

# Paleocoran: Virtuelle Rekonstruktion von Korankodizes mit IIF

## Pohl, Oliver

opohl@bbaw.de  
Berlin-Brandenburgische Akademie der Wissenschaften,  
Deutschland

## Marx, Michael

marx@bbaw.de  
Berlin-Brandenburgische Akademie der Wissenschaften,  
Deutschland

## Franke, Stefanie

stfranke@bbaw.de  
Berlin-Brandenburgische Akademie der Wissenschaften,  
Deutschland

## Artika, Farah

artika@bbaw.de  
Berlin-Brandenburgische Akademie der Wissenschaften,  
Deutschland

## Schnöpf, Markus

schnoepf@bbaw.de  
Berlin-Brandenburgische Akademie der Wissenschaften,  
Deutschland

## Mahmutovic, Edin

mahmutovic@bbaw.de  
Berlin-Brandenburgische Akademie der Wissenschaften,  
Deutschland

Das Projekt „Paleocoran“ untersucht Koran-Handschriftenfragmente aus dem siebten bis zehnten Jahrhundert aus der Amr ibn al-##s- Moschee in al-Fu##t (Alt-Kairo), die heute in verschiedenen Sammlungen weltweit verteilt aufbewahrt werden. Paleocoran sammelt kodikologische (buchgeschichtliche) und paläographische (schriftgeschichtliche) Daten zu den Handschriftenfragmenten, um die zu rekonstruieren, welche Fragmente ursprünglich einen Kodex gebildet haben. Die digitale Rekonstruktion der Kodizes – einem Puzzle mit ca. 25.000 Teilen zu vergleichen – wird durch IIF digital ermöglicht.

Das von F. Déroche (Paris) und M. Marx (Potsdam) geleitete DFG-ANR-Projekt „Paleocoran“ greift inhaltlich und methodisch an das Akademievorhaben „Corpus Coranicum“ der BBAW an: Datensätze aus den Corpus Coranicum-Datenbanken zu Handschriften und Koran-

Textvarianten werden in einem System zur wortgenauen Textstellenverortung von Koranpassagen verwendet. Die in „Paleocoran“ und „Corpus Coranicum“ generierten philologischen Daten werden in einem web-basierten System auf Grundlage des PHP-Frameworks Laravel in einer MySQL-Datenbank aufgezeichnet.

Für die Rekonstruktion der Korankodizes aus Alt-Kairo werden die Seiten der einzelnen Fragmente in der Datenbank erfasst und mit Textstellenkoordinaten (Sure-Vers-Wort) ausgezeichnet. Derzeit befinden sich ca. 1100 Koranhandschriften mit insgesamt über 25.000 Manuskriptseiten in der aus mehr als 40 Sammlungen Corpus Coranicum-Datenbank. Die hier vorgehaltenen Bilder werden über den IIF-kompatiblen Bildserver digilib ausgeliefert.

Anhand der Textstellenangaben wird der Text auf der Manuskriptseite mit dem Text der Koran-Druckausgabe Kairo 1924 nach Schreibvarianten, Textvarianten und Verszählung untersucht. In vielen Fällen enthalten die verschiedenfarbigen Tinten der Handschrift Vokalzeichen, die Textvarianten in die Handschriften eintragen. Auch diese werden als unterschiedliche Lesarten verzeichnet.

Außerhalb der philologischen Daten stellen Illuminationen und Ornamente eine wichtiges Kennzeichen für Manuskriptfragmente dar, die ursprünglich zusammengehörten. Die Form und farbliche Gestaltung der funktionalen Ornamente (Verstrenner und Kapitelüberschriften). Form und Layout (z.B. Pflanzenornamente oder geometrische Muster) der Ornamente oder deren farbliche Gestaltung lassen dabei auf eine gemeinsame Herkunft der Fragmente schließen.

Für den Vergleich und die Darstellung von Orthographiedifferenzen im Vergleich zum Text der Druckausgabe Kairo 1924 wurde im Rahmen des Paleocoran-Projekts die Programmierbibliothek „Rasmify“ entwickelt. Rasmify entfernt sämtliche buchstabendifferenzierenden Zeichen (Diakritika) und Vokalzeichen aus arabischen Zeichenketten, sodass nur das Konsonantenskelett (arabisch: rasm) der verarbeiteten Strings übrig bleibt. Dies ist wichtig, da die frühen Koranhandschriften sehr häufig Buchstaben undifferenziert schreiben. Durch die reduzierte Wortform wird es einfacher, Unterschiede zwischen einzelnen Koranhandschriftenfragmenten und der Kairiner Druckausgabe zu identifizieren. Das Programm „Rasmify“ wurde in PHP, Python 3 und JavaScript als freie Software veröffentlicht und kann einfach über die jeweiligen Abhängigkeitsverwaltungen composer, pip und npm nachgenutzt werden.

Ähnliche Muster bei Abweichungen in Textvarianten (Lesarten), Schreibvarianten- und Verssegmentierung weisen darauf hin, dass die betreffenden Fragmente ursprünglich demselben Korankodex stammen. Bei genügend vorliegenden Indizien werden die einzelnen Handschriftenfragmente bzw. Teile der Fragmente einem virtuellen Kodex zugeordnet. Durch die den einzelnen Handschriftenseiten zugeordneten Textstellen werden dann die dem virtuellen Kodex zugeordneten

Handschriftenseiten nach Textkoordinate sortiert, sodass letztendlich ein IIIF-Manifest für den virtuellen Kodex erstellt werden kann.

Auf der Projektwebseite [paleocoran.eu](http://paleocoran.eu) kann mittels des IIIF-Manifests der virtuell rekonstruierte Korankodex in seiner ursprünglichen Form im IIIF-Viewer Mirador digital abgebildet werden. Mirador bietet darüber hinaus Lichtschiffunktionalitäten, sodass sowohl einzelne Seiten desselben virtuell rekonstruierten Korankodex als auch unterschiedliche virtuell rekonstruierte Korankodizes miteinander gezielt verglichen werden können.

Zusätzlich werden Metadaten der zugeordneten Manuskriptfragmente sowie Metadaten zur kodikologischen und paleographischen Einordnung des rekonstruierten Kodex angezeigt. Weiterhin werden die Lesarten- und Orthographievarianten sowie Ornamente samt Wortkoordinate und zitierfähigem IIIF-Bildausschnitt angegeben, sodass Forschende die virtuelle Rekonstruktion nachvollziehen können.

Zum aktuellen Zeitpunkt wurden 338 virtuell rekonstruierte Kodizes bzw. Kodexteile angelegt und über 1500 Lesartenvarianten sowie über 2500 Orthographieunterschiede identifiziert. Durch die Anbindung an das Corpus Coranicum Projekt ist die langfristige Sicherung und Nachnutzung gewährleistet. Der Launch der Paleocoran-Projektwebseite soll Ende 2018 erfolgen.