SLUB
Wir führen Wissen.

Handreichung retrodigitalisierte Monographien

SLUB Dresden

Version 1.4.2, 2023-06-19

Inhaltsverzeichnis

Vorwort . . . 1
Allgemeines . . . 1
- Retrodigitalisate . . . 1
- Volltexte . . . 1

Abgrenzung . . . 2
Anforderungen Retrodigitalisate . . . 3
- Allgemein . . . 3
- Hinweise zu ICC-Farbprofilen . . . 13
- Validierung . . . 14

Volltexte . . . 15
- Verpflichtend . . . 15
- Gute Praxis . . . 15

Strukturinformationen . . . 17
- Verpflichtend . . . 17
- Gute Praxis . . . 17

Beispiel . . . 18
Tiefenerschliessung deskriptive Metadaten . . . 19
Bibliographie . . . 20

Release Note 1.4.2 vom 2023-06-19

- Ergänzung Strukturmetadaten und OCR
- Fix Tabellenauszeichnung
- Fix für Tolerierung ICC

Vorwort

Dieses Dokument richtet sich an Produzenten, die digitale Objekte in das SLUBArchiv.digital einliefern und diese langfristig benutzbar erhalten wollen.

Allgemeines

Retrodigitalisate

Die Digitalisierung der historischen Bestände (Monographien) der SLUB verfolgt mehrere wichtige Ziele:

- Schonung der Originalvorlagen
- weltweiter digitaler Zugriff

Der Prozess der Digitalisierung ist aufgrund der Bestandsgröße eine ressourcenaufwändige Aufgabe. Hinzu kommt, dass vom Zerfall bedrohte Originale zum Zwecke der Bestandserhaltung gesichert werden.

Das Tagged Image File Format (TIFF) bietet vielfältige Möglichkeiten, Bild- und Metadaten in Dateien abzulegen. Nicht alle davon sind aber langzeitarchivfähig. Einige der Funktionen stellen ein Risiko für die dauerhafte Aufbewahrung dar.

Grundsätzlich muss zwischen Bilddaten der Retrodigitalisierung und sogenannten born-digital Bildern unterschieden werden. Letztere sind Fotos, die ein Fotograf mit seiner Digitalkamera schießt, und die meist aus Vor- und Nachlässen stammen.

Volltexte

Obgleich zu den Retrodigitalisaten zugehörige Volltexte zu den Präsentationen zu zählen sind, erhalten sie

- durch den hohen Aufwand in der maschinellen Erzeugung (OCR)
- oder durch die manuelle Transkription
- sowie wegen der unter Umständen erfolgten manuellen Korrektur
- und aufgrund der Referenzierbarkeit der Textstelle

eine herausragende Stellung, die die Archivwürdigkeit begründet.

Für die Ablieferung von Bilddaten und zugehörigen Volltextdaten in ihrem Langzeitarchiv hat die SLUB sich deswegen auf eine verbindliche Untermenge der verfügbaren Funktionen festgelegt. Diese wird im Folgenden beschrieben.

Als Basis dient eine interne Analyse der Nutzungsszenarien und der daraus abgeleiteten, signifikanten Eigenschaften retrodigitalisierter Monographien.

> Vorbehaltlich gesonderter Absprachen mit dem SLUBArchiv sind von dieser Handreichung abweichende Änderungen **nicht** gestattet. Informationspakete, die diese Anforderung nicht erfüllen, sind nicht langzeitarchivfähig und werden vom SLUBArchiv zurückgewiesen.

Abgrenzung

Dieses Dokument beschreibt die Anforderungen des SLUBArchiv an retrodigitalisierte Monographien und vergleichbare Schriften, die vergleichbare Nutzungsarten aufweisen.

Nicht von diesem Dokument erfasst sind

- originär digital erstellte Publikationen
- originär digital erstellte (künstlerische) Fotos
- Pflichtexemplare
- digitale Vor- und Nachlässe
- sonstige retrodigitalisierte Objekttypen, sofern sie nicht explizit auf diese Handreichung verweisen

Digitalisierungsanforderungen an Dienstleister sind im Einklang mit diesem Dokument zu verfassen.

