



Metadata: a simple guide to describing pictures for use in online image libraries

About the author

Michael Wells is the founder and Managing Director of Third Light Ltd., a software company based in Cambridge, England. Third Light produce image library systems for corporate and online libraries, including search and metadata technology based around XMP.

Abstract

This paper looks at the purpose of adding metadata to images, and the evolution of keywording into XML formats for metadata. Simple models for keywording images and structuring metadata are discussed, with an emphasis on making effective use of the underlying database technology.

Introduction

Building a searchable library of images on the web is an everyday problem for photographers and picture managers in businesses of all sizes. Despite the time and cost involved in keywording collections, only an *effectively* indexed collection of images is inherently valuable.

I am often asked how to prepare images for search purposes, and it is striking how many different ideas and technologies have been brought to bear on this innocent-sounding problem.

Firstly, it is worth confirming that if you want your images to be found, you do need to index your collection and make it searchable. The idea of searching comes naturally to us all; we are adept at choosing search terms that give us specific results and we expect to use search tools, not merely browse thumbnails.

The difficulty is indexing the content in the first place. Many of us are required to perform this task with large collections of digital assets but with no prior knowledge of how to do it well. To step into the role of librarian is a step into the unknown. For example:

Which technology should be used; which is best?

How do you keyword a large number of pictures?

How can the labour involved be minimised, and re-work avoided in the future?

What are the pitfalls, and how do I avoid them?

It is understandable that trial and error creates frustrating results. You may already have some experience of this, either through someone else's web site or your own picture management tools.

In a few short steps it is easy to get down to the details. I will include liberal examples to help explain my points. Firstly, let us look at the underlying system, metadata, which allows images to be searched.

What is metadata?

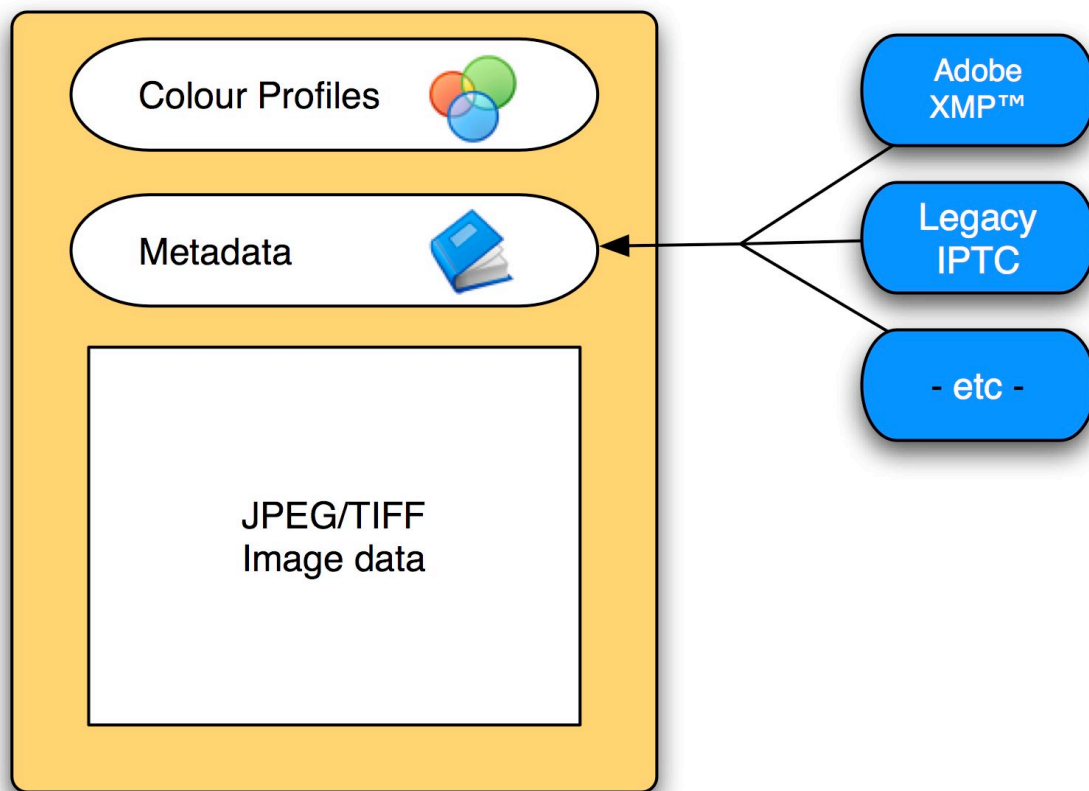
Metadata means 'data about data'. For example, a library catalog is metadata, describing books. Descriptions of images (such as keywords, captions and copyright information) are metadata, and the value of supplying this information cannot be understated. Here is why.

