

## Some thoughts on vendor lock-in, from the domestic to the institutional (is Apple Mac OS X evil?)

I spent a fair bit of time during a period of enforced physical inactivity in March sorting out the home music and picture collections, getting rid of stray MP3s that nobody wants and trying to work out how to start organizing our photos a bit better. This is another [exercise in self preservation, like the ongoing but currently stalled work I'm doing on my theses](#).

We have a couple of Macs around the house now, a new PowerBook that I recently bought, and sometimes the work laptops (a brand new Intel MacBook Pro and an old G4 PowerBook). The patched-together home PC now runs Linux (Ubuntu 7.10) exclusively, no Windows, but we do have a couple of Windows XP licenses if we really need them and XP is installed on the work PowerBook so I can test stuff in Windows when I absolutely can't find anyone else to do it.

So lets talk about **managing digital pictures**, and how we might start to organize them and add metadata. How can we label stuff so it's findable and sortable, and remains so years and years from now?

The basic requirement is that you can add metadata to images, in the form of tags and captions, and have them stored in the image in a standard way so that future software can use the tags and captions.

You need to be smart about this. I could tag an image "Peter Sefton" but does that mean that I took the picture, or I'm in it or I own the copyright? There are a couple of approaches to sorting this out:

1. Use tags which include a predicate and object, what the Flickr folks call [Machine Tagging](#), AKA Triple Tagging; if I took the picture I could tag it `dc:creator=Peter Sefton` or if I'm in it I might tag it `foaf:Person=Peter Sefton` or maybe `foaf:Person=pt@ptsefton.com`. If you're wondering what the DC and FOAF bits mean then you can [read up on it](#).
2. Use hierarchical tags. So if I'm in a picture it might be tagged `People/PeterSefton` or `People/Sefton/Peter` or suchlike. This kind of hierarchical tagging is built in to some tools. If you do this then it should be possible to [map these tags onto a more formal metadata system](#) later on, given the right software. (Note: don't use spaces in tags, while some tools support them not all do.)

Anyway, provided you can capture one of the above kinds of tag and make sure they stay with your images then that's a good start. So I looked for software which would:

1. Respect your [EXIF metadata](#) which is embedded in pictures that come from modern digital cameras, and not mess it up.
2. Write captions and tags into the EXIF metadata.
3. Deal with the [IPTC](#) standard at least for captions and tags.
4. Understand Adobe's [XMP](#) which is better than plain IPTC even if all you're doing is tagging because it deals with more than just ASCII characters.

I've oversimplified here I know – feel free to elaborate in the comments.

So what are the options available on the Mac?

One of the first impressions I got when I started using a Mac a couple of years ago is just how bad the operating system level support for images is. In the Finder you can look at thumbnails and that's about it.

Even on Windows XP there's a pretty useful built in photo viewer that lets you flip through a directory full of pictures, use keys to rotate them and print them and so on. Not so on the Mac. Even with the new CoverFlow feature in OS 10.5 you can look, but you can't touch. That is, you can't rotate an image or do anything useful from the Finder. And you can't mess with picture metadata in the *Get Info* dialogue either. There is basically no operating system level support for images.

Why is the Mac like this when it's supposed to be so great for multimedia?

I think it's to do with Apple's plans to lock you in to their hardware/software platform. You want photo management? Well, every new Mac comes with iPhoto. Shut up and use it. Better still fork out hundreds of dollars for their up-market photo software.

So what if I *were* to simply let iPhoto import all my images?

Two very bad things would happen.

1. By default on OS 10.5, *Leopard*, images all **disappear into a black hole** called *iPhoto Library* in the Pictures folder. It looks like this:



Click on *iPhoto Library* and iPhoto opens. But where are the files? Have they been eaten by iPhoto?

If you right-click you can Show Package Contents and see the files in a complicated directory structure that keeps track of originals and changes you have applied.

But good luck knowing which file to click if you want to grab one and do something with it without using iPhoto. And who knows what would happen if you changed one of the files in the 'package'.

2. iPhoto will steal my hard-earned metadata. There's a nifty looking tagging interface, but the tags live in the iPhoto database, wherever that is, not in the images themselves. So if you invest in tagging up the photos then you'd better be prepared to stick with iPhoto for the rest of our life.

From what I can tell there used to be scripts that could round trip metadata from the iPhoto database to your images themselves, and the other way round, but iPhoto '08 broke that and [I gather that Apple have taken away the scripting hooks that made it work](#). Oops, sorry valued customers.

OK, so there's a '[rudimentary](#)' export that will put the tags into your images into the IPTC metadata, but if you do that then iPhoto will just spew them all out into a single directory losing any directory structure you might have devised.