Anforderungen Retrodigitalisate

Allgemein

Verpflichtend

- Alle Digitalisate müssen grundsätzlich dem TIFF 6.0 [https://web.archive.org/web/20130204132534/http://partners.adobe.com/public/developer/en/tiff/TIFF6.pdf]-Baseline-Standard entsprechen; v. a. auch in Bezug auf Datentypen und deren Wertebereiche. SLUB-spezifische Erweiterungen werden explizit genannt.
- Es dürfen ausschließlich Datentypen aus dem TIFF 6.0 [https://web.archive.org/web/20130204132534/http://partners.adobe.com/public/developer/en/tiff/TIFF6.pdf]-Standard verwendet werden, jedoch nicht aus dem BigTIFF-Standard.
- Es sind ausschließlich Images in bitonal, grayscale oder RGB zugelassen (kein palette color, keine Transparency Mask). Die gewählte Qualität darf nicht unter der Qualität der Vorlage liegen.
- Es darf nur genau ein Image File Directory (IFD) vorhanden sein.
- Multipage-TIFF ist nicht erlaubt (d. h., dass nur genau das Master-Image darf enthalten sein darf; keine Vorschaubilder/Thumbnails, keine alternativen Ansichten etc.)
- Es gilt der Grundsatz, dass Metadaten möglichst mit den bereits vorhandenen Tags abgebildet werden sollten. Nur wenn dies nicht möglich ist, darf auf zusätzliche Metadatensätze (XMP [http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=57421], EXIF [http://www.cipa.jp/std/documents/e/DC-008-2012_E.pdf]) ausgewichen werden.
- Wenn die TIFF-Headerdaten über eine wie auch immer geartete nachgeschaltete Modifikation der TIFF-Header (genauer des IFD) im Nachhinein entfernt werden sollen, so muss sichergestellt sein, dass auch ggf. im Dateikörper referenzierte Datenblöcke mit entfernt werden und danach alle Offsets wieder korrekt sind.
- Tags vom Typ "ASCII" dürfen nur druckbare Zeichen aus dem ASCII-Zeichensatz, sowie Leerzeichen und NULL-Bytes enthalten. Aufeinanderfolgende NULL-Bytes sind nicht erlaubt.

Zwingend erforderliche Tags nach Baseline-TIFF6 plus SLUB-spezifische Erweiterungen

Diese Felder **müssen** in jeder TIFF-Datei enthalten sein. Die nachfolgende Tabelle dient der näheren Erläuterung.

Tag-ID	Tag-ID (hex)	Tag-Name	Pflichtfeld?	Bedingung
256	0100	ImageWidth	verpflichtend	beliebiger Wert zwischen 1 und 4.294.967.295 (TIFF-Typ LONG)

Tag-ID	Tag-ID (hex)	Tag-Name	Pflichtfeld?	Bedingung
257	0101	ImageLength	verpflichtend	beliebiger Wert zwischen 1 und 4.294.967.295 (TIFF-Typ LONG)
258	0102	BitsPerSample	optional / verpflichtend / verpflichtend	Tag muss den Wert „1“ (bitonal), „1“, „4“, „8“ (Graustufen) oder „8,8,8“ bzw. „16,16,16“ (RGB) enthalten. Optionales Feld für bitonal (Default: „1“), Pflichtfeld für Graustufen, Pflichtfeld für RGB
259	0103	Compression	verpflichtend	Wert „1“ bedeutet „no compression“; keine anderen Werte erlaubt; Default: „1“
262	0106	PhotometricInterpretation	verpflichtend	Wert „0“ bedeutet WhiteIsZero, Wert „1“ bedeutet BlackIsZero, Wert „2“ bedeutet RGB. Zugelassen sind NUR die Werte „0“ oder „1“ oder „2“. Wert „3“ bedeutet Palettenfarben (verboten), „4“ bedeutet Transparency Mask (verboten). Ausdrücklich KEIN Default!
273	0111	StripOffsets	verpflichtend	beliebiger Wert zwischen 1 und 4.294.967.295 (TIFF-Typ LONG)

