About Metadata

TABLE OF CONTENTS

Metadata

By simple definition, *metadata* is data about data. Metadata is structured information that explains, describes, or locates the original primary data, or that otherwise makes using the original primary data more efficient. A wide variety of industries use metadata, but for the purposes of digital imaging, there are currently only a few technical structures or *schema* that are being employed. A schema is a set of properties and their defined meanings, such as the type of value (date, size, URL, or any useful designation).

File Info

Prior to Adobe® Photoshop® CS software, the primary image metadata schema was File Info, which is nonimage data embedded in image files. Originally deployed by the International Press and Telecommunications Council (IPTC), this metadata contains only a few fields of limited text used to help organize and distribute photographic images for newspaper publishing. File Info, which is the Photoshop implementation of the IPTC specification plus additional data fields, defines both the storage format and the actual metadata. Text fields in the current specification include but are not limited to Caption, Caption Writer, Headline, Special Instructions, Keywords, Category, Supplemental Categories, Urgency, Byline, Byline Title, Credit, Source, Object Name, Date Created, City, Province-State, Country Name, Original Transmission Reference, and Preserve Additional Information. Mark As Copyrighted and URL are additional fields beyond the IPTC specification.

The IPTC fields allow only a scant amount of text information to be embedded within the file, but it is critical that this information moves through the newspaper publishing system locked together with the file. This embedded information is critical for the thousands of digital images used in major newspaper companies. Surprisingly, a large number of Photoshop users don't realize that these metadata fields already exist. Photographers routinely send out digital images without even marking them as Copyrighted or embedding simple ownership or contact information.

EXIF

Several years ago, a group of major photographic manufacturers, in conjunction with the Japan Electronic and Information Industries Association (JEITA), developed a metadata schema called *Exchangeable Image File Format* (EXIF). This schema is designed to embed in a digital capture—right at the moment of exposure—certain information relating to the camera's function, image parameters, and miscellaneous additional information.

Unfortunately, the various schemas currently used in image file metadata do not yet meet the explicit needs and desires of imaging artists and photographers. For EXIF or any additional schema to reach broad adoption and acceptance, the creators of digital images need to be brought into the process of designing and specifying what metadata can be used for.

Adobe

XMP metadata framework

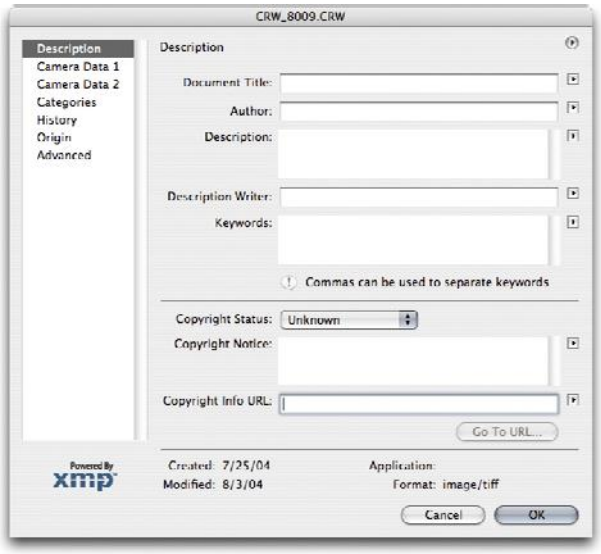
At Seybold San Francisco in the fall of 2001, Adobe Systems Incorporated announced a new technology initiative called *Extensible Metadata Platform* (XMP), a common metadata framework that standardizes the creation, processing, and interchange of document metadata across publishing workflows. Adobe has also committed to the World Wide Web Consortium (W3C) standards, including Extensible Markup Language (XML). Thus, XMP is XML compatible. For the future of metadata and its usefulness to photographers and imaging artists, XMP holds the promise of providing a significant and extensive platform for a wide array of various metadata implementations. Both public and private metadata schemas can use a widely adopted *framework schema* that will enable a rich and extensive use of metadata for critical technologies, such as Digital Rights Management (DRM) and Digital Asset Management (DAM).

The Creative Suite update to Photoshop, Adobe Illustrator®, and Adobe InDesign® as well as Adobe Acrobat® 6.0 has achieved broad support for XMP across most Adobe applications. Photoshop CS in particular offers an extensive deployment of XMP.

