaut ist, dient als primäre Plattform für Onlinepublikationen und garantiert damit die schnelle, kostenfreie Verbreitung der Forschungsergebnisse. Die Deutsche Nationalbibliothek hat den Webauftritt als Publikationsplattform registriert.¹⁴ In den Sektionen *Arbeitspapiere*, *Forschungsberichte* und *Projektberichte* werden fortlaufend Arbeiten zur Lexikographie, Entzifferung und Linguistik

¹² <http://creativecommons.org/licenses/by/4.0/>

¹³ <http://www.mayawoerterbuch.de>

¹⁴ ISSN 2366-5556

veröffentlicht, ferner jährliche Fortschrittsberichte, sowie Arbeits- und Konzeptpapiere zum Umgang mit den Werkzeugen der Digital Humanities in der Epigraphik. Für die nachhaltige Referenzierung und persistente Zitierbarkeit dieser digitalen Objekte wurde in Zusammenarbeit mit der SUB ein Workflow für die Registrierung eines eindeutigen Digital Object Identifier (DOI) etabliert, der auf das Objekt selbst und nicht den Speicherort (die URL) referenziert.

Neben speziellen Fallstudien ist ein weiterer Aspekt der Webseite die schnelle und breite Bekanntmachung von Forschungsdaten in der Sektion *Dokumentation* (z.B. Fundstättenliste, Zeichenkonkordanz, 3D-Meshes), sowie Ankündigungen und allgemeine Informationen über das Projekt.

Des Weiteren ist die Kommunikation mit dem Nutzer entscheidend: Für alle Forschungsdaten, die aus den Arbeitslisten entstanden sind, ist die Möglichkeit für Feedback gegeben, das vor allem der Qualitätssicherung der Daten dienen soll. Ein anderer Aspekt sind Diskussionen zu den Onlinepublikationen, diese können öffentlich kommentiert werden, um einen wissenschaftlichen Diskurs zu starten.

Neben der eigentlichen Website spielen Kanäle in sozialen Netzwerken eine große Rolle. Zeitgleich mit der Website gingen Kanäle auf Facebook¹⁵ und Twitter¹⁶ online. Bei Sketchfab¹⁷ ist nicht nur eine Plattform für die 3D-Modelle eingerichtet, sondern eigene Modelle und die anderer Nutzer können zudem geteilt und vernetzt werden.

Zudem soll ab dem 2. Quartal 2018 ein Portal entwickelt werden, welches die Daten aus TextGrid nicht nur präsentiert sondern auch feingranular durchsuchbar macht. Es wird über die Website erreichbar sein, ist aber technologisch unabhängig. Über das Portal werden später auch alle im Projekt erzeugten Forschungsdaten und die aktuellen Arbeitsversionen des Wörterbuchs verfügbar sein. Das Portal wird ebenfalls gezielte Suchfunktionen anbieten, welche komplexe Abfragen an das TextGrid-Rep mit allen in Abschnitt 3.2. genannten textuellen und nicht-textuellen Metadaten ermöglichen.

3.2. Nutzung und Anpassung der ConedaKOR-Bilddatenbank zur Erfassung der digitalisierten Materialien

Veröffentlichungen der erschlossenen und erstellten Daten sind ein wichtiger Bestandteil der Projektarbeit. Vor allem das aus Schenkungen stammende dokumentierte und digitalisierte Forschungsmaterial zu nicht-textuellen Objekten und Inhalten aus dem Bereich der Maya-Kultur (siehe Abschnitt 3.3.1.) soll nachhaltige Verwaltung und Archivierung finden und zur Recherche öffentlich zugänglich gemacht werden. Diese Aufgabe soll eine eigene Bild-Datenbankinstanz in ConedaKOR¹⁸ als weitere Publikationsplattform des Projektes erfüllen. In ConedaKOR werden die digitalisierten Archivalien wie z.B. unbeschriftete Alltagsgegenstände, Artefakte und Bauten, die dem kulturellen Bereich des Klassischen Maya angehören, zueinander in Beziehung gesetzt und abgebildet. Seit Ende 2016 werden Konzepte für ein Metadatenschema erstellt und die Erschließung der aufbewahrten Inhalte in einer projekteigenen ConedaKOR-Instanz anhand ausgewählter Digitalisate erprobt. Hierbei ist der Plan, die Modellierung eines nicht auf die Digitalisate sondern auf die abgebildeten Entitäten (wie Artefakte, Orte, Personen) ausgerichteten Metadatenmodells auf die gegebene Graphdatenbank abzubilden. Die Inhalte sollen neben der objektgeschichtlichen Beschreibung durch Metadaten ebenfalls mit bibliographischen Angaben verknüpft werden, wobei wiederum die Zotero-Bibliographie

¹⁵ <https://www.facebook.com/idiom.project/>

¹⁶ https://twitter.com/idiom_project/

¹⁷ <https://sketchfab.com/idiom-project/>

¹⁸ <https://kor.uni-frankfurt.de/>

des Projekts genutzt wird und mit der Datenstruktur technisch verlinkt werden wird. Innerhalb der DARIAH-Infrastruktur soll zukünftig ein Zugang zu unserer ConedaKOR-Instanz durch einen Web-Service ermöglicht werden.

