The International Dunhuang Project.
Die Berliner Turfansammlungen in der IDP database (IDP Berlin)

Dr. Simone-Christiane Raschmann
AdW Göttingen, KOHD - Arbeitsstelle Turfanforschung, Berlin

Andrea Schlosser
BBAW, Turfanforschung – DFG-Projekt Digitalisierung

Einführung

Seit mehr als 100 Jahren werden in Berlin die Funde von vier deutschen archäologischen Expeditionen, die am Anfang des 20. Jh. (1904-1912) nach Zentralasien aufbrachen und wie in einem Wettlauf mit anderen europäischen und japanischen Expeditionen die Ruinenstätten in Ostturkistan - vor allem in der Oase von Turfan – untersuchten, aufbewahrt und wissenschaftlich bearbeitet. Die Berliner Turfansammlungen des Museums für Asiatische Kunst, Kunstsammlung Süd-, Südost- und Zentralasien und der Berlin-Brandenburgischen Akademie der Wissenschaften beherbergen eine umfangreiche Anzahl von Kunstwerken vorislamischer Kulturen Zentralasiens, darunter Wandmalereien, Skulpturen aus Lehm, Holz und Metall, Miniaturen auf Papier und Stoff sowie Tempelfahnen und ca. 40.000 Textfragmente.

Kataloge, Digitalisierung und Datenbanken erleichtern heute den Zugang zu den unterschiedlichen Materialien in den zentralasiatischen Sammlungen weltweit oder führen sogar wieder zusammen, was im Ergebnis der zahlreichen archäologischen Untersuchungen an die verschiedensten Sammlungsorte verbracht wurde. Dieses Anliegen brachte im Oktober 1993 Vertreter der verschiedenen Sammlungen, darunter auch der Berliner Turfansammlung, zu einem ersten Treffen zusammen und das im Resultat dieses Treffens 1994 gegründete *International Dunhuang Project* (IDP) wurde zu einer gemeinsamen Plattform für die Durchsetzung dieser Pläne und Aufgaben. Die IDP-database, die eine gleichzeitige Recherche in mehreren zentralasiatischen Sammlungen gestattet, ist ein hervorragendes Beispiel für den Erfolg dieser Bemühungen.

„Anstatt eines Türkenlandes – dies ist ja der Sinn des Wortes Turkistan – entdeckten wir, dass überall an der Seidenstraße bis zur Mitte des 8. Jahrhunderts Völker indogermanischer Zunge, Iranier, Inder und sogar Europäer gesessen haben. Ihre Sprachen, die zum Teil dem Namen nach, zum Teil überhaupt nicht bekannt waren, wurden in zahlreichen Manuskripten gefunden. Sie wurden alle in Berlin entziffert, übersetzt und wissenschaftlich behandelt. Die Anzahl der Manuskripte ist außerordentlich groß. Es finden sich nicht weniger als 17 verschiedene Sprachen in 24 verschiedenen Schriftarten.“ [1]
Dieses Zitat von ALBERT VON LE COQ mag zu einer kurzen Beschreibung der Vielfalt der Textfragmente in der Berliner Turfansammlung dienen. Schwankt auch die Zahl der verschiedenen Sprachen und Schriften in den einzelnen Darstellungen, so übertrifft doch die Berliner Sammlung durch ihre Mannigfaltigkeit die Textfunde der englischen, französischen, russischen und japanischen Expeditionen bei weitem. Die Fülle der Schriften und Sprachen, in denen die Texte abgefasst wurden, machen zugleich den hohen Wert dieser Sammlung für die Wissenschaft aus. Die chinesischen, alttürkischen und die Sanskrit-Fragmente gehören zu den zahlenmäßig am stärksten vertretenen Sprachgruppen.

Nach der Gründung der Orientalischen Kommission im Mai 1912 wurden die Textfunde nach und nach zur Bearbeitung an die Berliner Akademie der Wissenschaften übergeben.

Der 2. Weltkrieg und die in der Folge eingetretenen politischen Entwicklungen in Deutschland bedeuteten auch einen tiefen Einschnitt in der Erschließung und Bearbeitung der Berliner Turfan-Sammlung. Im Herbst 1943 begann die Akademie, ihre Handschriften und Archive auszulagern. Die Turfantexte der Orientalischen Kommission kamen an verschiedene Orte, z.B. in die Bergwerke Winthershall, Solvayhall und Schönebeck/Elbe. Ein großer Teil der Materialien wurde auf diese

[1] ALBERT VON LE COQ: *Auf Hellas Spuren in Ostturkistan*. Berichte und Abenteuer der II. und III. Deutschen Turfan-Expedition. Leipzig 1926, 10.