Some assets (like web pages) contain text that can be easily consumed by search engines, but there is no such trick that can be applied to images. The visual contents of a picture may be obvious to the human eye, but they are well and truly beyond the reach of any search engine (so far). Without accurate metadata, pictures are only ever browseable by the most basic attribute of all – the filename – and are destined to be left untouched.

It is understandable that a great deal of time and effort has gone into devising methods for describing pictures. Various technologies are still evolving around metadata for images. Most of these are based on storing a set of structured information within the file itself.

The most common approach is to embed an XML (Extensible Markup Language) header into the file. So for instance, a JPEG image can have metadata attached to it, slightly increasing the size of the file.

Typical Image File Contents



For completeness, the diagram shows that colour profiles are also stored with the image: in fact they are a form of metadata, too.

Embedding the metadata in this way is a good idea. It means that the information travels with the file itself when it is emailed or written to a CD-ROM, and cannot be inadvertently separated from the picture. Compared to proprietary solutions which store metadata separately, embedded metadata is more dependable and does not create an unnecessary legacy or vendor lock-in.

Which standards to use?

It is worth stepping aside for a moment to take a closer look at which metadata standards are available to you. Some metadata standards (like EXIF) are for storing technical data from digital cameras, and ICC profiles describe colour, but our interest lies in the descriptive metadata such as captions and keywords.

The idea of embedded, descriptive metadata is well-established, and was quickly adopted in the late 1980s by international news agencies. The original standard, now obsolete but still widely used, is referred to as IPTC, named after the organisation which devised the standard, the International Press and Telecommunications Council. Many vendors have implemented this standard, making it almost ubiquitous.

IPTC is not a standard you should adopt today in 2006 or beyond, though. It has a number of pitfalls and is inflexible if you should wish to go beyond the basics. More adaptable, modern standards exist that out-perform IPTC.

For example, in 2001 Adobe announced that its core metadata would be stored in a new format, XMP – Extensible Metadata Platform – and started to support this standard in its products. At the same time, these Adobe products started to transparently write documents with XMP embedded. This means that if you open a file with IPTC data in PhotoShop, it will save it with IPTC and XMP simultaneously. Since the new format co-exists with the old format, this has gradually allowed the industry to move on. It is five years since this process began.

As its name suggests, XMP can be extended to store more information whenever necessary, so its widespread adoption is understandable.

The main point of this discussion is to encourage you to avoid legacy tools. We can look at the process of populating metadata with consistent, worthwhile information whether you choose XMP or another modern metadata standard – the points that follow are equally applicable.

Using the database to your advantage

It would be too easy to begin keywording a collection in an ad hoc manner, only to find that as the collection grows the structure of the library is becoming confused or inconsistent. The perils of writing metadata too hastily can be underestimated. The apparent simplicity of typing in a search term also tends to conceal some of the important details about how search engines work. Knowing some of this detail is informative. Here are some important points to consider.

The database contains an index of all of your metadata, and can search this information based on patterns, words or other more innovative heuristics. A modern image database will usually:

- Ignore **upper- and lower-case**, instead treating both as equivalent,
- Offer **partial matches** (for example, searching for ‘car’ will usually return ‘cars’ but could also find ‘caramel’),
- Increasingly, offer some intelligence for matching **like-sounding words**, for instance ‘colour’ and ‘color’,
- **Prioritise precise matches**, but also provide a large number of possible matches (therefore ensuring that ‘caramel’ is scored lower than ‘car’ or ‘cars’ if you search for ‘car’),
- Offer a **simple and advanced search** facility, to balance convenience against precision according to the situation.