3.3. Das Inschriftenarchiv in den Digitalen Sammlungen der ULB Bonn

In Zusammenarbeit mit der ULB Bonn soll ein Inschriftenarchiv des Klassischen Maya innerhalb der Digitalen Sammlungen der ULB¹⁹ entstehen. Das Archiv präsentiert die im Projekt dokumentierten Textträger mit Digitalisaten und objektbezogenen Metadaten sowie deren Inschriften mit Transkriptionen und Übersetzungen. Die Befüllung des Archivs mit Inhalten ist hierbei in der zeitlichen Projektabfolge abhängig von zuvor erarbeiteten Arbeitspaketen und Meilensteinen: diese beziehen sich sowohl auf die Schaffung von technischer Infrastruktur sowie auf die im Folgenden darin geleistete Erstellung von Inhalten.

Hierbei müssen z.B. die folgenden Voraussetzungen erfüllt werden:

1. Vollständige und dokumentierte Metadatenschemata für Text und Textträger sowie eine produktive Erfassungsumgebung;
2. unter Realbedingungen erzeugte Erfassungsdaten;
3. Realisierung eines Mappings vom internen Datenformat zum dem von den Digitalen Sammlungen genutzten Zielformat (METS/MODS);
4. Veröffentlichung der erfassten Daten im TextGrid-Rep, damit sie über die OAI-PMH Schnittstelle von TextGrid abrufbar sind.

Für den Bereich der Textträger konnten bereits aussagekräftige Testdaten erzeugt werden, da die Voraussetzungen 1), 2) und 3) von Seiten des Projekts seit Ende 2015 erfüllt sind. Mit diesen Testdaten kann die Firma Semantics (Dienstleister der ULB) eine erste Projektion des Inschriftenarchivs für die Textträger vornehmen. Sobald in der zweiten Projektphase die Technologien für den Bereich der Textauszeichnung und -analyse erfüllt sind, wird Voraussetzung 4) konsolidiert für Textträger und Texte erfüllt werden, und es können komplette Testdaten an Semantics geliefert werden. Ab diesem Zeitpunkt kann eine Implementierung des Inschriftenarchivs in den Digitalen Sammlungen der ULB und mit der sukzessiven Veröffentlichung der Inschriften von Campeche und den bereits geleisteten Arbeiten zu Tikal gerechnet werden kann.

3.4. Präsentation, Öffentlichkeitsarbeit und Vernetzung

Das Projekt hat sich von 2014 bis 2016 auf zahlreichen nationalen und internationalen Veranstaltungen mit Vorträgen, Workshops, Präsentationen oder Postern präsentiert²⁰ (siehe Übersicht im Anhang).

Ziel dieser Aktivitäten war neben der Projektpräsentation im eigenen Fachbereich und der breiten Öffentlichkeit auch die interdisziplinäre Vernetzung mit anderen Projekten in den Digital Humanities, die eine vergleichbare Fragestellung und Zielsetzung aufweisen. Von den 26 Veranstaltungen wurden 12 durch das Projekt, den Projektträger oder die Kooperationspartner (mit)organisiert. Ferner hat die Akademie in Düsseldorf eine Ausstellungssäule herstellen lassen, die bislang bei mehreren öffentlichen

¹⁹ <http://digitale-sammlungen.ulb.uni-bonn.de/>

²⁰ Nähere Informationen zu den einzelnen Veranstaltungen sind in den Jahresberichten des Projekts verfügbar (2014: <http://mayawoerterbuch.de/?p=4477>; 2015: <http://mayawoerterbuch.de/?p=6789>).

Veranstaltungen aufgestellt wurde und immer mit Mitarbeitern des Projekts besetzt wurde, u.a. im Düsseldorfer Landtag zur Auftaktpressekonferenz²¹.

3.5. Publikationen

Die Dokumentation der Konzeption sowie der versionskontrollierte Sourcecode des fertiggestellten Objekt-Metadatenchemas werden webbasiert durch den git-Service für Software-Entwicklungsprojekte der GWDG frei zur Verfügung gestellt werden.

Neben der Website werden die digitalen Dokumente für den Printbereich regelmäßig für einen Sammelband ediert, der im Print-on-Demand-Verfahren gedruckt wird. Die Projektpublikationen stehen damit auch gedruckt der Öffentlichkeit in Bibliotheken zur Verfügung und können weltweit über Internet-Shopping erworben werden. Das erste dieser Jahrbücher für die Berichtsjahre 2014-2015²² enthält die englische Übersetzung der online erschienenen Artikel, sowie ausgewählte Inhalte der Website.

Eine vollständige Publikationsliste für den Zeitraum 2014-2016 findet sich im Anhang.