Tag-ID	Tag-ID (hex)	Tag-Name	Pflichtfeld?	Bedingung
277	0115	SamplesPerPixel	optional / optional / verpflichtend	muss den Wert „1“ (bitonal, Graustufen) oder „3“ (RGB) enthalten. Baseline TIFF erlaubt für dieses Tag auch Werte über „3“; das ist aber für die Langzeitarchivierung NICHT erlaubt. Default: „1“
278	0116	RowsPerStrip	verpflichtend	beliebiger Wert zwischen 1 und 4.294.967.295 (TIFF-Typ LONG); Default: 2**32 - 1 (gesamtes Bild ist ein Strip)
279	0117	StripByteCounts	verpflichtend	beliebiger Wert zwischen 1 und 4.294.967.295 (TIFF-Typ LONG)
282	011A	XResolution	verpflichtend	Wert vom TIFF-Typ RATIONAL
283	011B	YResolution	verpflichtend	Wert vom TIFF-Typ RATIONAL

Optionale Tags nach Baseline-TIFF6 plus SLUB-spezifische Erweiterungen

Die nachfolgende Tabelle dient der näheren Erläuterung.

Tag-ID	Tag-ID (hex)	Tag-Name	Pflichtfeld?	Bedingung
254	00FE	NewSubfileType	optional mit Einschränkungen	Optional erlaubt. Wenn es verwendet wird, sind ausschließlich die Werte „0“ oder „2“ erlaubt. Default: „0“

Tag-ID	Tag-ID (hex)	Tag-Name	Pflichtfeld?	Bedingung
255	00FF	SubfileType	verboten	Darf nicht vorkommen (veraltet). Stattdessen muss Tag 254 NewSubfileType verwendet werden.
263	0107	Thresholding	verboten	Darf nicht vorkommen; nicht langzeitarchivfähig.
264	0108	CellWidth	verboten	Darf nicht vorkommen; nicht langzeitarchivfähig.
265	0109	CellLength	verboten	Darf nicht vorkommen; nicht langzeitarchivfähig.
266	010A	FillOrder	optional mit Einschränkungen	Optional erlaubt. Wenn es verwendet wird, ist ausschließlich Wert „1“ erlaubt. Default: „1“
269	010D	DocumentName	optional mit Einschränkungen	Optional erlaubt. Dokumentenname (TIFF-Typ ASCII). Darf (muss aber nicht) vorkommen. Wenn es aber vorkommt, muss es den Dateinamen des Dokumentes in ASCII-Text enthalten. Ein leerer String ist nicht erlaubt.

Tag-ID	Tag-ID (hex)	Tag-Name	Pflichtfeld?	Bedingung
270	010E	ImageDescription	optional mit Einschränkungen	Optional erlaubt. Bildbeschreibung (TIFF-Typ ASCII). Darf (muss aber nicht) vorkommen. Wenn es aber vorkommt, muss es eine Beschreibung des Bildinhaltes in ASCII-Text enthalten. Ein leerer String ist nicht erlaubt.
271	010F	Make	optional mit Einschränkungen	Optional erlaubt. Scannerhersteller (TIFF-Typ ASCII). Darf (muss aber nicht) vorkommen. Wenn es aber vorkommt, muss es den Scannerhersteller in ASCII-Text enthalten. Ein leerer String ist nicht erlaubt.
272	0110	Model	optional mit Einschränkungen	Optional erlaubt. Scannermodell (TIFF-Typ ASCII). Darf (muss aber nicht) vorkommen. Wenn es aber vorkommt, muss es das Scannermodell in ASCII-Text enthalten. Ein leerer String ist nicht erlaubt.

