

All About Metadata

An Introduction for Digital Photographers

By David Hanks • dh@davidhanks.com
February 2007

All content (except where noted) is ©2007 David Hanks. It may be used freely for educational purposes provided this notice is included and content is not altered. This article is available in pdf online at www.davidhanks.com > digital photo management

Contents

- Section 1: Definitions
- Section 2: Purpose of Metadata
- Section 3: Types of Metadata
- Section 4: Further Information

Section 1: Definitions

- 1a. The prefix *meta-*
- 1b. The word *data*

1a. Meta- (prefix): About itself. Greek origin.

The prefix meta- is used to mean about (its own category). For example, metadata are data about data (who has produced it, when, what format the data are in and so on). Similarly, any subject can be said to have a meta-theory, which is the theoretical consideration of its foundations and methods. *(from: <http://en.wikipedia.org/wiki/Meta->)*

1b. Data: Factual information, especially information organized for analysis or used to make decisions. Latin origin.

Any form of information whether on paper or in electronic form. Data may refer to any electronic file no matter what the format: database data, text, images, audio and video. Everything read and written by the computer can be considered data except for instructions in a program that are executed (software). *(from: <http://www.answers.com/topic/data>)*

Section 2: Purpose of Metadata

- 2a. Why it exists
- 2b. Photographic usage
- 2c. Reasons for use

2a. Why metadata exists

Metadata is cataloging done by people - its constructed information, which means that it is of human invention and not found in nature. A good example of constructed information is the use of longitude and latitude to describe the earth and points thereon. The real planet obviously does not have lines going around it, although we are by now very accustomed to seeing maps and globes with them, but the invention of longitude and latitude allows us to talk about locations on the planet and to navigate precisely across vast expanses with no landmarks to guide us. *(from: http://www.kcoyle.net/jal2_Metadata.html)*

2b. Why metadata is used in digital photography

Metadata is *data about data* (or, notes about the photo - the non-image data associated with a digital photo or scan containing information about the photograph). Metadata always includes information about the camera, lens, and strobe that captured the photo as well as their settings (see below: *EXIF*). Metadata can also include basic or thorough information about a photo, such as the keywords, people, location, and other descriptive facts within the photo as well as information about copyright, and terms of use (see below: *IPTC*). Metadata is like writing notes on the back of a photo, although it's usually much more thorough.

2c. Reasons for using photographic metadata

Virtually all news organizations and stock agencies now require the use of metadata. Many images are no longer being accepted if the photographer doesn't enter metadata. News organizations want the complete who/what/when/where/why and copyright info embedded into the photo for credibility and fact checking. Stock agencies want metadata entered for cataloguing and copyright information. Thorough keywording allows their users to search large inventories for images, resulting in increased sales for the agency (and of course, the photographer). There has even been proposed legislation that would create *Orphan Works* – images without metadata that would fall into the public domain. Entering metadata can be done for each image separately or via batch processing and will usually increase the size of the file by a less than a few KBs.

Section 3: Types of Metadata

3a. File Properties

3b. EXIF

3c. IPTC

3d. Proprietary

3a. File Property Metadata

This is metadata that the computer generates about the photo at various stages of its evolution from RAW to TIF or JPG, etc. This metadata is usually non-editable, although making adjustments to the photo (data) will result in the metadata being changed. File Property metadata can contain fields such as the following:

Filename

File Location (path to the photo on the hard drive)

Document Kind (RAW, TIF, JPG, etc)

Application (software that transferred the photo from camera to computer)

Date Created

Date Modified

File Size (KB, MB)

Pixel Dimensions

Resolution (300dpi, 72dpi, etc)

Bit Depth (8, 16, 24)

Compression (1:X)

Color Mode (RGB, CMYK, Grayscale, etc)

Color Profile (sRGB, Adobe 1998, etc)

Modification History

3b. EXIF Metadata (*Exchangeable image file format*)

This metadata is recorded by the camera about the image at the time of capture. It is usually non-editable and can contain fields such as the following:

<i>EXIF Version</i>	<i>Shutter Speed</i>	<i>Flash Function</i>	<i>Range</i>
<i>Capture Date</i>	<i>ISO Speed</i>	<i>Flash Red-eye</i>	<i>Focal Plane Resolution</i>
<i>Capture Time</i>	<i>Exposure Bias</i>	<i>Orientation</i>	<i>Global Positioning</i>
<i>Camera Make</i>	<i>Exposure Program</i>	<i>White Balance Mode</i>	<i>System (GPS)</i>
<i>Camera Model</i>	<i>Brightness Value</i>	<i>Sharpening Applied</i>	<i>Thumbnail</i>
<i>Lens Make</i>	<i>Metering Mode</i>	<i>Sensor Type</i>	<i>(for quick previewing)</i>
<i>Lens Model</i>	<i>Flash Fired</i>	<i>Sensing Method</i>	<i>User-entered fields</i>
<i>Focal Length</i>	<i>Flash Return</i>	<i>Light Source</i>	<i>(input via camera</i>
<i>Aperture</i>	<i>Flash Mode</i>	<i>Subject Distance</i>	<i>menus)</i>

