

# Papyrus Projekt

Umsetzung digitaler Papyrus-Sammlungen  
auf Basis der Content Management Anwendung  
des MyCoRe Arbeitskreises

von

Prof. Reinhold Scholl (Uni Leipzig)

Jens Kupferschmidt (Uni Leipzig)

Andy Wermke (Uni Leipzig)

Erik Körner (Uni Leipzig)

Tina Zimmermann (Akademie der Wissenschaften zu Göttingen)

Sarah Winkelmann (Akademie der Wissenschaften zu Göttingen)

u. a.

**Leipzig, 17.05.2022**

**Version 2020.06**

## Abstrakt

Diese Dokumentation beschreibt die Umsetzung des DFG Papyrus-Projektes der Universität Leipzig (und weiterer Sammlungen) auf informationstechnischem Gebiet. Inzwischen wird die Implementierung dieses Vorhabens für eine Reihe von Papyrus-Katalogisierungen genutzt. Neben dieser<sup>1</sup> vorrangig für griechische Papyri genutzten Anwendung gibt es auf eine für die KOHD betriebene Instanz für koptische Papyri<sup>2</sup> mit gleicher technischer Grundlage. Diese sind mit Inhalt der vorliegenden Dokumentation. Weiterhin soll dieses Dokument für Interessierte anderer Sammlungen den Einstieg zur Nachnutzung des Projekts ermöglichen.

Als Grundlage der Datenhaltung wird eine Applikation verwendet, die auf dem MyCoRe-Projekt (<http://www.mycore.de>), Release 2020.06 basiert und hierfür eine Referenzanwendung darstellt. MyCoRe ist eine Arbeitsgruppe verschiedener deutscher Universitäten, die einen Open Source Framework und darauf aufbauende Anwendungen zur Lösung von Aufgaben aus den Bereichen digitale Bibliotheken, Sammlungen und Digital Humanities anbietet. Diese Komponenten bieten für das Papyrus-Projekt unserer Ansicht nach langfristig die erforderliche Sicherheit bezüglich Soft- und Hardware-Anpassung. Der Einsatz dieser Software gestattet auch eine Erweiterung des Projektes auf andere Hochschulen und Institute im In- und Ausland. Die Anwendung unterliegt der GNU GENERAL PUBLIC LICENSE Version 3.

1 Siehe <https://papyri-collection.dl.uni-leipzig.de>

2 Siehe <https://coptica.kohd.adw-goe.de>

## Änderungen

Version	Datum	Autor
bis 1.4	04.03.2002	Jens Kupferschmidt, URZ der Uni Leipzig
bis 2.24	16.01.2006	Jens Kupferschmidt, URZ der Uni Leipzig Prof. Dr. R. Scholl, UBL der Uni Leipzig Ulrike Krönert, URZ Jena Mathias Hegner, URZ Jena Jürgen Vogler, Uni Leipzig
3.0.1	30.06.2007	Jens Kupferschmidt, Uni Leipzig
3.0.2	15.07.2007	Jens Kupferschmidt, Uni Leipzig
3.0.3	24.07.2007	Jens Kupferschmidt, Uni Leipzig
3.0.4	30.08.2007	Jens Kupferschmidt, Uni Leipzig
3.0.5	25.10.2007	Prof. Dr. R. Scholl, UBL der Uni Leipzig Jens Kupferschmidt, Uni Leipzig
3.0.6	06.11.2007	Jens Kupferschmidt, Uni Leipzig
3.0.7	03.12.2007	Marius Gerhardt, UBL der Uni Leipzig
3.0.6	06.12.2007	Jens Kupferschmidt, Uni Leipzig
4.0.1	31.08.2009	Jens Kupferschmidt, Uni Leipzig
4.0.2	07.10.2009	Frank Ursin, Uni Leipzig
4.0.3	13.11.2009	Jens Kupferschmidt, Uni Leipzig
4.2.0	21.01.2011	Andy Wermke, Uni Leipzig
4.2.1	12.05.2011	Andy Wermke, Uni Leipzig
4.5	01.09.2011	Andy Wermke, Uni Leipzig
4.5.2	10.01.2012	Frank Ursin, Uni Leipzig
4.5.3	10.02.2012	Jens Kupferschmidt, Uni Leipzig
4.8	22.10.2013	Erik Körner, Uni Leipzig
5.0	02.02.2016	Jens Kupferschmidt, Uni Leipzig
5.1	26.07.2016	Jens Kupferschmidt, Uni Leipzig
5.1	24.08.2016	Tina Zimmermann, Akademie der Wissenschaften zu Göttingen
5.1	26.08.2016	Jens Kupferschmidt, Uni Leipzig
5.1	01.11.2016	Tina Zimmermann, Akademie der Wissenschaften zu Göttingen
5.1	20.12.2016	Tina Zimmermann, Akademie der Wissenschaften zu Göttingen
5.1	21.12.2016	Jens Kupferschmidt, Uni Leipzig
5.2	15.02.2017	Tina Zimmermann, Akademie der Wissenschaften zu Göttingen
5.3	11.04.2017	Jens Kupferschmidt, Uni Leipzig
5.3	05.07.2017	Tina Zimmermann, Akademie der Wissenschaften zu Göttingen
5.4	09.03.2018	Jens Kupferschmidt, Uni Leipzig
5.4	29.03.2018	Jens Kupferschmidt, Uni Leipzig
5.4	16.07.2018	Tina Zimmermann, Akademie der Wissenschaften zu Göttingen
5.5	13.09.2018	Tina Zimmermann, Akademie der Wissenschaften zu Göttingen
5.5	24.09.2018	Tina Zimmermann, Akademie der Wissenschaften zu Göttingen
5.6	01.02.2019	Jens Kupferschmidt, Uni Leipzig
2019.06	06.03.2020	Jens Kupferschmidt, Uni Leipzig
2020.06	14.01.2021	Sarah Winkelmann, Akademie der Wissenschaften zu Göttingen

## Abkürzungen

<b>Kürzel</b>	<b>Erläuterung</b>
ACL	Access Control List – eine Technologie für die Zugriffskontrolle auf Daten.
API	Application Programming Interface – eine allgemeine Bezeichnung der Programmierschnittstellen in einem Projekt.
DC	Dublin Core – ein Metadaten-Schemata zur Beschreibung von Archivalien und sonstigen Dokumenten
DDBDP	Duke Databank of Documentary Papyri
EAD	Encoded Archival Description – ein XML-Metadaten-Austauschformat zur Beschreibung von Archivgut
ID	Identifikator, eine eindeutige Marke.
JAVA	Eine objektorientierte, weit verbreitete Programmiersprache.
JDBC	Java Database Connectivity – eine Schnittstelle zum einheitlichen Zugriff auf relationale Datenbanken.
JSP	JavaServer Pages, Web-Programmiersprache
KOHD	Die Katalogisierung der Orientalischen Handschriften in Deutschland (KOHD) ist ein Forschungsprojekt der Akademie der Wissenschaften zu Göttingen.
LIDO	Lightweight Information Describing Objects – ein XML- Standard zum Austausch von Metadaten
LZW	Lempel-Ziv-Welch-Algorithmus zur Datenkomprimierung
METS	Metadata Encoding & Transmission Standard – ein XML-Standard zur Strukturierung von digitalen Objekten
MODS	Metadata Object Description Schema – ein sehr gebräuchliches XML-Metadaten-Austauschformat
SQL	Structured Query Language – eine vereinheitlichte Abfragesprache für relationale Datenbanken.
UBL	Universitätsbibliothek Leipzig
URZ	Universitätsrechenzentrum der Universität Leipzig
XML	Extensible Markup Language – ein Standard zur Notation von Daten.
XPath	Ein Standard zum Zugriff auf XML-Daten.
XSL-FO	Eine standardisierte Auszeichnungssprache zur Gestaltung von Druckvorstufen in XML.
XSLT	Extensible Stylesheet Language for Transformation – eine Beschreibungssprache zur Transformation von XML Daten.

# Inhaltsverzeichnis

1	Projektbeschreibung.....	7
1.1	Sicherheitskonzept.....	9
1.1.1	Sicherung der Rohdaten.....	9
1.1.2	Zugriffssicherung in der Datenbank.....	10
1.2	Die Bilddaten.....	11
1.2.1	Digital Master.....	11
1.2.2	Derivate.....	12
1.2.3	Umsetzung der Bildaufnahme im Papyrus-Projekt Halle-Jena-Leipzig.....	12
1.2.3.1	Erfassung der Originaldaten (Scan).....	12
1.2.3.2	Bearbeitung der Daten für die Datenbank.....	13
2	Datenmodell.....	14
2.1	Begrifflichkeiten.....	14
2.1.1	Schematische Darstellung des Papyrus-Datenmodells.....	15
2.2	Datenstruktur für Informationen zum Schriftträger und ererbte Fragmente.....	16
2.3	Datenstruktur für Informationen zum Text und zu Publikationen.....	19
2.4	Datenstruktur für zusätzliche Dokumente.....	22
2.5	Vererbung der Schriftträgerdaten für fragmentarische Papyri.....	22
2.6	Klassifikationen.....	24
2.7	Exportformate.....	25
2.7.1	Mapping auf OAI Dublin Core.....	25
2.7.2	Mapping auf MODS.....	26
2.8	Service-Informationen.....	27
2.8.1	Datumsangaben.....	27
2.8.2	Bilddarstellung.....	27
3	Regeln zur Erstellung / Bearbeitung neuer Datensätze.....	28
3.1	Inventarnummern.....	28
3.2	Projektinterne ID's.....	29
3.3	Antike Datumsangaben.....	30
3.3.1	Festlegungen.....	30
3.3.2	Eingabefelder.....	31
3.3.3	Beispiele.....	31
3.3.4	Epochen bezogen auf Ägypten.....	32
3.4	Codierung von URL's.....	32
3.5	Literaturangaben.....	32
3.6	Eingabe von Lesungen.....	32
4	Dateneingabe direkt in XML.....	34
4.1	Erstellen der Daten.....	34
4.1.1	Laden der Daten.....	34
4.1.2	Update der Daten.....	34
4.1.3	Erweitern offener Klassifikationen.....	35
4.2	Verarbeitung der Bilddaten.....	35
4.2.1	Scannen und Sicherheitsverfilmung.....	35
4.2.2	Ausschneiden der Bilder für die Web-Präsentation.....	35
4.2.3	Erzeugen der Web-Bilder.....	35
4.3	Erstellen des Inventarbuches.....	36
4.4	Erstellung der Google Sitemaps.....	36
4.5	Sicherungsmaßnahmen.....	36
4.5.1	Dokumentation.....	36
4.5.2	Der Server in Leipzig.....	37
5	Installation / Nutzung eines Demo-Systems.....	38
5.1	Installation.....	38
5.1.1	Nutzung anderer Datenbanken.....	38

5.1.2	Zugang über HTTPS.....	38
5.2	Commandline Tools.....	38
5.2.1	Spezielle Kommandos der Anwendungen.....	38
5.2.2	Zur Verarbeitung von Metadatenobjekten.....	39
5.2.3	Zur Verarbeitung von digitalen Objekten (Derivaten).....	39
5.2.4	Zum Erzeugen der Inventarbücher.....	39
5.2.5	Reparieren der Indizierungen.....	39
5.2.6	Zum Erzeugen der userconfig-Datei.....	40
5.2.7	Überprüfen der Links.....	40
5.2.8	Zum Sichern der Daten.....	40
5.2.9	Sonstige Kommandos.....	40
6	Glossar.....	41
7	Anhang.....	42
7.1	Übersetzungstabelle deutsch – englisch.....	42
7.2	Übersetzungstabelle deutsch – französisch.....	45
7.3	Projektbeteiligte.....	49
7.3.1	Die wissenschaftliche Leitung.....	49
7.3.2	Für die Restaurierung.....	49
7.3.3	Für die technische Betreuung.....	49

# 1 Projektbeschreibung

An dem Ur-Projekt einer digitalen Papyrus-Sammlung arbeiteten seinerzeit verschiedene Einrichtungen der Universitäten Jena, Halle und Leipzig mit. Die Umfänge der einzelnen Papyrus-Sammlungen differierten dabei zwischen den Partnern stark. Trotzdem wurde ein umfassendes und verbindliches Rahmenkonzept zur Aufarbeitung, Sicherung und Präsentation geschaffen, das auch weiteren Kooperationspartnern Anreiz zur Mitarbeit sein soll. Es liegen einheitliche Vorgehensweisen und Datenmodelle vor, die in dieser Dokumentation näher erläutert werden sollen.

<b><i>Institut für Altertumswissenschaften der Friedrich-Schiller-Universität Jena und die Thüringer Universitäts- und Landesbibliothek</i></b>
<b><i>Institut für klassische Altertumswissenschaften der Martin-Luther-Universität Halle-Wittenberg</i></b>
<b><i>Universitätsbibliothek der Universität Leipzig</i></b>

Tabelle 1: An der Entwicklung beteiligte Einrichtungen

Im Rahmen einer Nachnutzung wurde das Pilotprojekt in eine allgemeine Beispielanwendung und spezielle Anwendungen für die einzelnen Sammlungen unterteilt. Hinzu kommt die softwaretechnische Bereitstellung der Daten für das Papyrus-Portal<sup>3</sup> sowie für die Papyrus-Sammlung koptischer Papyri der KOHD<sup>4</sup>. Derzeit nutzen neben den oben genannten noch folgende Kooperationspartner die Anwendung im Rahmen einer Sammelanwendung:

<b><i>Staats- und Universitätsbibliothek Bremen</i></b>
<b><i>Universitätsbibliothek Erlangen-Nürnberg</i></b>
<b><i>Universitätsbibliothek Gießen</i></b>
<b><i>Staats- und Universitätsbibliothek Hamburg</i></b>
<b><i>Institut für Altertumskunde der Universität Köln</i></b>
<b><i>Hessischen Staatsarchivs Marburg</i></b>
<b><i>Stadt- und Kulturgeschichtliches Museum Torgau</i></b>
<b><i>Universitätsbibliothek der Universität Würzburg</i></b>
<b><i>Akademie der Wissenschaften zu Göttingen (KOHD)</i></b>

Tabelle 2: Nachnutzer des Projektes

Nach erfolgter Restaurierung der einzelnen Papyri<sup>5</sup> wird jedes Exemplar mit bestmöglicher Auflösung digitalisiert und ggf. verfilmt. Dabei erstreckt sich diese Arbeit, bedingt durch den Zustand der Objekte, über einen längeren Zeitraum.

Parallel dazu erfolgt eine Katalogisierung der zugehörigen bibliographischen Daten (Metadaten) in der projektinternen Datenbank. Bedingt durch die asynchronen Arbeitsabläufe Restaurieren, Scannen und Erfassen der Metadaten, kann es vorkommen, dass nicht zu jedem bereits

3 siehe <http://www.papyrusportal.de>

4 siehe <https://papyri.kohd.adw-goe.de>

5 Eine genaue Beschreibung der Arbeiten finden Sie in der Dokumentation von Jörg Graf, UB Leipzig.

in der Datenbank erfassten Papyrus auch die entsprechenden Bilder vorhanden sind. Diese werden oft zu einem späteren Zeitpunkt nachgeladen.

Als Datenbasis für die im Projekt erfassten Metadaten und Bilder kommt eine auf dem Open Source Projekt MyCoRe<sup>6</sup> basierende Applikation zum Einsatz, die sowohl die interaktive Suche im Gesamtdatenbestand wie auch die interaktive Bearbeitung der Daten durch berechnigte Wissenschaftler gestattet. Die Anwendung ist als verteilte, mandantenfähige Applikation konzipiert, die jeder Einrichtung gestattet, die eigenen Daten zu bearbeiten und zu kontrollieren. Aus personellen Gründen ist ein verteiltes System praktisch ungünstiger. Eine Alternative bietet die zentral in Leipzig gehostete Sammelanwendung. Hier sind die Nutzereinrichtungen im Editormodus voneinander getrennt, jedoch wird eine Präsentationsebene genutzt. Momentan liegt die DV-technische Bereitstellung von Hard- und Software allein beim Universitätsrechenzentrum der Universität Leipzig. Dies schließt aber nicht aus, dass zukünftig mehrere Standorte als verteiltes System mit getrennten Verantwortlichkeiten betrieben werden können. Entwickelt wurde die hier vorgestellte Anwendung ebenfalls am URZ der Universität Leipzig unter der technischen Leitung von Jens Kupferschmidt, einem Mitentwickler des MyCoRe-Projektes.

Die folgende Grafik soll einen Überblick über die mögliche standortübergreifende Zusammenarbeit der einzelnen Instanzen des Projektes geben. Die standortunabhängige Nutzung ist dabei schon realisiert.