4. Projektphase II (2017 - 2019) - Textauszeichnung und -analyse zur Entwicklung des Wörterbuchs

Für die Korpusarbeit soll der epigraphische Workflow in der VRE möglichst genau abgebildet werden. Die wesentlichen Aspekte der Epigraphik sind dabei: Dokumentation des Textträgers, epigraphische Analyse mit Zeichenklassifikation, Transliteration und Transkription, morphologische Segmentierung und Glossierung, und linguistische Interpretation. Neben dieser Korpusanalyse als Grundlagenarbeit der Lexikographie untersuchen wir auch die Phonologie, Morphologie, Syntax, Semantik und Pragmatik des Klassischen Maya. Zudem rücken auch Fragen der historischen Linguistik, Untersuchungen zur Schriftgeschichte und des Schriftgebrauchs sowie der Schriftkonventionen in das Blickfeld der Forschung, soweit möglich in einer komparativen Sichtweise. Der integrierte Ansatz in einer VRE hat bereits in der Fachwelt eine positive Aufmerksamkeit erfahren.²³

Die zweite Projektphase umfasst neben den weiterlaufenden Arbeiten aus der ersten Phase schwerpunktmäßig drei Arbeitspakete: Zeichenkatalog, epigraphische Textauszeichnung und linguistische Analyse.

4.1. Zeichenkatalog

Basierend auf der Arbeitsliste der Zeichenkonkordanz bisheriger Zeicheninventare wird ebenfalls in TextGrid ein neuer digitaler Zeichenkatalog erstellt, welcher die für die nicht-textuellen Objekte programmierte Eingabemaske nachnutzen kann. Er besteht nicht nur aus der Konsolidierung bisheriger Listen, deren Unzulänglichkeiten er korrigieren soll, sondern wird ebenso laufend ergänzt, wenn bisher nicht inventarisierte Zeichen oder Graphvarianten in den Texten isoliert werden.

²¹ <http://bit.ly/29dvDmk>

²² ISBN 978-3-7392-4593-5

²³ So bescheinigen Stephen Houston und Simon Martin dem Projekt großes Potential durch "collecting and cross-referencing data." (S. HOUSTON & S. MARTIN. 2016. Through seeing stones: Maya epigraphy as a mature discipline. *Antiquity* 90(350): 443-455.

Die Modellierung des Metadatenschemas für den Zeichenkatalog und die Festlegung der Kriterien für eine Neuklassifizierung von Zeichen und deren Graphvarianten erfolgten ab der zweiten Hälfte 2015 und werden zum 3. Quartal 2016 abgeschlossen sein, so dass nach Abschluss einer Testphase zum Ende der ersten Projektphase der Zeichenkatalog mit realen Daten befüllt werden kann. Der Katalog unterscheidet in seiner ontologischen Struktur zwischen Zeichen (als Informationsträger) und Graph.

Jedes Zeichen kann multiple Transliterationswerte erhalten, etwa um polyphone Zeichen erfassen zu können. Viel wichtiger aber ist der Aspekt der Berücksichtigung von alternativen Lesungsvorschlägen, die zwangsläufig in einem teilweise entzifferten Schriftsystem auftreten. Um diese bewerten zu können, werden auf Logogramme, Silbenzeichen und Diakritika zugeschnittene Kriterien aufgestellt, mittels derer Lesungsargumente qualitativ im sprachlichen Kontext definiert und festgelegt werden können. Die Kriterien sind über Aussagelogik verknüpft²⁴, so dass bei unterschiedlichen Permutationen der definierten Merkmale automatisch eine entsprechende Konfidenz im Katalog eingetragen wird.

Eine vergleichende Untersuchung von Zeichen im Katalog und später kontextuell im Korpus kann somit anhand der Lesungssicherheit erfolgen und dabei helfen, bisherige Hypothesen zu bestätigen oder zu falsifizieren, oder idealerweise neue Entzifferungen zu ermöglichen. Damit können auch Lesungen mit höchstem Konfidenzgrad gefiltert werden und die Basis für die Wörterbucheinträge generieren, wenn die sprachliche Lesung gesichert ist. Für alle Lesungen mit geringerer Plausibilität können die Wörterbucheinträge dementsprechend gekennzeichnet werden, ebenso gleichberechtigte Alternativen unter verschiedenen Lemmata. Ebenso kann bei unentzifferten Zeichen falls möglich die semantische Domäne ebenfalls mit einem Konfidenzgrad erfasst werden. Damit wird das Wörterbuch erstmals den Umständen eines bisher nur partiell entzifferten Schriftsystems und einer nur teilweise erschlossenen Schriftsprache Rechnung tragen.

Jedes Zeichen ist mit mindestens einem Graphen verknüpft, bei dem eine standardisierte Abbildung als Referenz hinterlegt ist und der Variationstyp genannt wird. Während der Konzipierung des Katalogs konnte ein innovatives Klassifikationsschema erarbeitet werden, das Zeichenvarianten durch subgraphemische Segmentation und Variation erfasst. Auf der Graphenebene werden die Konkordanzen zu den bisherigen Katalogen erfasst, ebenso Informationen zum Graph-Ikon (über ein kontrolliertes Vokabular) und dessen Relationen zu den Graphen anderer Zeichen, so dass in einem gewissen Umfang auch Studien zur Zeichengenease ermöglicht werden (Ikon - Lautwert-Bezug).

