

THE ACADEMY OF PLUMBING 10 – GETTING ORGANISED: METADATA

‘Noun: a set of data that describes and gives information about other data,’ says the definition of metadata in my OS X Dictionary. Why is that of interest to us? Because, as explained last month, metadata is a pretty handy tool for finding, organising and sorting the data to which it refers: namely, our pictures. Also, short of burning a visible watermark into your image files, it’s the only way to let the world at large know who created and owns them. It, therefore, came as a bit of a shock to me to discover that at an Alamy contributors’ meeting an AOP member had asked the question of another attendee: ‘What’s metadata?’ It’s time to put this lack of knowledge right, because if we don’t it’s bye-bye to our cherished images and income derived from them when some kind of Orphan Works spawn-of-the-Devil finally finds its way into UK or US copyright law.

EYES, RIGHT?

Sadly not. Library search engines look for keywords, not pictures. Clients read descriptions and, one hopes, copyright information and usage licences. Potential commissioners read contact details. All of this information is metadata about your image and should be embedded in it if it is to be discovered, properly used, not stolen and you are to be paid for its usage. Last month, we looked at how iView MediaPro (to give it its 2006 moniker: <http://www.iview-multimedia.com>) could use metadata embedded by default in image files to search your image archive and find images based on a range of criteria. This month, we’re going to go one better and see what you should add to your pictures to refine searches and make your images more findable, no matter who is looking for them: a web researcher, a library search or you rifling your own picture archive.

When it comes to what to put in and where, David Riecks’ Controlled Vocabulary website (<http://www.controlledvocabulary.com>) is authoritative and handsomely repays the half-hour it takes to read what you really need to know, in particular the bare minimum if you wish to consider your images to be properly annotated. With permission, I’ve reproduced his Sample IPTC Fields table here. Note that IPTC Core stored as XMP by CS3 and other recent applications doesn’t have these character limitations, though if you intend to share images with those using older applications, they are still useful guidelines to follow. With permission, I’ve reproduced his © Sample IPTC Fields table below.

IPTC Field	Description	Example	Max Length (# characters)*
Caption	Long form description of Subject and related subject data in a Natural Language caption	Dictator Simms addresses a crowd at the first futbol match of the season.....etc	2000
Keywords	List of subject data by discrete words / phases	POLITICIAN, DICTATOR, FUTBOL, SOCCER	64 ‡
Credit	photographers name, agency name, etc	Juanita Doe/Parador Times	32
Copyright	photographers name, agency name, rights	© Juanita Doe/Parador Times, all rights reserved	128
Object Name**	similar to filename; may include unique identifier	pe10102r	64
Created Date	date image created	03/16/2003	16
City	location of image	Chicacata	32
Province State	location of image	Paranita	32
Country	location of image	Parador	64
Special Instructions	list any restrictions, embargo, rights info	MAGS OUT, INTERNET OUT, NO THIRD PARTY SALES	256
Byline	Precise text of Byline	© Juanita Doe/Parador Times	64
Category ***	Category Code (see ANPA Wire Categories below)	I = International	3
Headline	Subject data, in short form	Dictator Simms at futbol match 03/16/2003	256
Source	provider of image	Parador Times	32

Notes: *According to strict interpretation, some applications may allow for much longer entries, but you need to test to make sure additional data is not lost in other IPTC editors.

** Formerly known as the ‘slug’ (keyword) field in the old standard, but expanded in newer versions.

*** This field has been expanded from the current single alphabetic character under NAARTNDA to a maximum of three, in common use under IPTC. It is likely, however, that many news agencies will continue the current practice of using one character in the NAARTNDA implementation of the standard and three characters in the IPTC implementation. Subject to agreement in North America, it may be possible to use the second character position for the RTNDA version of category codes where they differ from NAA. Like many datasets in the new standard, it can be repeated – meaning a story can have more than one supplemental category. The American Newspaper Publishers Association (ANPA) wire category codes are available in PDF from the Newspapers Association of America site (<http://tinyurl.com/fcnj3>).

