

Car History, Archives and Digital Technology

Bob Schmitt

(This article was first submitted in December 2013 to "Beaded Wheels", the magazine of the Vintage Car Club of New Zealand. It was revised and resubmitted in January 2016 and updated in April 2016)

This article describes methods that individual car owners, a branch or club can use to identify and organize their photos and documents, especially in digital formats. A brief review of database (and similar) software used for organization is included, but the primary focus is creating personal or branch/club archives with the free, open-source Greenstone Digital Library software, developed at the University of Waikato.

Car Owners and Car Club Members

Why should the owner of a classic/collector car, or a car club care about the digital world? The short answer is: "It's all around us!" A more profound answer is that the "digital world" provides many tools to preserve, promote and explain history. Even if a particular photo or document never "becomes digital", knowing the process to go from "analog to digital" should encourage all in the car hobby to look at "our archives" and take a few steps towards better identification and organization of photos and documents.

Mystery Photo(s)

Frequently we read in a magazine (or Internet) article of a "mystery car", sometime with people in a partly familiar setting. The photo may have a caption or story: "I found this in my late Uncle Fred's garage. Can anyone identify the car and the family standing next to it?" And don't we all think, "Why didn't Uncle Fred pencil something on the back of the photo?"

Recommendation #1: Identify your photos, slides and negatives! If not now, when? If not you, who will do it?

Each of us now has (or soon will) thousands, or tens of thousands digital photos. If these are not identified, we have an immense problem that will only worsen in the future. Digital photos already have two obvious items of identification: a file date stamp and file name. Putting photos in dated or named folders is a good step. Changing the photo's file name to provide descriptive data has pros and cons – it can better identify the photo, but the original file identity can be lost.

Metadata to the Rescue!

Many photo organizing and editing programs provide tools to put captions and keywords (tags) on digital photos. PDF, Word and Excel documents can be better identified in the "file properties" sections of those editing programs. But what exactly is a "caption" or "keyword" on a digital photo? These are "embedded metadata" stored inside the photo (or document) in hundreds of categories. A digital camera creates a great amount of

“embedded metadata” in each photo. Much of this is well-known to professional photographers who use Adobe products (Creative Suite, Lightroom, Photoshop) and other software products to add copyright and other copyright management information to their photos.

However anyone can use the free Picasa program, the low-cost Photoshop Elements or other programs to add captions, tags and location data to digital images. This is a great first step and further explained at <http://CarLibrary.org/CarLibrary-Basic.htm>

Another very good free program to identify photos is the ExifTool. To make this tool more user friendly, Bogdan Hrastnik has written a GUI (Graphical User Interface, Windows only) for the ExifTool, which allows very easy access to many of the ExifTool functions. The ExifToolGUI can identify hundreds of photos in a single step by embedding metadata of many types.

There’s a introduction to functions and complexity of the ExifTool and the GUI on the <http://CarLibrary.org/CarLibrary-Metadata.htm> webpage. There is much more about the program on Phil Harvey’s (the author) website:

<http://www.sno.phy.queensu.ca/~phil/exiftool/>

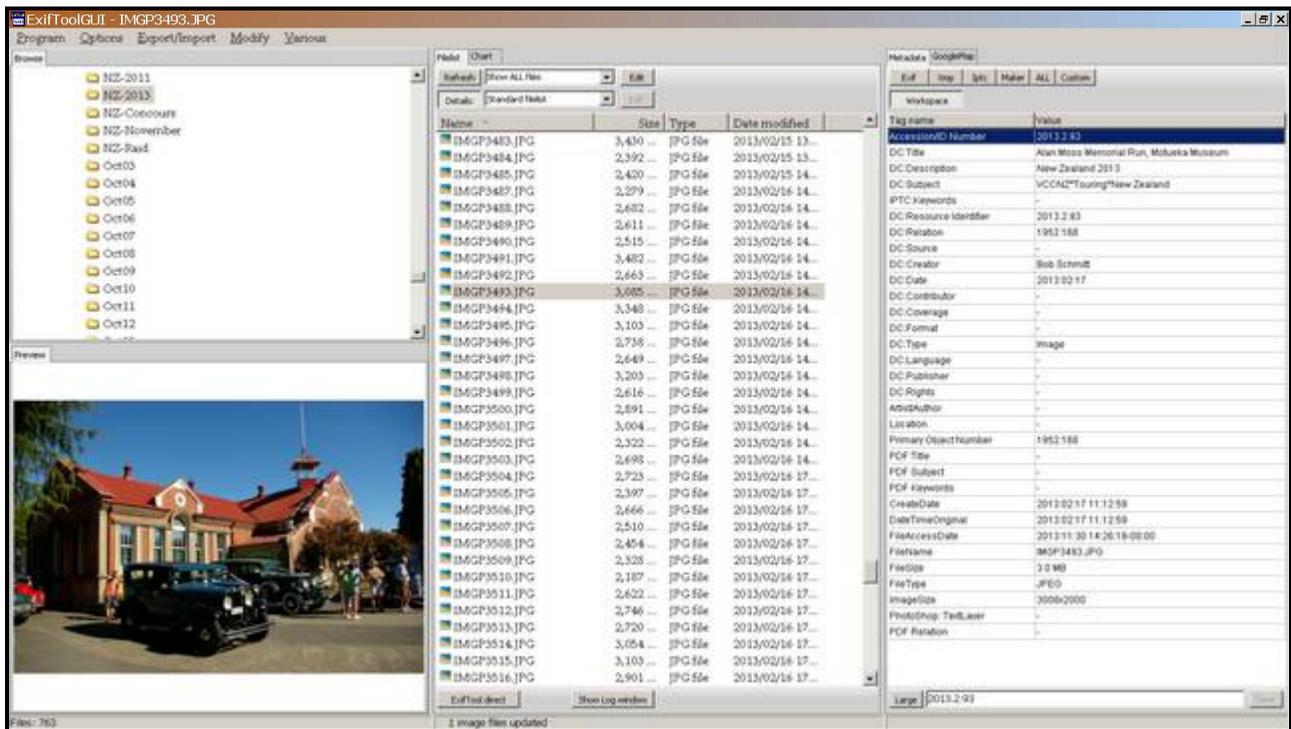


Figure 1 –Adding metadata with the ExifToolGUI to a photo of a Nelson NZ branch event in Motueka

Recommendation #2: Use Picasa, the ExifTool or other software to put captions, keywords and location data on your digital photos.

Digital Photos, Embedded Metadata and Lists

Are we better off with photos that now have captions and keywords? Yes, because the ExifTool can also generate an Excel-ready comma-delimited file of entire folders (or hard-drives!) of photos showing all, or selected, metadata in those photos such as file name, file size, file type, original photo date, subject, title, etc. All can all be displayed and printed. The ExifTool can also do the reverse: an Excel/comma-delimited file can be used to write data into hundreds of photos in a single step. The commands to do this are tricky, but the results are great. See: <http://CarLibrary.org/CarLibrary-Metadata.htm#WritingMetadata>

A Greenstone Archive

The Greenstone Digital Library software now comes into the picture. Greenstone's standard method to make a collection/archive is to "Gather" (import) the photos or documents into the program and then add "external metadata" to each object with the program's "Enrich" function. This is not difficult, but can be tedious for archives with hundreds or thousands of objects.

However, there are two other methods of importing hundreds of objects and classifying them in a few steps in Greenstone. The first is an "import" function that uses an Excel/comma-delimited file which contains the photo/object identifying information – which we described creating in the previous step – and "explode" it into individual records, corresponding to the individual photos.

A second, and perhaps more efficient method, is to use the embedded metadata directly. Photos or documents that have embedded metadata, such as "title", "subject", or keywords, can use this data directly in Greenstone for very complete classification, searching and display.

Car Club Archives

Although I have no experience with Club branch operations, I suspect most are like US and UK car clubs – a few members manage membership and car lists, probably using Excel files or self-developed databases. This type of software works well for clerical operations, but is not suitable to create a "branch archive" or "digital library". Greenstone is designed to do this, and can be used in a highly distributed, democratic fashion.