Weise gerettet. Nach dem Ende des Krieges wurde der Teil der Sammlung, der sich an Orten befand, die der Sowjetischen Besatzungsbehörde unterstanden, nach Berlin zurückgebracht und im August 1946 in das alte Akademiegebäude Unter den Linden zurückgeführt. Ein kleinerer Teil der Sammlung, der sich in Westdeutschland befand, kam erst im Laufe des Jahres 1947 nach Mainz und wurde dem ehemaligen Direktor der Preußischen Akademie der Wissenschaften H. SCHEEL zur Verwahrung übergeben.

Die sog. "Mainzer Sammlung" wurde später an die Staatsbibliothek der Stiftung Preußischer Kulturbesitz übergeben und kam über die Zwischenstation Marburg in den 70er Jahren nach Westberlin.

Die Arbeit der "Orientalischen Kommission" wurde ab 1947 im Rahmen des Instituts für Orientforschung an der Deutschen Akademie der Wissenschaften in Berlin weitergeführt. 1965 wurde zur Intensivierung der Arbeit an den Turfantexten auf Initiative von Akademiemitglied WOLFGANG STEINITZ und des ungarischen Turkologen GYÖRGY HAZAI eine Turfan-Forschungs-gruppe gegründet.

Erst nach der Wiedervereinigung der deutschen Staaten 1990 wurden auch die Handschriften und Blockdrucke der Turfan-Sammlung wieder zusammengeführt.[2] Als Eigentum der Berlin-Brandenburgischen Akademie der Wissenschaften (BBAW) befinden sie sich nun als Depositum in der Orientabteilung der Staatsbibliothek zu Berlin - Preußischer Kulturbesitz, der die konservatorische und bibliothekarische Verantwortung obliegt. Die etwa 8500 türkischen und 4000 iranischen Textfragmente befinden sich zur weiteren Bearbeitung im Archiv der Arbeitsstelle des Akademienvorhabens Turfanforschung in der BBAW, wo die

wissenschaftliche Bearbeitung der bisher unpublizierten türkischen und iranischen Turfantexte und ihre Edition erfolgt.

Seit 1993 befindet sich in der BBAW auch eine Arbeitsstelle Turfanforschung der "Katalogisierung der Orientalischen Handschriften in Deutschland", das seit 1990 ein Vorhaben der Akademie der Wissenschaften zu Göttingen ist. Im Rahmen dieses Vorhabens werden neben Berlin auch in weiteren Arbeitsstellen in Deutschland (Göttingen, Kassel) sowie in ehrenamtlicher Mitarbeit Katalogbände mit Beschreibungen der Fragmente der Berliner Turfansammlung erarbeitet und in der Reihe "Verzeichnis der Orientalischen Handschriften in Deutschland" (VOHD) publiziert. Die Leitung des Akademienvorhabens ist in der Orientabteilung der Staatsbibliothek zu Berlin angesiedelt.

Ziele der Digitalisierung

- Archivierung: Bewahrung der auf den Fragmenten enthaltenen unersetzlichen Informationen
- Schutz: Reduzierung der Arbeit mit den Originalen, um eine Qualitätsminderung bzw. Beschädigung derselben zu vermeiden; eine Zusammensetzung von mehreren Fragmenten erfolgt nur noch digital
- Zugang: Erschließung der Fragmente für einen Gelehrtenkreis weltweit durch Abrufbarkeit im Internet
- Kontext: Präsentation der Digitalisate mit Metadaten im Kontext mit gleichwertigem Material anderer Zentralasiensammlungen weltweit

Als erste Textgruppe der Berliner Sammlung wurden im Rahmen des TITUS-Projektes der Universität Frankfurt a. M. und mit Unterstützung der Tamai-Foundation seit 1995 die Tocharica digitalisiert.[3] In einem Datenbankformat sind diese Digitalisate, verbunden mit einer Textbearbeitung und formalen Angaben zu jedem Fragment, im Internet präsent.

---

[2] Vgl. dazu: SIMONE-CHRISTIANE RASCHMANN: *Berlin's Re-united Collections*. In: IDP News: Newsletter of the International Dunhuang Project 3 (1995), 1f.

[3] Für weitere Informationen siehe http://titus.fkidg1.uni-frankfurt.de/texte/tocharic/tht.htm.