Present and future possibilities of metadata

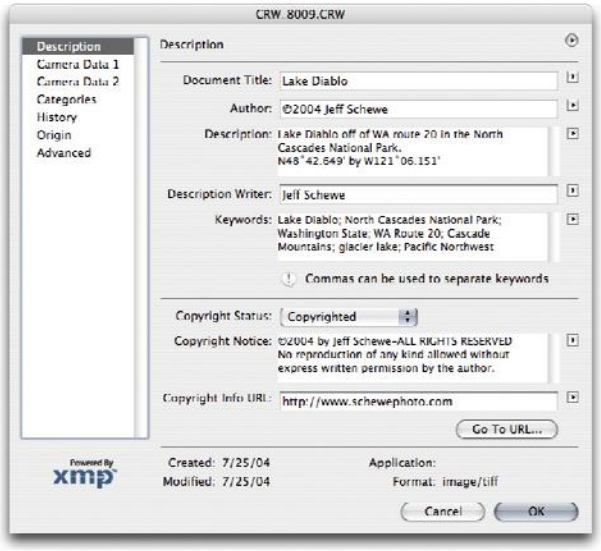
Because of the lack of coordination between manufacturers and organizing bodies, much of the potential of metadata remains unrecognized. Following are examples of metadata and how it can help photographers, both now and in the future.

You can use the File Info dialog box to view and append or replace metadata in image files.



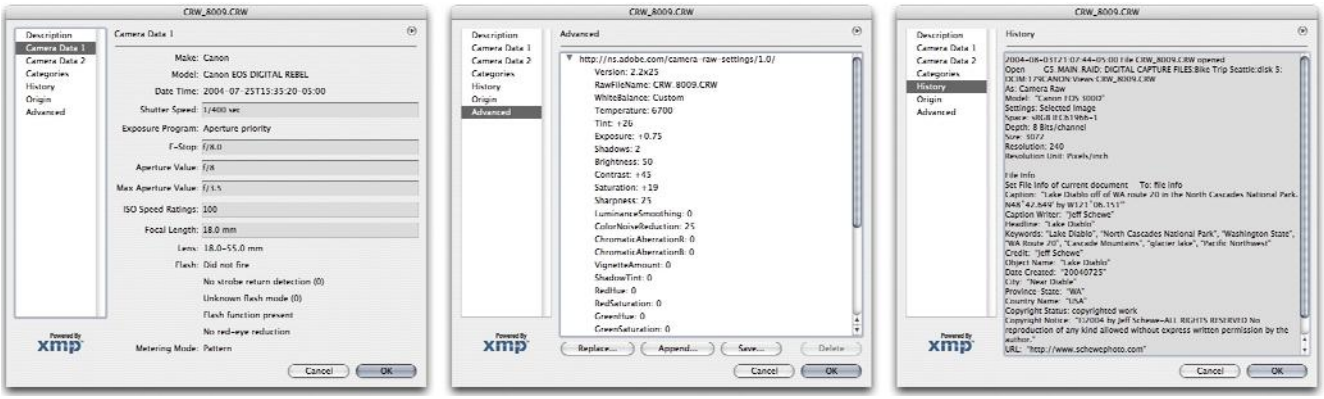
The basic File Info dialog box in Photoshop CS

Embedded metadata can be in XMP form and is also aliased to other schemas for backward compatibility. Some of the information, such as the IPTC metadata, is editable; other information is read-only. The creator of an image file should adopt a policy regarding whether subsequent users should replace or simply append or add to existing metadata. For example, it would not be appropriate for users who didn't create the image file to remove or replace the Author information or to take an image marked Copyrighted and change it to Unknown.



The creator of the image file should place his or her own information in the file. Users downstream should respect this metadata.