These two points make me think that this is a very **deliberate, cynical, contemptuous move by Apple**. Unsuspecting Mac users are being treated like crayfish and being ever so gently boiled alive. By the time you wake up it will be 2015 and Apple will own your entire digital

life and you'll be sliced in half and served to Steve Jobs for supper. I am beginning to wonder whether the Finder is such a useless file manager because Apple want to actually move as many of your files as possible into 'packages' so you have to use Apple software to manage them. Just wondering.

Mark Pilgrim wrote in 2006 why he was leaving the Mac:

I'm creating things now that I want to be able to read, hear, watch, search, and filter 50 years from now. Despite all their emphasis on content creators, Apple has made it clear that they do not share this goal. Openness is not a cargo cult. Some get it, some don't. Apple doesn't.

<http://diveintomark.org/archives/2006/06/02/when-the-bough-breaks>

Me, I think Apple get it. They know what's open. They can read blogs and forums and know that users want iPhoto to write metadata like tags back to their images. I mean who wouldn't if you explained it to them? But if Apple can get away with not doing it then they will. This is the story of Microsoft Word's file-format lockin all over again. Ten years from now we can enjoy the fireworks when Apple's media database standards go through ISO, OOXML style. Tim Bray [will write an essay](#).

Take iTunes for another example. It does everything it can to tie you to the iTunes store, and it wants to organize your files for you. It actually does write metadata changes back into the music files, unlike iPhoto, but for such a slick application it has some surprising dumb spots, such as the way it can't detect changes to your music library if you add or change files using another tool. If you delete a bunch of songs from underneath iTunes it gets awfully confused and if you add new ones it won't notice until you point it back at its own library and tell it to 'import' a directory into itself.

Why? I think it's just designed to wear you out so you give in and let iTunes manage your stuff. I wonder how long it will be before the iTunes library is an impenetrable package like the iPhoto library.

To summarize the state of just these two media offerings from Apple, there are really, really obvious features that are in the customer's interest that aren't there. And there's no excuse.

For me, the Mac's still a worthwhile platform, and I can even live with iTunes, but I'm not letting iPhoto boil me alive however tasty the stock I get boiled in.

The really funny part is that Microsoft's much maligned Windows Vista [sounds like it has really good, really open built in image support](#). I don't have the constitution to go down that particular rabbit hole and invite yet another version of Windows into my home, but no doubt we'll have a look at work and see how it all works.

Outside of installing Vista, and without spending money on pro tools that may have their own lock in, looks like the solution is [Digikam](#). It's open source and [Mark Pilgrim likes it](#) so it must be good. And it plays nice with metadata.

Digikam can be installed on OS X to run under the X Windows system, if you're very persistent and learn how to compile and use the [Fink](#) package manager. A native version should be available [soon](#). I also got Digikam running in a virtual machine using Virtualbox and Ubuntu. That worked but it was a bit sluggish.

Plan B is [Mapivi](#). It's far from pretty but it seems to be a serviceable photo manager. I have not bothered to get it running on OS X yet. (In testing this I worked out how to make tags that will interoperate nicely between Digikam and Mapivi, I'll post a how-to at some stage.)

Plan C is Google's non-free but very slick [Picasa](#). Runs on Linux and is rumored to be coming to OS X this year. I have managed to get the old Windows 98 compatible version to run on the Mac using Wine, but it doesn't do tagging the right way.

I'm talking about the home situation here, but lots of this applies to institutional contexts too; it may be that the repository you're using, or about to be using suffers from some of the same problems as consumer-level software.

Is your repository keeping some of your important data or metadata in its own database and not giving lossless export? Maybe researchers in your university are having trouble looking after their data. Some of the same kinds of commercial forces I describe below may be at work to subvert your interests.

What does this mean for the institutional context?

Two main things:

1. We need to make sure that repository software vendors/developers are not pulling a Steve Jobs on us. If you enter metadata into your repository to make a collection is it written back to your storage layer in a way that you can re-use it in future? Likewise for access control information, and any other 'feature' that seems like you really must have it.

I won't name names, but repository vendors and open source developers you know who you are and you know *we are watching you*.

2. For the forthcoming Australian National Data Service we turn our attention to what researchers do with all their digital stuff. A lot of that stuff is photos that they're pulling off digital cameras – how should they look after it all? How to add metadata?

I [mentioned](#) the [Field Helper application](#) here before, from the University of Sydney Archaeological Computing Laboratory. But I'm beginning to wonder if we might not be able to go a long way by tying together existing tools, like Digikam or Picasa, to help out our researchers. Lots more work is needed in this space, but if you're using iPhoto to look after important research data then I suggest you seek help before it's too late. Maybe a switch to Windows Vista, which [cares](#) about your metadata would be in order?