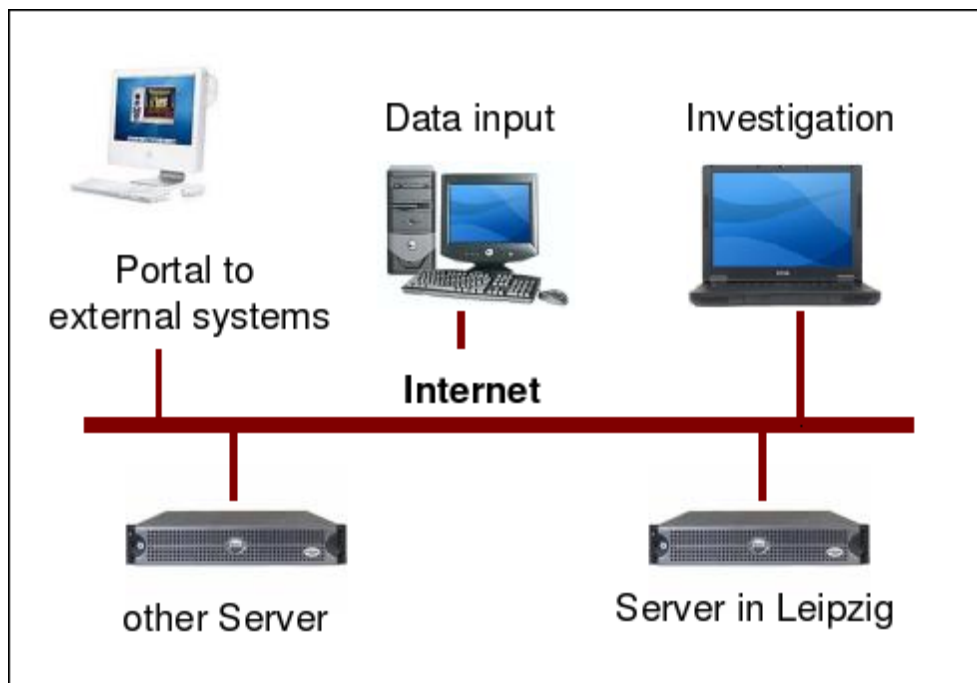


Abbildung 1: Das Client-Server-System

Da die meisten Nachfolgeprojekte zur Katalogisierung weiterer Bestände über die Deutsche Forschungsgemeinschaft (DFG) finanziert werden, orientiert sich die Entwicklung der Software stark an den aktuellen Richtlinien der DFG.

6 siehe <http://www.mycore.de>



## 1.1 Sicherheitskonzept

### 1.1.1 Sicherung der Rohdaten

Für die Sicherheit der erzeugten Scans und Fotos gibt es verschiedene Verantwortlichkeiten, die auch von der Bearbeitungsphase abhängt. Viele Teile dieses Kapitels beziehen sich auf die Vorgehensweise in Leipzig, sind aber auch für andere Standorte adaptierbar.

#### **Phase 1 – Erzeugen der Scans und Fotos**

Beim Erzeugen der Scans und Fotos als Rohdaten ist auf eine ausreichende Qualität (z. B. gemäß Vorgabe der DFG) zu achten. Für jeden Papyrus sind die optimalen Scan-Daten zu ermitteln. Die Scans sind entsprechend den Vorgaben mit Dateinamen zu versehen und zu speichern. Die Rohdaten auf der Scan-Station dürfen erst gelöscht werden, wenn die Daten beim Hauptverteiler (z. Z. Leipzig) angekommen sind und dort fehlerfrei auf ein Filesystem und ggf. DVDs kopiert wurden. Die entstanden Fotos sind vor Auslieferung an die Projektteilnehmer auf Qualität zu prüfen.

#### **Phase 2 – Erzeugen der Bildausschnitte für die Datenbank**

Nachdem die Scans erfasst und auf DVD gebrannt wurden bzw. in zentralen Storages abgelegt und gesichert sind, werden Sie per Hand mit einem Bildbearbeitungsprogramm ausgeschnitten und auf 300 dpi reduziert. Dabei erfolgt noch einmal eine visuelle Kontrolle aller Scans zur Fehlererkennung. Die bearbeiteten Bilder werden auch sicher gespeichert und an den Hauptverteiler (z. Z. Leipzig) ausgeliefert.

#### **Phase 3 – Verteilung der Datenträger**

Beim Hauptverteiler werden die eingehenden Bilddateien auf Lesbarkeit geprüft und vervielfältigt. Dabei wird eine vollständige Sicherungskopie erzeugt. Weiterhin werden für die anderen beteiligten Einrichtungen Kopien ihrer Daten gemacht und an diese ausgeliefert. Ggf. werden die zusätzlich angefertigten Sicherheitsfilme kontrolliert und auf die beteiligten Einrichtungen aufgeteilt.

#### **Phase 4 – Langzeitarchivierung**

Da eine Langzeitarchivierung technisch noch nicht vollständig ausgereift ist, beschränkt sich die Aufbewahrung auf die Speicherung auf zentralen, ausfallsicheren Speichersystemen sowie deren zentrale Backup Sicherung. Ggf. müssen die Daten auf neuere Systeme umkopiert werden.

#### **Phase 5 – Arbeit mit der Datenbank**

Als erster Schritt werden die Pflichtmetadaten für alle Papyri (Schriftträger und Texte) erfasst und in die Datenbank eingegeben. Die für die Datenbank aufbereiteten Bilder werden in das System eingestellt. Zyklisch (wöchentlich / monatlich) werden alle Daten aus dem System ausgelesen und auf ein separates Plattensystem auf dem hochverfügbaren Fileserver des URZ in XML-Form abgelegt. Die SQL-Datenbank wird weiterhin mit einem internen Backup gesichert. Anschließend werden alle beteiligten Systeme über das zentrale Backup-System des URZ auf Tape gesichert.

#### **Phase 6 – Metadaten-Verteilung**

Einmal Jährlich im Dezember werden die ausgespielten Daten inklusive der Inventarbücher an die beteiligten Einrichtungen in geeigneter Form übertragen.

## Phase 7 – Systemsicherheit

Für die Sicherheit des Systems, auf dem die Datenbank und die darauf basierende WEB-Anwendung läuft, ist der jeweilige Betreiber (derzeit Leipzig) zuständig und verantwortlich. Hier sind alle erforderlichen Arbeiten durchzuführen.

### 1.1.2 Zugriffssicherung in der Datenbank

Die systemseitige Sicherheit der Projektdaten (Programme, Metadaten und konvertierte Bilder) und der Logins des Betriebssystems zur Arbeit mit der Datenbasis liegt beim jeweiligen Betreiber (z. Z. das URZ der Uni Leipzig).

Darüber hinaus hat jede Instanz des Projektes noch Applikationsbenutzer, welche interaktiv bestimmte Operationen auf die Daten ausführen dürfen. Die Sicherheit basiert auf dem User2-System des MyCoRe-Kerns und ist für das Papyrus-Projekt speziell definiert:

- a) Systemadministratoren als Mitglied der Gruppe **admin**: Diese dürfen alle Funktionen in einer MyCoRe-Instanz ausführen.
- b) Die Editoren der einzelnen Einrichtungen sind Mitglied der Gruppen **editorhalgroup**, **editorjengroup**, **editorlpzgroup**, **editor...group**. Sie haben Zugriff auf alle Daten des Projektes bzw. ihrer Lokation. Sie können neue Daten einstellen und bestehende Daten der Lokation ändern und lesen. Es dürfen alle als '**readintern**' eingestuft Daten (z. B. Standorte der Medien) gelesen werden. Sie haben weiterhin die Rechte c und d.
- c) Papyrologen und Studentengruppen der entsprechenden Einrichtungen, d. h. des einzelnen Projektteilnehmers wie Jena, Halle oder Leipzig, haben als Mitglieder der Gruppen **internhalgroup**, **internjengroup**, **internlpzgroup**, **intern...group** lesenden Zugriff auf alle Daten (interne Anmerkungen), die als '**readintern**' gekennzeichnet sind. Sie haben weiterhin alle Rechte des Benutzers **guest**.
- d) Der Benutzer **guest** (Interessierte und Papyrologen weltweit) darf nur alle ungeschützten (**public**) Daten lesen. Für unveröffentlichte Papyri werden Images in geringerer Auflösung (72 dpi) angeboten.

Grundsätzlich gilt, dass die Bilder der Papyri in voller Größe **nur dann** gezeigt werden, wenn der Nutzer das entsprechende Recht (Permission) hat. Ausgenommen sind alle Bilder, für die im Datensatz im Feld ST24 (Publikationsart) **unpubliziert - frei** steht. Dieses Feld ist für Recto und Verso gesondert anzugeben!

Das Wechseln des Benutzerkontextes erfolgt über einen Zugangs- bzw. Login-Mechanismus, der im Projekt implementiert ist. Der Schutz der einzelnen Daten erfolgt durch Filtern der Daten in einem Server-seitigen XSLT-Prozess. Dieser arbeitet abhängig von der Mitgliedschaft des Nutzers in einer Gruppe. Dabei werden nicht darzustellende Datenteile aus der Präsentation entfernt.

## 1.2 Die Bilddaten

Je nach Verwendungszweck existieren unterschiedliche Formate für die Bilder eines Digitalisierungsprojektes. Grundsätzlich ist zwischen dem *digital master* (Master) und den Derivaten zu unterscheiden. Der Master liegt in einem Archivformat vor, welches verlustfrei ohne Komprimierung arbeiten sollte. Es ist ein standardisiertes, plattformunabhängiges Format zu wählen, um die Bearbeitung mit Mac- und Windows-Systemen zu gewährleisten.

Die Derivate sind verkleinerte Kopien des Masters für die Darstellung im Internet. Sie liegen in einem komprimierten Format vor, das ebenfalls standardisiert und plattformunabhängig angezeigt werden kann. Im Gegensatz zum Master darf das Derivat bearbeitet werden,<sup>7</sup> um eine optimale Darstellung des wesentlichen Bildinhalts, wie etwa der Schrift auf den Papyri, zu gewährleisten. Nachfolgend die derzeit üblichen Formate im Überblick:

<b>digital master</b>	<b>Extension</b>	<b>Komprimierung</b>	<b>Bemerkung</b>
TIFF	*.tif	nein, verlustfrei	Datenaustausch mit Programmen und Systemwelten wie Mac und Windows, schnelles Öffnen und Speichern durch unkomprimierte Datei, Quasi-Standard
<b>Derivate</b>			
JPEG	*.jpg	ja, einstellbar, Verluste	für Dateikompression mit Datenverlust
JPEG2000	*.jp2	ja, einstellbar, verlustarm	reduzierte Artefaktbildung, seit 02.01.2001 ein ISO-Standard, Wasserzeichen speicherbar, bei einmaligem Durchlauf eines Wavelet-Algorithmus quasi-verlustfreie Kompression, kostenpflichtig
PNG	*.png	ja	Vom W3C empfohlen
GIF	*.gif	Ja, hoch, verlustreich	schnelle Datenübertragung von 256-Farbenbilder

Tabelle 3: Grafikformate I

### 1.2.1 Digital Master

In diesem Projekt wird das TIFF-Format(\*.tif) ohne Komprimierung für die Archivierung der Farbbilder verwendet. Damit wird der Empfehlung der DFG gefolgt, auf mit LZW komprimierte TIFFs zu verzichten, da schon eine geringe Beschädigung der Datei das gesamte Bild beeinträchtigen kann.<sup>8</sup> Da der Master der Ausgangspunkt für die weitere Bearbeitung ist, muss er eine entsprechende Qualität aufweisen. Einfluss auf die Qualität eines Bildes haben Farbtiefe und Auflösung. 24 Bit Farbtiefe ist ein guter Kompromiss zwischen Dateigröße und Qualität. Es besteht dabei immer noch Spielraum für eventuelle Tonwert-Korrekturen in der weiteren Verarbeitungen zum Derivat. Deshalb werden, der Empfehlung der DFG folgend in diesem Projekt für den Scan und Master 24 Bit Farbtiefe verwendet.

Die DFG empfiehlt für „Handschriften und Kartenwerke mit feinen Strichen oder Beschriftungen“<sup>9</sup> eine Auflösung von maximal 400 dpi. Papyri haben die Besonderheit, dass die Struktur des Beschreibstoffes sehr wichtig ist und auch beim Master erkennbar bleiben muss. So er-

7 Vgl. DFG-Praxisregeln Digitalisierung, S. 7.

8 Vgl. ebd., Punkt 1.3.1.3

9 Vgl. DFG-Praxisregeln Digitalisierung

möglichst eine Auflösung von 600 dpi das Zusammenführen von versprengten Fragmenten, kleinste Tintenspuren können über die Konstitution von Texten entscheiden und Faltsuren können an Papyrusbriefen identifiziert werden. Seit Projektbeginn wurde mit 600 dpi gescannt und der Master erstellt, womit auch sehr gute Derivate erzeugt werden konnten. Ein Wechsel zu einer niedrigeren Auflösung ist nicht angebracht.

## **1.2.2 Derivate**

„Für die Publikation im Internet empfehlen sich wegen ihrer großen Verbreitung JPEG und PNG.“<sup>10</sup> Diese Formate bieten die Möglichkeit der Kompression, was für einen schnellen Zugriff sorgt. Dabei können allerdings Bildverluste auftreten. Es ist ein Kompromiss zwischen Bildqualität und vertretbarer Dateigröße zu finden.

Wichtig ist auch, ob die Bilder nur angezeigt werden oder auch druckbar sein sollen. Bei normaler Anzeige reicht eine Auflösung zwischen 70 und 100 dpi, da PC-Monitore ohnehin nur ca. 75 dpi anzeigen können. Soll gedruckt werden (1:1 .. 1:n), sollte die Auflösung 300 dpi betragen. Es bietet sich also an, ein Bild in mehreren Auflösungen anzubieten.

## **1.2.3 Umsetzung der Bildaufnahme im Papyrus-Projekt Halle-Jena-Leipzig**

### **1.2.3.1 Erfassung der Originaldaten (Scan)**

Beim Scan und der Verfilmung (Hybridverfahren) muss das aufzunehmende Objekt eine Farbreferenz und eine Maßleiste aufweisen. Beide sind so ins Bild zu bringen, dass die eigentliche Darstellung der Papyri für eine virtuelle Zusammenführung extrahiert werden kann. Die gescannten Master werden auf DVD gebrannt und entsprechend den Sicherheitsrichtlinien weiterverarbeitet.

Für das Papyrus-Projekt wird für die Langzeitarchivierung das TIFF Format unkomprimiert mit 600 dpi im Farbmodus verwendet. Der Rohscan wird schon geschnitten, um Speicherplatz zu sparen. Die Größe der Dateien beträgt je nach Größe des Papyrus zwischen 50 und 500 MB. Wenn die durch die Auflösung bedingte maximale Größe eines Papyrus von 46,05 x 34,54 cm überschritten ist, müssen mehrere Scans angefertigt werden, die später digital zusammengeführt werden.

Der Scanner ist ein „HIT ReproScanner“, der mit der Aufnahme-Software „Sinar CaptureShop“ bedient wird, welche auf einem Mac läuft.

Die Ostraka wurden mit einer Mamiya RZ67 Pro II, einem Mamiya M 4/65 mm L-A (Mamiya M 1:4 / f=65mm L-A) Objektiv und einem Sinar Digitalback fotografiert. Die verwendete Aufnahme-Software war Sinar Capture Shop 4.1. Von jedem Ostrakon wurden jeweils ein Farbbild und ein Infrarot-Bild aufgenommen. Das Equipment war dabei identisch außer einem Gelatine-IR-Filter für die Infrarot-Aufnahmen.

Parameter der Kamera für die Farbbilder

Belichtung 1/15

Blende 16

Parameter der Kamera für IR-Bilder  
Belichtung 4  
Blende 8

Die Ostraka wurden auf ein schräg liegendes Schaumstoffkissen gebettet, damit die Kamera von schräg oben noch die nicht unwesentliche Tiefe eines Ostrakons erfassen konnte. Der Abstand der Kamera zum Ostrakon betrug ca. 40 cm. Zur Beleuchtung dienten zwei Halogen-Lampen, die die Ostraka zur Minderung des Schattenwurfes beidseitig halbfrontal beschienen.

### 1.2.3.2 Bearbeitung der Daten für die Datenbank

Das JPEG Format wurde für die Darstellung im Internet gewählt. Die Erstellung der JPEG-Derivate erfolgt in mehreren Schritten mit dem Bildbearbeitungsprogramm Photoshop CS3.

Einstellungen für Papyri

1. Modus auf 8 Bit wechseln
2. Auflösung auf 300 dpi
3. Helligkeit/Kontrast: + 50 / + 50
4. Selektive Farbkorrektur:
5. Rot: Cyan + 100
6. Gelb + 100
7. Weiß: Schwarz - 100

Für die Bearbeitung der Ostraka entfällt bei der selektiven Farbkorrektur bei sonst gleicher Behandlung wie die Papyri die Anpassung des Rot- und Gelb-Kanals. Eine weitere Ausnahme ist, dass bei manchen Ostraka die Luminiszenz etwas erhöht wurde um eine gleichmäßige Ausleuchtung des weißen Hintergrundes zu gewährleisten. Es hat sich zudem als nützlich erwiesen im Schwarz-Kanal bei der selektiven Farbkorrektur den Wert etwas anzuheben um die Schrift deutlicher hervortreten zu lassen.

Diese Schritte können mit PhotoShop automatisiert werden, sodass ein Kommando reicht, um sie alle auszuführen. Die Selektive Farbkorrektur geschieht in Abgleich mit dem Original und dient der Reduzierung der Rottöne, die beim Scan hervorgehoben werden. Gleichzeitig wird der weiße Hintergrund normalisiert, was keinerlei Einfluss auf das Farbverhalten des Papyrus selbst hat. Anschließend werden von den Bildern die Farbskala und die Maßleiste entfernt. Die aus dieser Bearbeitung per Hand gewonnenen Abbildungen sind im JPEG-Format mit 300 dpi und TrueColor zu speichern und für die weitere Bearbeitung bereitzustellen.

Bei der Einarbeitung der Derivate in die Datenbank sind durch den Mitarbeiter alle Daten zu den Sicherheitsverfilmungen und Scans zu erfassen und im Datenfeld ST31 (Kopie in Medien) einzuspeichern. Dies muss sowohl für die Recto- wie auch für die Verso-Daten erfolgen. Die Erstellung der im Internet präsentierten Bilder erfolgt automatisch aus den nachbearbeiteten 300 dpi Dateien.