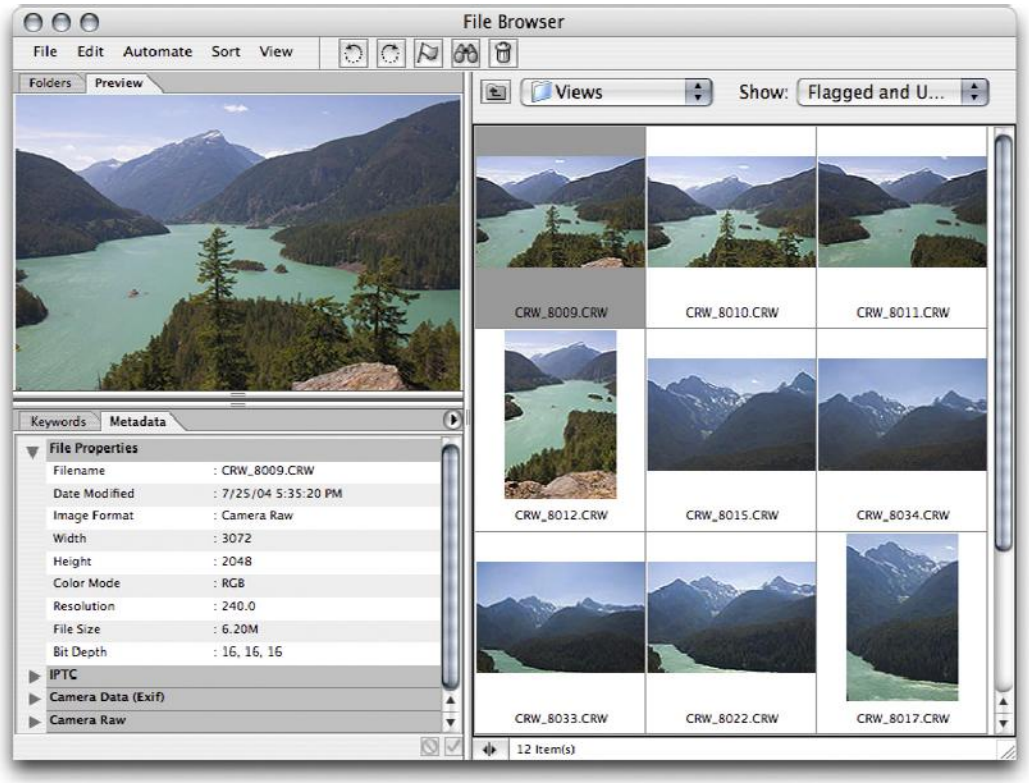
You can view a variety of metadata for any file in the File Info dialog box, including IPTC, EXIF, Camera Raw settings (for raw files), and more. You can even view the saved edit history of an image if you use the history tracking feature of Photoshop CS software, and the history is embedded in the file as metadata.



EXIF fields (left) Camera Raw settings (center) File History (right)

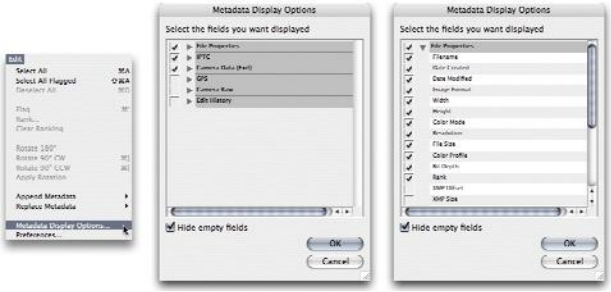
Some metadata is known as *public data* and some of it is referred to as *private data*. While there are standards for all public metadata fields, the private fields are largely undocumented. That is one of the problems associated with some of the private maker notes embedded in raw digital camera files.

Embedding metadata image by image isn't particularly productive. While IPTC metadata can be embedded in Actions and Batching operations from Photoshop, it's useful and more productive to use the File Browser in Photoshop CS to add metadata to multiple files at once.