4.2. Epigraphische Texterfassung und -auszeichnung

Die Verknüpfung eines Texts mit seiner Originalschreibung wird in TextGrid mittels des Text-Bild-Link-Editors (TBLE) erreicht, der Bereiche in einem Digitalisat auszeichnen kann. Die Dokumentation der Originalschreibung ist grundlegend für die epigraphische Arbeit mit syllabischen und logo-syllabischen Hieroglyphen- und Keilschriftsystemen, denn aus der Umschrift der Texte geht die Originalschreibung nicht mehr hervor. Die Auszeichnung geschieht auf Graphenebene. Linguistisch betrachtet ist das Graph als kleinste graphische Einheit eines Schriftsystems noch nicht einem Graphem als Zeichen zugeordnet worden. Demgegenüber hat das Graphem als kleinste bedeutungsunterscheidende Einheit in einem Schriftsystem eine Zuordnung erfahren. Jeder als Graph identifizierte Bereich wird mit dem Zeichenkatalog verknüpft, der damit Voraussetzung und Startpunkt für die epigraphische Texterfassung und -auszeichnung ist (Arbeitspaket "Textauszeichnung I"). Jedem ausgezeichneten Bereich können damit die Graphnummer und der Transliterationswert semi-automatisch zugewiesen

²⁴ Die Kriterien folgen damit Boolescher Algebra und können wahr oder falsch sein, etwa „ist im Landa-Manuskript zu finden“, „ist eine homophone Substitution des Logogramms gegeben“, „das semantische Feld ist bekannt“, etc.

werden.²⁵ Dieser Wert ist dabei eher als Zeichenname zu werten (vorbehaltlich einer weiteren Analyse)²⁶ und soll zudem die Zeichenklasse (traditionell durch Groß- / Kleinschreibung indiziert) anzeigen können.

Zur Indikation der Graphotaktik existieren noch weiterzuentwickelnde Operatoren um die räumliche Relation der Zeichen im Block anzuzeigen. Grundlage dieser mit dem TBLE geschaffenen Textauszeichnung ist ein Metadatenschema auf Basis von TEI und EpiDoc, welches noch entwickelt werden muss. Es soll ebenso Informationen zu Textstruktur, Erhaltungszustand (basierend auf dem Leidener System) oder zur Farbe von Graphen aufnehmen können.

Die so erfolgte Texterfassung und -auszeichnung ist epigraphisch rein deskriptiv und macht das Digitalisat damit maschinenlesbar. Die Auszeichnung entbehrt noch jeder linguistischen Analyse, ist jedoch Voraussetzung für ein digitales Korpus und die darauf aufbauende, mehrstufige analytische Annotation.

Die Kriterien zur Auszeichnung werden gegenwärtig konzipiert und können zusammen mit notwendigen Anpassungen in TextGrid ab 2017 in XML umgesetzt werden.

4.3. Linguistische Textannotation und -analyse

Neben der deskriptiven Textauszeichnung erfolgt in paralleler Entwicklung, jedoch späterer Implementierung (produktiv voraussichtlich im 4. Quartal 2018), das Arbeitspaket "Textauszeichnung II", welches "Textauszeichnung I" technologisch voraussetzt, darauf aufbaut und erweitert. Damit sind dann auch die Voraussetzungen für eine komplette Analyse und Erstellung des digitalen Korpus gegeben, das dann zur Verwendung in den Digitalen Sammlungen der ULB gemappt werden kann und auch auf dem eigenen Portal publiziert wird.

Ab 2017 wird in Zusammenarbeit mit Dr. Cristina Vertan (Hiob Ludolf Centre for Ethiopian Studies, Universität Hamburg) ein von ihr für das Silbenschriftsystem Fidal der altäthiopischen Sprache (Ge'ez) entwickeltes XML-basiertes Tool²⁷ nachgenutzt, indem es für die epigraphische und linguistische Analyse der Maya-Hieroglyphentexte angepasst und in die VRE implementiert wird. Die technischen Möglichkeiten dieses Tools wurden im April 2016 positiv evaluiert und eine Kooperation vereinbart. Das Tool erlaubt sowohl eine Annotation mit Korrektur, als auch eine Mehrebenen-Annotation. Es ermöglicht weiter die Anlage der Wörterbucheinträge bzw. der Lemmata aus der Analyse heraus.

Vor den technologischen Arbeiten ist die Definition des mehrstufigen Analyseschemas und die Erarbeitung der grammatikalischen Regeln für das Klassische Maya zu leisten (unterstützt durch die Arbeitsliste zu grammatikalischen Morphemen). Genau wie der Zeichenkatalog muss das Tool so flexibel gestaltet sein, dass es neue Forschungen und Erkenntnisse zur Grammatik und Morphosyntax des Klassischen Maya berücksichtigen und bisherige Analysen korrigieren kann.

