SLUB
Wir führen Wissen.

Handreichung retrodigitalisierte Monographien

SLUB Dresden

Version 1.3.14, 2022-07-26

Inhaltsverzeichnis

Allgemeines . . . 1
Abgrenzung . . . 1
Anforderungen Retrodigitalisate . . . 2
- Allgemein . . . 2
- Hinweise zu ICC-Farbprofilen . . . 12
- Validierung . . . 13

Release Note 1.3.14 vom 2022-07-26

- Aktualisierung und Klarstellung SampleFormat
- Tag 296 (ResolutionUnit) nun optional, da es in der TIFF 6.0 [https://web.archive.org/web/20130204132534/http://partners.adobe.com/public/developer/en/tiff/TIFF6.pdf]-Spezifikation einen Default-Wert "2" gibt.
- Verschärfung *Lino* als preferred-CMM-type in ICC-Farbprofilen.
- Klarstellung Tabellenüberschrift
- Links aktualisiert, ua. zu checkit_tiff
- Klarstellung zu Tag-Typ ASCII
- Anpassung Link zu Repository
- Hinweis zu ICC Farbprofilen der Version 2.4.0
- Beschreibung der anzuwendenden Objekttypen
- Dokumentenstruktur verbessert

Allgemeines

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 aus Gründen 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.

Für die Ablieferung von Bilddaten 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.

Abgrenzung

Dieses Dokument beschreibt die Anforderungen des SLUBArchiv an retrodigitalisierte Monographien und vergleichbare Schriften, die ähnliche 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)
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“

Tag-ID	Tag-ID (hex)	Tag-Name	Pflichtfeld?	Bedingung
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)
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)

Tag-ID	Tag-ID (hex)	Tag-Name	Pflichtfeld?	Bedingung
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“
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.

Tag-ID	Tag-ID (hex)	Tag-Name	Pflichtfeld?	Bedingung
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.
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.

Tag-ID	Tag-ID (hex)	Tag-Name	Pflichtfeld?	Bedingung
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.
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“

Tag-ID	Tag-ID (hex)	Tag-Name	Pflichtfeld?	Bedingung
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.
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“

Tag-ID	Tag-ID (hex)	Tag-Name	Pflichtfeld?	Bedingung
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.
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“.

Tag-ID	Tag-ID (hex)	Tag-Name	Pflichtfeld?	Bedingung
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.
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“

Tag-ID	Tag-ID (hex)	Tag-Name	Pflichtfeld?	Bedingung
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.
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.

Tag-ID	Tag-ID (hex)	Tag-Name	Pflichtfeld?	Bedingung
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.
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

Vorbehaltlich gesonderter Absprachen mit dem SLUBArchiv sind andere als die in den oben genannten Tabellen aufgeführten TIFF-Tags und Erweiterungen NICHT gestattet. Dateien, die dieser oder kommender Handreichungen nicht entsprechen, können in der Regel nicht im SLUBArchiv langzeitarchiviert werden.

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).

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).

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/v1.4.1] geprüft werden und sind in der mitgelieferten Konfigurationsdatei 'cit_tiff6_baseline_SLUB.cfg' (Git-Release *stable, 1.4.1*) hinterlegt.