

Metadata Embedded in Image Files

iBase Professional Digital Asset Management

1 What is embedded metadata in the context of digital asset management?

Metadata is the information associated with a digital asset (i.e. an image, video, audio file etc...) for the purpose of describing it and its various attributes including descriptive, technical and administrative.

Metadata can be embedded within the digital asset file, or stored separately in an associated database, often both methods are used. This article is specifically about metadata embedded (stored) within an image file.

2 Technical outline

Metadata is embedded separately from the pixel data that makes up the actual image itself. Adobe's TIFF format set out the original method for embedding metadata in image files, which has since been adopted by others. The schema field data (IPTC-IIM and/or EXIF) is stored as blocks and this is referred to as Image Resource Block (IRB) format. Sets of IRB data can be nested together allowing multiple schemas in the same file. There can be drawbacks, however, in terms of size limits within the file header.

To get around this, XMP was introduced by Adobe in 2001. Based on XML (Extensible Mark-up Language), it is a more flexible storage method and offers more space for information.

Unlike Image Resource Block format, there are no limits on languages characters or data size because of the way in which XMP carries the information alongside its accompanying image in what's termed a 'sidecar' file. Storing metadata together with image data in this way provides complete encapsulation (akin to the glue that's used to stick it to the back of the photograph), which means both types of data can be shared and exchanged reliably as one unit. Metadata that is stored in the image file format is referred to as embedded metadata.