Die fachwissenschaftlichen Voraussetzungen werden seit 2016 in Bonn erarbeitet und werden ab 2017 sukzessive in die Programmierung einfließen. Im Gegensatz zur bisherigen Tradition, auf die Transliteration die Transkription (ggf. mit morphologischer Segmentation) und Übersetzung folgen zu lassen, wird das Analyseschema insgesamt acht Schritte umfassen, die granular und transparent nicht

²⁵ Ggf. muss der Bearbeiter entscheiden, welcher polyphone Lautwert zu verwenden ist. Lesungsalternativen der gleichen Zeichenfunktion (Logogramm / Silbenzeichen) werden automatisch in absteigender Plausibilität hinterlegt.

²⁶ Zum Beispiel: **CHAN**, das kontextuell in 1) **CHAN-na** und 2) **CHAN-nu** auftauchen kann, dort aber als 1) *chan*, "Schlange" oder 2) *cha'an*, "Fänger" zu lesen ist.

²⁷ Im Rahmen des EU-Forschungsprojekts TraCES: <https://www.traces.uni-hamburg.de/>

nur die Anforderungen an die Maschinenlesbarkeit widerspiegeln, sondern allgemein die Verständlichkeit einer Analyse erhöhen. Nach den deskriptiven Schritten 1) Klassifikation und 2) graphematische Transliteration aus "Textauszeichnung I" wird darauf aufbauend die analytische Annotation die folgenden Schritte beinhalten: 3) phonemische Transliteration, 4) morphologische Transkription, 5) morphophonemische Transkription, 6) morphologische Glossierung, 7) konsolidierte Transkription, und 8) Übersetzung.²⁸

4.4. Wörterbuch

Gleichzeitig mit der Erstellung des digitalen Inschriftenarchivs war laut Antrag auch eine laufend aktualisierte Arbeitsversion des Wörterbuches vorgesehen. Wie in Abschnitt 4.3. dargelegt, müssen jedoch erst die technischen Voraussetzungen geschaffen werden, insbesondere der Textanalyse. Die Konzipierung und Entwicklung lexikologischer Funktionen im Annotationstool kann daher erst parallel mit "Textauszeichnung II" erfolgen.

Es erfolgten aber bereits Vorarbeiten, um die Kompilation des Wörterbuchs zu begleiten. 2014 bis 2015 haben die SHKs 114 Wörterbücher und Grammatiken von über 20 Mayasprachen digitalisiert. Viele davon wurden bereits mit OCR durchsuchbar gemacht und sollen langfristig ausgezeichnet werden, um Kognate und Etyma für das Klassische Maya im Annotationstool auffindbar zu machen.

Mit der Lemma-Arbeitsliste wurde eine Konkordanz diverser Wortlisten erstellt, um die Transkriptionen verschiedener Autoren miteinander zu vergleichen. Die verschiedenen Konventionen werden ebenfalls der Forschungsgeschichte wegen im Wörterbuch eingespeist.

Publikationsliste 2014-2016

Projektbezogene Publikationen seitens der Mitarbeiter und externer Autoren sowie anderweitige Publikationen, in die Projektmitarbeiter involviert sind, umfassen folgende Beiträge (mehrsprachige Artikel sind mit den Abkürzungen DE, EN, bzw. ES gekennzeichnet). (Siehe auch Abschnitt 4.3. Publikationen.)

1. DIEDERICHS, K. 2015. Unsere „Open Science“-Strategie. Elektronisches Dokument *Textdatenbank und Wörterbuch des Klassischen Maya Working Paper 1*. <http://mayawoerterbuch.de/?p=4334>. DE | EN
2. DIEDERICHS, K., C.M. PRAGER, E. WAGNER, S. GRONEMEYER & N. GRUBE. 2015. Textdatenbank und Wörterbuch des Klassischen Maya. Elektronisches Dokument *Wikipedia*. https://de.wikipedia.org/wiki/Textdatenbank_und_Wörterbuch_des_Klassischen_Maya.
3. DIEDERICHS, K., C.M. PRAGER, E. WAGNER, S. GRONEMEYER & N. GRUBE. 2015. Text Database and Dictionary of Classic Mayan. Elektronisches Dokument *Wikipedia*. https://en.wikipedia.org/wiki/Text_Database_and_Dictionary_of_Classic_Mayan.
4. EBERL, M. & S. GRONEMEYER. 2016. Organización política y social, in M. Eberl & C.M. Vela González (Hg.) *Entre reyes y campesinos: investigaciones recientes en la antigua capital maya de Tamarindito (Paris Monographs in American Archaeology, 45)*: 137-46. Oxford: Archaeopress.
5. GRONEMEYER, S. 2014. E pluribus unum: Embracing Vernacular Influences in Classic Mayan Scribal Tradition, in C. Helmke & F. Sachse (Hg.), *A Celebration of the Life and Work of Pierre Robert Colas (Acta Mesoamericana, 27)*: 147-62. München: Verlag Anton Saurwein.
6. GRONEMEYER, S. 2015. Class Struggle: Towards a Better Understanding of Maya Writing Using Comparative Graphematics, in H. Kettunen & C. Helmke (Hg.) *On Methods: How We Know What We*

²⁸ Im Antrag sind im Beispiel einer Analyse in Anlage 11 insgesamt 19 Analyseschritte gezeigt. Diese sind nicht obsolet, aber die Tools erlauben es durch das Analyseschema, einige dieser Punkte zu subsumieren und anderweitig zu erfassen. Ein Beispiel für die Analyseschritte wäre: 1) 740st.683br:126bb > 2) **SIH.ja:ya** > 3) **sih-ya=ja** > 4) *si{h}y-aj* > 5) *siy-aj-∅* > 6) 'gift-INCH-3s.ABS' > 7) *siyaj* > 8) "be born".