Folgende Bildgrößen werden bereitgestellt:

- 300 dpi und 100 dpi des ausgeschnittenen Bildes als JPEG
- 72 dpi des ausgeschnittenen Bildes als JPEG mit reduzierter Geometrie und niedriger Qualität und mit nicht lesbarer Schrift für die Präsentation nicht veröffentlichter Papyri
- 72 dpi Thumbnail mit 100 Pixel auf der längsten Seite des ausgeschnittenen Bildes als JPEG

# 2 Datenmodell

## 2.1 Begrifflichkeiten

Das Datenmodell basiert auf den im Anhang des ersten DFG Antrages (2002) aufgeführten Metadaten. Im Laufe der Entwicklung und des Ausbaus des Projektes wurde das Datenmodell dann weiter verfeinert und angepasst. Um die komplexe Struktur der einzubringenden Informationen programmtechnisch abbilden und andererseits diese mit geeigneten Suchmechanismen wieder auffinden zu können, wird auf objektorientierte Techniken wie Vererbung zurückgegriffen. In den Kindobjekten werden dabei Daten teils geerbt und teils durch eigene Daten ersetzt. In MyCoRe sind diese Techniken implementiert.

Um ein einheitliches Verständnis der verwendeten Terminologie zu erreichen, soll diese hier kurz erläutert werden.

<b>Bezeichnung</b>	<b>Beschreibung</b>
Papyrusobjekt	Bezeichnet die Gesamtheit von physisch und inhaltlich zusammengehörenden Teilen eines antiken Exemplars. Es beinhaltet die betreffenden Fragmente oder Fragmentteile, Textdokumente und die bibliographischen Daten zu diesen Objekten.
Schriftträger	Beschreibt alle physikalischen Merkmale des Schrift tragenden Objektes. Nach der Terminologie des zugrunde liegenden DFG-Antrags werden diese Daten in der Hauptoberfläche dargestellt. Die Daten für die Vorder- und Rückseite (Recto/Verso) können dabei teilweise unterschiedlich sein. (z.B. Schriftrichtung)
Schriftträger-Fragment	Besteht der Schriftträger aus mehreren Bruchstücken, die noch nicht zusammengesetzt sind oder die daneben weiter einzeln gespeichert werden sollen, können spezielle Daten dieser Fragmente in einem zusätzlichen Datensatz gespeichert werden. Dabei werden nur erweiternde oder individuelle Informationen (zum Beispiel Größe) gespeichert, alle anderen Daten erbt das Schriftträger-Fragment vom Schriftträger.
Schriftträger-Fragmentteil	Spezieller Datensatz zur Speicherung von Daten großer Schriftträger-Fragmente (z. B. Papyrus Ebers). Dieser erbt wiederum vom Fragment.
Text	Text, der sich auf einem oder mehreren Schriftträgern oder aber auch nur auf Teilen davon befindet. Hinzu kommen Informationen zur jeweiligen Bearbeitung des Textes. Diese Daten werden in der im DFG-Antrag beschriebenen ‚Textinformationsoberfläche‘ dargestellt. Auch die Texte und die dazugehörigen Bearbeitungen können in einem Vererbungsbaum existieren. So kann z. B. ein Text mehrere Bearbeitungen und/oder Publikationen haben. Ebenfalls denkbar sind mehrere verschiedene Textlesungen und Bearbeitungen.
Dokumente	Bezeichnen zusätzliches Material über die Papyrus-Objekte (z. B. Restaurierungsberichte). Das System gestattet deren Speicherung und Recherchierbarkeit.

Tabelle 4: Terminologieübersicht

## 2.1.1 Schematische Darstellung des Papyrus-Datenmodells

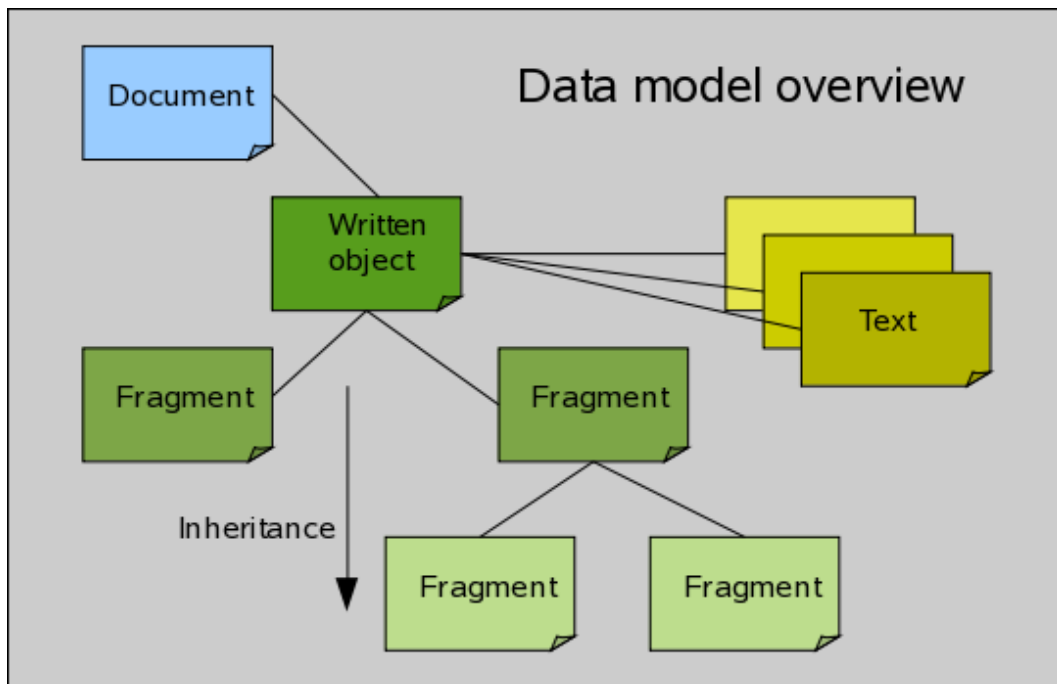


Abbildung 2: Datenmodellschema

Zwischen Schriftträger und Texten sowie zwischen Fragmenten bzw. Teilfragmenten und Texten können bidirektionale Links bestehen. Zu einem Schriftträger existieren meist Publikationen und/oder Lesungen in Form von Texten, die logisch untrennbar mit dem Schriftträger verbunden sind. Auch einem Fragment oder Fragmentteil können Texte zugeordnet werden, z. B. wenn nicht klar ist, wie die Zugehörigkeit zu einem Papyrus ist. Die Metadaten der Texte enthalten Informationen mit dem Verweis zum eigentlichen Schriftträger, Fragment oder Fragmentteil. Logisch sind diese Links als bidirektional aufzufassen, in der Umsetzung im Projekt sind sie einseitig gerichtet (vom Text zum Schriftträger).

Um komplexes Suchen zu ermöglichen, aber z.B. nicht jedem Fragment auch noch alle Metadaten des logisch „über“ ihm stehenden Schriftträgers mitteilen zu müssen, werden Vererbungsmechanismen auf Metadaten-Ebene benutzt.

Ein besonders schönes Beispiel ist der Leipziger Ebers-Papyrus. Der Titel ‚*Papyrus Ebers*‘ gilt für alle Tafeln dieses geteilten Papyrus. Jede Tafel erbt also diesen Titel und eine Suche danach sollte auf diese oberste Hierarchieebene als Treffer zeigen. Die nächste Schriftträgererebene sind die einzelnen Tafeln. Deren Metadaten-Informationen erben wiederum die Scans der einzelnen Spalten. Diese Stückelung ist notwendig, da nur so eine sinnvolle Präsentation des Objektes in der Datenbank möglich ist.

Darüber hinaus sind mit den Schriftträgern externe und interne Dokumente oder andere Datenobjekte wie Webseiten verbunden. Der Schriftträger enthält die Verweise auf die externen Referenzen. Im Dokumentenbereich des Projektes kann jedem Dokument auch ein interner Link zu einem Schriftträger zugeordnet werden. Auch dieser Link ist als bidirektional aufzufassen, so dass eine Verknüpfung zwischen dem Dokumentenobjekt und dem Schriftträger abgebildet wird.

## 2.2 Datenstruktur für Informationen zum Schrifträger und ererbende Fragmente

Nr.	Feld	Beschreibung	Zuordnung	Pflicht	wiederholbar	View	Suche	Metadaten Typ
1	Titel	Sprechender Titel	Identifizierung	ja	r / v max 10	allgemein	ja	MCRMetaPap LangText
2	Inventarnummer	Inventarnummer des Schrifträgers oder Schrifträgerfragments (z.B. P.Jen.Inv. 34 R)	Identifizierung	ja	nein	allgemein	ja	MCRMeta LangText
3	Andere Ordnungskriterien	Alt-, Sonderinventare bzw. -bereiche (z.B. Knüllfunde)	Identifizierung	nein	max 5	allgemein	ja	MCRMeta LangText
4	Sammlung	Leipzig, Halle, Jena, ggf. mit vermisst	Identifizierung	ja	nein	allgemein	ja	MCRMeta Classification 00000001
5	Unterbringung	Lokalisierung in Sammlung	Identifizierung	nein	nein	intern	nein	MCRMeta LangText
6	Erwerbung	Fundort, Datum, Finder Kauforte, Kaufdatum, Verkäufer	Identifizierung	ja	nein	allgemein	ja	MCRMetaPap Acquisition
7	Publikation	Verknüpfung mit Publikationsbeschreibung	Identifizierung	nein	max 5	allgemein	nein	MCRMeta Link
8	Publikationsnummer	z. B. publiziert P.Hal. 1	Identifizierung	nein	r / v max 20	allgemein	ja	MCRMetaPap LangText
9	Material	Papyrus, Pergament, Hadernpapier, Ostrakon/Ton, ... (vgl. Klassifikations-Liste)	Information zum Schrifträger	ja	nein	allgemein	ja	MCRMeta Classification 00000002
10	Farbe (grob)	Farbklassifikation des Objektes	Information zum Schrifträger	nein	r / v je 1	allgemein	ja	MCRMetaPap Classification 00000010
11	Farbe (verbal)	Farbbeschreibung	Information zum Schrifträger	nein	r / v je 1	allgemein	nein	MCRMetaPap LangText
12	Maße	Höhe u. Breite, ggf. Stärke	Information zum Schrifträger	nein	nein	allgemein	ja	MCRMeta Number
13	Zustand	karbonisiert, verschmutzt, Schimmel, Anhaftungen, ... (vgl. Klassifikations-Liste)	Information zum Schrifträger	ja	max 5	allgemein	ja	MCRMeta Classification 00000003
14	Komplett	ja / nein /fast/ nicht zu entscheiden	Information zum Schrifträger	ja	nein	allgemein	ja	MCRMeta Classification 00000004
15	Aufbewahrung	in Glas etc. (vgl. Klassifikations-Liste)	Information zum Schrifträger	ja	nein	intern	ja	MCRMeta Classification 00000005
16	Erhaltung	Konservierungsmaßnahmen	Information zum Schrifträger	nein	max 5	intern	nein	MCRMeta LangText
17	Ausführung	Codex, Rolle, Blatt, ... (vgl. Klassifikations-Liste)	Information zum Schrifträger	ja	nein	allgemein	ja	MCRMeta Classification 00000006
18	rekonstruierter Umfang	Länge der Rolle, Blattgröße, bei Codex Blattzahl	Information zum Schrifträger	nein	nein	intern	nein	MCRMeta LangText
19	Klebungen	Anzahl und Maße von links gemessen (vgl. Regeln)	Information zum Schrifträger	nein	max 10	allgemein	ja	MCRMeta Number
20	phys. Zugehörigkeit, publik	verbal, zu fremden Stücken, sofern diese außerhalb des Projektes sind	Information zum Schrifträger	nein	max 5	allgemein	nein	MCRMeta LangText
21	Fremdlink	Link zu fremden Stücken, sofern diese außerhalb des Projektes sind	Information zum Schrifträger	nein	max 5	allgemein	nein	MCRMeta Link
22	phys. Zugehörigkeit (interne Vermerke)	eigene + fremde Signatur; unpublizierte Vermerke	Information zum Schrifträger	nein	max 5	intern	nein	MCRMeta LangText
23	phys. Zugehörigkeit (interne Links)	eigene + fremde Links; unpublizierte Vermerke	Information zum Schrifträger	nein	max 5	intern	nein	MCRMeta Link
24	Publikationsart	Auswahlfeld	Information zum Schrifträger	ja	r / v max 2	allgemein	ja	MCRMetaPap Classification 00000014
30	Abbildung, publiziert	Bibliogr. Verweis (vgl. Regeln)	Dokumentation / Archivierung	nein	r / v max 5	allgemein	nein	MCRMetaPap LangText
31	Kopie in Medien	Photoplatten, Negative, Ausdrücke, Dias, CD etc.	Dokumentation / Archivierung	nein	r / v max 10	intern	nein	MCRMetaPap Media
32	Bemerkungen, extern		Textüberblick	nein	max 5	allgemein	nein	MCRMeta LangText
33	Bemerkungen, intern		Textüberblick	nein	max 5	intern	nein	MCRMeta LangText
34	Weitere Links		Textüberblick	nein	max 5	allgemein	nein	MCRMeta Link
35	Bearbeiter	Die Bearbeiter des Datensatzes	Administration	ja	max 10	intern	ja	MCRMeta



Nr.	Feld	Beschreibung	Zuordnung	Pflicht	wiederholbar	View	Suche	Metadaten Typ
								LangText
36	Lizenz	Lizenz des Datensatzes	Administration	nein	nein	allgemein	nein	licenses
37	alternative ID	URN, PPN usw.	Administration	nein	max 5	intern	nein	MCRMeta LangText
38	interne ID	interne IDs von Datenübernahmen	Administration	nein	max 5	intern	nein	MCRMeta LangText
	Status	Der Status des Datensatzes	Administration	ja	nein	intern	nein	statusOfWork- flow

Tabelle 5: Datenstruktur Schrifträger

## Erläuterungen zu den einzelnen Feldern des Schrifträgers:

Nr.	Feld	Kommentar	Länge
1	Titel	Notwendig zur Identifizierung der Daten in der Präsentation. Es sollte ein ‚sprechender‘ Titel gewählt werden.	128
2	Inventarnummer	Können nicht vereinheitlicht werden, siehe Abschnitt 2.4	128
3	Andere Ordnungskriterien	Diese Angaben sind je nach Lokation unterschiedlich.	128
4	Sammlung	Enthält Aufbewahrungsort, vgl. Klassifikations-Liste	
5	Unterbringung	Für Leipzig: Kartons (Großbuchstaben) Sonstiges (Für Leipzig alte Standort- und Fachnummer)	128
6	Erwerbung	<i>Erwerbung</i> gliedert sich in die Teilfelder <ul style="list-style-type: none"> <li>o Erwerbungsstyp</li> <li>o Fundort (z.B.: el-Bahnasa (Oxyrhynchos))</li> <li>o Funddatum (in der Form tt.mm.jjjj) (z.B.: 01.11.1903)</li> <li>o Finder</li> <li>o Kaufdatum (in der Form tt.mm.jjjjj)</li> <li>o Kaufort</li> <li>o Käufer (z.B.: Müller und Meier)</li> <li>o Verkäufer</li> <li>o Verlosungstag (in der Form tt.mm.jjjjj)</li> <li>o Kiste (in der Form Ort, Zinnkiste, Nr.) (z. B.: IX a)</li> <li>o Bemerkung (Für Leipzig der oben genannte Karton: Umbettung; Kiste nicht identifizierbar) (z.B.: November 1903 ; unbekannter Verkäufer)</li> </ul> Bei fehlenden Daten bitte dies in <i>Bemerkung</i> notieren.	
7	Publikation	URL für eine digitale Publikation oder Edition. Verlinkung auf digitalisierte/gescannte Papyrustexte – wie z.B. literarische Papyri, die nicht in der DDBDP aufgenommen sind.	128
8	Publikationsnummer	Obsolet. Stattdessen TE42 verwenden. Wie Dokumentation, Syntax: “publiziert Publ.Nummer [weitere Publ.Nummer(n)]“ (nach Checklist) oder “unpubliziert”. Bei Texten, die keine eigene Publikationsnummer haben (in Zeitschriften, Festschriften, usw.) ist Publikationsort mit Seitenzahl und eventuell der dortigen Nummerierung anzugeben, und zwar nach den Regeln der Angaben für Sekundärliteratur. Bei längeren Titeln sind Abkürzungen mit Verweis auf ed. pr. und weitere Editionen möglich.	128
9	Material	Vgl. Klassifikations-Liste	
10	Farbe (grob)	Vgl. Klassifikations-Liste	
11	Farbe (verbal)	Freitext mit Bestimmung nach Michel Farbtafel	256
12	Maße	<i>Maße</i> gliedert sich in die Teilfelder (alle Angaben in cm) <ul style="list-style-type: none"> <li>o Höhe</li> <li>o Breite</li> <li>o Stärke (nur für Ostraka)</li> </ul> Bei fehlenden Angaben wird als Standard 0 eingetragen.	
13	Zustand	Vgl. Klassifikations-Liste	
14	Komplett	Vgl. Klassifikations-Liste	
15	Aufbewahrung	Vgl. Klassifikations-Liste	
16	Erhaltung	Die Informationen zu den Erhaltungsmaßnahmen müssen in einem zweiten Arbeitsschritt ergänzt werden. Grund ist der Arbeitsablauf bei der Erfassung der Metadaten.	4096
17	Ausführung	Vgl. Klassifikations-Liste	
18	rekonstruierter Umfang	Hier sollte soweit möglich die ursprüngliche Größe des Stückes mit ungefähren Maßen in cm eingegeben werden. Bei literarischen Fragmenten (z. B. Homer) kann man manchmal feststellen, wie viele Verse zwischen den Fragmenten fehlen, ohne dass man angeben kann, wie groß die Rolle ursprünglich gewesen ist.	4096
19	Klebungen	Klebungen werden von links auf der Recto-Seite und parallel zum Schriftverlauf in cm gemessen. Ausnahme: <i>transversa charta</i> .	
20	phys. Zugehörigkeit, publik	Verbale Angaben der Zugehörigkeit zu fremden Stücken, sofern diese außerhalb des Projektes sind.	128
21	Fremdlink, publik	Eine URL, welche auf Online-Daten der Beschreibung von ST20 verweist.	
22	phys. Zugehörigkeit (interne Vermerke)	Verbale Angaben der Zugehörigkeit zu eigenen, fremden Stücken, sofern diese außerhalb des Projektes sind.	128