For comparison, let us also summarise how human beings usually work:

- We prefer typing short words, often just one keyword, into a simple box,
- We tend to search by broader concepts first (eg. “holiday”, not “madeira”),
- We work by refining our searches in stages,
- We misspell or mistype words surprisingly often, but usually get the ‘sound’ right,
- We avoid advanced search forms if we possibly can,
- We don’t think about plurals, case or ambiguity very carefully.

Between the human being and the database, your influence as the author of the metadata is vital. You can pre-empt your user if you put yourself in their shoes, and then shape your metadata so that the database can be effective.

It also becomes apparent that you can save a great deal of time by studying the behaviour of your users, and most powerful image library solutions will allow you to browse a history of actual search terms used for this reason.

Creating Captions

Captions are short, descriptive sentences which give your users an accurate statement about the picture's contents or meaning. Captions are factual and will normally be seen by the end user underneath or near to the image. A good caption contains carefully selected words, ideally remaining focused on the actual contents of the image to provide the highest quality of search results.

You may find the following approach gives you a sensible framework.

1. **Name** any people in the photograph, and state their roles. For example, "*Prime Minister Tony Blair and Chancellor Gordon Brown*".
2. If the photograph is a stock photograph, it can help your clients select appropriate images if you include the **age group and gender** of the subjects, For example, "*Male student*", or "*Retired couple*".
3. A surprising amount of detail is uncovered if you describe the **subject and its attributes**, or layout. For instance, "*Small yacht with sails against storm clouds in rough seas*".
4. **Location** information. If you consistently include the location of photographs, a geographical search will be possible. For example, "*Students graduating at the Senate House, Cambridge University*".
5. **Colours and visual attributes** can make it easier for designers to find a match. "*Large blue beach ball*" or "*Black cat with green eyes*".
6. **Timing information**. Whether this is necessary depends on the subject, but you could include reference to the season, time of day or a specific instant. Examples include "*Autumn forest*", "*The London Eye by night*" or "*September 11 2001*".
7. The meaning or **purpose of the image**; an indication of its significance. For example, "*The first Eurostar train leaving Waterloo Station, London, for Paris*".

It is possible to produce a caption which still reads clearly, while being laden with information. This is the art of caption writing. However, it is not just a matter of following rules. The purpose of the caption is to assist a search but also to describe the picture in a coherent way.

If you find that you want to add more information but cannot make the caption flow correctly, you should move onto keywords.

Creating keywords

Compared to writing captions, keywording is more logical and should be simpler to do. Keywords are usually single words, or very short phrases, which are attributes of the image. Between **five and ten keywords** is a sensible aim.

When keywording images, you can choose to adopt a **controlled vocabulary**, which is a list of pre-defined keywords that you have consistently applied (in a sense, making keywords into categories). Image library software will usually allow keywords to be hyperlinked, giving surprisingly helpful results that will surpass simple "folder browsing".

Another structure you may find helpful is a **taxonomy**, which is a tree-like structure of keywords which becomes more specific as you go down the tree. A taxonomy is often represented left to right in this form: "*Europe -> Italy -> Rome -> Colosseum*".

I have found that using either a controlled vocabulary or a taxonomy is all-but essential if you

work in a team, since the likelihood of different approaches will grow with each additional person working with you.

Here is a keywording strategy that should succeed:

- **Index by concept and by subject**, starting broad and becoming more specific. eg. “*sport, cricket, england, ashes*”. Remember, users search iteratively.
- If you can, adopt a **controlled vocabulary or taxonomy**, but be prepared to extend this list very liberally at first, based on searches that your users are actually performing.
- **Include synonyms and use plurals** where possible, since this increases the chance of a match. Remember that many plurals will catch singular terms automatically, like ‘*table*’ and ‘*tables*’, although some plurals (eg. ‘*city*’ and ‘*cities*’) will need to be specified separately.
- **Do not use prepositions** or conjunctions like *at, from, to, with*. These do not add information to your picture and will normally be discarded by databases during searches. If you only include concepts, you will save yourself time.
- **Do not add tenuous, long lists of keywords** which are for concepts that the picture does not contain. It sounds obvious, but this has become quite common practice and has caused considerable tension (it is akin to polluting a search engine). Adding vaguely related keywords is not the same as adding closely related keywords. It degrades the quality of all search results and wastes time and money.