- Think We Know about the Maya. Proceedings of the 17th European Maya Conference, 2012 (Acta Mesoamericana, 28)*: 101–17. Markt Schwaben: Verlag Anton Saurwein.
7. GRONEMEYER, S. 2016. Textos jeroglíficos de Tamarindito, in M. Eberl & C.M. Vela González (Hg.) *Entre reyes y campesinos: investigaciones recientes en la antigua capital maya de Tamarindito (Paris Monographs in American Archaeology, 45)*: 107-22. Oxford: Archaeopress.
 8. GRONEMEYER, S. 2016. Filling the Grid? More Evidence for the <t'a> Syllabogram. Elektronisches Dokument *Textdatenbank und Wörterbuch des Klassischen Maya Research Note 4*. <http://mayawoerterbuch.de/?p=7118>
 9. GRONEMEYER, S., C.M. PRAGER, E. WAGNER, K. DIEDERICHS, N. GRUBE, F. DIEHR & M. BRODHUN. 2015. Website. Elektronisches Dokument *Textdatenbank und Wörterbuch des Klassischen Maya*. <http://mayawoerterbuch.de/>. DE | EN | ES
 10. GRONEMEYER, S., C.M. PRAGER & E. WAGNER. 2015. Evaluating the Digital Documentation Process from 3D Scan to Drawing. Elektronisches Dokument *Textdatenbank und Wörterbuch des Klassischen Maya Working Paper 2*. <http://mayawoerterbuch.de/?p=3405>.
 11. GRUBE, N. & C. A. LUIN. 2014. A drum altar from the vicinity of Yaxchilan. *Mexicon XXXVI* (2): 40-48.
 12. GRUBE, N. 2016. The logogram JALAM. Elektronisches Dokument *Textdatenbank und Wörterbuch des Klassischen Maya Research Note 3*. <http://mayawoerterbuch.de/?p=2221>.
 13. GRUBE, N. & K. DELVENDAHL. 2014. The Monuments of the 8 Ajaw House: New Evidence on the Calakmul-Uxul Relationship, in C. Helmke & F. Sachse (Hg.), *A Celebration of the Life and Work of Pierre Robert Colas (Acta Mesoamericana, 27)*: 75-95. München: Verlag Anton Saurwein.
 14. GRUBE, N. & C.M. PRAGER. 2015. Textdatenbank und Wörterbuch des Klassischen Maya, in N.-W. Akademie der Wissenschaften und der Künste (Hg.) *Jahrbuch 2015*: 160–64. Düsseldorf: Nordrhein-Westfälische Akademie der Wissenschaften und der Künste.
 15. GRUBE, N., C.M. PRAGER, K. DIEDERICHS, S. GRONEMEYER, E. WAGNER, M. BRODHUN, F. DIEHR & P. MAIER. 2015. Jahresabschlussbericht 2014. Elektronisches Dokument *Textdatenbank und Wörterbuch des Klassischen Maya Project Report 2*. <http://mayawoerterbuch.de/?p=4477>. DE | EN | ES
 16. MAIER, P. 2015. Ein TEI-Metadatenchema für die Auszeichnung des Klassischen Maya. Elektronisches Dokument *Textdatenbank und Wörterbuch des Klassischen Maya Working Paper 3*. <http://mayawoerterbuch.de/?p=5067>. DE | EN
 17. PRAGER, C.M. 2015. Das Textdatenbank- und Wörterbuchprojekt des Klassischen Maya: Möglichkeiten und Herausforderungen digitaler Epigraphik, in H. Neuroth, A. Rapp, & S. Söring (Hg.) *TextGrid: Von der Community – für die Community: Eine Virtuelle Forschungsumgebung für die Geisteswissenschaften*: 105–24. Glückstadt: Werner Hülsbusch.
 18. PRAGER, C.M. (Hg.). 2016. *Jahrbuch • Year Book • Anuario 2014 – 2015: Textdatenbank und Wörterbuch des Klassischen Maya, Arbeitsstelle der Nordrhein-Westfälischen Akademie der Wissenschaften und der Künste*. 284 pp. Norderstedt: Books On Demand. ISBN 978-3739245935.
 19. PRAGER, C.M., S. GRONEMEYER & E. WAGNER. 2015. A Ceramic Vessel of Unknown Provenance in Bonn. Elektronisches Dokument *Textdatenbank und Wörterbuch des Klassischen Maya Research Note 1*. <http://mayawoerterbuch.de/?p=2221>.
 20. PRAGER, C.M & E. WAGNER. 2016. A Possible Logograph XAN “Palm” in Maya Writing. Elektronisches Dokument *Textdatenbank und Wörterbuch des Klassischen Maya Research Note 5*. <http://mayawoerterbuch.de/?p=7206>
 21. SACHSE, F. & M. DÜRR. 2015. Morphological Glossing of Mayan Languages under XML: Preliminary Results. Elektronisches Dokument *Textdatenbank und Wörterbuch des Klassischen Maya Working Paper 4*. <http://mayawoerterbuch.de/?p=2122>.
 22. VELA GONZÁLEZ, C.M., S. LEVITHOL, A. DÍAZ, S. GRONEMEYER & M. EBERL. 2016. Excavaciones extensivas, in M. Eberl & C.M. Vela González (Hg.) *Entre reyes y campesinos: investigaciones recientes en la antigua capital maya de Tamarindito (Paris Monographs in American Archaeology, 45)*: 79-106. Oxford: Archaeopress.
 23. VELA GONZÁLEZ, C.M., S. LEVITHOL, L. VELÁSQUEZ, A. DÍAZ, J.M. PALOMO, S. GRONEMEYER & M. EBERL. 2016. Excavaciones de pozos de sondeo, in M. Eberl & C.M. Vela González (Hg.) *Entre reyes y campesinos: investigaciones recientes en la antigua capital maya de Tamarindito (Paris Monographs in American Archaeology, 45)*: 21-77. Oxford: Archaeopress.
 24. WAGNER, E., S. GRONEMEYER & C.M. PRAGER. 2015. Tz’atz’ Nah, a “New” Term in the Classic Mayan Lexicon. Elektronisches Dokument *Textdatenbank und Wörterbuch des Klassischen Maya Research Note 2*. <http://mayawoerterbuch.de/?p=3827>.