Nr	Feld	Kommentar	Länge
23	Fremdlink, intern	Eine URL, welche auf Online-Daten der Beschreibung von ST22 verweist.	
24	Publikationsart	Auswahlliste für den Publikationsstatus (recto / verso einzeln) (vgl. Klassifikations-Liste)	
30	Abbildung, publiziert	Ort der Publikation nach den Regeln für Literaturangaben. Abkürzungen sind erlaubt, wenn ein Verweis auf ed. pr. oder weitere Editionen erfolgt, wo dann die vollständige Angabe dazu steht.	128
31	Kopie in Medien	<p><i>Abbildungen in Sammlungen</i> gliedert sich in die Teilfelder</p> <ul style="list-style-type: none"> <li>o Typ (Foto, Negativ, Dia, Scan, ...)</li> <li>o Ordnungsnummer (z.B. CD 3; Film 26, ...)</li> <li>o Unterordnungsnummer (z. B. Bild 4, Dateiname, ...)</li> <li>o Datum (in der Form tt.mm.jjjj)</li> <li>o Aufbewahrungsort</li> <li>o Bemerkung</li> </ul> <p>Der Eintrag erfolgt für die aktuellen Scans/Fotos aus technischen Gründen in einem zweiten Schritt. Hier kann auch eine Liste mit der angegebenen Struktur stehen, d. h. die Struktur kann ggf. mehrmals wiederholt werden.</p>	
32	Bemerkungen, extern	Hier auch Vermerk ob Siegel, Stempel, Verschnürung usw.	4096
33	Bemerkungen, intern	Weitere Anmerkung zur internen Arbeit.	4096
34	Weitere Links	Dieser Eintrag enthält eine URL.	
35	Bearbeiter	Hier stehen Einträge aus einer Liste von Bearbeitern.	128
36	Lizenz	Hier ist aus einer Listen von Lizenzen eine für den Datensatz auszuwählen.	
37	alternative ID	URN, PPN usw.	128
38	interne ID	interne alte IDs für Datenübernahmen	128
	Status	Auswahl aus der Liste der Bearbeitungszustände	

Tabelle 6: Ausfüllhinweise Schrifträger

## 2.3 Datenstruktur für Informationen zum Text und zu Publikationen

Nr.	Feld	Beschreibung	Zuordnung	Pflicht	wiederholbar	View	Suche	Metadaten Typ
1	Titel	Sprechender Titel	Identifizierung	ja	max 5	allgemein	ja	MCRMeta LangText
2	Inventarnummer	Inventarnummer des Schriftträgers oder Schrifträgerfragments (z.B. P.Jen. Inv. 34 R)	Identifizierung	ja	nein	allgemein	ja	MCRMeta LangText
3	Erhaltungszustand des Textes	Feste Begriffe (vgl. Klassifikations-Liste)	Textbeschreibung	ja	max 5	allgemein	ja	MCRMeta Classification 0000022
4	Erhaltungszustand des Textes verbal	Ggf. weitere Angaben	Textbeschreibung	nein	max 10	allgemein	ja	MCRMeta LangText
5	Textart	Literarisch, dokumentarisch, andere, unbestimmbar	Textbeschreibung	ja	nein	allgemein	ja	MCRMeta Classification 0000021
6	Textthema	Epos, Homer, Steuerwesen; vgl. Klassifikations-Liste	Textbeschreibung	ja	max 5	allgemein	ja	MCRMeta Classification 0000026
7	Inhaltsangabe	Digest (Personen, Orte, Sachbegriffe)	Textbeschreibung	nein	max 5	allgemein	ja	MCRMeta LangText
8	Abfassung und Bestimmung	Ortsname	Textbeschreibung	ja	max 5	allgemein	ja	MCRMeta Classification PapyriPlace
8b	Abfassung und Bestimmung	Gauname	Textbeschreibung	ja	max 5	allgemein	ja	MCRMeta Classification PapyriNome
9	Antikes Archiv	Antikes Archiv etc. Link nach Leuven	Textbeschreibung	nein	max 5	allgemein	ja	MCRMeta Link
10	Datierung, numerisch	Vgl. Regeln	Textbeschreibung	ja	max 5	allgemein	ja	MCRMetaHist oryDate
11	Datierung ausführlich	Kriterien (explizit, Prosopographie, Paläographie, etc.)	Textbeschreibung	nein	max 5	allgemein	nein	MCRMeta LangText
12	Datierungsvermutung	interne Angabe	Textbeschreibung	nein	max 5	intern	nein	MCRMeta LangText
13	Inhaltliche Verknüpfung	Verweis zu externen Quellen, zu Duplikaten, Parallelen	Textbeschreibung	nein	max 5	allgemein	nein	MCRMeta LangText
14	Kolumnenzahl		Kodikologisches	nein	max 5	allgemein	ja	MCRMeta Number
15	Zeilenzahl	Je Kolumne	Kodikologisches	nein	max 5	allgemein	ja	MCRMeta Number
16	Tinte	Farbe, Konsistenz	Kodikologisches	ja	max 5	allgemein	ja	MCRMeta Classification 0000025
17	Buchstabenhöhe		Kodikologisches	nein	max 5	je nach Fall	ja	MCRMeta Number
18	Zeilenlänge		Kodikologisches	nein	max 5	je nach Fall	ja	MCRMeta Number
19	Zeilenabstand		Kodikologisches	nein	max 5	je nach Fall	ja	MCRMeta Number
20	Kolumnenhöhe		Kodikologisches	nein	max 5	je nach Fall	ja	MCRMeta Number
21	Schriftart	mit Merkmalen, z.B. Apices, verschiedene Hände	Kodikologisches	nein	max 5	je nach Fall	ja	MCRMeta LangText
22	Schriftrichtung	parallel zur Faser	Kodikologisches	ja	nein	allgemein	ja	MCRMeta Classification 0000023
23	antike Gliederungszeichen, -angaben		Kodikologisches	nein	max 5	allgemein	ja	MCRMeta LangText
24	Illustrationen	ggf. beschreibende Stichworte	Kodikologisches	nein	max 5	allgemein	ja	MCRMeta LangText
25	Seitenfolge	Rückseite Fortsetzung bzw. Adresse / getrennt	Kodikologisches	nein	max 5	allgemein	nein	MCRMeta LangText
26	Editio princeps	Vgl. Regeln	wissenschaftliche Bearbeitung	nein	max 5	allgemein	nein	MCRMeta LangText
27	Weitere Editionen	Vgl. Regeln	wissenschaftliche Bearbeitung	nein	max 5	allgemein	nein	MCRMeta LangText
28	Textsammlungen und Referenzwerke	Pack, van Haelst, Leuvenener Datenbank, DDBDP, Perseus, Heidel. Gesamtverz.	wissenschaftliche Bearbeitung	nein	max 5	allgemein	nein	MCRMeta Link
29	Text publiziert		wissenschaftliche Bearbeitung	nein	max 100	allgemein	nein	MCRMeta LangText
30	Text intern		wissenschaftliche Bearbeitung	nein	max 100	intern	nein	MCRMeta LangText
31	Übersetzung publiziert	Hier kann Text der Übersetzung eingegeben werden.	wissenschaftliche Bearbeitung	nein	max 100	allgemein	ja	MCRMeta LangText
32	Übersetzung intern		wissenschaftliche	nein	max 100	intern	ja	MCRMeta

Nr.	Feld	Beschreibung	Zuordnung	Pflicht	wiederholbar	View	Suche	Metadaten Typ
			Bearbeitung					LangText
33	Berichtigung publiziert		wissenschaftliche Bearbeitung	nein	max 100	allgemein	nein	MCRMeta LangText
34	Berichtigung intern		wissenschaftliche Bearbeitung	nein	max 100	intern	nein	MCRMeta LangText
35	Literatur allgemein		wissenschaftliche Bearbeitung	nein	max 5	allgemein	nein	MCRMeta LangText
36	Literatur intern		wissenschaftliche Bearbeitung	nein	max 5	intern	nein	MCRMeta LangText
37	Editoren	Ersteditor, weitere Bearbeiter	Benutzer	nein	max 5	allgemein	ja	MCRMeta Classification 00000024
38	Publikationserlaubnis	Auch Überlassung von Fotos usw.	Benutzer	nein	max 5	allgemein	nein	MCRMeta LangText
39	Bemerkungen, extern		Textüberblick	nein	max 5	allgemein	nein	MCRMeta LangText
40	Bemerkungen, intern		Textüberblick	nein	max 5	intern	nein	MCRMeta LangText
41	Publikation	Verknüpfung mit Publikationsbeschreibung	Identifizierung	nein	max 5	allgemein	nein	MCRMeta Link
42	Publikationsnummer	z. B. publiziert P.Hal. 1	Identifizierung	nein	max 5	allgemein	ja	MCRMeta LangText
43	WOID	ID des Schrifträgers	Administration	ja	nein	unsichtbar	ja	MCRMeta LinkID
44	Bearbeiter	Die Bearbeiter des Datensatzes	Administration	ja	max 10	allgemein	nein	MCRMeta LangText
46	Schrift(en)	Vgl. Klassifikationsliste	Textbeschreibung	ja	max 5	allgemein	ja	MCRMeta Classification 00000035
47	Sprache(n)	Vgl. Klassifikationsliste	Textbeschreibung	ja	max 5	allgemein	ja	MCRMeta Classification 00000036
48	Schrift / Sprache verbal		Textbeschreibung	nein	max 5	allgemein	nein	MCRMeta LangText
	Status	Der Status des Datensatzes	Administration	ja	nein	intern	nein	statusOfWorkflow

Tabelle 7: Datenstruktur Text

## Erläuterungen zu den einzelnen Feldern des Textes

Nr.	Feld	Kommentar	Länge
1	Titel	Notwendig zur Identifizierung der Daten in der Präsentation.	128
2	Inventarnummer	Können nicht vereinheitlicht werden, vgl. Regeln	128
3	Erhaltungszustand des Textes	Die Angaben beziehen sich sowohl auf den Schriftträger als auch auf den Text. Aber wesentlich ist die Bedeutung für den Erhaltungszustand des Textes. Hier kann auch eine Liste mit mehreren Angaben stehen. (vgl. Klassifikations-Liste)	
4	Erhaltungszustand des Textes (verbal)	Hier könnte frei formuliert das stehen, was bei P. Lips. II in petit quasi als mise en page erscheint. (u.a. Faltungen, Ränder)	4096
5	Textart	Dokumentarisch, literarisch, ... (vgl. Klassifikations-Liste)	
6	Textthema	Epos, Homer, ... (vgl. Klassifikations-Liste)	
7	Inhaltsangabe	Knappe Inhaltsangabe ; andere Angaben bei unveröffentlichten Papyri zu Nebenpersonen, Orten, usw. können ins Feld TE40 (Bemerkungen intern)	4096
8	Abfassung und Bestimmung	<i>Abfassung und Bestimmung</i> gliedert sich in die Teilfelder <ul style="list-style-type: none"> <li>o Ortsnamen (vgl. Klassifikations-Liste)</li> <li>o Gaunamen (vgl. Klassifikations-Liste)</li> </ul> Da es ein Pflichtfeld ist muss hier ggf. ‚unbekannt‘ eingetragen werden.	
9	Antikes Archiv	Nach Möglichkeit mit Link auf die Leuener Sammlung. Hier kommt eine URL hin. Vgl. auch Regeln. Diese kann auch Web-Seiten oder Dokumente des Papyrus-Projektes referenzieren.	
10	Datierung, numerisch	Vgl. Regeln	
11	Datierung, ausführlich	Hier sind erläuternde Texte für die angegebene Datierung einzutragen. u. U. Zitat eines anderen Papyrus.	4096
12	Datierungsvermutung	Hier sind erläuternde Texte für die vermutete Datierung einzutragen.	4096
13	Inhaltliche Verknüpfung	Welche Parallelen angegeben werden, wird für jeden Fall entschieden. Hier genügen Publikationsnummern zu Duplikaten, externen Quellen, etc.	4096
14	Kolumnenzahl	Nur erhaltene Kolumnen angeben. Bei einer einzigen Kolumne kann das Feld auch frei bleiben.	
15	Zeilenzahl	Nur erhaltene Zeilen zählen. Bei Unsicherheit kann dies im Feld Bemerkungen erläutert werden. Bei mehreren Kolumnen wird eine Liste der Zeile pro Kolumne angegeben.	
16	Tinte	Vgl. Klassifikationsliste	
17	Buchstabenhöhe	Es handelt sich um einen Durchschnittswert. Interessant sind hier die Ausreißer mit ungewöhnlichen Größen. Die Angabe erfolgt in cm.	
18	Zeilenlänge	Soweit erhalten erfolgt die Angabe in cm.	
19	Zeilenabstand	Es handelt sich um einen Durchschnittswert. Interessant sind hier die Ausreißer mit ungewöhnlichen Größen. Die Angabe erfolgt in cm.	
20	Kolumnenhöhe	Die Angabe erfolgt in cm.	
21	Schriftart	Hier erfolgt eine möglichst knappe Beschreibung oder Zuordnung der Schrift: "Biblische Unziale", "Strenger Stil", "Buchschrift" "Kanzleischrift" "Akzente", "Punktuaton" "Kursiv". Auffällige Neigung nach rechts oder links. (vgl. Schubart, Paläographie). Dieses Feld dient auch für eine spätere Edition. Bei Papyri, die verloren sind, ist man für solche Beschreibungen dankbar.	1024
22	Schriftrichtung	Vgl. Klassifikationsliste. Ggf. sollte bei Codizes die Schriftrichtung auf der ersten erhaltenen Seite angegeben werden.	
23	antike Gliederungszeichen, -angaben	Beispiele: Paragraphos, Leerraum, Einrückung (Eisthesis), Auszug (Ekthesis), Koronis	1024
24	Illustrationen	Hier stehen die Illustration beschreibende Stichworte.	1024
25	Seitenfolge	z. B. Rückseite Fortsetzung oder Adresse bei einem Brief	1024
26	Editio princeps	Hier erfolgt nur die Angabe der ersten vollständigen Edition. Descripta und weitere Editionen werden in TE27 genannt. Ausnahme: der Papyrus ist bisher nur unvollständig publiziert/beschrieben. Vgl. Regeln	1024
27	Weitere Editionen	Verbale Angaben zu weiteren Editionen. Vgl. Regeln	1024
28	Textsammlungen und Referenzwerke	Hier kommen URLs hin. Anmerkungen sollten unter Bemerkung extern stehen. (z. B. URL-Verweise auf die ‚Heidelberger Gesamtverzeichnis‘, ‚Leuener Database of Ancient Books‘ oder die ‚Duke Databank of Documentary Papyri‘). Link auf Mertens-Pack immer angeben nach der Syntax: "Mertens-Pack [Nr.]: [URL]", auch wenn der Link nur auf die Startseite führt. Siehe auch den Abschnitt ‚Codierung von URLs‘.	
29	Text publiziert	Hier kommt der publizierte originale Text hin.	4096
30	Text intern	Hier kommt der unpublizierte originale Text hin.	4096
31	Übersetzung publiziert	Hier kommt die publizierte deutsche (und ggf. englische) Übersetzung hin.	4096
32	Übersetzung intern	Hier kommt die unpublizierte deutsche (und ggf. englische) Übersetzung hin.	4096
33	Berichtigung publiziert	Hier kommt die publizierte deutsche (und ggf. englische) Berichtigung hin. Hier steht auch der Verweis auf die Berichtigungsliste (später kommt der Text aus der Berichtigungsliste hinzu). Sollten im „Heidelberger Gesamtverzeichnis“ zusätzliche Berichtigungen aufgeführt sein, ist darauf zu verweisen. Andere Berichtigungen und Neulesungen werden nach den Regeln der Literaturangaben gemacht.	4096
34	Berichtigung intern	Hier kommt die unpublizierte deutsche (und ggf. englische) Berichtigung hin.	4096
35	Literatur allgemein	Vgl. Regeln der Literaturangaben	1024
36	Literatur intern	Vgl. Regeln der Literaturangaben	1024
37	Editoren	Verzeichnet alle Editoren (Erstedition, Neuedition etc.), aber keine Autoren von Berichtigungen. (vgl. Klassifikationsliste)	
38	Publikationserlaubnis	Interne Informationen zur erteilten Publikationserlaubnis.	1024
39	Bemerkungen, extern	verbaler Text	4096
40	Bemerkungen, intern	verbaler Text	4096
41	Publikation	URL für eine digitale Publikation oder Edition. Verlinkung auf digitalisierte/gescannte Papyrustexte – wie z.B. literarische Papyri, die nicht in der DDBDP aufgenommen sind.	
42	Publikationsnummer	Wie Dokumentation: "publiziert Publ.Nummer [weitere Publ.Nummer(n)]" (nach Checklist) oder "unpubliziert". Bei Texten, die keine eigene Publ.sn. haben (in Zeitschriften, Festschriften, usw.) ist Publikationsort mit Seitenzahl und evtl. der dortigen Nummerierung anzugeben (nach den Regeln der Angaben für Sekundärliteratur). Bei längeren Titeln sind Abkürzungen mit Verweis auf ed. pr. und Editionen möglich.	128
43	WOID	ID des Schriftträgers	
44	Bearbeiter	Hier stehen Einträge aus einer Liste von Bearbeitern.	128
46	Schrift(en)	Auswahl der Objektschrift (vgl. Klassifikations-Liste)	
47	Sprache(n)	Auswahl der Objektsprache (vgl. Klassifikations-Liste)	
48	Schrift/Sprache verbal	verbale Beschreibung	

Status	Auswahl aus der Liste der Bearbeitungszustände	
--------	--	--

Tabelle 8: Ausfüllhinweise Text

## 2.4 Datenstruktur für zusätzliche Dokumente

Für die Abbildung der ergänzenden Dokumente wurde das standardisierte MODS-Format<sup>11</sup> genutzt. Die entsprechenden Datenzuordnungen sind gemäß der Spezifikation umgesetzt. Dabei wird jedoch nur ein abgespecktes Modell für MODS genutzt, welches die relevanten Daten enthält.