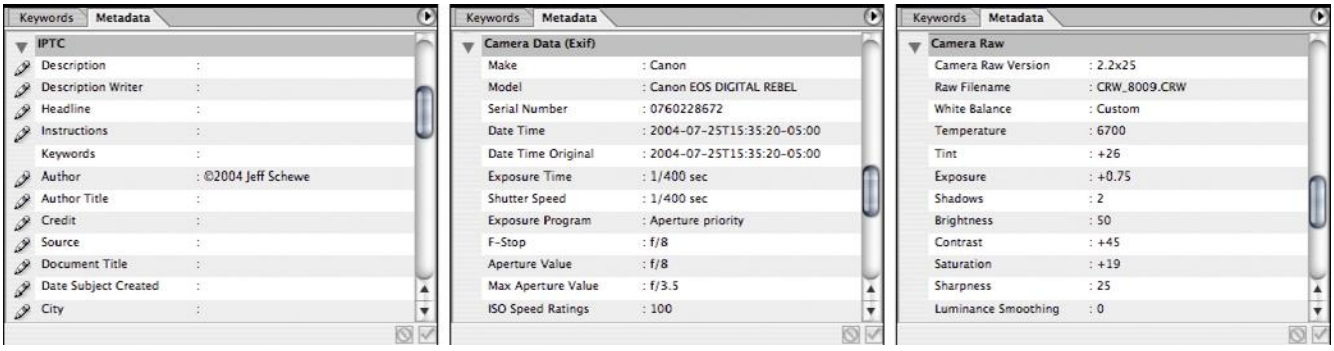


Selected image in the File Browser

In the File Browser, there are movable panes that contain the fields of metadata that you have chosen to show. You can control the visible fields by selecting them from the Metadata Display Options in the Edit menu of the File Browser.



Selecting fields in the Metadata Display Options

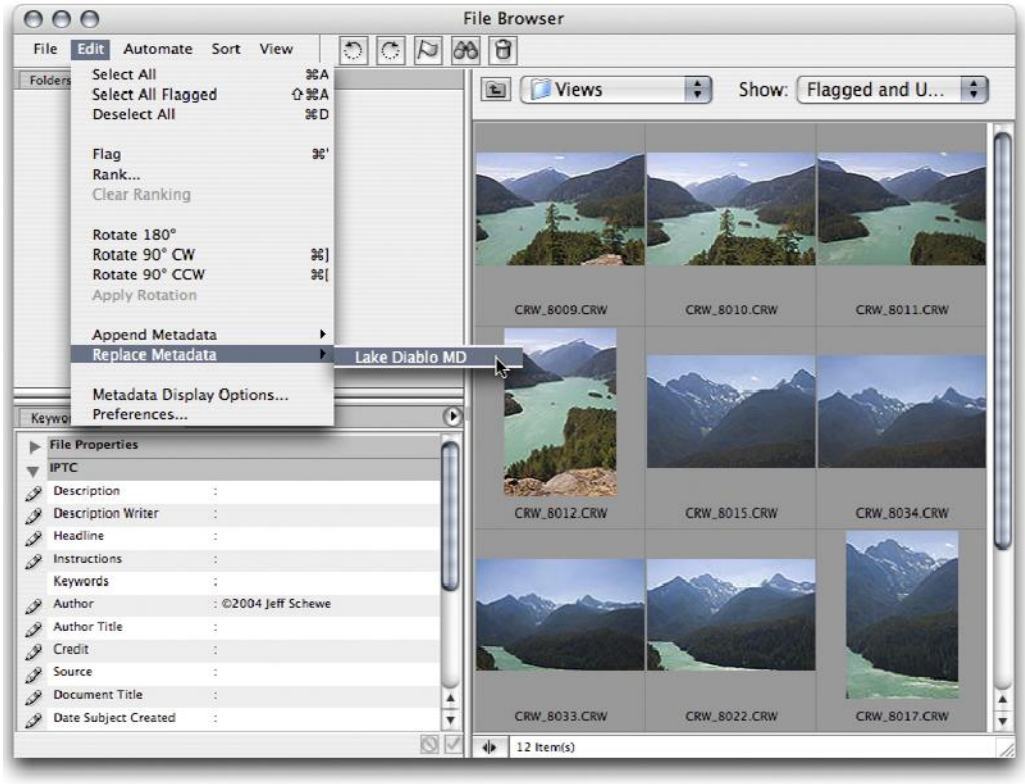


Selected fields appear as a scrollable list in the Metadata tab of the File Browser.

You can edit metadata directly in the fields that allow editing (such as IPTC) either image by image or in selected groups. However, it's also productive if you create and save custom metadata templates.

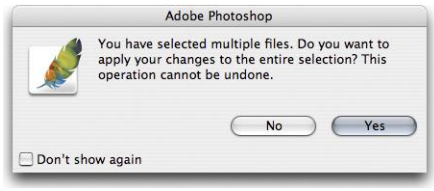
Saving a custom metadata template

To create a custom template, you need to have an image open and to display the File Info dialog box. Enter in the fields that can be applied to a wide variety of the images you'll be working on, such as Copyright Status, Author, and Copyright Notice. After you enter the information, but before closing the dialog box, choose Save Metadata Template from the pop-up menu. You can name the template in the Save dialog box that appears. The template will then be saved and will appear in the pop-up menu of the File Info dialog box and also in the File Browser.