25. WAGNER, E. 2014. The Sculptured Frieze of Structure A-6-2nd, Xunantunich, Belize: A Preliminary Iconographic Analysis, in C. Helmke & F. Sachse (Hg.), *A Celebration of the Life and Work of Pierre Robert Colas (Acta Mesoamericana, 27)*: 97-146. München: Verlag Anton Saurwein.
26. BRODHUN, M., H. RIEBL & M. RODZIS 2014. tgFormsMayaImpl [RDF-Inputmask Framework in CoffeeScript], Available at <https://projects.gwdg.de/projects/tgformsmayaimpl>
27. BRODHUN, M. 2014. Mayan Database [RDF-Inputmask in CoffeeScript], Available at <http://git.projects.gwdg.de/mayandatabase.git>

Präsentationsliste 2014-2016

Tabellarische Auflistung der Präsentationen (siehe auch Abschnitt 4.5. Präsentation, Öffentlichkeitsarbeit und Vernetzung).

Datum	Beitrag	Veranstaltung (Ort)
1.2.2014	Projektvorstellung, Vorträge: 1) <i>Ziele und Aufgaben des Wörterbuchprojektes IDIOM</i> , 2) <i>Konzept, Umsetzung und Forschungsstrategie des Wörterbuchprojektes IDIOM</i> , 3) <i>Paläographie und Ikonologie im Rahmen des Wörterbuchprojektes IDIOM</i>	Tagung: <i>XVII. Mesoamerikanisten-Tagung</i> (Basel)
4.-5.6.2014	Ausstellungssäule	Veranstaltung: <i>Die Projekte der AWK-NRW im Landtag</i> (Düsseldorf), organisiert durch die AWK-NRW
14.9.2014	Ausstellungssäule	Veranstaltung: <i>Tag des offenen Denkmals</i> in der AWK-NRW (Düsseldorf), organisiert durch die AWK-NRW
14.10.2014	Präsentation: <i>Text Database and Dictionary of Classic Mayan</i>	Workshop: <i>1st Annual Workshop of the Text Database and Dictionary of Classic Mayan Project</i> (Düsseldorf), organisiert durch das Projekt
26.11.2014	Vortrag: <i>Epigraphische Texte und Objekte im Projekt „Textdatenbank und Wörterbuch des Klassischen Maya“</i>	Workshop: <i>TextGrid-Nutzertreffen ‘Norm- und Metadaten’</i> (Essen), organisiert durch TextGrid
25.2.2015	Vortrag: <i>Text – Bild – Inschrift: Hieroglyphenschrift und Sprachen der Maya annotieren</i>	Tagung: <i>Jahrestagung Digital Humanities im deutschsprachigen Raum</i> (Graz)
3.-5.3.2015	Poster: <i>Text Database and Dictionary of Classic Mayan</i>	Tagung: <i>Digital Humanities Summit</i> (Berlin), organisiert durch DARIAH-DE und TextGrid
10.3.2015	Vortrag: <i>Text Database and Dictionary of Classic Mayan</i>	Workshop: <i>Semantic Web Applications in the Humanities</i> (Göttingen)
20.4.2015	Vortrag: <i>Text Database and Dictionary of Classic Mayan: Some examples of the Classic Mayan writing system, of inscriptions and steps of analysis</i>	Workshop: <i>Epidoc</i> (London)
11.5.2015	Vortrag: <i>Open Access vs. Copyright im Projekt ‘Textdatenbank und Wörterbuch des Klassischen Maya’</i>	Workshop: <i>Urheberrecht und Open Access in den digitalen Geisteswissenschaften</i> (Göttingen)
11.5.2015	Ausstellungssäule	Veranstaltung: <i>Akademientag 2015</i> (Berlin), organisiert durch die Akademiunion

