Guidelines for TIFF Metadata
Recommended Elements and Format
Version 1.0

February 10, 2009

Tagged Image File Format (TIFF) is a tag-based file format for the storage and interchange of raster images. It serves as a wrapper for a variety of encoded bit-mapped images, and is a popular format to use as a sustainable master in cultural heritage digital imaging projects. Created in the mid-1980s, TIFF was designed to be cross platform, backward compatible and, whenever possible, forward compatible. The most recent version of TIFF is 6.0, published in 1992. Adobe Systems controls the TIFF specification.

Metadata is an essential component of the TIFF format; many TIFF images dating back to the 1980s still can be displayed in modern TIFF readers through the interpretation of the tagged metadata set. The tag set is extensible through a system of "private tags" and a framework for new complete systems of tags. Aware Systems' TIFF Tag Reference lists 36 "Baseline" tags, 60 "Extension" tags, 74 "Private" tags, plus a set of 58 Exchangeable image file format (Exif) tags standardized by the Japan Electronics and Information Technology Industries Association (JEITA), a large group of camera manufactures. Other large extension sets of tags have been developed for GPS, GeoReference, and medical informatics purposes, as well as for Adobe's Digital Negative (DNG) specification. The Library of Congress also maintains a complete tag list, along with significant additional information on the sustainability of numerous digital file formats at http://www.digitalpreservation.gov/formats/content/tiff_tags.shtml.

While the tag system has been extremely successful, the proliferation of tags and tag sets complicates metadata extraction; most TIFF programs only locate and display the baseline, extension, and a few private tags. Only the Exif tag system is widely supported. A second difficulty is that the extracted data is difficult to use and store because of the different data types for the various tagged fields, and the lack of any systematic data structures and formats.

One response to these TIFF metadata issues has been the development and adoption of Adobe's Extensible Metadata Platform (XMP). XMP defines a standard way to express standardized TIFF metadata in RDF-based XML. Dublin Core and International Press Telecommunications (IPTC) description sets, the Exif technical set, and any other formally defined metadata sets are also structured into RDF XML, and the entire metadata package is embedded in a single TIFF tagged field.

These TIFF tag guidelines, submitted by the Still Image Working Group, present a minimal set of embedded metadata for digital imaging of historical and cultural heritage materials. These guidelines assume that all baseline tags required by the TIFF Specifications V6.0 are properly in place so that all images will be accepted by any TIFF image validator. The first table, the "Suggested Minimal TIFF Tag Set," specifies those baseline tags considered particularly important to the cultural heritage community. Some, such as XResolution (Tag 282), may be inaccurately entered. PhotometricInterpretation (Tag 262) is important for grayscale images which are not supported by the Exif standard. Also included on this list are those extension tags

which present critical information, such as DocumentName (Tag 269), or which contain mandatory data, when applicable, such as ExtraSamples (Tag 338), needed to mark an alpha channel.

The tags included in the second table, "Notes on Additional Recommended Metadata," are recommended for cultural heritage or historical digitization projects, but may not be considered essential by all. These tags are drawn from the entire TIFF tag set including the Exif, IPTC, and Dublin Core private tag sets. Examples include the Make (Tag 271), giving the scanner manufacturer; IPTC copyright information (contained in Tag 33723); and LightSource (Tag 37384), which is particularly important when a target is included within the image. While these fields are categorized as "recommended," their inclusion adds value to the workflow and lifecycle chains of the images.

These recommendations for TIFF tags are presented as an interim guideline. A new recommended guideline encompassing a richer metadata set in a more flexible format, with wider file format support, is expected to be developed by the Embedded Metadata Sub-group.

References

Adobe Systems Incorporated, *TIFF Revision 6.0*, Final – June 3, 1992
   <http://partners.adobe.com/public/developer/en/tiff/TIFF6.pdf>.

Adobe Systems Incorporated, *XMP Specification*, September 2005
   <http://www.adobe.com/devnet/xmp/pdfs/xmp_specification.pdf>.

Dublin Core Metadata Initiative, *Dublin Core Metadata Element Set, Version 1.1*, January 14, 2008 <http://www.dublincore.org/documents/2008/01/14/dces/>.

IPTC, *IPTC – NAA Information Interchange Model Version 4*, Version No. 4, Rev. 1, July 1999
   <http://www.iptc.org/std/IIM/4.1/specification/IIMV4.1.pdf>.

JEITA CP-3451, *Exchangeable Image File Format for Digital Still Cameras: Exif Version 2.2*, April, 2002 <http://exif.org/Exif2-2.PDF>.