Digitalisierungsprojekte (DFG)

I. (3 Teilprojekte)

1997–2005 Digitalisierung der alttürkischen, mitteliranischen und mongolischen Fragmente. Online im Digitalen Turfan-Archiv (DTA I) abrufbar (insgesamt ca. 14.500 Fragmente).

II. (2 Teilprojekte)

2005–2008 Digitalisierung der chinesischen (ca. 4200) und tibetischen (ca. 150) Turfanfragmente, Einbindung in IDP database.

2008–2012 Digitalisierung der syrischen (ca. 400) und Sanskrit-Fragmente (ca. 14.000), Einbindung in IDP database.

DTA I **http://www.bbaw.de/forschung/turfanforschung/dta/index.html**

DTA II / IDP **http://idp.bbaw.de/**

Weitere Digitalisierungsprojekte

2008–2009 Im Rahmen des von der Europäischen Union (GD Bildung und Kultur, Programm „Kultur") geförderten Projekts IDP – CREA (International Dunhuang Project – Cultural Routes of Eurasia): Überführung ausgewählter Bilddaten von DTA I in IDP database inkl. Verknüpfung mit neu hinzugefügten Metadaten (manichäische und christlich-sogdische Fragmente in nestorianischer Schrift).

Im Rahmen von IDP-CREA wurde das Museum für Asiatische Kunst (Berlin-Dahlem) neue Partnerinstitution des IDP und digitalisierte in diesem Zeitraum 1.500 historische Fotografien und Skizzenzeichnungen aus seinem Turfanarchiv, die nunmehr in der IDP database zur Verfügung stehen.

2011 Im Rahmen der Akademienvorhaben KOHD, AdW Göttingen und Turfanforschung, BBAW: Digitalisierung der vorhandenen Diafilme der tocharischen Turfanfragmente im TIFF-Format (ca. 6.800, vgl. TITUS), Einbindung der Bilddaten mit Basis-Metadaten in IDP database

IDP-CREA **http://idp.bl.uk/idp_crea/index.htm**

TITUS **http://titus.fkidg1.uni-frankfurt.de/texte/tocharic/thtframe.htm**

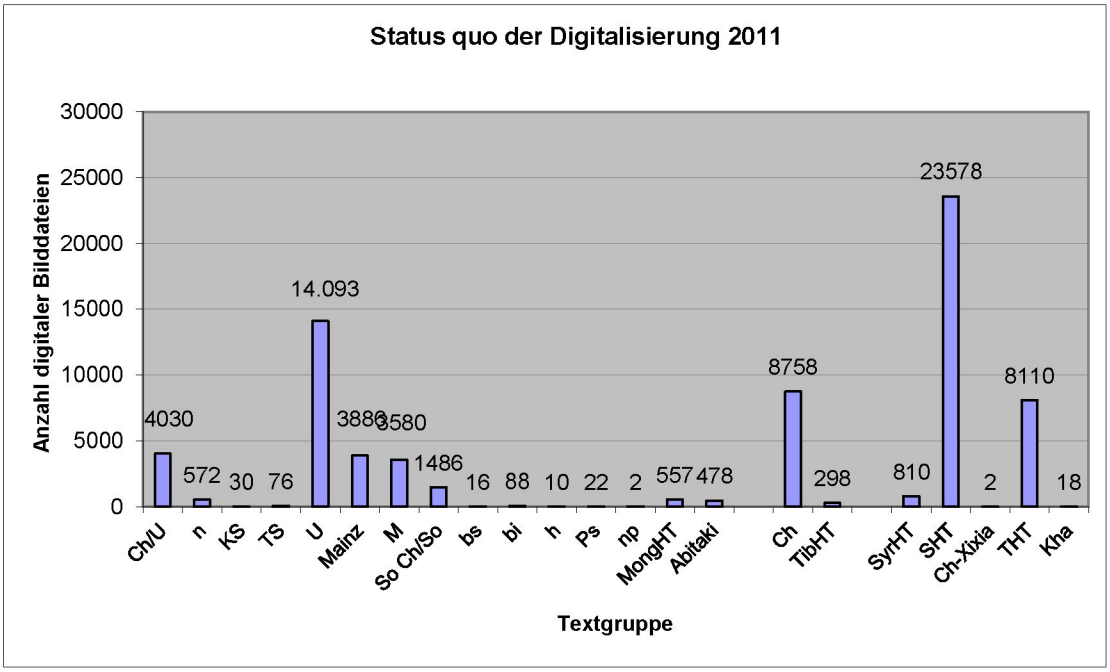


Abb. 1 Übersicht digitalisierter Daten 1997 bis 2011.