It would not be overly complex to setup a Branch archive with the Greenstone software. Initially, the program can be installed on a single PC (or Mac) and the "built-in" Greenstone metadata categories enhanced with car-specific ones, such as "car.make", "car.model", "car.year", "car.owner", "car.event", "car.event-date" and "car.event-type". Any Branch photos or documents can be used initially to start the archive. After some experience, standard "car" categories for embedded metadata can be set, notices sent to the Branch members and photos solicited. Whether delivered by email or on a CD, the members' photos can be easily imported into the nascent Branch archive and a Digital Archive/Library is formed.

Greenstone allows separate collections to be joined ad hoc for searching and display, so it's feasible to keep member data ("car.owner" and similar details) separate from other car and photo data. Greenstone supports password protection of specific documents or an entire collection/archive.

The Branch Archive can be shared with members on the Internet (password protected if so desired) or put on a CD, DVD, USB drive for distribution.

Advanced: The Greenstone software can also be installed on a "Linux Live" USB drive, 8 GB or larger. This process is described at <http://carlibrary.org/CarLibrary-PortableGreenstone.htm> and results in a full version of both Linux and Greenstone that can be plugged into any PC. The USB drive can be easily cloned for use by Branch members to create their own archives or as a starter kit for a training seminar.

One caveat: Greenstone is a very capable system to make and manage a digital library/collection/archive. However, its ability to export data is very limited. It cannot generate reports or lists, which are strong features of database software. Perhaps future versions will add these functions, so it's unlikely Greenstone will replace any software now being used by a club or branch to make newsletters, contact members or schedule events.

Another Software Type

Except for export, Greenstone has a full range of functions that perform very well, but true "collections management" is not one of those. Collections Management Software (CMS) is software category with more than 30 commercial systems and a few open-source programs. This category is summarized at <http://carlibrary.org/CarLibrary-CollectionManagementSystems.htm>. On that page is a link to a table of the system's and vendors: <http://carlibrary.org/CarLibrary-CMS-Table.htm>

Auckland's Vernon Systems has provided a high-end CMS since 1985 to customers worldwide. More recently, Vernon's eHive software is targeted to small museums and collections as an online service. Small collections up to 200 images are free and collections up to 2,000 images are \$99 annually. See <http://ehive.com/> for more details.

PastPerfect is a low-cost commercial system used worldwide by thousands of museums and collections. I have created three car marque and museum prototype collections with it and have received excellent support even as a "mere volunteer". My experiences, which include two videos, are at: <http://carlibrary.org/CarLibrary-PastPerfect.htm>

Further Archive Recommendations

Recommendation #3: Traditional photographs - Scan at least at 300 dpi; some archives recommend 400-600 dpi. Use Picasa or other software, as described for digital photos above, to add captions, tags, geo-tags to each image.

Recommendation #4: Slides and Negatives- Scanning a slide or negative directly will give better results than scanning a photo if the scan is made at 2800-4000 dpi (for the common 35 mm negative/slide format). Then identify the digital image as above.

Recommendation #5: Documents - If text recognition or text searching is not a concern, scan the documents as images, described above. Otherwise, scan at 300 dpi or higher and process the document with optical character recognition (OCR) software such as Adobe Acrobat, ABBYY FineReader or similar. Add identifying information with the OCR software or use the ExifToolGUI to later add metadata to the PDF file.

Background - How Did I Get Here?

More than three years ago, after visiting several car museums and private collections, I was curious how the collection owners and museum managers keep track of the collection's cars, documents, photos, and books. I looked at my modest "collection" - a few cars, Vespa scooters, books, magazines and photos, slides, and negatives. And thousands of digital photos! After discussions with other collectors and Internet research, I learned that ledgers, file cabinets, Excel spreadsheets, Access and SQL databases are all being used for organization.

I reviewed and updated a rudimentary Access database project made years ago using sample car club records - both members and cars - to see how well it would connect the cars with all their owners, photos, documents and events (tours, shows and races). (See sample Access table, Figure 2) It was encouraging, but it seemed the best that could be expected was a tidy "inventory" of all. Not too bad, but what else was possible?