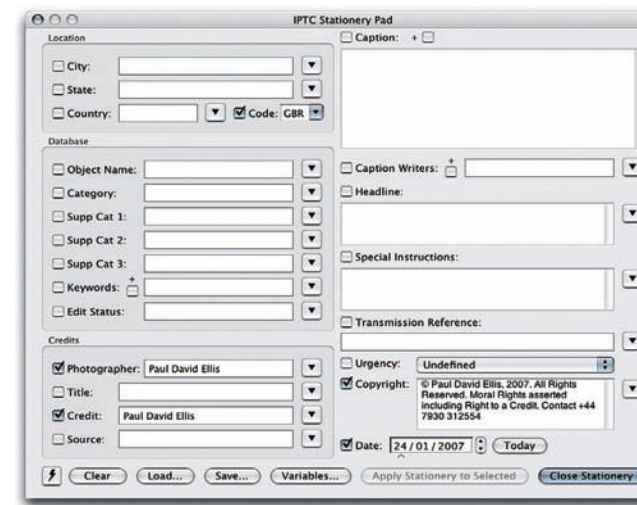
‡ The keyword field is a ‘multi-value’ field. You can have unlimited numbers of keywords and phrases but no single keyword or key phrase can exceed this 64-character maximum limit.

GET IT ON

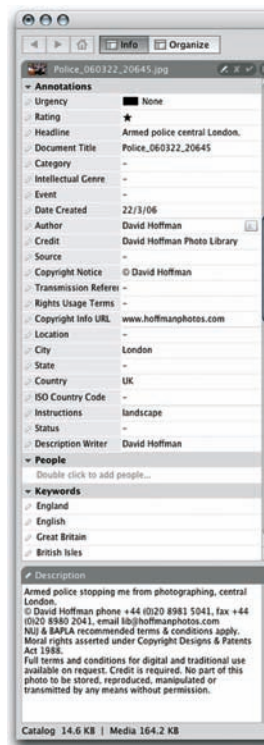
OK, how do we get this info into our pictures without turning ourselves into a typing pool? The answer is templates, which several applications support. Amongst the most useful are Apple Aperture (<http://www.apple.com/aperture/>), iView MediaPro, CameraBits’ Photo Mechanic (<http://www.camerabits.com>) and even Adobe Photoshop. David Riecks maintains a page of links to applications that support IPTC reading, editing and embedding. I’ll look primarily at iView MediaPro and Photo Mechanic because they’re the ones I use and are both cross-platform.

Photo Mechanic is your friend, not only because it’s such a fast RAW viewer, great picture editor and can copy files from your CF card to two destinations simultaneously, giving you an instant backup, but also because it features extensive libraries and templates of metadata which it can embed into your image files for you as you download them, saving you from ever having un-annotated images. Here’s Photo Mechanic’s IPTC Stationery Pad window (above right).

Every field in this window can have its information stored in a list accessible from the arrow button next to it, making it very quick

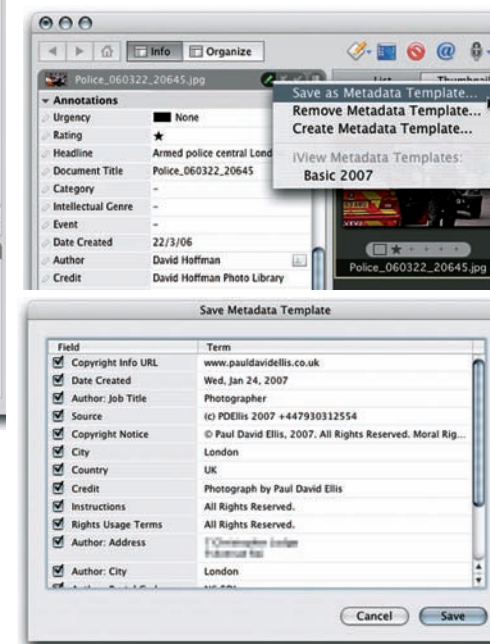


to set up the window for batch-annotating many images at once. Furthermore, once set up, the entire window can be saved as a preset. I know of photojournalists who literally cannot live without this functionality: take Photo Mechanic away from them and they’d never get their jobs submitted to their picture desks on time. Funnily enough, Photoshop CS2 attempts something similar in its File Info window, with arrow buttons allowing you to save field entries for re-use. Although Photoshop can batch-apply metadata to images by using Actions, it’s still nowhere near as quick and easy as Photo Mechanic in this regard. However, it’s definitely better than nothing and if Photoshop is all you have, make sure you make proper use of this functionality.



