

Metadata

By Robert Edwards

Besides the actual photo, digital images also contain metadata; data about data, including EXIF and IPTC. EXIF is the information about how the image was captured (shutter speed, lens, time/date, etc). IPTC is used to enter information to describe what the image is about (the who, what, when, why, how).

If and when the Orphan Works legislation is passed by the US Congress, IPTC data may become even more important. More on this later.

EXIF

EXIF (Exchangeable Image File Format) records a lot of useful, technical information about how an image is captured. As mentioned this includes the camera settings such as camera model, shutter speed, colour temperature, color space, and also includes when the photo was taken, where it was shot (with a GPS enabled camera), etc. With the exception of date and time EXIF data is read-only.

Apart from geek appeal why is this stuff useful to professional photographers? Well, if you use multiple cameras on a job and want to quickly sequence the images later, most image browsers allow you to view by time captured (Fig 1). From there it's simply a matter of batch renaming using a number sequence.

Some applications use the EXIF to automate post-production. Noise reduction software for example can detect what camera and ISO you've used, and use the appropriate settings without needing intervention. Similarly Photoshop plugins that correct lens distortion and chromatic aberrations can use EXIF to set default parameters automatically.

If you suspect your equipment or technique is at fault it's easy to find all images taken with a particular camera/lens/aperture/color temperature, etc. With film you would have had to shoot a lot of film and take careful notes.

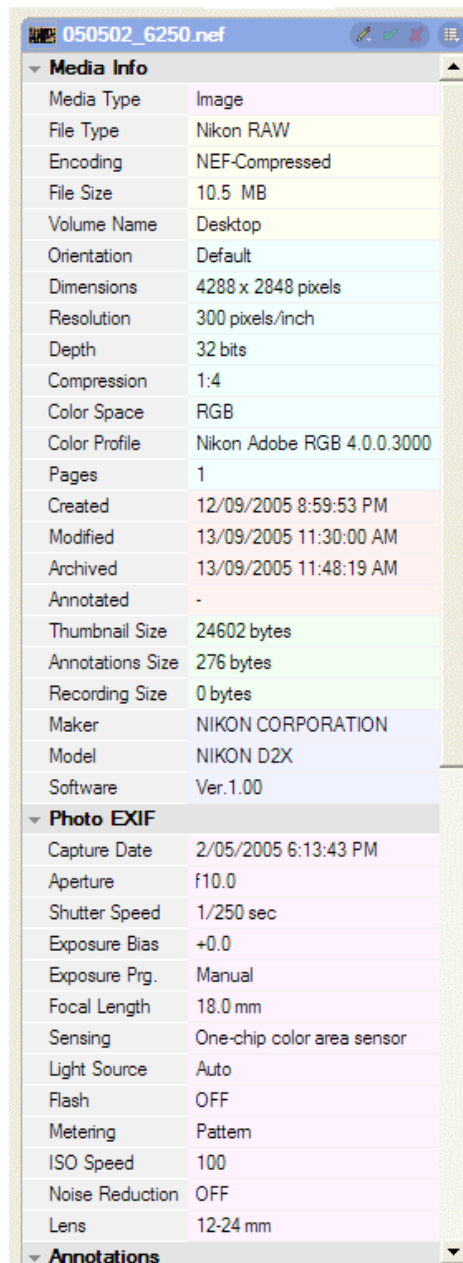


Fig 1. EXIF as displayed in Expression Media.

Some photographers may not want clients accessing all this EXIF information. They may not necessarily like others to know when an image was made, or the time taken between shots, or even the type of camera used.

Software is available that can strip EXIF data out. The Adobe Studio Exchange (www.adobe.com/studio) have a free JavaScript for Photoshop called StripExif that will remove all metadata including IPTC.

IPTC

In the mid-seventies the IPTC (International Press Trade Council) devised a set of metadata for images that were “wired” to media organizations to contain information about the subject of the photograph. IPTC fields were standardised and include Caption, Author, and Copyright. Other fields were set up to record codes relevant to news organisations such as, Country codes and Urgency.

Using IPTC photographers can store inside their photos information about copyright, their name and URL, and anything pertinent to the image. Applications such as Adobe Photoshop and the latest computer operating systems can display this information to anyone looking for it (Fig 2).