Workflow

DTA I (1997 bis 2005)

Alttürkische, mitteliranische und mongolische Fragmente

In der ersten Etappe des Digitalisierungsprojekts wurde das Material „doppelt“ archiviert, sowohl analog als Diafilm als auch digital als Bilddatei. Diese Entscheidung erfolgte auf Basis der Erfahrung des TITUS-Projekts in Frankfurt a.M., das 1996–1997 die tocharischen Fragmente digitalisierte. Ausschlaggebend waren zum damaligen Zeitpunkt die bessere Qualität der Bilddaten sowie die langfristig sicherere Archivierung.

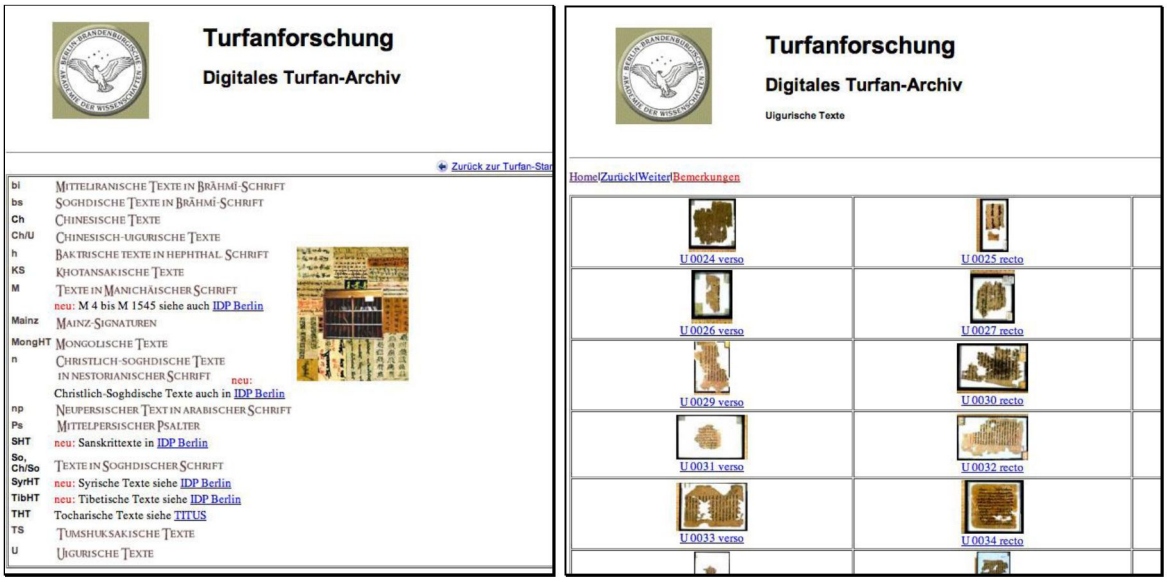


Abb. 2 und 3: Screenshot vom Digitalen Turfan-Archiv (DTA I) mit Übersicht über die verschiedenen Signaturen-/Textgruppen. Durch einen Klick auf eine Gruppe gelangt man zu einer tabellarischen Darstellung von Thumbnails, die wiederum zu einer größeren Abbildung führen.

Der Digitalisierungsprozess in dieser ersten Phase verlief wie folgt:

- Herstellung von Diapositiven der verglasten Fragmente (Fotostelle der Staatsbibliothek zu Berlin).
- Scannen der Dias und Speicherung der TIFF-Datei unter einem Dateinamen, der im Kern mit der Standortsignatur identisch ist: Textgruppe, Fragmentnummer, eventuell vorhandene alphabetische Benennung einzelner Fragmente in Sammelplatten, Bestimmung von Vorder- und Rückseite. Zum Beispiel: »U0255a-b_recto« (U steht für die uigurische/alttürkische Textgruppe).
- Brennen der Scans auf CD-ROM (doppelte Kopie und Lagerung an verschiedenen Orten), ab 2004 zusätzliche Speicherung auf externen Festplatten, da eine reine CD-Speicherung zu unsicher war.
- Protokollierung der Verfilmung und Digitalisierung durch eine tabellarische Datenbank nach Signatur, darin Eintrag der Film- und CD-Nr. sowie weitere Informationen über die digitale Archivierung.
- Bildbearbeitung: Drehen der Scans in die richtige Leseposition, Zuschneiden der Detailaufnahmen, Verbesserung der Lesbarkeit z.B. durch Änderung von Helligkeit und Kontrast.
- Dateiformatierung: Komprimierung der TIFF-Dateien zu hoch- wie niedrigauflösenden JPG-Dateien für die Darstellung im Internet.
- Internetpräsentation in Thumbnail-Tabellen nach Signaturengruppen geordnet (DTA I).