Tag-ID	Tag-ID (hex)	Tag-Name	Pflichtfeld?	Bedingung
274	0112	Orientation	optional mit Einschränkungen	Optional erlaubt. Wenn es verwendet wird, ist ausschließlich Wert „1“ erlaubt. Default: „1“
280	0118	MinSampleValue	optional	Optional erlaubt. Default: „0“
281	0119	MaxSampleValue	optional	Optional erlaubt. Default: 2^BitsPerSample-1, sh. TIFF 6.0 [https://web.archive.org/web/20130204132534/http://partners.adobe.com/public/developer/en/tiff/TIFF6.pdf] Spezifikation, Seite 35
284	011C	PlanarConfiguration	optional mit Einschränkungen	Optional erlaubt. Wenn es verwendet wird, ist ausschließlich Wert „1“ erlaubt. Default: „1“
288	0120	FreeOffsets	verboten	Darf nicht vorkommen; nicht langzeitarchivfähig.
289	0121	FreeByteCounts	verboten	Darf nicht vorkommen; nicht langzeitarchivfähig.

Tag-ID	Tag-ID (hex)	Tag-Name	Pflichtfeld?	Bedingung
290	0122	GrayResponseUnit	optional mit Einschränkungen	Optional, nur sinnvoll, falls es sich um ein Graustufenbild mit nichtlinearer Kennlinie handelt (Tag 258 = 4/8). Die Werte müssen im Bereich 1-5 liegen, falls das Tag verwendet wird. Default: „2“, Empfohlen: „3“
291	0123	GrayResponseCurve	optional	Optional, nur sinnvoll, falls es sich um ein Graustufenbild handelt (Tag 258 = 4/8).
296	0128	ResolutionUnit	optional	Wert „2“ (= Inch). Keine anderen Werte erlaubt. Default: „2“
297	0129	PageNumber	optional mit Einschränkungen	Optional erlaubt. Der erste Wert beschreibt die aktuelle Seite und muss „0“ lauten, der zweite Wert beschreibt die Gesamtanzahl der Seiten in der vorliegenden Datei und muss „1“ lauten.

Tag-ID	Tag-ID (hex)	Tag-Name	Pflichtfeld?	Bedingung
305	0131	Software	optional mit Einschränkungen	Name und Version der Erzeugungssoftware. Darf (muss aber nicht) vorkommen. Wenn es aber vorkommt, muss es die Erzeugungssoftware (möglichst mit Versions- und/oder Buildnummer) in ASCII-Text enthalten. Ein leerer String ist nicht erlaubt.
306	0132	DateTime	optional	Datum der Bilderzeugung im Format „YYYY:MM:DD HH:MM:SS“.
315	013B	Artist	verboten	Darf nicht vorkommen; nicht langzeitarchivfähig, weil es potentiell personengebundene Daten als Freitext enthält. Außerdem gibt es bei Scans keinen „Artist“ im eigentlichen Sinne, weil kein Künstler ein Foto geschossen hat, sondern ein Scanoperator lediglich eine Kopie eines bestehenden Werkes erstellt hat.

Tag-ID	Tag-ID (hex)	Tag-Name	Pflichtfeld?	Bedingung
316	013C	HostComputer	verboten	Darf nicht vorkommen; nicht langzeitarchivfähig.
318	013E	WhitePoint	optional	Optional erlaubt.
319	013F	PrimaryChromaticities	optional	Optional erlaubt.
320	0140	ColorMap	verboten	Darf nicht vorkommen; nicht langzeitarchivfähig.
338	0152	ExtraSamples	verboten	Darf nicht vorkommen; nicht langzeitarchivfähig.
339	0153	SampleFormat	optional mit Einschränkungen	Wenn gesetzt, dann für RGB-TIFFs nur „1,1,1“, für Graustufen-TIFFs „1“
700	02BC	XMP [http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=57421]	optional mit Einschränkungen	Optional erlaubt. Beliebige XMP-Metadaten (TIFF-Typ BYTE[]). Darf (muss aber nicht) vorkommen. Wenn es aber vorkommt, muss es XMP-Metadaten enthalten.