IPTC works sort of like the label we would stick on 35mm slide mounts. Unfortunately just as with other metadata, IPTC can be stripped from a photo. Then again, those slide labels could just as easily be removed.

Since the rise of digital photography in the nineties people have been using IPTC to store information not originally intended for particular fields, such as recording the usage licensed in the Special Instructions field. This prompted the IPTC update the IPTC fields to better reflect today’s requirements.

In 2004 IPTC Core was released. It offers many more useful fields aimed at areas such as rights management and specific contact details. Based on XMP it allows users to make custom IPTC fields and share them as templates in other XMP compliant software. Currently only Adobe fully supports XMP. Like anything new there are teething problems.

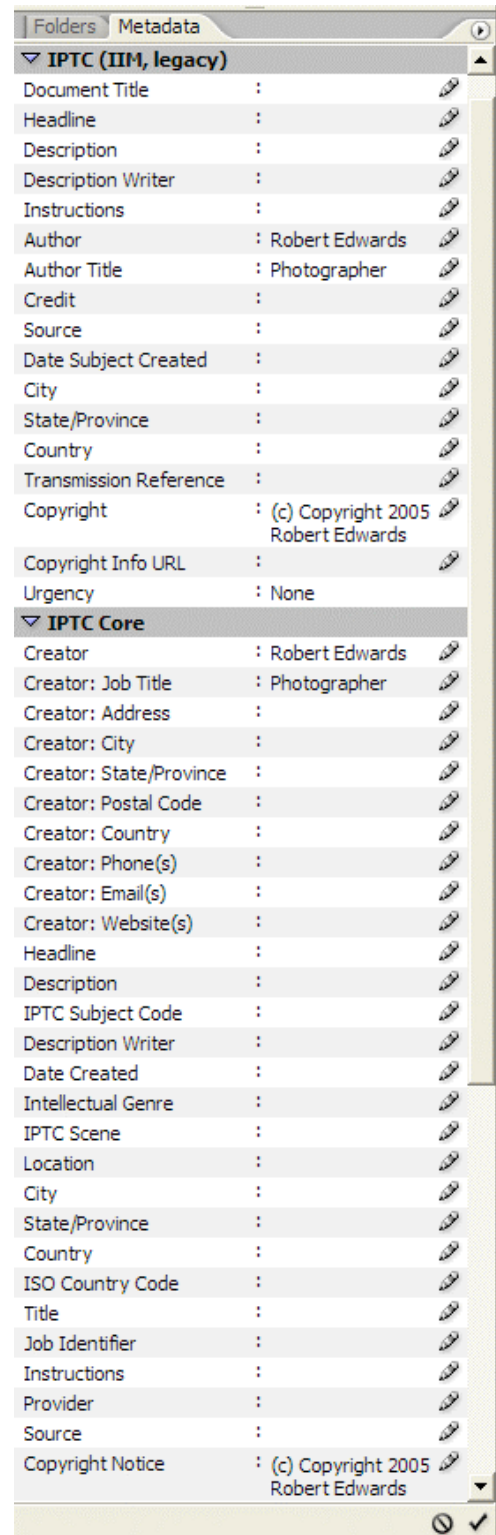


Fig 2. Adobe Bridge displaying IPTC. Note it can show both the IIM and IPTC Core fields.

Most application vendors that support IPTC do so using the older specifications that stores data in the Information Resource Block (IRB) based on Information Interchange Model (IIM). Bringing software up to speed with XMP will take time.

Issues have arisen with the change to XMP. When one application has written IPTC using IRB into a file and a second program updates that data using XMP it's possible that two sets of metadata now exist in the file, IRB and XMP.

Furthermore, older image browsers may not be able to see the updated XMP information. The question is, which information is current and which should be ignored? Remember your clients may not be using the latest software.

Until all mainstream applications use XMP it may be best to stick to either IRB and IIM, or XMP and IPTC Core. Essentially this means only caption in Adobe products, or never caption in Adobe products.

If you caption before Photoshop it will honor legacy IPTC metadata and copy it to any derivative image. Alternatively discipline yourself to only use Adobe Bridge, Lightroom or Photoshop for writing and appending IPTC.