DTA II / IDP (2005 bis 2008)

Chinesische und tibetische Fragmente

Auch im 1. Teilprojekt der zweiten Etappe erfolgte die Digitalisierung indirekt (zweistufig) mittels der Erstellung von Diapositiven und dem anschließenden Scannen der Dias.

Die Bilddaten wurden seitdem aber nicht mehr im DTA I zur Verfügung gestellt, sondern in die IPD-Datenbank eingebunden. Dabei handelt sich um eine internationale Plattform zur Präsentation von Manuskripten, Gemälden, Textilien und Artefakten aus dem Gebiet der Seidenstraßen.4 Mit Hauptsitz in London (British Library) wurden im Laufe der Jahre mehrere Partnerinstitutionen in Paris, Berlin, St. Petersburg, Beijing, Kyoto, Seoul und Dunhuang eingerichtet, die Webseiten in ihren jeweiligen Landessprachen betreuen. Alle Inhalte wie Bilder, Metadaten und Kataloge werden regelmäßig synchronisiert, sodass sie von allen nationalen Benutzeroberflächen aus zugänglich sind.

Durch die bereits etablierte Grundstruktur von IDP bot sich die Möglichkeiten, zusätzlich zu den Bildern mit ihren Signaturen weitere Metadaten wie z.B. Fundort, Fundortsigel, Form, Material oder Größe aufzunehmen (zu einer Übersicht über alle verfügbaren Metadaten siehe Tab. 1 und Abb. 4).

[4] »Die IDP-Datenbank ist ein geordneter, relationaler und indizierter Speicher von Informationen zu über 50.000 Manuskripten, Malereien, Artefakten und Fotografien von archäologischen Fundorten, datiert von 100 v. Chr. bis 1200 n. Chr. in Ost-Zentralasien/an der Seidenstraße (hauptsächlich in den heutigen chinesischen Provinzen Xinjiang und Gansu).« (http://idp.bbaw.de/pages/help.a4d#2)

Tab. 1: Basisdaten/Metadaten

Basisdaten	Press-mark	Frühere Nr.	Fundort (Site ID)	Sigel	Exp.	Institution	Höhe/ Breite	Material	Form	Art	Sprache/ Schrift	Thema	Kurztitel
Daten-bankein-trag	+	+	+	+	+	+	+	+	+	+	+	+	+
Sichtbar in IDP	+		+	+		+	+	+	+		(+)		
Aktuell suchbar in IDP	+		+			+			+	+	+	+	+

Weiterhin wurden die Informationen zu einem Fragment durch Online-Kataloge im XML-Format ergänzt, die zusätzliche Informationen wie Titel, Textstellenangabe und Bibliographie zur Verfügung stellen. Die XML-Dateien werden nach Vorgaben von IDP London, basierend auf dem TEI--Standard, erstellt (die Vorgaben können z.B. unter http://idp.bbaw.de/pages/technical_resources.a4d#1 heruntergeladen werden).



Abb. 4: Screenshot einer Datenbankabfrage (Suchwert »Ch 2341«) mit Metadaten (»Generelle Informationen«) und Katalogeintrag (Bsp. Thilo 1985) mit Angabe des Texttitels und seiner Stelle im chinesischen Kanon (Taishō-Ausgabe).

```
<msDescription type="KK" n="143">
  <msIdentifier>
    <settlement>Berlin</settlement>
    <repository>Staatsbibliothek, Orientabteilung</repository>
    <collection>Turfan collection</collection>
    <idno type="current">Ch 2341</idno>
  </msIdentifier>
  <msContents>
    <msItem n="1">
      <title lang="chi-Latn">Da bo re bo luo mi duo jing</title>
      <title lang="chi-Hant">大般若 波羅蜜多經</title>
      <note>
        <p>Taishō 220 = 5, 534c, 11 - 16</p>
      </note>
      <listBibl>
        <bibl>
          <ref target="Thilo_1985"/>
        </bibl>
      </listBibl>
    </msItem>
  </msContents>
</msDescription>
```

Abb. 5: Screenshot der zu Abb. 4 gehörigen XML-Datei als Beispiel für die Struktur eines Online-Katalog-Eintrags.