Tag-ID	Tag-ID (hex)	Tag-Name	Pflichtfeld?	Bedingung
33432	8298	Copyright	optional mit Einschränkungen	Optional erlaubt. Beliebige Copyright-Informationen in Textform (TIFF-Typ ASCII). Der Inhalt des Feldes sollte frei konfigurierbar sein (falls er das in der bisherigen Software nicht schon ist). Darf (muss aber nicht) vorkommen. Wenn es aber vorkommt, muss es Copyright-Informationen in ASCII-Text enthalten. Ein leerer String ist nicht erlaubt.
33723	83BB	IPTC	verboten	Darf nicht vorkommen; nicht langzeitarchivfähig.
34377	8649	Photoshop	verboten	Darf nicht vorkommen; nicht langzeitarchivfähig.
34665	8769	EXIF [http://www.cipa.jp/std/documents/e/DC-008-2012_E.pdf] IFD	optional	Pointer auf EXIF-Metadaten. Darf (muss aber nicht) vorkommen. Wenn es aber vorkommt, muss es auf EXIF [http://www.cipa.jp/std/documents/e/DC-008-2012_E.pdf]-Metadaten verweisen.

Tag-ID	Tag-ID (hex)	Tag-Name	Pflichtfeld?	Bedingung
34675	8773	ICC [http://www.color.org/specification/ICC1v43_2010-12.pdf] Profile	optional / verpflichtend	Pointer auf Farbprofildaten im RGB-Farbraum. Pflichtfeld für Ablieferung an die SLUB falls es sich um ein Farbbild handelt (Tag 258 = 8,8,8 / 16,16,16).

Zusammengefasst[1] sind für Ablieferungen an die SLUB verpflichtend:

- die Tags 256, 257, 258, 259, 262, 273, 277, 278, 279, 282, 283
- für RGB-Bilder zusätzlich das Tag 34675

Von Seiten der SLUB empfohlen werden desweiteren die Tags 271, 272, 305, 306, 700 und 34665.

[1] Ausschlaggebend bleiben immer obige Tabellen

Hinweise zu ICC-Farbprofilen

Verpflichtend

In der Vergangenheit wurden große Bestände an TIFF Dateien bekannt, die im ICC-Farbprofil als Preferred CMM Type den Wert "Lino" enthalten. Nach der ICC [http://www.color.org/specification/ICC1v43_2010-12.pdf] Spezifikation ist dieser Wert nicht zulässig. Die Einlieferung von TIFFs mit diesen ICC-Farbprofilen ist nicht möglich, da eine Garantie für den Erfolg zukünftiger Formatmigrationen nicht gegeben werden kann.

Im übrigen sind nur Farbprofile zulässig, die die ICC-Spezifikationen ICC 1:2022 oder ICC 1:2010 implementieren (siehe http://www.color.org/icc_specs2.xalter).

Für Bestandsdaten wird die Farbprofilversion ICC 1:2001-04 noch toleriert. Für Neudigitalisierungen ist ICC 1:2001-04 nicht mehr zulässig.

Toleriert werden darüber hinaus ICC.1:2004-10, ICC.1:2004-10 (ISO 15076:2005), ICC.1:2004-04, ICC.1:2003-09 (Version 4.1.0) und ICC.1:2001-12 (Version 4.0.0).

Die Farbprofilversion ICC 1:2001-04 ist durch die ICC seit 2019 als veraltet markiert. Bitte stellen Sie ihre Workflows, Gerätepark und Software rechtzeitig um! Das SLUBArchiv wird in einer der nächsten Releasezyklen Dateien mit ICC [http://www.color.org/specification/ICC1v43_2010-12.pdf]-Profilen älter als ICC 1:2001-12 nicht mehr zulassen.

Gute Praxis

Es wird empfohlen, darauf zu achten, dass die Farbprofile

- auf proprietäre "preferred CMM" verzichten
- im "Copyright" keine Einschränkungen vorsehen
- "CalibrationDatetime" enthalten und aktuell ist

Auf Anfrage berät das SLUBArchiv zu möglichen Reparaturoptionen.

Validierung

Die dargestellten Einschränkungen können mit der Software checkit_tiff [https://git.fsfe.org/art1pirat/checkit_tiff/releases/tag/v1.4.4] geprüft werden und sind in der mitgelieferten Konfigurationsdatei 'cit_tiff6_baseline_SLUB.cfg' (Git-Release *stable, 1.4.4*) hinterlegt.