## 2.5 Vererbung der Schrifträgerdaten für fragmentarische Papyri

Um hierarchische Strukturen zur Darstellung von fragmentarischen Papyri in das System einbringen zu können, muss festgelegt werden, welche Daten des definierten Datenmodells vererbbar sein sollen. Dabei ist zu beachten, dass immer nur ganze Metadatentypen gemäß MyCoRe-Definition vererbbar oder von der Vererbung auszuschließen sind.

Bei der Auswahl des Standardverhaltens der einzelnen Felder wurden folgende Kriterien beachtet:

- Beschreibt das Datenfeld einen Zustand oder eine Eigenschaft, welche nur dem speziellen Fragment anhaftet (z. B. Maße), so ist das Datenfeld **nicht** vererbbar.
- Handelt es sich bei dem oben beschriebenen Datenfeld um ein Pflichtfeld, so sind hier entweder die für den gesamten Papyrus geltenden Werte (z. B. Maße) oder ein Verweis auf die Fragmente (z. B. Zustand) anzugeben.
- Ist das Datenfeld für alle Fragmente gültig, so wird es nur im Schrifträger angegeben (z. B. Sprachen).
- Der Schrifträger ist immer mit allen Pflichtfeldern zu versehen. Dabei ist ggf. auch ein Verweis auf die Fragmente einzutragen.
- Fragmente müssen nicht in allen Pflichtfeldern ausgefüllt werden, da sie diese teilweise vom Schrifträger erben (siehe Tabelle).

Die nachfolgende Tabelle beschreibt die Festlegungen und Möglichkeiten der Dateneingabe bei strukturierten Schrifträgern für fragmentarische Papyri.

Nr.	Feld	Pflicht St.	vererbbar	Pflicht Frag.	Anmerkung
1	Titel	ja	ja	nein	
2	Inventarnummer	ja	nein	ja	Im St. als Liste der einzelnen Inv. Nr.
3	Andere Ordnungskriterien	nein	nein	nein	
4	Sammlung	ja	nein	ja	Für St. Klassifikation 1 # 5 angeben
5	Unterbringung	nein	nein	nein	
6	Erwerbung	ja	nein	ja	Im St. nur unter Bemerkung einen Hinweis auf die Fragmente.
7	Publikation	nein	nein	nein	
8	Publikationsnummer	nein	nein	nein	
9	Material	ja	ja	-	
10	Farbe (grob)	nein	nein	nein	
11	Farbe (verbal)	nein	nein	nein	
12	Maße	nein	nein	ja	Es sind jeweils die entsprechenden Werte anzugeben.
13	Zustand	ja	nein	ja	Für Schrifträger Klassifikation 3 # 14 angeben.
14	Komplett	ja	ja	nein	
15	Aufbewahrung	ja	nein	ja	Für Schrifträger Klassifikation 5 # 6 angeben.

11 <http://www.loc.gov/standards/mods/>

Nr.	Feld	Pflicht St.	vererbbar	Pflicht Frag.	Anmerkung
16	Erhaltung	nein	nein	nein	
17	Ausführung	ja	ja	-	
18	rekonstruierter Umfang	nein	nein	nein	
19	Klebungen	nein	nein	nein	
20	phys. Zugehörigkeit, publik	nein	ja	nein	
21	Fremdlink	nein	ja	nein	
22	phys. Zugehörigkeit (interne Vermerke)	nein	ja	nein	
23	phys. Zugehörigkeit (interne Links)	nein	ja	nein	
24	Publikationsflag	ja	nein	ja	
30	Abbildung, publiziert	nein	ja	nein	
31	Kopie in Medien	nein	nein	nein	
32	Bemerkungen, extern	nein	ja	nein	
33	Bemerkungen, intern	nein	ja	nein	
34	Weitere Links	nein	ja	nein	
35	Bearbeiter	ja	nein	ja	
36	Lizenz	nein	nein	nein	
37	alternative ID	nein	nein	nein	
38	interne alt. ID	nein	nein	nein	

*Tabelle 9: Vererbungsmatrix*

Im Normalfall beziehen sich alle Textbeschreibungen auf den gesamten Schriftträger. Für Ausnahmen wie den Papyrus Ebers, der noch in Kolonnen eingeteilt wird, kommen später noch ergänzende Regeln hinzu. Bei Papyri, deren Fragmente auf mehrere Lokationen verteilt sind, müssen die Fragmente aus technischen Gründen in jeder Server-Instanz gespeichert sein.

## 2.6 Klassifikationen

Die gemeinsamen Klassifikationen bilden eine wesentliche Stütze des Projektes, da über sie eine einheitliche Benennung von Sachverhalten so möglich ist, dass alle Beteiligten bei der Dateneingabe auf dieselben Werte referenzieren. Dies hat besonders bei den Orts- und Gaunamen Bedeutung, da hier für die Suche eine einheitliche Schreibweise festgelegt werden kann. Im Projekt sind offene und geschlossene Klassifikationen bekannt. Während für die Erweiterung einer offenen Klassifikation lediglich eine Mitteilung an den Administrator erforderlich ist, muss für eine Erweiterung der geschlossenen Klassifikationen die Zustimmung der Projektleitung vorliegen. Die Werte werden in einer MyCoRe-konformen XML-Datei abgelegt. Die aktuellen Klassifikationen sind über den Leipziger GitLab-Server unter folgendem Link zu finden:

[https://git.sc.uni-leipzig.de/mycore\\_applications/papyri/dptbase-papdata/tree/master/src/main/resources/classifications](https://git.sc.uni-leipzig.de/mycore_applications/papyri/dptbase-papdata/tree/master/src/main/resources/classifications)

Für die Schrifträger sind folgende Klassifikationen vergeben:

Nr. innerhalb der Metadaten	Art	Dateiname	offen/geschlossen
04	Sammlungen (XML-Datei in den jeweiligen Anwendungen zu finden)	Papyri_class_00000001	geschlossen
09	Material	Papyri_class_00000002	geschlossen
10	Farben	Papyri_class_00000010	geschlossen
13	Zustand	Papyri_class_00000003	geschlossen
14	Komplett	Papyri_class_00000004	offen
15	Aufbewahrung	Papyri_class_00000005	geschlossen
17	Ausführung	Papyri_class_00000006	geschlossen
24	Publikation	Papyri_class_00000014	geschlossen

*Tabelle 10: Schrifträger Klassifikationen*

Für die Texte sind folgende Klassifikationen vergeben:

Nr. innerhalb der Metadaten	Art	Dateiname	offen/geschlossen
03	Erhaltungszustand	Papyri_class_00000022	geschlossen
05	Grobbestimmung (Textart)	Papyri_class_00000021	geschlossen
06	Textthema	Papyri_class_00000026	geschlossen
08a	Ortsname	Papyri_class_00000031	offen
08b	Gauname	Papyri_class_00000032	offen
16	Tinte	Papyri_class_00000025	geschlossen
22	Schriftrichtung	Papyri_class_00000023	geschlossen
37	Editoren	Papyri_class_00000024	offen
46	Schrift	Papyri_class_00000035	geschlossen
47	Sprache	Papyri_class_00000036	geschlossen

*Tabelle 11: Text Klassifikationen*



Einige Klassifikationen werden für alle Anwendungen genutzt und befinden sich in übergeordneten Modulen:

[https://git.sc.uni-leipzig.de/mycore\\_applications/common/dptbase-common/tree/master/src/main/resources/classifications](https://git.sc.uni-leipzig.de/mycore_applications/common/dptbase-common/tree/master/src/main/resources/classifications)

Name in den Metadaten	Art	Anmerkung	Dateiname	offen/geschlossen
servstate	Status	Betrifft Schriftträger und Texte	statusOfWorkflow	geschlossen
servflag type="license"	Lizenzen	Betrifft Schriftträger und Fragmente	licenses	geschlossen

Tabelle 12: sonstige Klassifikationen

## 2.7 Exportformate

### 2.7.1 Mapping auf OAI Dublin Core

Um den Zugriff auf die Daten auch über eine OAI-Schnittstelle<sup>12</sup> zu gestatten, muss ein Mapping der Schriftträgerdaten auf Exportdatenmodell erfolgen. Geplant sind Exporte nach LIDO und EAD. Bereits implementiert ist das Mapping auf Dublin Core (DC) und sowie MODS. Die nachfolgende Tabelle beschreibt diese Zuordnung.

Dublin Core Feld	Feld	Anmerkung
dc:title	ST01 - Titel	
dc:creator	ST35 - Bearbeiter	
dc:description	TE07 - Inhalt	
dc:subject	TE05 - Textart TE06 - Textthema	
dc:publisher	ST04 - Sammlung	
dc:contributor		Fest: Prof. Dr. Reinhold Scholl; Universitätsbibliothek Leipzig Jens Kupferschmidt, Universitätsrechenzentrum der Universität Leipzig
dc:date	servdate[createdate]	
dc:type	ST09 - Material	
dc:identifier	ST02 - Inventarnummer @ID	
dc:format	ST17 - Ausführung	
dc:language	TE47 - Sprache	
dc:rights	servflag[license]	
source and relation		nicht belegt
dc:coverage	TE08 - Ort TE08b - Gau	

Tabelle 13: Mapping Schriftträger nach Dublin Core

<sup>12</sup> <http://www.openarchives.org/>

## 2.7.2 Mapping auf MODS

Beschreibung der Zuordnung nach MODS (Vers 3.7, Januar 2018). Der Schriftrager ist hierbei das Hauptelement und wird durch das Datenobjekt Text im Element relatedItem erweitert.

<i>MODS</i>	<i>MyCoRe</i>	<i>Beispiel und Bemerkung</i>
<modsCollection>		
<titleInfo> <title>	ST01 - Titel	Schluß einer Schenkungsurkunde
<name> <namePart> <role> <roleTerm> <roleTerm>	ST35 - Bearbeiter  → „creator“ → „author“	Quenouille
<typeOfResource>	→ „text“	
<physicalDescription> <digitalOrigin> <form> @type=“material“ @type=“color“ @type=“condition“ @type=“completeness“ @type=“retention“ @type=“preservation“ @type=“format“ @type=“reconstruction“ @type=“gluing“ <extent>	→ „reformatted digital“  ST09 - Material ST10 - Farbe ST13 - Zustand ST14 - Vollstandigkeit ST15 - Aufbewahrung ST16 - Erhaltung ST17 - Ausfuhrung ST18 - rekonstruierter Umfang ST19 - Klebung ST12 - Mae	Papyrus ( <b>type=material</b> ) hellbraun ( <b>type=colour</b> ) verschmutzt ( <b>type=state of preservation</b> )  in Glas  Codex  Mae (height / width / thickness)
<note> @type=“publications“ <note> @type=“internal“ @type=“external“	ST08 - Publikationsnummer  ST33 - Bemerkungen, intern ST32 - Bemerkungen, extern	P.Hal. 1
<originInfo> @eventType=“acquisition“ <place/placeTerm> <dateOther> @eventType=“excavation“ <place/placeTerm> <dateOther>  <edition> <place> <placeTerm> <publisher>	ST06 Kaufort ST06 Kaufdatum/Verlosungstag  ST06 Fundort ST06 Funddatum  → „[Electronic ed.]“	<b>(Sammlungsbezeichnung je nach Eintragung in ST04 - Sammlungen)</b> Bonn Abteilung fur alte Geschichte
<relatedItem type="constituent" > <titleInfo> <title> <name> <namePart> <role> <roleTerm> <roleTerm> <originInfo> <dateIssued> <dateIssued> <place> <placeTerm> <language> <languageTerm> <scriptTerm> <physicalDescription> <form> @type=“ink“ @type=“writing direction“ @type=“condition“ @type=“columns“	TE01 - Titel  TE37 - Editoren  → „creator“ → „author“  TE10 - Datierung, numerischen TE11 - Datierung, ausfuhrlich  TE08 - Abfassungsort  TE47 - Sprachen TE46 - Schrift  TE16 - Farbe TE22 - Schriftrichtung TE03 - Erhaltungszustand TE14 - Kolumnenzahl	Papyrus Ebers  Quenouille  1. Jh. n. Chr.  Theben  Griechisch griechisch  hellbraun

@type="lines" @type="letter size" @type="line length" @type="line spacing" @type="column height" @type="writing type" @type="structuring signs" @type="illustration" @type="page order" <abstract> <subject> <genre> <topic> <geographic> <relatedItem> @type="isReferencedBy" <location> <identifier> @type="local" <relatedItem> @type="isReferencedBy" <location> <url> <note> @type="internal" @type="external"	TE15 - Zeilenzahl TE16 - Buchstabenhöhe TE17 - Zeilenlänge TE18 - Zeilenabstand TE19 - Kolumnenhöhe TE21 - Schriftart TE23 - antike Gliederungszeichen TE24 - Illustrationen TE25 - Seitenfolge TE07 - Inhaltsangabe TE05 - Textart TE06 - Textthema TE08b - Abfassungsgau  TE41 - Publikation  TE28 - Textsammlungen  TE40 - Bemerkungen, intern TE39 - Bemerkungen, extern	dokumentarisch Weiheschriften Bompae
<relatedItem type="host" > <titleInfo> <title> <typeOfResource> <genre> <subject> <topic> <location> <url> <name>	→ „Papyrus- und Ostrakasamm- lung“ → „software, multimedia“ → „online system or service“  → „Handschrift“  \$WebApplicationBaseURL → „contributor“	<a href="https://papyri-collection.dl.uni-leipzig.de/">https://papyri-collection.dl.uni-leipzig.de/</a> Kupferschmidt, Scholl
publi<identifier>	ST37 – alternative ID	
<identifier>	→ URL	
<location> <physicalLocation> @type="collection" @type="storage" <shelfLocator> <url>	ST04 - Sammlung ST05 - Unterbringung ST02 - Inventarnummer ST07 → xlink:href	Leipzig, Cod. Tischendorfianus  P.Lips.Inv. F. 35 <b>Publikationsbeschreibung</b>
<accessCondition>	→ ST36, „All rights by the owner and / or publisher of the collection.“	
<recordInfo> <recordIdentifier> <recordCreatingDate> <recordChangeDate>	→ URL → createdate → modifydate	

Tabelle 14: Mapping Schriftträger nach MODS

### 2.7.3 Mapping auf LIDO

Beschreibung der Zuordnung nach LIDO (Vers 1.1).

LIDO	MyCoRe	Beispiel und Bemerkung
<lidoWrap> <lido>		
<lidoRecID> @source="Universität Leipzig, UBL Papyrussammlung" @type="local"	ST02 – Inventarnummer	
lidoRecID>	@label	

<pre>@source="Universität Leipzig – UBL Papyrussammlung" @type="sortable"</pre>		
<pre>&lt;objectPublishedID&gt;</pre>	ST37 – alternative ID	
<pre>&lt;descriptiveMetadata&gt; &lt;objectClassificationWrap&gt; &lt;objectWorkTypeWrap&gt; &lt;objectWorkType&gt; &lt;term&gt; ST09 - Material  &lt;objectIdentificationWrap&gt; &lt;titleWrap&gt; &lt;titleSet&gt; &lt;appellationValue&gt; ST01 – Titel  &lt;repositoryWrap&gt; &lt;repositorySet&gt; &lt;repositoryName&gt; &lt;legalBodyName&gt; &lt;appellationValue&gt; ST04 – Sammlung  &lt;workID&gt;  &lt;objectDescriptionWrap&gt; &lt;objectDescriptionSet&gt; @type="color" &lt;descriptiveNoteValue&gt; ST10 - Farbe  &lt;objectDescriptionSet&gt; @type="condition" &lt;descriptiveNoteValue&gt; ST13 – Zustand  &lt;objectDescriptionSet&gt; @type="completeness" &lt;descriptiveNoteValue&gt; ST14 – Vollständigkeit  &lt;objectDescriptionSet&gt; @type="form" &lt;descriptiveNoteValue&gt; ST17 – Ausführung  &lt;objectDescriptionSet&gt; @type="texttype" &lt;descriptiveNoteValue&gt; TE05 – Textart  &lt;objectDescriptionSet&gt; @type="script" &lt;descriptiveNoteValue&gt; TE46 – Schrift  &lt;objectDescriptionSet&gt; @type="language" &lt;descriptiveNoteValue&gt; TE47 – Sprache  &lt;objectDescriptionSet&gt; @type="remark" &lt;descriptiveNoteValue&gt; ST32 – Bemerkung  &lt;objectMeasurementsWrap&gt; &lt;objectMeasurementsSet&gt; @type &lt;objectMeasurements&gt; &lt;measurementsSet&gt; &lt;measurementType&gt; &lt;measurementUnit&gt; &lt;measurementValue&gt; ST12 – Maße: @dimension ST12 – Maße: @dimension ST12 – Maße: @measurement ST12 – Maße: Text  &lt;inscriptionsWrap&gt; &lt;inscriptions&gt; &lt;inscriptionTranscription&gt; @xml:lang TE29 – Text publiziert  &lt;inscriptionDescription&gt; @type="translation" &lt;descriptiveNoteValue&gt; TE31 – Übersetzung publiziert  &lt;inscriptionDescription&gt; @type="condition" &lt;descriptiveNoteValue&gt; TE03/TE04 – Erhaltungszustand des Textes</pre>		