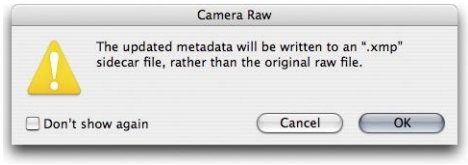
A saved template appears in the Edit menu of the File Browser.

After you select the template, you'll be alerted that you are applying metadata to multiple selected files. You can choose Don't Show Again if this alert bothers you.



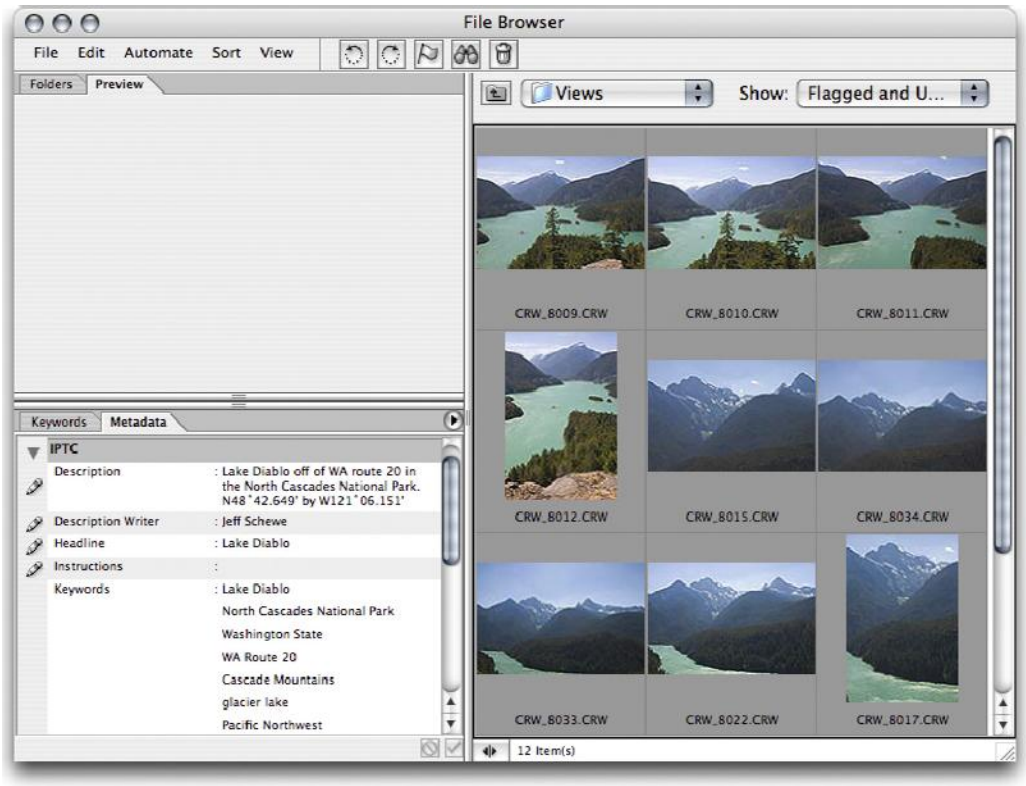
Alert message about selecting multiple files

If you are adding metadata to raw files, note that the data will be stored only in separate sidecar files in the same folder as the file or in the Camera Raw database, depending on your preferences.