Volltexte

Verpflichtend

Die Volltexte werden jeweils als AltoXML-Dateien kodiert. Jeder Volltextdatei ist eine Bilddatei zuortenbar. Sie besitzen den gleichen Dateinamensprefix wie die zugehörigen Bilddateien. Eine separate Ordnerstruktur mit je einem eigenen Unterverzeichnis für Bild- und Volltextdateien ist ausdrücklich erlaubt.

> Die Verwendung von Volltextdateien als Bestandteil der digitalen Objekte ist **nicht verpflichtend**. Digitale Objekte, die nur Bilddateien und Strukturinformationen enthalten sind ausdrücklich möglich.

Die Volltextdateien müssen folgende Bedingungen erfüllen:

- Sie sind wohlgeformt und valide hinsichtlich der AltoXML (v2.0) [https://www.loc.gov/standards/alto/v4/alto-2-0.xsd]-Spezifikation
- die Positionsangaben beziehen sich auf die zugeordnete und mitarchivierte Bilddatei und sind
  - entweder in 1/10mm
  - oder in Pixel zu kodieren.
- leere Elemente sind nicht statthaft

> Zum gegenwärtigen Zeitpunkt wird nur AltoXML (v2.0) [https://www.loc.gov/standards/alto/v4/alto-2-0.xsd] unterstützt, da die Aufwände für ein Upgrade auf AltoXML (v4.3) [https://www.loc.gov/standards/alto/v4/alto-4-3.xsd] durch die Fachabteilungen noch nicht abschliessend beurteilt werden konnte. Die AltoXML (v2.0) [https://www.loc.gov/standards/alto/v4/alto-2-0.xsd] Dokumente lassen sich mit hoher Wahrscheinlichkeit allein durch Änderungen des Namespace von
>
> ```
> <alto xmlns="http://www.loc.gov/standards/alto/ns-v2#"
> xmlns:xlink="http://www.w3.org/1999/xlink"
> xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
> xsi:schemaLocation="http://www.loc.gov/standards/alto/ns-v2#
> http://www.loc.gov/standards/alto/alto-v2.0.xsd">
> ```
>
> auf
>
> ```
> <alto xmlns="http://www.loc.gov/standards/alto/ns-v4#"
> xmlns:xlink="http://www.w3.org/1999/xlink"
> xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
> xsi:schemaLocation="http://www.loc.gov/standards/alto/ns-v4#
> http://www.loc.gov/standards/alto/alto-v4.3.xsd">
> ```
>
> anpassen.

Gute Praxis

- Positionsangaben sollten in 1/10mm erfolgen, da
  - das Top-level Element *description* entfallen kann (1/10mm ist Voreinstellung in AltoXML)
  - diese sich ohne Umrechnungen auf das Ausgangsmaterial beziehen
- Sinnvoll sind funktionale Angaben (Leserichtung, Lesereihenfolge, ...)
- auf das Top-level Element *style* ist zu verzichten, wenn
  - ausgezeichnete Elemente nicht vorhanden sind
  - markierte Stellen nicht zwingend zum Textverständnis beitragen
- Erforderlich sind alle Informationsobjekte, die den Volltext logisch strukturieren (Seite, Absatz, ...)
- Angaben sind verzichtbar, wenn sie
  - nicht bedeutungstragend sind
  - identisch zu Voreinstellungen sind
  - zu Missverständnissen und Widersprüchen führen könnten

Strukturinformationen

Verpflichtend

Die Zuordnung der Retrodigitalisate und Volltexte erfolgt durch eine **minimale** Kodierung in METS [https://www.loc.gov/standards/mets/mets.xsd]. Diese muss folgende Bedingungen erfüllen:

- Die METS-XML-Datei ist wohlgeformt und valide hinsichtlich der METS [https://www.loc.gov/standards/mets/mets.xsd] Spezifikation
- die Zuordnung der TIFFs (Retrodigitalisate) zu den AltoXML-Dateien (Volltexte) ist vollständig. Sie erfolgt durch relative Links innerhalb der intellektuellen Einheit (IE).
- die Zuordnung der TIFFs (Retrodigitalisate) zu den physischen Vorlagen ist vollständig.
- die *fileSec* darf nur die *fileGrp* enthalten, die für die genannten Zuordnungen notwendig sind. Sie sind einheitlich und sprechend zu wählen. Andere *fileGrp* sind nicht erlaubt.
- die *structMap* ist erforderlich und beschreibt:
  - mindestens die Zuordnung zwischen Bilddatei und physischer Vorlage.
  - die Zuordnung zwischen den Bild- und Volltextdateien.
- Eingebettete AltoXML-Elemente sind **nicht** zulässig
- IE-spezifische Links, die auf Quellen außerhalb des IEs verweisen sind **grundsätzlich unzulässig**.
- die Top-level Elemente *dmdSec*, *amdSec*, *behaviourSec* sind **nicht** zu verwenden