<pre> &lt;inscriptionDescription&gt;   @type="content"   &lt;descriptiveNoteValue&gt;  &lt;inscriptionDescription&gt;   @type="ink"   &lt;descriptiveNoteValue&gt;  &lt;inscriptionDescription&gt;   @type="writing"   &lt;descriptiveNoteValue&gt;  &lt;inscriptionDescription&gt;   @type="literature"   &lt;descriptiveNoteValue&gt;  &lt;inscriptionDescription&gt;   @type="remark"   &lt;descriptiveNoteValue&gt;  &lt;eventWrap&gt;   &lt;eventSet&gt;   &lt;event&gt;   &lt;eventType&gt;   &lt;conceptID&gt;   @type="production"   @source="http://terminology.lido-schema.org/eventType/production"   &lt;eventDate&gt;   &lt;displayDate&gt;   &lt;eventPlace&gt;   &lt;displayPlace&gt;    &lt;eventSet&gt;   &lt;event&gt;   &lt;eventType&gt;   &lt;conceptID&gt;   @type="excavation"   @source="http://terminology.lido-schema.org/eventType/excavation"   &lt;eventActor&gt;   &lt;actorInRole&gt;   &lt;actor&gt;   &lt;nameActorSet&gt;   &lt;appellationValue&gt;   &lt;roleActor&gt;   &lt;conceptID&gt;   @type="lidoActorRoles"   @source="http://mycore.uni-leipzig.de/classifications/lidoActorRoles"   &lt;eventDate&gt;   &lt;displayDate&gt;   &lt;eventPlace&gt;   &lt;displayPlace&gt;  &lt;objectRelationWrap&gt;   &lt;subjectWrap&gt;   &lt;subjectSet&gt;   &lt;displaySubject&gt; </pre>	<p>TE07 – Inhaltsangabe</p> <p>TE16 – Tinte</p> <p>TE21; TE22 – Schriftart; Schriftrichtung</p> <p>TE35 – Literatur</p> <p>TE39 – Bemerkungen, external</p> <p>„production“</p> <p>TE10/TE11 – Datierung</p> <p>TE08/TE08b – Ort; Gau</p> <p>„excavation“</p> <p>ST06 – Erwerbung → Finder</p> <p><u>lido_actorRole_excavator</u></p> <p>ST06 – Erwerbung → Datierung</p> <p>ST06 – Erwerbung → Fundort</p> <p>TE06 – Textthema</p>	
<pre> &lt;administrativeMetadata&gt;   &lt;rightsWorkWrap&gt;   &lt;rightsWorkSet&gt;   &lt;rightsHolder&gt;   &lt;legalBodyName&gt;   &lt;appellationValue&gt;    &lt;recordWrap&gt;   &lt;recordID&gt;   @type="MyCoreID"    &lt;recordType&gt;   &lt;conceptID&gt;   @type="recordType"   @source="http://terminology.lido-schema.org/recordType"    &lt;recordSource&gt;   @type="editor"   &lt;legalBodyName&gt;   &lt;appellationValue&gt; </pre>	<p>„Universität Leipzig, UBL Papyrus-sammlung“</p> <p>MyCoRe-ID</p> <p>item-level_record</p> <p>ST35 – Bearbeiter</p>	

<pre> &lt;recordSource&gt;   @type="repository"   &lt;legalBodyName&gt;     &lt;appellationValue&gt;    &lt;legalBodyWebLink&gt;  &lt;recordRights&gt;   &lt;rightsType&gt;   &lt;conceptID&gt;     @source     @type   &lt;term&gt; </pre>	<pre> „Universität Leipzig, UBL Papyrus- sammlung“ Link zur Sammlung  ST36 – Lizenz → @categid Mycore Klassifikation (@classid) @categid ST36 – Lizenz → Text </pre>	
---	--	--

Tabelle 15: Mapping Schrifträger nach LIDO

## 2.8 Service-Informationen

Neben dem eigentlichen Datenmodell führt das System noch eine Reihe von Informationen in jedem Datensatz mit, die zur Überwachung der Arbeit erforderlich sind. Diese Daten werden im Service-Teil des Datensatzes abgelegt. Dies sind Datumsangaben, Zugriffslisten und Flags.

### 2.8.1 Datumsangaben

Hierbei handelt es sich um das Ersteinstellungsdatum (,createdate') und das letzte Änderungsdatum (,modifydate'). Beide Angaben werden beim Laden bzw. beim Update des Datensatzes automatisch erzeugt.

### 2.8.2 Bilddarstellung

Die Zugriffsrechte spielen eine wichtige Rolle bei der Präsentation der Daten. Hier wird abhängig davon, ob ein Benutzer ein bestimmtes Zugriffsrecht hat, die Darstellung des Suchergebnisses gefiltert. Zusätzlich spielt dabei für die Bereitstellung der Bilder in den verschiedenen Auflösungen auch das Feld ST24 (Publikationsart) eine wichtige Rolle. Es wird dabei zwischen **Recto** und **Verso** unterschieden. Bilder, bei denen ST24 als unpubliziert-gesperrt markiert ist, werden nur in besserer Auflösung (300 dpi) angezeigt, wenn der Benutzer das entsprechende Zugriffsprivileg hat. Der Status **unpubliziert-gesperrt** ist die Standardannahme. In Fall von **publiziert** oder **unpubliziert-frei** kann die veröffentlichte Seite in 300 dpi gezeigt werden, während die unveröffentlichte oder gesperrte Seite dem Gast nur in 72 dpi zur Verfügung steht, dem entsprechenden Inhaber der Rechte jedoch in höherer Auflösung.

# 3 Regeln zur Erstellung / Bearbeitung neuer Datensätze

Die hier aufgeführten Regeln dienen der Qualitätssicherung und sollten dringend beachtet werden, um u.a. Identifizierungen zu gewährleisten.

- Der Verweis von den Texten erfolgt immer auf die interne MCRObjectID des Schriftträgers.
- Die Standardsprache ist deutsch, optional könnten Informationen noch in Englisch und Französisch abgelegt werden.
- Orts- und Personennamen sind in lateinischer Schrift anzugeben. Griechische Namen behalten ihre griechische Form (Oxyrhynchos; Ptolemaios). Das griechische „ou“ wird mit „u“ geschrieben (Panobchunis, Patus); Ausnahme: bei Römern wird „ou“ zu „v“ (Flavios). Es heißt „Hermupolis“, aber der Gau „Hermopolites“.
- Orts und Gaunamen werden als offene Klassifikationen gespeichert.
- Unsichere Ortsnamen werden durch , ?' (Leerzeichen+Fragezeichen) erweitert.
- Kolumne wird als „Kol.“ abgekürzt mit Zählung in römischen Zahlen (Kol. V).
- Die Berichtigungsliste wird „BL“, der Band mit römischen Zahlen und die Seitenzahl mit Komma abgetrennt in arabischen Zahlen (BL VIII, 233).
- Papyri und Ostraka werden nach der Checklist unter <http://www.papyri.info/docs/checklist> zitiert.

## 3.1 Inventarnummern

Die Inventarnummern sind in jeder Lokation aus historischen Gründen sehr unterschiedlich und bilden somit ein Problem für einheitliche Such- und Sortierstrategien. Daher werden diese Informationen derzeit als Text-String gespeichert. Der individuelle Aufbau der Inventarnummern gestattet es leider nicht, diese in systemseitig sortierter Reihenfolge auszugeben. Daher werden die Inventarnummern ein zweites mal in sinnvoll sortierbarer Notation im Label-Feld des Datensatzes abgelegt. Für jedes Projektmitglied sind die wichtigsten Schemata der Inventarnummern im Folgenden verzeichnet:

---

Leipzig	P.Lips.Inv. Nummer [R V] (Bsp.: <i>P.Lips.Inv. 333R</i> ) P.Lips.Inv. Nummer Fr. Nummer P.Lips.Inv. Nummer, A, Blatt oder Doppelblatt Nummer und Seite (z.B. für das Konvolut) P.Lips.F. Nummer (Bsp.: <i>P.Lips.Inv. F 1</i> )(so die später als <i>P.Leipz. Veröffentlichten Papyri</i> ) FF Nummer (Bsp.: <i>P.Lips.Inv. FF 28</i> ) (=Förderer und Freunde der Universität Leipzig) FF Nummer Kleinbuchstaben (Bsp.: <i>P.Lips.Inv. FF37b</i> ) P.Ebers Taf. I-XXIX, [R/V] Kolumnen Nummer * Direkt zusammenhängende Fragmente werden als ein Datensatz angelegt.
Jena	P.Jen.Inv. Nummer [R/V] [Teilnummer]
Halle	Es existieren zwei verschiedene Inventarisierungssysteme.

---

---

### Papyri der Sammlung Kurth

Römische Ziffer + Kleinbuchstabe [+ arabische Ziffer] [R|V]

### Alle anderen Papyri

P.Hal.Inv. + arabische Ziffer [R|V] [+ Teilbezeichnung: Buchstabe oder Zahl]

---

Würzburg P.Würzb.Inv. [+ arabische Ziffer]

---

Köln P.Köln.Inv. [+ arabische Ziffer]

---

Erlangen P.Erl.Inv. [+ arabische Ziffer]

---

Gießen P.Gies. [+ arabische Ziffer] – Papyri Giessensis  
P.B.U.G. [+ arabische Ziffer] – Papyri bliothecae universitatis Giessensis  
P.Iand. [+ arabische Ziffer] – Papyri Iandanae  
O.Gies. [+ arabische Ziffer] – Ostraca Giessensis

---

Bremen P.Brem. [+ arabische Ziffer]

---

Hamburg P. Ham. graec.

---

KOHD Copti- P [+ arabische Ziffer] [+ Teilbezeichnung]  
ca

---

## 3.2 Projektinternet ID's

Die projektinternen Identifikationsnummern dienen der einheitlichen Bezeichnung des Datensatzes unabhängig von der Inventarnummer oder anderen Merkmalen. Die ID entspricht der im MyCoRe-Projekt verwendeten MCRObjectID. Diese ist für jeden Datensatz eindeutig zu vergeben. Der grundlegende Aufbau ist:

### **projectid\_type\_number**

Für die bisher an diesem Gemeinschaftsprojekt beteiligten bzw. nachnutzenden Einrichtungen wurden folgende **projectids** festgelegt:

<b>Einrichtung</b>	<b>projectid</b>
Basel	<i>BslPapyri</i>
Bonn	<i>BonnPapyri</i> <sup>13</sup>
Bremen	<i>BrmPapyri</i>
Erlangen	<i>ErlPapyri</i>
Gießen	<i>GiePapyri</i>
Halle	<i>HalPapyri</i>
Hamburg	<i>HamPapyri</i>
Jena	<i>IAwJPapyri</i>
Köln	<i>KlnPapyri</i>

13 nur ein Stück



Leipzig	<i>UBLPapyri</i>
Marburg	<i>MarPapyri</i> <sup>14</sup>
Torgau	<i>TorPapyri</i> <sup>15</sup>
Würzburg	<i>WrzPapyri</i>
KOHD	<i>KOHD Papyri</i> <sup>16</sup>

Der **type** ist *schrift* für die Schriftträger-Daten, *fragment* für Fragmente, *text* für die Text-Daten, *mods* für die zusätzlichen Dokumente und *derivate* für alle digitalen Objekte. Die Nummerierung erfolgt für alle Komponenten in 10er-Schritten. Die Nummern sind immer **achtstellig mit Vornullen** anzugeben. Innerhalb des Editor-Dialoges werden die Nummern automatisch vergeben. Die Vergabe der Nummern für die Bilder ist weiter unten in diesem Dokument beschrieben. Seit der Projektversion 3.0 werden die Projekt-IDs automatisch beim Neuanlagen eines Datensatzes vergeben.

Die ID's werden im Projekt zur eindeutigen Verknüpfung zwischen den Daten benötigt und sind bei Bedarf (z. B. beim Ausfüllen eines Text-Formulars) entsprechend den angegebenen Regeln anzugeben. In der Anzeige der interaktiven Suche werden die IDs angezeigt.

### 3.3 Antike Datumsangaben

Antike Datumsangaben bilden ein besonderes Problem bei der Speicherung. Um im Projekt auf eine sehr aufwendige Konvertierung und Normalisierung zu verzichten, wurde eine Aufgliederung der Datumsangabe eingeführt.

*Datierung* gliedert sich in die Teilfelder

- Datierungstext - verbale textliche Beschreibung, z. B. '2. Jh. v. Chr.' oder '1.2.135 v. Chr.'
- Datumsangabe - numerisch **von** Datum
- Datumsangabe - numerisch **bis** Datum
- Kalender – ist optional und derzeit immer **gregorianisch**
- Integerwerte für die **von** und **bis** Angaben

#### 3.3.1 Festlegungen

- Datumsangabe in der numerischen Form: [-][[[**t**].]**m**]**m**.] [**yyy**]**y** [**BC|AD**]
- Standarddatum für den **von**-Wert ist 01.01.4713 BC
- Standarddatum für den **bis**-Wert ist 31.12.4000 AD
- Alle Datumsangaben für **von** und **bis** erfolgen im **Gregorianischen Kalender**. Werte vor der Gültigkeit dieses Kalenders werden entsprechend umgerechnet.
- Bei ante- und post-Datierungsangaben ist immer ein **bis**- bzw. ein **von**-Datum anzugeben.
- In der Präsentation sollten die Datumsangaben in der gefälligeren Form (2. Jh. v. Chr. usw.) erscheinen.
- Im Datumstextfeld werden die Ptolemäerkönige griechisch (Kleopatra, Ptolemaios, Alexandros), die römischen Kaiser lateinisch (Constantius, Hadrianus) geschrieben.
- Diese Syntax gilt für die Felder TE10 bis TE12.

14 Nur zwei Stücke

15 Nur zwei Stücke

16 Eigenständige Anwendung

### 3.3.2 Eingabefelder

verbal - verbaler Text  
von - Anfangsdatum  
bis - Enddatum

### 3.3.3 Beispiele

- verbal - **100 v. Chr. bis 213 n. Chr.**  
von - 100 BC 01.01.0100 BC  
bis - 213 31.12.0213 AD
- verbal - **22.01.103 n. Chr.**  
von - 22.01.0103 AD  
bis - 22.01.0103 AD
- verbal - **24.07.130 n. Chr.**  
von - 24.07.0130 AD  
bis - 24.07.0130 AD
- verbal - **2. Jh. v. Chr.**  
von - 199 01.01.0199 BC  
bis - 100 31.12.0100 BC
- verbal - **2. oder 3. Jh. n. Chr.**  
von - 100 01.01.0100 AD  
bis - 299 31.12.0299 AD
- verbal - **Anfang 4. Jh. n. Chr.**  
von - 300 01.01.0300 AD  
bis - 320 31.12.0320 AD
- verbal - **Mitte 4. Jh. n. Chr.**  
von - 335 01.01.0335 AD  
bis - 364 31.12.0364 AD
- verbal - **Ende 4. Jh. n. Chr.**  
von - 380 01.01.0380 AD  
bis - 399 31.12.0399 AD
- verbal - **378/379 n. Chr.**  
von - 378 01.01.0378 AD  
bis - 379 31.12.0379 AD
- verbal - **römisch**  
von - -30 01.01.0030 BC  
bis - 284 31.12.0284 AD
- verbal - **unbekannt**  
von - 30 01.01.4713 BC  
bis - 284 28.01.4000 AD

### 3.3.4 Epochen bezogen auf Ägypten

<b>Epoche</b>	<b>von</b>	<b>bis</b>
ptolemäisch	323 v. Chr.	30 v. Chr.
römisch	30 v. Chr.	284 n. Chr.
byzantinisch	284 n. Chr.	641 n. Chr.
arabisch	nach 641 n. Chr.	heute

Tabelle 16: Geschichtliche Epochen

## 3.4 Codierung von URL's

Bei der Angabe von URL's ist dringend darauf zu achten, dass Hexadezimalcodierungen durch die entsprechenden Zeichen ersetzt werden. (%5f mit \_ oder %3A mit : oder %2F mit /) z.B.

Papyrus und Ostraka Projekt

<http://papyri.uni-leipzig.de/content/disclaimer.xml?XSL.lastPage.SESSION=/content/disclaimer.xml>

## 3.5 Literaturangaben

Literaturangaben sollten in der folgenden Form angegeben werden.

Monographien: Name, Vorname abgekürzt, Titel, Ort, Jahr (Reihe), Seiten

Zeitschriften: Name, Vorname abgekürzt, Titel, Titel der Zeitschrift (nicht abgekürzt), Band, Jahr, Seiten

Sammelbände: Name, Vorname abgekürzt, Titel des Beitrages, in: Titel des Sammelwerkes, hrsg. von Vorname abgekürzt, Name, Ort, Jahr, (Reihe) Seiten

## 3.6 Eingabe von Lesungen

Das Projekt erlaubt es, die gelesenen Texte direkt mit in der Datenbank zu speichern. Voraussetzung dafür ist, dass sich die Schriftzeichen in einem Unicode-Zeichensatz darstellen lassen.<sup>17</sup> Dazu wurde in das System das Tool „Multi Language Keyboard“ integriert, mit dem Sie Zeilen mit Unicode erstellen können. Aktuell werden **arabisch, armenisch, hebräisch, altgriechisch, javanesisch, koptisch, persisch, diverse Sonderzeichen und Urdu** unterstützt. Eine Erweiterung auf andere im Unicode darstellbare Schriften ist jederzeit möglich.