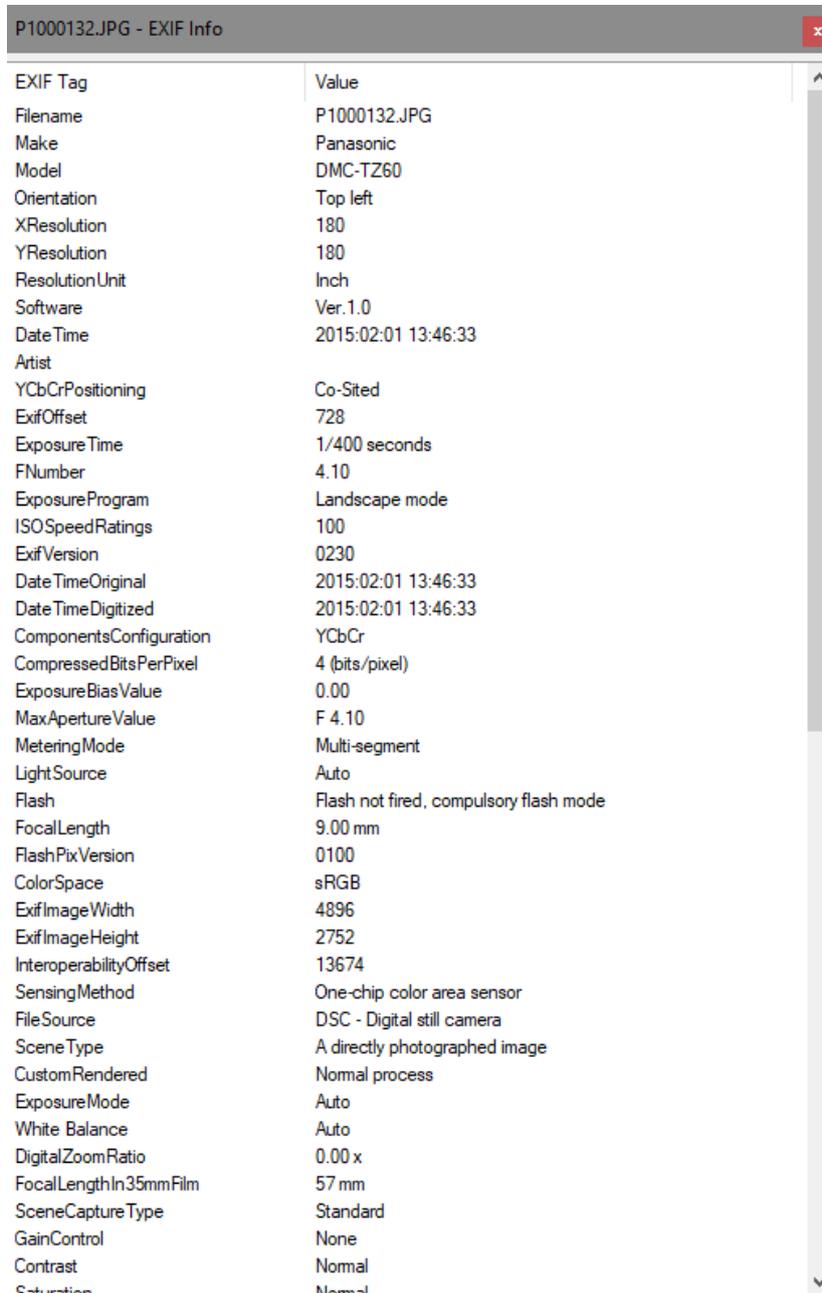
3 Embedded metadata standards

There are several embedded metadata standards in use –

- IIM (Information Interchange Model) which was created by the International Press and Telecommunications Council to aid news and media organisations when captioning and cataloguing early digital images. It was later adopted by Adobe for use in Photoshop.
- IPTC-IIM is the most widely used and includes fields identifying an image's creator and/or Rights holder, capture time, location, caption, headline, title, copyright notices and so on.
- IPTC Core and IPTC Extension build on this by including more descriptive and administrative information, a robust data format in the shape of XMP (Extensible Metadata Platform) and fields supporting the needs of stock photography and cultural heritage organisations.

- EXIF, which is both a storage format and a schema, includes technical information about an image and capture method, such as exposure settings, capture time, GPS location and model of camera (or device) that took it.

Here's a screen shot of just part of some typical EXIF data for a photograph taken with a Panasonic DNC-TZ60 camera.



EXIF Tag	Value
Filename	P1000132.JPG
Make	Panasonic
Model	DMC-TZ60
Orientation	Top left
XResolution	180
YResolution	180
ResolutionUnit	Inch
Software	Ver.1.0
DateTime	2015:02:01 13:46:33
Artist	
YCbCrPositioning	Co-Sited
ExifOffset	728
ExposureTime	1/400 seconds
FNumber	4.10
ExposureProgram	Landscape mode
ISOSpeedRatings	100
ExifVersion	0230
DateTimeOriginal	2015:02:01 13:46:33
DateTimeDigitized	2015:02:01 13:46:33
ComponentsConfiguration	YCbCr
CompressedBitsPerPixel	4 (bits/pixel)
ExposureBiasValue	0.00
MaxApertureValue	F 4.10
MeteringMode	Multi-segment
LightSource	Auto
Flash	Flash not fired, compulsory flash mode
FocalLength	9.00 mm
FlashPixVersion	0100
ColorSpace	sRGB
ExifImageWidth	4896
ExifImageHeight	2752
InteroperabilityOffset	13674
SensingMethod	One-chip color area sensor
FileSource	DSC - Digital still camera
Scene Type	A directly photographed image
CustomRendered	Normal process
ExposureMode	Auto
White Balance	Auto
DigitalZoomRatio	0.00 x
FocalLengthIn35mmFilm	57 mm
SceneCaptureType	Standard
GainControl	None
Contrast	Normal
Saturation	Normal

4 Why is embedded metadata useful?

First and foremost, most of it can be captured automatically with no effort required whatever, and it stays with the image permanently unless deliberately removed.

In the context of digital asset management and online libraries, good systems will be able to extract these data and map them to fields in the database to enable them to be searchable and displayed as

required. Similarly, it should be possible from within a digital asset management system to overwrite or create new embedded metadata.

Frequently used EXIF data includes -

- Date and time the image was created.
- Width and height in pixels.
- File type.
- Colour space.
- Place at which the image was created - from GPS latitude and longitude.
- Resolution.
- Capture device e.g. camera type.

iBase digital asset management systems can display any embedded metadata required along with curatorial metadata for an image. Similarly, curatorial metadata can be embedded back into the image.

[Contact us](#) by email or phone for more information or to request a free system.