Die Arbeitsschritte im Digitalisierungsprozess sind seither wie folgt:

Digitalisierung:

- Erstellen der Diapositive (Fotostelle der Staatsbibliothek zu Berlin)
- Scannen der Dias

Bildbearbeitung:

- Bildoptimierung (Korrektur von Kontrast, Tonwert, Farbe), bei Sammelverglasungen digitale Beschriftung der einzelnen Fragmente
- Entfernung Farbskala, evtl. Drehung, Positionierung Lineal
- Stapelverarbeitung zur Erstellung von JPGs (large, middle, thumb)
- Hochladen auf den Server und Einbindung in die IDP-Datenbank

Metadaten:

- Sammeln der Daten in einer Excel-Tabelle
- Importieren der Daten in die IDP-Datenbank

Kataloge:

- Importieren von Teildaten aus der Excel-Tabelle in eine XML-Datei
- Hinzufügen von weiteren Informationen
- Importieren der XML-Datei in die IDP-Datenbank

DTA II / IDP (2008 bis 2012)

Syrische und Sanskrit-Fragmente

In der letzten Projektphase erfolgte die Digitalisierung teils indirekt, teils direkt. So wurden die syrischen Fragmente noch im zweistufigen Verfahren (Diapositiv, Diascan) digitalisiert. Die Sanskrit-Fragmente hingegen wurden größtenteils direkt mithilfe eines Großscanners/Buchscanners durch eine beauftragte Fremdfirma innerhalb weniger Monate zu Beginn des Projekts Anfang 2009 gescannt. Nur im Falle von Großformaten oder bei restaurierungsbedürftigen Fragmenten, die erst später zur Verfügung standen, erfolgte die Digitalisierung noch auf dem bisherigen Wege. Die Arbeitsschritte blieben im Übrigen dieselben.

Vergleich der Bildqualität

- Diafilm/Scan: schärfer, sattere Farben, teilweise aber auch Probleme beim Einscannen (Scharfstellung bei Einzelscans, Kratzer/Beschädigungen der Filmoberfläche, Fingerabdruck); Vorteil: längere, risikofreiere Archivierung
- Direktscan: schnellerer Bearbeitungsprozess, schnellere/platzsparende Speicherung bei nicht gravierend schlechterer Bildqualität
- Digitalfotografie: zu geringe Erfahrungswerte, bisher mit den anderen Verfahren vergleichbar gute Ergebnisse, leicht andere Farbwerte

Vorteile der IDP-Datenbank

- Vernetzung:
  Bei den Sanskrit-Fragmenten (SHT) z.B. verweisen Links auf andere Fragmente, die zu demselben Blatt oder derselben Handschrift gehören. Durch die Vernetzung der Datensätze ist so ein schneller Überblick über zusammengehörige Fragmente möglich, welche ansonsten auf verschiedene SHT-Nummern über mehrere gedruckte Katalogbände verteilt sind (Verzeichnis orientalischer Handschriften Deutschlands, Band X, 1–11).

- Bündelung:
  Der Online-Katalog der Sanskrit-Fragmente liefert erstmals zu jedem einzelnen Fragment alle verfügbaren Informationen – auch zu den äußerst umfangreichen Katalognummern der ersten SHT-Teilbände, die dort wenig detailliert dokumentiert wurden. Alle Informationen zu einem Fragment sind somit an einem Ort gebündelt und zudem auf dem aktuellsten Stand der Forschung.
- Bibliographie:
  Eine vollständige Bibliographie zu den Sanskrit-Fragmenten kann in der IDP-Datenbank mittels der bibliographischen Suche nach dem Schlüsselwort „SHT“ angezeigt werden.
- Erweiterte Suche:
  Über die erweiterte Suche von IDP in der Rubrik »Kurztitel« ist ein schnelles Auffinden von Fragmenten, die den gleichen Text enthalten, möglich (z.B. Bhiksuni-pratimoksa-vibhanga oder Saddharmapundarikasutra).
  Auch kann nach bestimmten Kriterien gesucht werden, wie z.B. Blockdrucke in chinesischer Schrift, die sich im Besitz der Berlin-Brandenburgischen Akademie befinden.
  Weitere Suchen wie z.B. nach Sanskrit-Handschriften, die aus Kucha stammen, sind möglich und eröffnen neue Perspektiven auf Inhalt und Kontext der Handschriftensammlungen.