Like any new standard time will eventually sort this out. When applications can read and write IPTC together it is often referred to as roundtrip metadata. Asset management software Extensis Portfolio 8, Microsoft Expression Media, and MediaDex can both read and write XMP, but currently only Portfolio can work with IPTC Core custom panels.

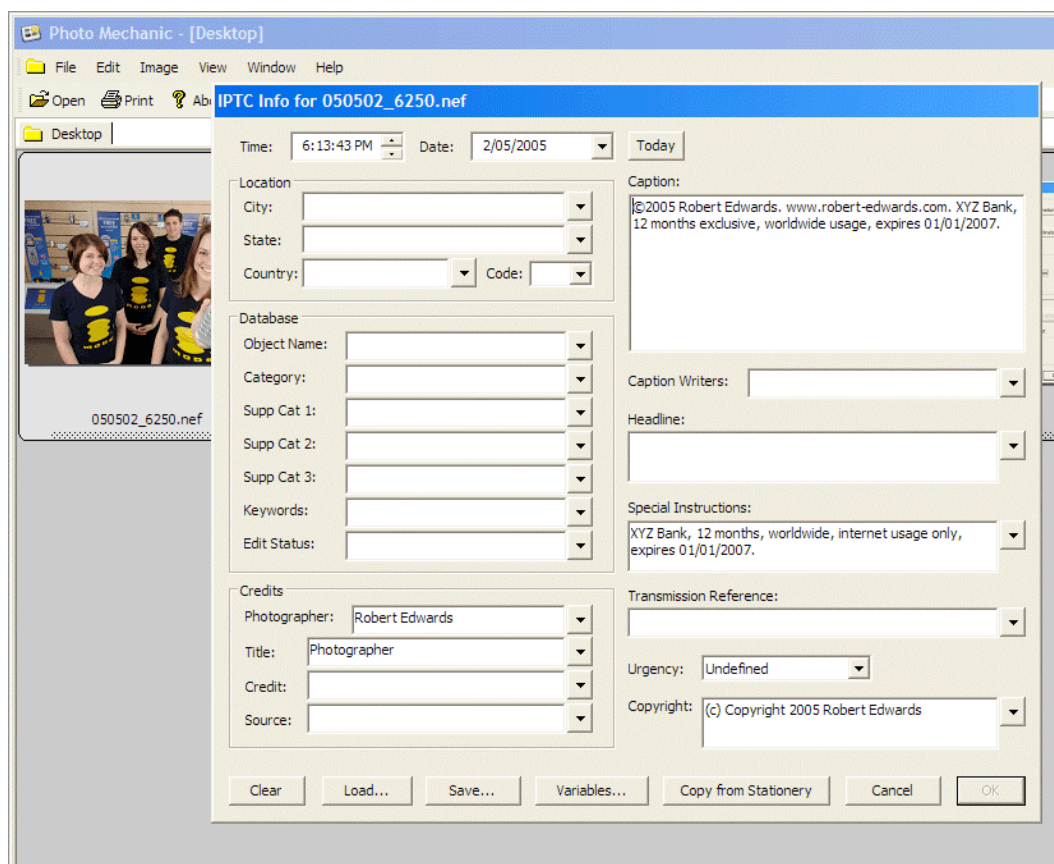


Fig 3. Photo Mechanic maintains IPTC standards and allows you to append or replace metadata.

Which IPTC fields should I use?

You can use whichever fields you like. As a minimum I would suggest:

- *Copyright:* © 2005 Robert Edwards
- *Author:* e.g. Robert Edwards
- *Author Title:* Photographer
- *Special Instructions:* XYZ Bank, 12 months, worldwide, internet usage only, expires 01/01/2007.

Most applications that write IPTC allow you to save data entered into the various fields as templates. This saves time and prevents mistakes. Your first template should be one that has all your copyright and contact details. Then it's a simple matter of adding this metadata as a batch to any number of images. Ignoring Caption and Special Instructions, Figure 3 demonstrates what a basic copyright template for Photo Mechanic can include. Using IPTC Core allows you to enter your telephone number, address, and email, and more.

If you supply photo libraries check how they use IPTC. Most have strict guidelines and a few offer custom templates and metadata panels. It's important to enter the information correctly so potential customers searching for images find your photos. These taxonomies are essential for librarians, however as a photographer you don't need to know the underlying science. Using a limited set of words, or a *controlled vocabulary* to describe images ensures standardisation. What you call a landscape to others may be a mountain, hillside, country, farm...