iView MediaPro’s approach is not dissimilar to Photo Mechanic’s but its presentation is different. All of a file’s metadata can be displayed in a single column by clicking on the Info button. Here’s how the tenacious David Hoffman has annotated an image of a policeman attempting to prevent him from photographing a street scene (left).

Much like Photo Mechanic, iView MediaPro allows you to set up preset template data, apply presets to single images or groups of images and save existing image metadata into a new preset.



All of the file types we generally use (DNG, TIFF, JPEG, PSD, PDF) support the embedding of metadata. Sadly, many RAW file formats are unlikely to allow direct embedding without creating problems immediately (corruption) or down the road (metadata not readable in your choice of RAW processor) – another good reason to move away from closed, proprietary file formats and towards DNG. Properly

Remote viewing

There are times when you really need to know what your IP addresses are. Addresses? Yes: your IP address as far as the Internet is concerned, and your address on your Local Area Network behind your router. Think of them as your main phone number and internal extension number. Why might you need to know this? Well, maybe you left your studio Mac on while you went on location and suddenly really need to get a file from it. Or, maybe you need a crafty plumbing type to take remote control of your Mac and check or fix it for you. Here’s how.

- Get a copy of BwanaDik (I know: I didn’t name it: <http://tinyurl.com/at46a>) and make sure it runs at login. It will put a cheery green globe in your menu bar, tell you both your external and internal IP addresses and, optionally, warn you or email you (or someone else of your choosing) when those addresses change. All for nothing.
- Download Lighthouse (<http://tinyurl.com/yo9hl8>), a clever little application to open the necessary control ports in your router. I recommend iGet for remote file transfer (<http://www.fivespeed.com>). Enable Remote Login in your Sharing/Services System Preference, run the Ssh profile in Lighthouse and Bob’s your uncle: fast, secure file transfer to and from your Studio Mac from anywhere in the world. If you’d like remote plumbing services, make contact and I’ll send you a Lighthouse profile to enable Apple Remote Desktop control of your Mac.

apply that little lot and your images need never enter Mr Gowers’ Orphanage. Unless, that is, someone strips out your metadata. Sadly, with the exception of PDF, no image format currently in widespread use allows you to embed and lock your metadata such that it can’t be removed or tampered with. Take a carefully annotated image, Save for Web using CS2’s default settings (which no web designer has ever bothered to alter) and kiss goodbye to all your work. This is a serious problem and cleverer, prettier heads than mine are being scratched in the process of searching for an answer to it. An obvious answer might be to gently move our clients towards accepting PDFs of our work, which we can apply DRM to, but even then, once the client opens the file with the password and then resaves as a TIFF, the metadata is still at risk. In addition, ‘locked’ PDFs can be cataloged by most image database products. Some clients might well accept PDF delivery; but for the vast majority such a proposal will merely elicit a blank look followed by a call to the next photographer on their list. Ho hum.

THE KEY TO WORDS

Of course, none of the above touches upon how to describe and keyword your images so that they will be easily found, or when they are found, will be relevant to what the searcher was looking for. Keywording is a bit of a black art which has recently become somewhat blacker thanks to Alamy’s introduction of its new AlamyRank search algorithm for ranking images in search returns. Controlled Vocabulary is, of course, a place to look for information and advice on keywording and David Riecks has made available for license a set of keywords (the Controlled Vocabulary Keyword Catalog) that works with many popular image cataloguing programs (BreezeBrowser, iView MediaPro, Image Info Toolkit, FotoStation, Stockview and more), but it strikes me that this is a subject for an article, if not a series of articles, all of its own. Any takers?

Next month we’ll get back to a bit of Mac nerdiness applied to the tedious but necessary process of keeping your admin productive and in order.