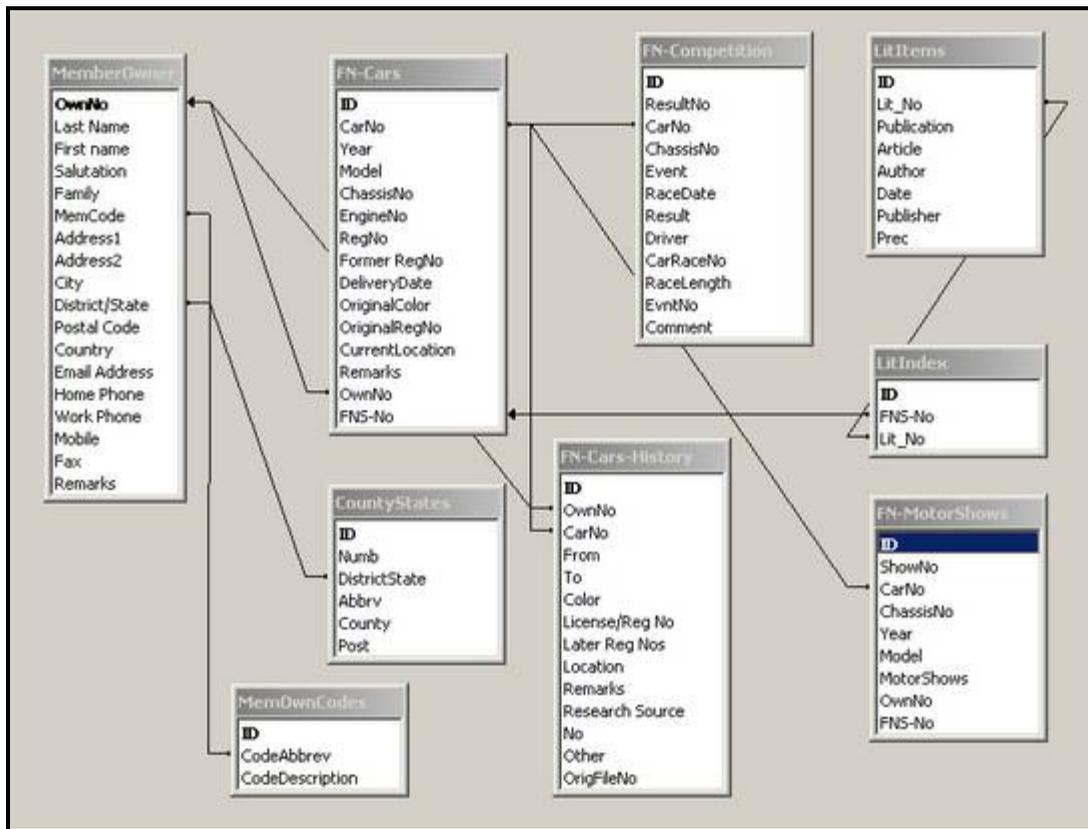


Figure 2 – Access “Car” database tables showing relations.

I reviewed “content management systems” (such as Joomla and Drupal), and other current methods to manage digital content but then chanced upon the open-source Greenstone Digital Library software. It was developed in New Zealand and my car was mostly restored in New Zealand – the stars aligned and pointed me to further investigation.

Starting the Project

The Greenstone software, the workshops and tutorials were downloaded from the Greenstone.org website. In a short time I had made a “collection” (archive) of about 200 personal car photos and scanned documents. I created another collection, using photos taken at the Petersen Automotive Museum (Los Angeles), classifying all by exhibit theme, car make, model, year, etc. Finally, I made a collection of 50 newsletters from a Southern California car club, distributed in PDF format. The all of the classifications were straightforward, but Greenstone also indexes every word in such PDF text documents and searching could be done by any name or word.

All this worked efficiently on my older Windows PC, but another Greenstone feature was also impressive – collections can be exported to a CD (or DVD) which is “self-contained” and runs without need of other software on any Windows PC. I made CDs of the sample collections to share with the car clubs and other collectors, hoping to spark a move towards digital car archives.