Orphan Works

The aforementioned Orphan Works US legislation now makes embedding IPTC essential for all professional photographers. This bill requires users of your images to only make a "good faith, reasonably diligent search" to locate you - the legal copyright owner. If they can't locate you then they are essentially free to use your work.

If you discover their use and the copyright infringer ceases using your images they are not required to pay you a cent. Even if they have profited from using your images before you notified them! Don't underestimate the power of this bill, its effect would spread beyond the USA. So start writing IPTC information into your files so people know you own the copyright and how they may contact you. Willfully removing copyright metadata is an offense under the Digital Millennium Copyright Act.

How do I write IPTC?

Some information can be added automatically (Fig 4). For example most applications that transfer your files from the CF card to the hard drive can write the general information such as copyright, author and author title at the same time. Many applications can batch caption so you need not open one image at a time.

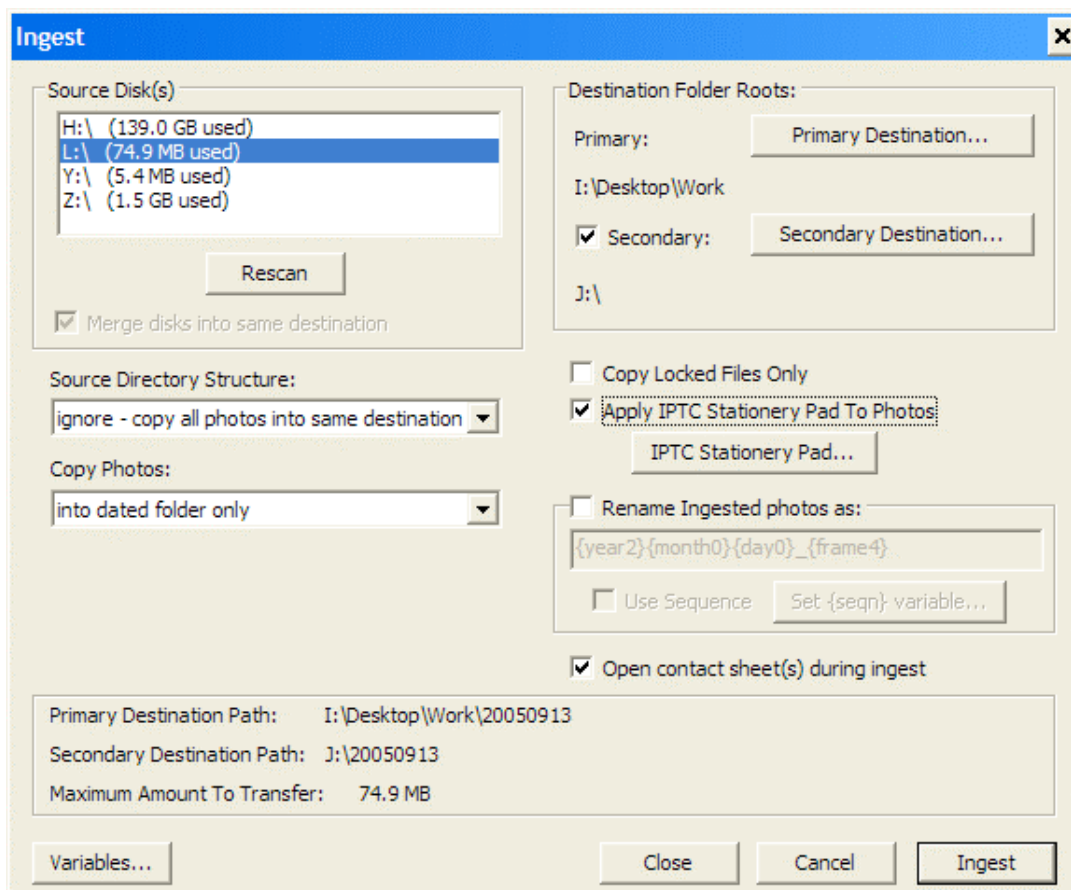


Fig 4. Photo Mechanic is shown here using the ingest feature. Once set it can automatically transfer files, back up, rename and enter IPTC information.

Now the caveats.