Die Verwendung von Strukturinformationen ist **in jedem Fall verpflichtend**, auch wenn nur Retrodigitalisate archiviert werden sollen. Der Hintergrund ist, dass nur mithilfe der Strukturinformationen sichergestellt ist, dass eine eindeutige Zuordnung von physischen Vorlagen und Bilddateien ermöglicht wird.

Gute Praxis

- Als *fileGrp* für die Bilddateien wird *"digital_preserved_image"* empfohlen
- Als *fileGrp* für die Volltextdateien wird *"digital_preserved_text"* empfohlen
- *structLink* kann für den Erhalt der Navigationsfähigkeit zwischen physischer und logischer Strukturierung genutzt werden.

Beispiel

Minimalbeispiel (Verzeichnis IE)

```
IE/
  ├── 1.tiff
  ├── 1.alto.xml
  └── mets.xml
```

Minimalbeispiel (METS Datei mets.xml)

```xml
<?xml version="1.0"?>
<METS:mets
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xmlns:METS="http://www.loc.gov/METS/"
  xmlns:xlink="http://www.w3.org/TR/xlink"
  xsi:schemaLocation="http://www.loc.gov/METS/
  https://www.loc.gov/standards/mets/mets.xsd">
  <METS:fileSec>
    <METS:fileGrp USE="digital_preserved_image">
      <METS:file ID="FILE_0000_digital_preserved_image">
        <METS:FLocat xmlns:xlink="http://www.w3.org/1999/xlink" LOCTYPE="URL" xlink:href="file://1.tif"/>
      </METS:file>
    </METS:fileGrp>
    <METS:fileGrp USE="digital_preserved_text">
      <METS:file ID="FILE_0000_digital_preserved_text">
        <METS:FLocat xmlns:xlink="http://www.w3.org/1999/xlink" LOCTYPE="URL" xlink:href=
"file://1.alto.xml"/>
      </METS:file>
    </METS:fileGrp>
  </METS:fileSec>
  <METS:structMap TYPE="LOGICAL">
    <METS:div ID="LOG_0000" LABEL="Beispiel1" TYPE="monograph">
      <METS:div ID="LOG_0001" LABEL="Beispiel2" TYPE="title_page"/>
      <METS:div ID="LOG_0002" LABEL="Beispiel3" TYPE="chapter"/>
    </METS:div>
  </METS:structMap>
  <METS:structMap TYPE="PHYSICAL">
    <METS:div ID="PHYS_0000" TYPE="physSequence">
      <METS:div ID="PHYS_0001" ORDER="1" ORDERLABEL="1" TYPE="page">
        <METS:fptr FILEID="FILE_0000_digital_preserved_image"/>
        <METS:fptr FILEID="FILE_0000_digital_preserved_text"/>
      </METS:div>
    </METS:div>
  </METS:structMap>
  <METS:structLink>
    <METS:smLink xmlns:xlink="http://www.w3.org/1999/xlink" xlink:to="PHYS_0001" xlink:from="LOG_0000"/>
  </METS:structLink>
</METS:mets>
```

Tiefenerschliessung deskriptive Metadaten

Die Einführung der Kodierung deskriptiver Metadaten auf Strukturebene ist zur Zeit noch nicht abschliessend gelöst.

In dieser Spezifikation wird daher auf die Kodierung von deskriptiven Metadaten und deren Verknüpfung via logischer Structmap **ausdrücklich** verzichtet.

Dies kann sich in einer der nächsten Releases ändern.

Bibliographie

bibliography::[]