Lastly, a common mistake when keywording is to repeat the same information that is in the caption. If you use a controlled vocabulary, it’s certainly worth picking out the appropriate words to support hyperlinks, but the database will probably have already ‘hit’ an image if the information is in the caption, so you should aim to be brief. Instead, try using keywords to include a slightly broader list of concepts than you can achieve with the caption alone.

Example images

The following images have been captioned and keyworded according to my suggested framework.



Caption: Crew members of the Royal National Lifeboats Institute (RNLI) wearing helmets and

floatation vests conduct a training exercise at sea to lift an rigid inflatable boat into the water using an orange crane fitted to the Severn Class lifeboat “Katie Hannan” in August 2004.

Keywords: RNLI, portrush, lifeboats, sea, ocean, coastguard, cranes, emergency, maritime



Caption: Close up of Fluid RI-1393 preparation being injected onto a blue 10-sample ImmuGlow slide for microscopy analysis, using a Cambridge Life Sciences Ltd model C robotic tray handler.

Keywords: bioscience, model C tray handler, RI-1393, slide, needle

This image is an example of a very specific, technical caption which is designed for the benefit of knowledgeable members of staff.



Caption: Simon Wood driving his silver Mini Cooper around the “Esses” with three other competitors, Paul Morris, Gareth Neville and Steven Illingworth in pursuit at the Mighty Minis championship, Snetterton, Norfolk on 23 May 2004

Keywords: racing, cars, motorsports, simon wood, 43, snetterton, mighty minis

Other metadata fields, and grouping metadata

When you are filling in metadata for your images, depending on your software you will probably have the opportunity to supply some other information such as creation notes, copyright, special instructions, location details and so on. Some arcane fields like the Original Transmission Reference (OTR) and – distractingly – the Category Field also appear, although this is actually a three digit code from the IPTC standard and is not as useful as it might first seem.

Here is what these additional fields are designed to achieve.

- **Filename** – the filename of the source file, eg. “ocean_23.jpg”
- **Object name** – the name of the subject item, eg “Ben Nevis”
- **Category** – Three characters identifying a category, eg. “AAA”
- **Author** – the author or photographer’s name, eg. “Joe Cornish”.
- **Author’s position** – the author or photographer’s position, eg. “Editor”
- **Credit** – Credit information for the image – not necessarily the Author, eg. “Royal Photographic Society”.
- **Source** – Identifies the original owner / creator who provided the source, eg. “National Trust Picture Library”.
- **Copyright notice** – The copyright notice to be displayed with the image, eg. “Copyright 2005 Third Light Ltd”.
- **Date created** – a date, normally in the form ‘YYYYMMDD’ , eg. 20060123

- **City** – eg. “Moscow”
- **State** – eg. “New Jersey”.
- **Country** – eg. “South Africa”.
- **Headline** – Similar to the caption (description) field, used as the title for the article, eg. “Man rescued from sea by RNLI crew”.
- **Special Instructions** – eg. “No reproduction before 1/11/2006”.
- **Transmission Reference** – a code identifying where the object was transmitted from, eg. “UK5”. This field is highly specific to agencies.

Notice the distinct emphasis on press and news agencies. At this point you may feel that enough information is included in the caption and keywords, and that the time required to complete every field is excessive (or not relevant).

The trick is to fill in the relevant data, and use a data entry tool which makes it easy for you to carry the same information into the next picture, so avoiding re-entry of the same details for every image. It might even be that every picture has the same keywords and caption in some cases. In other words metadata can be grouped; never enter unnecessary information.

Summary

It is wise to begin adding metadata with a structured approach, and I hope I have shown that being consistent will also make your life easier. Tools such as a controlled vocabulary or taxonomy can make keywording more predictable and durable when your library grows. More importantly, your forethought should improve the results your users receive and make your image collection valuable to them.

I have advocated that using a future-proof technology like XMP is essential, since this means that your metadata will not be lost as you move from one application to another.

Although keywording your images will take time, it should be time well-spent, and by using standards you can certainly avoid re-entering data in the future. If you need to justify the time and money required to start a formal library in your organisation, it is reasonable to propose that adding metadata to a library is an investment in the time and effort made by everyone who searches the library in perpetuity.

Acknowledgements

Adobe and XMP are trademarks of Adobe Inc. Photographs featured are © Third Light Ltd. and may not be reproduced without prior consent. Captions and keywords used are devised for illustration purposes only.