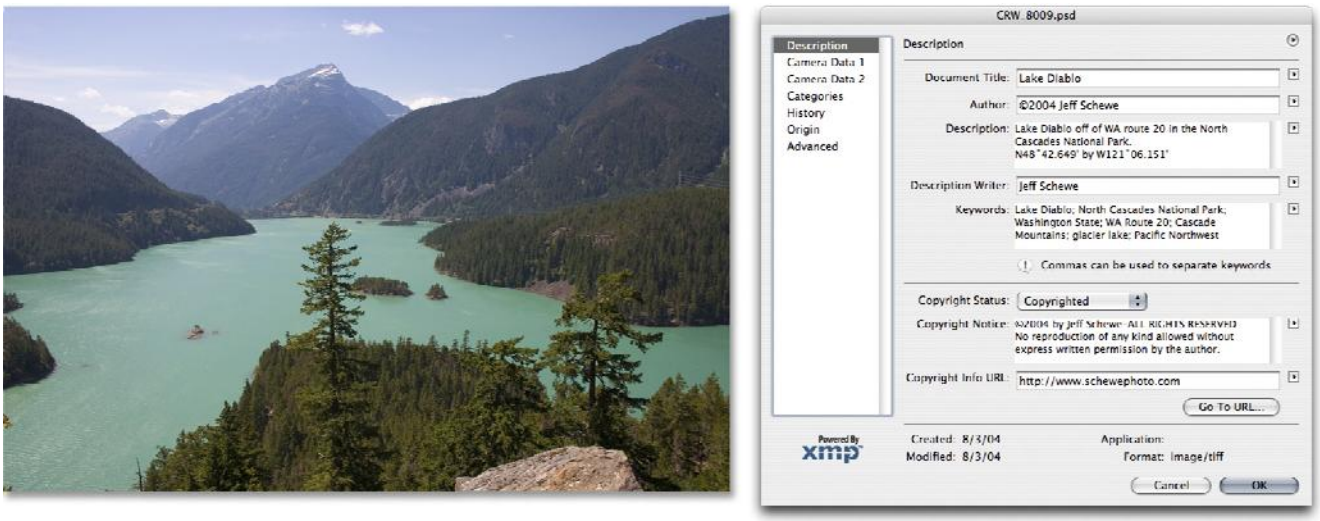
Alert message that updated metadata will be stored in a sidecar file.

Once you actually process the raw file, Camera Raw and Photoshop CS embed the sidecar file metadata directly in the processed file when it is saved. Note also that all public metadata fields will also be in the file, including EXIF, Camera Raw settings, and image history (if you selected that option).



Embedded sidecar file metadata

After you apply the template, all of the images have updated metadata from the template. You can still edit individual fields on an image-by-image basis as well as entering in keywords.



This processed raw file has had the common metadata entered and saved as part of the raw processing.

Metadata has phenomenal potential for the photography industry. For example, suppose two photographers are on assignment and shoot from the identical position with identical lighting. For all practical purposes, the images would be identical, and you would be unable to differentiate who shot which image. By using metadata, you would be able to see the serial number of the camera body and the owner's copyright, providing conclusive proof of ownership.

The ability to apply keywords by using extensive metadata currently exists. The ability to embed keywords means that placing an image into a stock library could eliminate or reduce the large amount of time currently spent on applying keywords to images. The current cost of placing a film-based image into a stock library and duplicating it can be as high as $140.00 per image. Reducing the amount of time that an agency spends applying keywords, along with solving other image identification problems, could drastically reduce the actual cost of placement within a system. Adding rich keywords and additional information actually makes an image more valuable.

Photographers and imaging artists can take advantage of metadata for many aspects of their imaging businesses. They can incorporate licensing information for an image, along with the password security for opening the file, into a metadata schema. Imagine being able to embed a unique image identification that will travel wherever the image goes—Photoshop CS does that right now. Users could generate and complete invoices and transactions by using metadata. Users could employ metadata for tracking billable time while image processing. Users could also track exactly what was done to an image file, and when.

Additional information resources about metadata

Adobe XMP
www.adobe.com/products/xmp/main.html

Dublin Core Metadata Initiative
www.dublincore.org

Japan Electronic and Information Industries Association (EXIF)
www.jeita.or.jp/

International Press Telecommunications Council (IPTC)
www.iptc.org

World Wide Web Consortium (XML)
www.w3.org

Jeff Schewe
Jeff, a summa cum laude graduate of Rochester Institute of Technology, has been an advertising photographer in Chicago for over 25 years. He has been doing digital imaging for over 14 years and is widely known and respected in the digital imaging community as a leading pioneer in the field. Jeff is a feature consultant and alpha tester for Adobe Photoshop.

Adobe Systems Incorporated • 345 Park Avenue, San Jose, CA 95110-2704 USA • *www.adobe.com*

Adobe