3c. IPTC Metadata (*International Press Telecommunications Council*)

The IPTC is a consortium of the world's major news agencies and maintains technical standards for improved news exchange. Its standards have been adopted in whole or part by virtually every major news organization, stock agency, and photography-related business in the world (*from: wikipedia*). IPTC metadata is information that the photographer should enter into the photo during post-production. It can be embedded into the photo, stored as a data file that travels with the photo (see below: *XMP Sidecar File*), or stored in a software application or program that was built as a repository of metadata. IPTC metadata is always editable and can contain fields such as the following:

<i>Creator/Author</i>	<i>Description</i>	<i>Product</i>	<i>Copyright Info URL</i>
<i>Creator Job Title</i>	<i>IPTC Subject Code</i>	<i>ISO Country Code</i>	<i>Copyright Notice</i>
<i>Creator Address</i>	<i>Description Writer</i>	<i>Title</i>	<i>Rights/Usage Terms</i>
<i>Creator City</i>	<i>Description Date</i>	<i>Job Identifier</i>	<i>Custom Fields</i>
<i>Creator State/Province</i>	<i>Created</i>	<i>Instructions</i>	<i>Keywords</i>
<i>Creator Country</i>	<i>Intellectual Genre</i>	<i>Provider</i>	<i>People</i>
<i>Creator Postal Code</i>	<i>IPTC Scene</i>	<i>Transmission</i>	<i>Categories</i>
<i>Creator Phone</i>	<i>Location</i>	<i>Reference</i>	<i>Events</i>
<i>Creator Email</i>	<i>Location City</i>	<i>Urgency level</i>	<i>Locations</i>
<i>Creator Website</i>	<i>Location</i>	<i>Source</i>	<i>Caption</i>
<i>Headline</i>	<i>State/Province</i>	<i>Credit</i>	<i>Annotations</i>
<i>Document Title</i>	<i>Location Country</i>	<i>Copyright Status</i>	

3d. Proprietary metadata (also XMP Sidecar Files)

This is metadata that the photographer can modify through proprietary software which can usually only be read via the software that was used to make the adjustments - although some companies are working to make their standards more universal. Virtually all photographic software has metadata fields that might *not* be able to be read by another piece of software.

Processing settings (ie: Adobe's *Camera Raw* or Phase One's *Capture One Pro* adjustments for temperature, tint, exposure, shadows, brightness, contrast, saturation, etc, etc) will probably only be able to be read by another computer with the same software.

Organizational indexes (ie: iView's *MediaPro* or Adobe's *Bridge* tools for Catalog Sets, Color Labels, Star Ratings, etc) might only be able to be read by another computer with the same software.

Although proprietary metadata has limitations, it's also necessary to use - and can be extraordinarily useful if both the photographer and client are using the same software. Although metadata can be entered directly into TIF, JPG, etc files, RAW files are unalterable because there is no industry standard for RAW files. When entering metadata for a file in it's RAW format, metadata must be stored in a separate text file that needs to travel with the RAW image file. This metadata is usually written in Adobe's XMP (Extensible Metadata Platform) and is often called a *sidecar file*. Although Adobe's XMP is open-source, not all software makes use of it with the same functionality.

Section 4: Further Information

- 4a. General
- 4b. Stock
- 4c. News/Press

4a. General Metadata Information

IPTC - Software Supporting IPTC Photo Metadata Standards
<http://www.iptc.org/photometadata/software-supportlist1.php>

Adobe - About Metadata
http://www.adobe.com/digitalimag/pdfs/phscs2ip_metadata.pdf

Imaging Society - Tagging Photos With GPS Metadata
http://www.imagingsociety.com/weblog/2006/10/gps_photo_metad_1.html

MicroSoft - Problems With Metadata in Windows Vista (02/07 – currently important, but hopefully fixed soon)
<http://support.microsoft.com/kb/927527/en-us>

UPDIG - Universal Photographic Digital Imaging Guidelines
<http://www.updig.org/guidelines/metadata.php>

4b. Stock Photography Metadata Information

Stock Artists Alliance - Metadata Manifesto
<http://www.stockartistsalliance.org/info/news/news.htm#manifesto>

Stock Asylum - Metadata Growing Up
<http://www.stockasylum.com/text-pages/articles/a6sm072006-metadata.htm>

PLUS - Picture Licensing Universal System
<http://www.useplus.com>

PACA - Image Metadata Standards for Distributors
http://www.stockphototalk.com/the_stock_photo_industry_/2006/07/paca_releases_i.html

4c. Photographic News & Press Gathering Metadata Information

IPTC - International Press Telecommunications Council
<http://www.iptc.org>

AP - Associated Press Style Guide for IPTC Metadata
<http://www.ap.org/apserver/userguide/codes.htm>

The Perils of Metadata: The Washington Post Outs a Confidential Source by Not Removing Metadata
http://fishbowl.pastiche.org/2006/02/19/the_perils_of_metadata