Some applications when appending IPTC delete any previous information. This is especially true with captions. Few apps can be set to *add* more info to a caption.

For example, say you photographed Nicole Kidman. On downloading the images to your computer it automatically adds all the relevant copyright and contact details, plus you tell it to also include in the caption; “Nicole Kidman”.

Later you want to append the Caption on some images to read “Nicole Kidman with her children, Isabella and Conor”. The easiest way is to select those images and add “ with her children, Isabella and Conor”. Some applications may not append the “Nicole Kidman” part and only write “ with her children, Isabella and Conor”.

Many applications including Photoshop don’t like carriage returns (line returns) in IPTC. So avoid using the return key when entering captions.

A more problematic issue arises for photographers using RAW files. The benefit captioning a RAW file is all conversions made from it will parse the IPTC data along, saving you having to caption the same photos multiple times.

With the exception of DNG, all RAW files are proprietary to their camera manufacturers who don’t document where to store IPTC. If it’s written to the wrong area of a RAW file it may irreversibly damage the file. Apple users are reasonably “safe” because the Mac file system allows you to safely store data like IPTC into the resource fork. While PC users can read IPTC written to the Mac resource fork it can accidentally be stripped simply by moving or emailing the image.

Many applications will write metadata to so-called sidecar files. These files sit alongside your image files and often have the same filename but different extension, eg. *040506_005.cr2* is the image and *040506_005.xmp* is the sidecar. When you write IPTC information it's stored in this sidecar leaving your image untouched.

This is good because you're not modifying undocumented files like RAW images. The downside is sidecar files are easily orphaned losing all the important information you invested time in entering. Some applications that use sidecar files to store IPTC for RAW files are Adobe Photoshop, PhaseOne C1DSLR, and Bibble.

Unfortunately none of their sidecar files compatible with one another. It also requires you to use the application that wrote the sidecar file to move, rename and delete images otherwise you risk losing them. Again it is best to enter IPTC at the beginning into RAW files, or at the end into the converted TIFF or JPEG files that are documented formats.

Digital asset management (DAM) applications can often sync metadata between files with the same name but different extension. Once you've captioned your converted images they will copy the metadata to the RAW file or vice versa. This information is stored in DAM database and doesn't modify the RAW file, unless you choose to do so.

Test, test, test.

Invest a little time to trial applications and procedures on disposable images before adopting metadata into your work environment. Experiment and decide which IPTC method best suits your workflow.

The first decision should be whether to write metadata into RAW files. This will determine when and how to write IPTC. Thoroughly test your IPTC through the whole digital workflow from download, RAW conversion, to archiving. In particular test any application you intend using to write metadata into RAW file formats.

Once established and adhered to the goal of writing metadata is save time and increase sales in the long term by making images easy to find, and protecting your rights.

Resources

www.DAMsimple.com : Digital Asset Management made simple.

www.adobe.com : Adobe Photoshop is the industry standard pixel editor. Photoshop CS3 includes a much improved version of the file browser Bridge.

www.camerabits.com : Photo Mechanic is the de facto industry standard image browser and IPTC editor.

www.controlledvocabulary.com : David Rieck's Controlled Vocabulary website is a mine of information on IPTC for photographers.

www.extensis.com : Extensis Portfolio is similar to Expression Media with the option for multi-user (network) access.

www.iptc.org : The official International Press Trade Council website for up to date information on IPTC standards including XMP4IPTC.

www.microsoft.com/Expression/ : Microsoft Expression Media (aka iView MediaPro) is a popular, user friendly digital asset management program.

www.mediadex.com : Formally the single user version of Canto Cumulus, MediaDex is another cross platform asset management application.

www.thedambook.com : The DAM Book by Peter Krogh is the authoritative source for information on digital asset management.

www.copyright.gov/orphan : Orphan Works allows people to profit from your work if they are unable to locate you.

Robert Edwards is a photographer, writer and educator based in Sydney Australia, a Board Member of the AIPP (Australian Institute of Professional Photography) and a member of the ACMP (Association of Australian Commercial and Media Photographers). He is chief author of APDIG (Australian Photographic Digital Imaging Guidelines) and a founding member and contributor to UPDIG (Universal Photographic Digital Imaging Guidelines).

© 2006 Robert Edwards