5.6.2015	Vortrag: <i>Textdatenbank und Wörterbuch des Klassischen Maya</i> (Projektvorstellung)	Workshop: <i>Abschluss Verbundprojekt MayaArch3D</i> (Bonn)
21.7.2015	Vortrag: <i>Textdatenbank und Wörterbuch des Klassischen Maya</i> (Projektvorstellung)	Veranstaltung: <i>Digital Humanities</i> (Bonn), organisiert von Jan Kenter, ULB
11.9.2015	Ausstellungssäule	Veranstaltung: <i>Tag des offenen Denkmals</i> in der AWK-NRW (Düsseldorf), organisiert durch die AWK-NRW
16.9.2015	Vortrag: <i>Dokumentation, Analyse und Edition der Hieroglyphentexte der Klassischen Maya in der Virtuellen Forschungsumgebung TextGrid</i>	Workshop: <i>Historische Semantik und Semantic Web</i> (Heidelberg), organisiert durch die Akademienunion
29.9.2015	Vortrag: <i>Dokumentation, Analyse und Edition der Hieroglyphentexte der Klassischen Maya in der virtuellen Forschungsumgebung TextGrid</i>	Workshop: <i>Digitales Arbeiten in den Geisteswissenschaften ermöglichen!</i> (Göttingen), organisiert durch DINI und DARIAH-DE
2.10.2015	Vortrag: <i>Digitale Epigraphik am Beispiel des Wörterbuchs des Klassischen Maya</i>	Workshop: <i>Digital Humanities Bilder</i> (Bonn)
4.11.2015	Vortrag: <i>Dokumentation, Analyse und Edition der Hieroglyphentexte der Klassischen Maya in der virtuellen Forschungsumgebung TextGrid</i>	Tagung: <i>Digitale Metamorphose: Digital Humanities und Editionswissenschaft</i> (Wolfenbüttel)
5.10.2015	Vortrag: <i>Maya Hieroglyphic Writing</i>	Workshop: <i>Digitale Epigraphik: XML/TEI und EpiDoc für die epigraphische Forschung zu nicht-alphabetischen Schriftsystemen</i> (Bonn), organisiert durch das Projekt
13.12.2015	Vortrag: <i>Digital Epigraphy – The Text Database and Dictionary of Classic Mayan Project</i>	Tagung: <i>20th European Maya Conference</i> (Bonn)
14.12.2015	Präsentation: <i>The Text Database and Dictionary of Classic Mayan Project</i>	Workshop: <i>2nd Annual Workshop of the Text Database and Dictionary of Classic Mayan Project</i> (Bonn), organisiert durch das Projekt
15.1.2016	Vortrag: <i>Digitale Epigraphik - Die Erforschung der Hieroglyphentexte und Bildbotschaften der Maya in der Virtuellen Forschungsumgebung TextGrid</i>	Tagung: <i>XIX. Mesoamerikanisten-Tagung</i> (Berlin)
29.1.2016	Vortrag: <i>A Virtual Research Environment to Document and Analyze Non- Alphabetic Writing</i>	Tagung: <i>EAGLE 2016 International Conference on Digital and Traditional Epigraphy in Context</i> (Rom)
1.3.2016	Vortrag: <i>Die Maya im digitalen Zeitalter</i>	Veranstaltung: <i>Museum der Kulturen</i> (Basel)
8.4.2016	Vortrag: <i>Graphemik des Klassischen Maya</i>	Tagung: <i>Ägyptologische ‚Binsen‘-Weisheiten III: Formen und Funktionen von Edition und Paläographie altägyptischer Kursivschriften</i> (Mainz)
18.5.2016	Ausstellungssäule	Veranstaltung: <i>Akademientag 2016</i> (Hamburg), organisiert durch die Akademienunion
14.9.2016	Vortrag: <i>Of Codes and Kings: Approaches in the Encoding of Classic Maya Hieroglyphic Inscriptions</i>	Tagung: <i>FORGE 2016 Forschungsdaten in den Geisteswissenschaften</i> (Hamburg)

6.10.2016	Vortrag: <i>Of Codes, Glyphs, and Kings: Tasks, Limits, and Approaches in the Encoding of Classic Maya Hieroglyphic Inscriptions</i>	Tagung: <i>ESTS 2016 / DiXiT 3: Digital Scholarly Editing: Theory, Practice, Methods</i> (Antwerpen)
17.11.2016	Vortrag: <i>Edificios como reliquias ancestrales. Un caso de Copán, Honduras</i>	Tagung: <i>1o Congreso Internacional de Arquitectura e Iconografía Precolumbina</i> (Valencia)
13.12.2016	Vortrag: <i>Of Codes, Glyphs and Kings: Tasks, Limits and Approaches in the Encoding of Classic Maya Hieroglyphic Inscriptions</i>	Seminar: <i>Digital Classicist: Digitale Methoden in den Altertumswissenschaften</i> (Berlin)



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/>