Library of Congress, *Sustainability of Digital Formats Planning for Library of Congress Collections, Updated March 6, 2008*
   <http://www.digitalpreservation.gov/formats/content/tiff_tags.shtml>

Suggested Minimal TIFF Tag Set

Metadata Set	Tag or identifier	Name	Description	Sample Values	Note
TIFF tag, baseline	256	ImageWidth	The number of pixels per row	3616	Typical scanner size
	257	ImageLength	The number of rows of pixels in the image	4418	Typical scanner size
	258	BitsPerSample	Number of bits per component	8 8 8 8	Grayscale 24-bit color
	259	Compression	Compression scheme used on image data	1 = Uncompressed 4 = CCITT Group 4	
	262	PhotometricInterpretation	The color space of the image data.	0 = WhiteIsZero. 1 = BlackIsZero. 2 = RGB.	Additional color spaces are possible: CMYK (5), YCbCr (6), CIE L*a*b* (8), and others.
	277	SamplesPerPixel	The number of components per pixel	1 3	Grayscale 24-bit RGB color
	282	XResolution	Horizontal pixel count per resolution unit (inches, centimeters)	2400000/10000	240 ppi: Rational data type - 240 pp(unit) is often displayed this way
	283	YResolution	Vertical pixel count per resolution unit (inches, centimeters)	629145600/2097152	300 ppi: see above
	296	ResolutionUnit	Unit of measurement for X and Y Resolution (inches, centimeters)	1 2 3	None: pixel is irregularly shaped Inches Centimeters
	306	DateTime	Date and Time image was scanned	2008:07:23 17:45:21	Note the space between the date and time and the use of the 24 hour clock. Ambiguous – no indication of time zone Possibly unreliable as it relies on the local clock of the scan station
	315	Artist	Used for ImageProducer	Library of Congress; Datatrac	Image referenced: Institution; Contractor Document referenced: Document author
		This tag must be present if applicable, for example if transparency data is present			
	338	ExtraSamples	Description of extra components	0 = Unspecified data 1 = Associated alpha data (with pre-multiplied color)	Associated alpha is generally interpreted as true transparency information..

		One and/or both of these tags below shall be incorporated into the dataset:			
TIFF tag, extended	269	DocumentName	Document Name	path/filename	Image referenced: Constructed to uniquely identify the file Document referenced: Document title
	270	ImageDescription	A text string that describes the subject of the image	Path/filename Agency uid	A baseline tag that must be respected by all applications. It may serve the same purpose as the 269 field.
	42016	ImageUniqueID	A unique file identifier	Uuid:01AF8BA45…	ASCII 128-bit UUID
		Additional tags are encouraged but not required by this guideline.			

The following table lists tags that are highly recommended, but not required.

Notes on Additional Recommended Metadata

Metadata Set	Tag or identifier	Name	Description	Sample Values	Note
TIFFf tag, baseline	271	Make	The scanner manufacturer	Sinar	Simple ASCII text string
TIFF tag, baseline	272	Model	The scanner model name or number	54H	Simple ASCII text string
TIFF tag, baseline	305	Software	Name and version of the software package(s) used to create the image	Stokes Software Inc. IWS - Version 02.04.01.01	Simple ASCII text string
TIFF tag, baseline	33432	Copyright	Copyright notice	John Smith, 20xx. All rights reserved	This field is inadequate for a complete expression of rights; consider IPTC, PLUS or external alternatives.
Exif	34665	Exif IFD	Pointer to collection of all Exif Metadata. Exif uses *field names* rather than tags to indicate the field content.	Scanners and cameras automatically capture and place significant technical metadata in Exif fields. Some software may duplicate *some, but **not all**,* Exif metadata into TIFF tags.	
Exif	37384	LightSource	The kind of light source.	0 = Unknown 1 = Daylight 21 = D65	Known light sources are appropriate when using targets containing reference color components.
Exif	40961	ColorSpace	sRGB specification	1 65535	sRGB uncalibrated
IPTC	33723	IPTC	IPTC metadata – descriptive		This does not have a tag series as Exif does. Data is usually written in BYTE format.

DC		Dublin Core	Descriptive metadata		Dublin Core data does not have a TIFF Tag location but may be written into XMP.
ICC	34675	ICC Profile	Color profile data		Undefined data type
XMP	700	XMP	Comprehensive metadata set	XML packet. Adobe products support "File Info" forms for Exif, IPTC, DC, and user defined metadata, which may be stored in the corresponding metadata fields and tags and in XMP.	
		Additional Metadata such as GeoReference tags are encouraged but not required by this guideline			