Zur Notation der Lesungen sollen folgende Regeln gelten:

<sup>17</sup> <http://www.unicode.org/>

<b>Notation</b>	<b>Erklärung</b>
[ ]	Lücke
[...]	Vermutliche Anzahl der fehlenden Buchstaben in der Lücke
[[ ]]	Tilgung durch den Schreiber
`αβγ'	Von einem Schreiber über die Zeile nachgetragene Buchstaben
< >	Ergänzung oder Änderung durch den Herausgeber
abg	Beschädigte Buchstaben, deren Lesung unsicher ist
...	Nicht lesbare Buchstabenreste
[]	Vom Schreiber unbeschrieben gelassenes Spatium
( )	Aufgelöste Abkürzungen
	Beginn einer neuen Zeile im Papyrus
→	Faserverlauf horizontal
↓	Faserverlauf vertikal
r	recto
v	verso
?	Zahlenzeichen
?	Unsichere Lesung von Zahlenzeichen

*Tabelle 17: Zeichentabelle für Notation*

# 4 Dateneingabe direkt in XML

## 4.1 Erstellen der Daten

Für eine alternative Dateneingabe ist es möglich, auf der Basis der bereitgestellten Beispiele im Modul **papyri-demo** die direkten XML-Daten zu bearbeiten und Datensätze zu erstellen. Als Template können die im Verzeichnis *papyri-demo/modules/demo/content/[document|glossary|schrift|fragment|text]* stehenden XML-Dateien verwendet werden. Diese werden in die entsprechenden WorkBaskets unter *papyri/data/workflow* bzw. das im Property *MCR.SWF.Directory.base* angegebene Verzeichnis kopiert und gemäß den Namensvorgaben für die MCRObjektID umbenannt. Achten Sie darauf, dass Sie **keine ID verwenden, welche schon im System benutzt wird**. Die in den WorkBasket kopierten Daten sind dann auch aus dem interaktiven Workflow heraus sichtbar und können dort weiterverarbeitet werden.

Ebenso können Sie Dateien, welche auf interaktivem Wege im Workflow erzeugt wurden, über die Kommandozeile Ihres Systems bearbeiten. Die XML-Dateien können mit jedem beliebigen Text-Editor bearbeitet werden. Achten Sie aber darauf, dass das **Abspeichern der Daten im UTF-8 Encoding-Format** erfolgen muss!

### 4.1.1 Laden der Daten

Nach erfolgter Eingabe können die Daten nun auch per Kommandozeile in das System geladen werden. Dazu **muss** die Web-Applikation gestoppt sein. Die erforderlichen Kommandos stehen unter *papyri/build/bin*. Benutzen Sie die für Ihre Instanz bereitgestellten Scripts.<sup>18</sup> Folgende Arbeitsschritte sind durchzuführen:

- Die Daten sind nochmals mittels `WorkflowObjectCheck.sh [schrift|fragment|text|document|glossary]` zu prüfen. Die hier angegebenen Fehler sind zu beheben.
- Sind die Daten soweit fertig, können sie mit `WorkflowObjectLoad.sh [schrift|fragment|text|document|glossary]` geladen werden. Datensätze, welche bereits im Server vorhandene IDs haben, werden als Fehler markiert und nicht mit geladen. **Achtung ! Zuerst die Daten der Schriftträger laden!!!**
- Löschen Sie nach erfolgreichem Laden den Datensatz aus dem WorkBasket.

### 4.1.2 Update der Daten

Sind im WorkBasket noch Daten, welche bereits im System existieren und sollen diese durch die per Hand modifizierte Form ersetzt, so können Sie diese mit den Kommandos `WorkflowObjectUpdate.sh [schrift|fragment|text|document|glossary]` einbringen. Auch hier kann ein Check der Daten vor dem Update hilfreich sein. Wie beim Laden müssen Sie die Daten anschließend aus dem entsprechenden WorkBasket löschen.

### 4.1.3 Erweitern offener Klassifikationen

<sup>18</sup> Eine Gesamtübersicht der Kommandos finden Sie weiter hinten.

Im Projekt sind einige Klassifikationen als offen gekennzeichnet. Diese stehen unter *papyri/modules/papyri/content/classifications* und können mit einem XML-fähigen Editor entsprechend erweitert werden. Anschließend kann mit dem Kommando `ant create.class` ein Update erfolgen. Bei gleichzeitiger Nutzung des interaktiven Systems kann es auf Grund der Cache-Zeiten eine Weile dauern, bis die Klassifikation auch da geändert erscheint. Diesen Weg sollten Sie aber nur gehen, wenn es Probleme mit der interaktiven Variante des Klassifikationseditors gibt. **Beachten Sie bitte auf jeden Fall das Encoding der Daten!**

## 4.2 Verarbeitung der Bilddaten

### 4.2.1 Scannen und Sicherheitsverfilmung

Scannen und Sicherheitsverfilmung erfolgen mittlerweile an verschiedenen Standorten der Sammlungen. Die Vorgaben hierfür sind oben (Projektbeschreibung - Bilddaten) angegeben. Nach diesen Arbeiten ist eine erste Kontrolle der erzeugten DVD's und Filme sowie der mitgelieferten Inhaltsliste durchzuführen!

Dabei ist zu beachten:

- dass die Dateinamen mit der Inventarnummer des Objektes übereinstimmen,
- dass die Dateinamen mit denen auf der DVD übereinstimmen und
- dass die Scans der Papyri in ihrer Orientierung richtig herum sind.

### 4.2.2 Ausschneiden der Bilder für die Web-Präsentation

Nachdem die Originaldaten die Phase 1 (Scan) verlassen haben, werden sie nach folgenden Kriterien mittels eines geeigneten Programms (z. B. Photoshop) ausgeschnitten, auf 300 dpi reduziert und anschließend der Weiterverarbeitung zugeführt. Als Sicherungskopie können hierfür CD's genutzt werden. Dabei sollte der Inhalt mehrerer DVDs nach dem Verkleinern und Schneiden einer auf eine CD gebrannt werden, um Platz zu sparen.

Folgende Arbeiten sind durchzuführen:

- die Farbskala und der große Maßstab sind wegzulassen,
- der Rahmen und sonstige Schildchen und Beschriftungen sind wegzulassen oder zu entfernen (ggf. mit Mitteln der digitalen Retusche).
- Der Papyrus soll im Idealfall nur mit dem eigens hinzugelegten Maßstab zu sehen sein.
- Das geschnittene Bild ist im Abgleich mit dem Original mittels Photoshop aufzuhellen (Helligkeit/Kontrast) und die Rottönung zu reduzieren (Selektive Farbkorrektur). Gegebenenfalls sind weitere Parameter so zu verändern, dass die Farbe des Papyrus auf dem Bildschirm der des Originals möglichst nahe kommt.

### 4.2.3 Erzeugen der Web-Bilder

Nach Abschluss der Phase 2 (Ausschneiden) werden von den Bildern die 300 dpi und die 72 dpi JPEG-Versionen erstellt. Dies erfolgt im interaktiven Prozess 'on the fly' mit dem unter

Punkt 2.5.7 beschriebenen Upload. Die Dateinamen der Derivate-XML-Files entsprechen in den Nummern denen der referenzierten Schrifträger. Insgesamt werden pro Schrifträger 6 Derivate erzeugt:

Der referenzierte Schrifträger hat die ID `Project_schrift_0050.xml`

- |           |                 |                           |
|-----------|-----------------|---------------------------|
| 1. ID+0 → | 300 dpi recto → | Project_derivate_0050.xml |
| 2. ID+3 → | 72 dpi recto →  | Project_derivate_0053.xml |
| 3. ID+4 → | 300 dpi verso → | Project_derivate_0054.xml |
| 4. ID+7 → | 72 dpi verso →  | Project_derivate_0057.xml |

## 4.3 Erstellen des Inventarbuches

Um die eingespeicherten Daten prüfen und dokumentieren zu können, wurde die Erzeugung eines Inventarbuches in das Projekt integriert. Hier werden alle eingegebenen Daten aufgelistet und die entsprechenden Thumbnails angezeigt. Das Inventarbuch kann in drei Varianten erstellt werden:

1. **overview** – Es werden alle ausgewählten Schrifträger in Sortierung der Inventarnummern mit den zugehörigen Texten und Fragmenten ausgegeben. Weiterhin werden die eingestellten Bilder als Thumbnails dargestellt. Von den Metadaten werden nur Titel, Inventarnummer, Änderungsdatum und MCRObjektID angegeben.
2. **alldata** – Es werden alle Daten des Schrifträgers inklusive der zugehörigen Texte und Fragmente sowie der Thumbnails aufgelistet.

Da besonders Punkt 3 sehr umfangreiche Dokumente produziert, sind die Inventarbücher je Lokation in Bände von maximal 250 Inventarnummern aufgeteilt. Die Erstellung der PDF-Dateien erfolgt durch Aufruf von Scripts `BuildInventoryBook.sh` [`overview` | `listdata` | `alldata`] aus der Kommandozeile.

## 4.4 Erstellung der Google Sitemaps

Die Implementierung enthält auch ein Tool zur Erzeugung von Google-konformen Sitemaps<sup>19</sup>. Diese gestatten zusammen mit der `robots.txt` – Datei ein gezieltes Angebot der Daten an Suchmaschinen. Die Erstellung der Google Sitemaps mittels des Kommandos `BuildGoogleSitemap.sh` sollte regelmäßig, am besten per cron-Job erfolgen.

## 4.5 Sicherungsmaßnahmen

### 4.5.1 Dokumentation

Zur Überwachung der Scans und deren Verteilung sind folgende Schritte durchzuführen:

- Von allen Scans-enthaltenden DVDs wird eine Kopie auf einen Plattenbereich des File Servers des URZ der Universität Leipzig angefertigt.
- Von allen CDs mit den beschnittenen Bildern wird eine Kopie auf einen Plattenbereich des File Servers des URZ der Universität Leipzig angefertigt.
- Die CDs bzw. DVDs werden kopiert, siehe Phase 1 / 2.
- Bei Übergabe der Datenträger wird ein Übergabeprotokoll angefertigt.

<sup>19</sup> <https://www.google.com/accounts/ServiceLogin?service=sitemaps&hl=de&continue=https%3A%2F%2Fwww.google.com%2Fwebmasters%2Ftools%2Fsiteoverview&nui=1>

- Das Original wird in einer Brandschutzsektion der jeweiligen Einrichtung aufbewahrt.
- Die Kopie wird im brandschutzgesicherten Bereich des URZ der Universität Leipzig aufbewahrt.
- Die Bilddaten werden auf einem Zentralen Plattensystem des URZ inklusive zentralem Backup abgelegt.

#### 4.5.2 Der Server in Leipzig

In Leipzig erfolgen die nachstehenden Datensicherungsmaßnahmen für alle an diesem Standort gehaltenen Daten.

- Vom Server-System erfolgt täglich um 17 Uhr ein Backup auf das zentrale Sicherungssystem (alle 4 Wochen FULL, sonst INCR).
- Freitags werden um 8 Uhr alle Daten der Datenbasis in ein gesondertes Verzeichnis auf einen Plattenbereich ausgelesen (Script `build/bin/Save....sh`) und danach auf den externen File-Server kopiert.
- Dieser File-Server wird auch wöchentlich freitags um 22 Uhr gesichert (alle 4 Wochen FULL, sonst INCR.).
- Alle ausgelesenen Daten werden zusätzlich in größeren Abständen zu definierten Zeitpunkten auf DVD-R gesichert. Hier erfolgt eine Trennung nach der Einrichtung.
- Die Verteiler-DVDs der Sicherungen enthalten weiterhin den aktuellen Stand des Inventarbuches der jeweiligen Einrichtung. Somit ist auch eine visuelle Kontrolle der eingegebenen Daten möglich.



# 5 Installation / Nutzung eines Demo-Systems

## 5.1 Installation

Unter [https://git.sc.uni-leipzig.de/mycore\\_applications/papyri/papyri-demo/tree/master](https://git.sc.uni-leipzig.de/mycore_applications/papyri/papyri-demo/tree/master) befindet sich eine Demo-Anwendung, die genutzt werden kann. In der README.md sind die entsprechenden Installationsanweisungen. Folgende Informationen sind dennoch zu beachten:

Erster und wesentlicher Punkt für die Produktion, ist die Trennung von Daten und Programmen auf dem Server. Hierfür sollten die Verzeichnisse *papyri/data* und *papyri/save* durch zwei Verzeichnisse außerhalb von `PAPYRI_HOME` ersetzt werden. In der Konfiguration sind dann die beiden entsprechenden Properties anzupassen. Werden dazu noch die vier Konfigurationsdateien entsprechend gesichert (z. B. in *~/config*) so kann nach einem Update der Sourcen die Anwendung problemlos wieder hochgefahren werden.

### 5.1.1 Nutzung anderer Datenbanken

Neben Postgres können auch andere Datenbanken verwendet werden. So z. B. DB2 oder MySQL. Dank der Java Persistence API ist das möglich. Sie müssen lediglich die Konfiguration in der Dateien *config/persistence.xml* austauschen und den Install-Prozess neu durchlaufen (`ant resolve, create...`). Um auch Arabisch unterstützen zu können, MÜSSEN Sie die Datenbank mit UTF8-Encoding anlegen.

### 5.1.2 Zugang über HTTPS

Neben dem normalen ungeschützten Zugängen zur Anwendung über HTTP kann auch eine Zugang über HTTPS erfolgen. Dies ist für Produktivsysteme zu empfehlen.

## 5.2 Commandline Tools

Grundsätzlich kann über die Kommandozeile auf alle verfügbaren MyCoRe-Kommandos zugegriffen werden. Eine Übersicht bekommt man mit `build/cli/bin/mycore.sh help`. Weitere Hinweise sind auf der MyCoRe-Website zu bekommen.

### 5.2.1 Spezielle Kommandos der Anwendungen

Das System hat ein Verzeichnis, in dem die Shell-Scripts zur Arbeit vorgehalten werden. Im Verzeichnis *build/cli/bin* stehen alle Scripts zur Administration und Konfiguration des Systems. Für den Import der Daten werden die Verzeichnisse des Workflow genutzt, die über die Properties *MCR.SWF.Directory....* in der Konfigurationsdateien angegeben sind.

### 5.2.2 Zur Verarbeitung von Metadatenobjekten

Kommando	Parameter	Verzeichnis	Funktion
----------	-----------	-------------	----------

WorkflowObjectCheck.sh	leer <nummer>	schrift, fragment, text, document, glossary	Prüft die Syntax eines oder mehrerer Schriftträger-Datensätze
WorkflowObjectDelete.sh	<nummer> <nummer> <nummer>		Löscht einen oder einen Bereich von Schriftträger-Datensätzen
WorkflowObjectLoad.sh	leer <nummer>	schrift, fragment, text, document, glossary	Laden einer oder mehrerer Schriftträger-Datensätze
WorkflowObjectUpdate.sh	leer <nummer>	schrift, fragment, text, document, glossary	Überschreibt einen oder mehrere Schriftträger-Datensätze

Tabelle 18: Kommandos zur Arbeit mit Metadatenobjekten

### 5.2.3 Zur Verarbeitung von digitalen Objekten (Derivaten)

Kommando	Parameter	Verzeichnis	Funktion
WorkflowDerivateDelete.sh	<nummer> <nummer> <nummer>		Löscht einen oder einen Bereich von Derivaten
WorkflowDerivateLoad.sh	leer <nummer>	schrift, fragment, document	Laden einer oder mehrerer Derivate
WorkflowDerivateUpdate.sh	leer <nummer>	schrift, fragment, document	Überschreibt einen oder mehrere Derivate

Tabelle 19: Kommandos zur Arbeit mit digitalen Objekten

### 5.2.4 Zum Erzeugen der Inventarbücher

Kommando	Parameter	Verzeichnis	Funktion
BuildInventoryBook-ext.sh	[overview   listall   alldata]	inventory	Dieses Script erzeugt ein vollständiges Inventarbuch.

Tabelle 20: Kommandos zum Erzeugen eines Inventarbuches

### 5.2.5 Reparieren der Indizierungen

Kommando	Parameter	Verzeichnis	Funktion
RepairIndexFragment.sh RepairIndexGlossary.sh RepairIndexMODS.sh RepairIndexSchrift.sh RepairIndexText.sh		schrift, fagment, text	Diese Scripte reparieren den entsprechenden Index.

Tabelle 21: Kommandos zum Reparieren der Indizierung

## 5.2.6 Zum Erzeugen der userconfig-Datei

Kommando	Parameter	Verzeichnis	Funktion
BuildFopUserconfig.sh		build/webapps/WEB-INF/classes	Der Script erzeugt die userconfig Datei.

Tabelle 22: Kommando zum Erzeugen der userconfig-Datei

## 5.2.7 Überprüfen der Links

Kommando	Parameter	Verzeichnis	Funktion
CheckPapyriLinks.sh CheckDocumentLinks.sh	leer		Dieses Scripte überprüfen die Links der jeweiligen Datensätze.

Tabelle 23: Kommandos zur Überprüfung der Daten

## 5.2.8 Zum Sichern der Daten

Kommando	Parameter	Verzeichnis	Funktion
SavePapyri.sh	BasePrefix, Type	save	Sichert alle Datensätze, welche im System eingestellt sind.

Tabelle 24: Kommando zur Sicherung der Daten

## 5.2.9 Sonstige Kommandos

Das MyCoRe-Kern-System liefert standardmäßig noch weitere Kommandos. Diese sind Commandline-Shell, in der auch alle oben angegebenen Kommandos analog durchgeführt werden können. Weiterhin gibt es noch einige zusätzliche Kommandos, die für den allgemeinen Nutzer **NICHT** bestimmt sind und Administrationszwecken dienen.

Kommando	Parameter	Verzeichnis	Funktion
mycore.sh			Das MyCore-Standard-Commandline-Tool
jettystart	leer		Startet den Jetty Daemon
jettystop	leer		Stoppt den Jetty Daemon

Tabelle 25: Kommandos zur Administration

## 6 Glossar

### **Derivate**

Bezeichnet die Einheit von einem digitalen Objekt und beschreibenden Daten über Import/Export-Informationen. Dabei kann ein Derivate ein oder mehrere digitale Objekte beinhalten, welche ggf. in einem Verzeichnis zusammengefasst sind.

### **Jetty**

Unter Jetty versteht man einen auf Java basierenden Webserver bzw. Servlet-Container, der den JSP-Compiler Jasper des Apache Tomcat nutzt.

### **Maven**

Apache Maven ist ein Build-Managementsystem der Apache Software Foundation basierend auf Java, welches u.a. die Entwicklung von Software durch diverse Automatisierungen unterstützt.

### **Metadaten**

Unter Metadaten sind alle zum eigentlichen Objekt (in diesem Falle Papyri oder Dokumente) gehörenden beschreibenden Daten zu physikalischen und inhaltlichen Angaben zu verstehen.

### **Ostraka**

Ostraka sind Scherben eines zerbrochenen Gefäßes.

### **Recto**

Bezeichnet die Vorderseite eines Dokumentes.

### **Servlet-Engine**

Servlets sind Programmteile, welche mit einem Web-Browser kommunizieren können und dynamisch Web-Inhalte erzeugen. Die Servlet-Engine steuert den Zugriff auf die darin konfigurierten Servlets. Typische Vertreter sind die Projekte Tomcat und Jetty.

### **Solr**

Apache Solr ist ein Suchmaschinen-Servlet, welches in Apache Lucene enthalten ist. Das Erfassen und Erstellen von Dokumenten erfolgt über das Hypertext Transfer Protocol. Des Weiteren bietet das Servlet u.a. entsprechende Container für Apache Tomcat.

### **Thumbnail**

Bezeichnet kleine digitale Bilder, die der Bildvorschau dienen.

### **Tomcat**

Der Webserver und Webcontainer der Apache Software Foundation erlaubt es, dass in Java basierte Web-Anwendungen auf Servlet-Basis ausgeführt werden können.

### **Verso**

Bezeichnet die Rückseite eines Dokumentes.

### **Workflow**

Unter dem Workflow verstehen wir die Gesamtheit der Arbeitsabläufe, die zum Einstellen, Speichern, Korrigieren und Präsentieren der Daten notwendig sind.

# 7 Anhang

## 7.1 Übersetzungstabelle deutsch – englisch

<i>deutsch</i>	<i>english</i>	<i>remarks</i>
Papyrusobjekt	papyrus object	means the unity/ensemble of physically and textually interconnected parts of an ancient object. It includes the related fragments, text documents and bibliographical data regarding the papyrus object.
Schriftträger	written object	describes all physical features of the object that carries the writing. Data about the front and back (recto/verso) can differ (e. g. script direction: along or across the fibres)
Schriftträger-Fragment	written object fragment	If the written object consists of several scraps that have not yet been joined together or should be recorded as single items, you can record certain data about these fragments in an additional record. In this way you must record only the additional or peculiar information because the written object fragment inherits the other data from the written object.
Schriftträger-Fragment-Teil	part of a fragment of a written object	special record for large written object fragments (e. g. Ebers Papyrus). This inherits again from the fragment.
Text	text	text which stands on one or more written objects or parts of them. Also included is information about editors, publications and use of the item in academic works.
Dokument	document	means additional material about the papyrus-objects (e. g. restauration reports). The item is recordable and researchable.
ST01/TE01 Titel	title	descriptive title
ST02/TE02 Inventarnummer	inventory number	inventory number of the written object or fragment
ST03 Anderes Ordnungskriterium	other ordering systems	old or special order criteria/numbers
ST04 Sammlung	collection	The collection name (e.g. the city name of the institution)
ST05 Unterbringung	placement in collection	internal information about housing within the collection.
ST06 Erwerbung	acquisition	findspot, date, finder, place of purchase, date of purchase, dealer, lottery, remarks
ST07 Publikation	publication	link to one or more online-publications of the written object
ST08 Publikationsnummer	publication number	e. g. published P.Hal.1 or unpublished, see top 1.7.2
ST09 Material	writing material	writing material (taken from a classification list)
ST10 Farbe (grob)	colour (approximate)	colour classification of the object.
ST11 Farbe (verbal)	colour (verbal)	The detailed description of the colour in own words

<i>deutsch</i>	<i>english</i>	<i>remarks</i>
		based on the 'Michel Farbtafel'.
ST12 Maße	dimensions/size	Height, width, thickness (only for ostraca)
ST13 Zustand	state of preservation	state of preservation according to a classification (carbonised, dirty, mould, adhesion etc.)
ST14 Komplet	Complete	This is a simple classification with values no, yes, almost, undecidable.
ST15 Aufbewahrung	Custodial information	This describes where the object is housed (in glass, paper, etc.)
ST16 Erhaltung	Conservation	Information about conservation
ST17 Ausführung	Format	This can be roll, sheet, codex etc.
ST18 rekonstruierter Umfang	reconstructed size	If it is possible you can enter the original extent of the written object with approximate measures. In the case of literary texts sometimes you can find out how many verses are missing between the fragments, but without defining the original length of the roll.
ST19 Klebung	colleseis	Colleseis are measured in cm from the left on the recto and parallel with the script direction. Exceptions are transversa charta.
ST20 phys. Zugehörigkeit	physical joins	joins to items that are outside collections in the project
ST21 Fremmlink	external link	URL link to physically related items (ST20) outside the project
ST22 phys. Zugehörigkeit (interne Vermerke)	physical joins (internal remarks)	internal remarks on physically related items in collections that are not included in the project
ST23 phys. Zugehörigkeit (interne Links)	physical joins (internal links)	internal URL link to physically related items (ST22) outside of the project
ST24 Publikationsart	kind of publication	according to a classification list (published, unpublished free, unpublished blocked, unpublished reserved)
ST30 Abbildung, publiziert	images, published	according to the rules for biographical references and abbreviations
ST31 Abbildung in Sammlungen	images in collections	This field includes some subdivisions like type (film negative, photos, slides, scan, CD, DVD etc.), the order number (CD number ...), the suborder number (record number ...), the date, the keeping/housing and remarks.
ST32 Bemerkung	remarks	some general remarks about this written object.
ST33 Bemerkung intern	remarks, internal	some internal remarks about this written object.
ST34 Weitere Links	additional links	some additional URL entries.
ST35 Bearbeiter	author of record	person who made the record.
ST36 Lizenzen	licenses	
ST37 alternative ID	alternate ID	alternate Ids like URN, PND
ST38 interne ID	internal ID	internal old Ids for data import
TE03 Erhaltungszustand des Textes	preservation of text	This information refers both to the written object and the text, but with a focus on the preservation of text.
TE04 Erhaltungszustand des Textes (verbal)	preservation of text in your own words	free text, which appears in P.Lips. for instance in small letters as 'mise en page' (layout) (e. g. foldings, margins)

<i>deutsch</i>	<i>english</i>	<i>remarks</i>
TE05 Textart	genre	according to a classification list (literary, documentary, others, undecidable)
TE06 Textthema	type of text	type of text or subject matter according to a classification list, which follows the categories of the 'Sammelbuch' and allows the possibility of a hierarchical structure of the data.
TE07 Inhaltsangabe	content	a short summary of the content. Other information for unpublished papyri (e.g. persons, places) can be entered in field TE40 (remarks internal).
TE08 Abfassungsort	place of writing, place of destination	Provenance (ancient) is divided in two subfields: places and district (nomos) names. Because this field is obligatory, you must record 'unknown' where appropriate.
TE09 Antikes Archiv	ancient archives	URL link to Trismegistos Database, URL (cf. rules) or references to web pages or to 'documents' of our papyrus project
TE10 Datierung numerisch	date numeric	For the input and presentation of dates see section 1.4.6.
TE11 Datierung ausführlich	date detailed	explanation for the date, e.g. citation of a papyrus in which the same person is mentioned
TE12 Datierungsvermutung	suggested dating	explanation for the suggested date
TE13 Inhaltliche Verknüpfung	textual connections	Parallels-- number of edition (check list) is sufficient
TE14 Kolumnenzahl	number of columns	only preserved columns. When there is only one column, the field can be left blank.
TE15 Zeilenzahl	number of lines	only preserved lines. If you are not certain, you can explain this in field 'Remarks'. When you have two or more columns you can make a list.
TE16 Tinte	ink	according to a classification list
TE17 Buchstabenhöhe	letter height	average height; here it is worth mentioning unusual sizes; given in cm
TE18 Zeilenlänge	line length	only (completely) preserved lines, given in cm.
TE19 Zeilenabstand	line distance/space between the lines	average distance between the lines. Here it is worth mentioning unusual sizes; given in cm.
TE20 Kolumnenhöhe	column height	only (completely) preserved columns; given in cm.
TE21 Schriftart	palaeographical description	short description or classification of the script: 'Biblical Majuscule', 'Severe Style', 'book hand', 'chancery or office style', 'accents', 'punctuation', cursive, inclination, scribal idiosyncracies (cf. Schubert, Paläographie; Turner/Parsons, Greek manuscripts of the Ancient World). This information could be useful for future editions. For missing papyri the description contained in the edition is the only source of information.
TE22 Schriftrichtung	script direction	according to a classification list. For codex give the script direction of the first preserved page.
TE23 Layout Besonderheiten	layout features	e.g. paragraphos, blank spaces, eisthesis and ekthesis, coronis.
TE24 Illustrationen	illustrations	use keywords
TE25 Seitenfolge	page flow	Recto/Verso; relation between front and back; verso



<i>deutsch</i>	<i>english</i>	<i>remarks</i>
		continuation or address of letters.
TE26 Editio princeps	editio princeps	Here you must enter only the first complete edition. Descriptions and other editions will be enumerated in TE27. Exception: the papyrus is only incompletely published or described by now.
TE27 Weitere Editionen	other editions	verbal information about other editions.
TE28 Textsammlungen	text collections	This field is only for URLs (e.g. Heidelberger Gesamtverzeichnis, Leuven Database of Ancient Books or Duke Databank of Documentary Papyri). The syntax for Mertens-Pack is always: 'Mertens-Pack [Nr.]::[URL]', even if the link refers to the home page. See chapter 'Encoding URLs'. Comments should be written in field 'Remarks external'.
TE29 Text publiziert	text published	Here you can write/insert the published text in original script.
TE30 Text intern	text internal	Here you can write/insert the unpublished text in original script.
TE31 Übersetzung publiziert	translation published	Here you can write/insert the published translation.
TE32 Übersetzung intern	translation internal	Here you can write/insert the unpublished translation.
TE33 Berichtigung publiziert	correction published	Here you can write/insert the published correction.
TE34 Berichtigung intern	correction internal	Here you can write/insert the unpublished correction.
TE35 Literatur allg.	bibliographical references	cf. Rules
TE36 Literatur intern	bibliographical references internal	cf. Rules
TE37 Editoren	editoWählen Sie den aktuellen Bearbeitungsstatus des Datensatzes aus.rs	all editors (first edition, new edition etc.) but not authors of corrections.
TE38 Publikationserlaubnis	permission to publish	Internal information regarding publication permission
TE39 Bemerkungen	remarks	remarks as text
TE40 Bemerkungen intern	remarks internal	internal remarks as text
TE41 Publikation	publication	see ST07
TE42 Publikationsnummer	publication number	see ST08
TE44 Bearbeiter	author of record	person who made the record
TE46 Schrift	script/writing	according to a classification list
TE47 Sprache	language	according to a classification list

Tabelle 26: Übersetzung deutsch – englisch

## 7.2 Übersetzungstabelle deutsch – französisch

<i>deutsch</i>	<i>französisch</i>	
Papyrusobjekt		
Schriftträger		
Schriftträger-Fragment		

<i>deutsch</i>	<i>französisch</i>	
Schriftträger-Fragment-Teil		
Text		
Dokument		
ST01/TE01 Titel		
ST02/TE02 Inventarnummer		
ST03 Anderes Ordnungskriterium		
ST04 Sammlung		
ST05 Unterbringung		
ST06 Erwerbung		
ST07 Publikation		
ST08 Publikationsnummer		
ST09 Material		
ST10 Farbe (grob)		
ST11 Farbe (verbal)		
ST12 Maße		
ST13 Zustand		
ST14 Komplett		
ST15 Aufbewahrung		
ST16 Erhaltung		
ST17 Ausführung		
ST18 rekonstruierter Umfang		
ST19 Klebung		
ST20 phys. Zugehörigkeit		
ST21 Fremmlink		
ST22 phys. Zugehörigkeit (interne Vermerke)		
ST23 phys. Zugehörigkeit (interne Links)		
ST 24 Publikation		
ST30 Abbildung, publiziert		
ST31 Abbildung in Sammlungen		
ST32 Bemerkung		
ST33 Bemerkung intern		
ST34 Weitere Links		
ST35 Bearbeiter		
ST36 Lizenz		
ST37 alternative ID		
ST38 interne ID		

<i>deutsch</i>	<i>französisch</i>	
TE03 Erhaltungszustand des Textes		
TE04 Erhaltungszustand des Textes (verbal)		
TE05 Textart		
TE06 Textthema		
TE07 Inhaltsangabe		
TE08 Abfassungsort		
TE09 Antikes Archiv		
TE10 Datierung numerisch		
TE11 Datierung ausführlich		
TE12 Datierungsvermutung		
TE13 Inhaltliche Verknüpfung		
TE14 Kolumnenzahl		
TE15 Zeilenzahl		
TE16 Tinte		
TE17 Buchstabenhöhe		
TE18 Zeilenlänge		
TE19 Zeilenabstand		
TE20 Kolumnenhöhe		
TE21 Schriftart		
TE22 Schriftrichtung		
TE23 Layout Besonderheiten		
TE24 Illustrationen		
TE25 Seitenfolge		
TE26 Editio princeps		
TE27 Weitere Editionen		
TE28 Textsammlungen		
TE29 Text publiziert		
TE30 Text intern		
TE31 Übersetzung publiziert		
TE32 Übersetzung intern		
TE33 Berichtigung publiziert		
TE34 Berichtigung intern		
TE35 Literatur allg.		
TE36 Literatur intern		
TE37 Editoren		
TE38 Publikationserlaubnis		
TE39 Bemerkungen		

<i>deutsch</i>	<i>französisch</i>	
TE40 Bemerkungen intern		
TE41 Publikation		
TE42 Publikationsnummer		
TE44 Bearbeiter		
TE46 Schrift		
TE 47 Sprache		

*Tabelle 27: Übersetzung deutsch – französisch*

## 7.3 Projektbeteiligte

An der Konzeption und Umsetzung des Gesamtprojektes waren die nachfolgenden Personen maßgeblich beteiligt. Die Liste berücksichtigt jedoch nicht die Mitwirkenden an den einzelnen Instanzen, welche das Projekt nachnutzen.

### 7.3.1 Die wissenschaftliche Leitung

- **Prof. Dr. Reinhold Scholl**,  
Universitätsbibliothek Leipzig, Universität Leipzig
- **Prof. Dr. Jürgen Hammerstaedt**,  
Institut für Altertumskunde - Klassische Philologie, Universität zu Köln
- **Prof. Dr. Burkhard Meißner**,  
Professur für Alte Geschichte  
Helmut-Schmidt-Universität - Universität der Bundeswehr

### 7.3.2 Für die Restaurierung

- **Jörg Graf**,  
Papyrusrestaurator,  
Universitätsbibliothek Leipzig, Universität Leipzig

### 7.3.3 Für die technische Betreuung

- **Jens Kupferschmidt**,  
IT Verantwortlicher des Gesamtprojektes,  
Universitätsrechenzentrum, Universität Leipzig
- **Britta Kapitzki**,  
Designerin. Design bis 2011  
Thüringer Universitäts- und Landesbibliothek Jena
- **Philip Ackermann**  
Designer. Design ab 2011
- **Andy Wermke**  
Studentische Hilfskraft (Februar 2010 - Februar 2012)  
Universitätsbibliothek, Universität Leipzig
- **Erik Körner**  
Studentische Hilfskraft (seit September 2012)  
Universitätsbibliothek, Universität Leipzig
- **Tina Zimmermann**  
Studentische Hilfskraft (seit Oktober 2016)  
Universitätsrechenzentrum, Universität Leipzig

## Tabellenverzeichnis

Tabelle 1: An der Entwicklung beteiligte Einrichtungen.....	7
Tabelle 2: Nachnutzer des Projektes.....	7
Tabelle 3: Grafikformate I.....	11
Tabelle 4: Terminologieübersicht.....	14
Tabelle 5: Datenstruktur Schrifträger.....	17
Tabelle 6: Ausfüllhinweise Schrifträger.....	18
Tabelle 7: Datenstruktur Text.....	20
Tabelle 8: Ausfüllhinweise Text.....	22
Tabelle 9: Vererbungsmatrix.....	23
Tabelle 10: Schrifträger Klassifikationen.....	24
Tabelle 11: Text Klassifikationen.....	24
Tabelle 12: sonstige Klassifikationen.....	25
Tabelle 13: Mapping Schrifträger nach Dublin Core.....	25
Tabelle 14: Mapping Schrifträger nach MODS.....	27
Tabelle 15: Geschichtliche Epochen.....	32
Tabelle 16: Zeichentabelle für Notation.....	33
Tabelle 17: Kommandos zur Arbeit mit Metadatenobjekten.....	39
Tabelle 18: Kommandos zur Arbeit mit digitalen Objekten.....	39
Tabelle 19: Kommandos zum Erzeugen eines Inventarbuches.....	39
Tabelle 20: Kommandos zum Reparieren der Indizierung.....	39
Tabelle 21: Kommando zum Erzeugen der userconfig-Datei.....	40
Tabelle 22: Kommandos zur Überprüfung der Daten.....	40
Tabelle 23: Kommando zur Sicherung der Daten.....	40
Tabelle 24: Kommandos zur Administration.....	40
Tabelle 25: Übersetzung deutsch – englisch.....	45
Tabelle 26: Übersetzung deutsch – französisch.....	48

## **Abbildungsverzeichnis**

Abbildung 1: Das Client-Server-System.....	8
Abbildung 2: Datenmodellschema.....	15