

Adventures in geocoding part 1: The Toowoomba BUG Cycle Hazard Investigation Team does Ruthven Street

One of the things we're thinking about at [USQ](#) is how researchers might integrate data into their publications. This will be key to the Australian National Data Service.

I have posted a few things here before about stuff like embedding chemistry, maps, and graphing. I'll start [keeping track](#) of those and other posts under the del.icio.us tag [DataIntegrationForDocuments](#) for want of a better tag.

(You know, I don't bother with WordPress categories, I use del.icio.us as an outboard organizer for my site and have one less thing to worry about when I migrate it. Hmm – I wonder if my automated backup of my del.icio.us tags is still working).

I think one of the big kinds of data integration we need to work on is getting geo-spatial data hooked in to documents. This is potentially useful for lots of disciplines and there are lots of tools out there waiting to be mashed up.

But as usual there are issues. Just like the [stuff I talked about recently regarding metadata for images and potential vendor lockin](#) we need to be very careful about how we encode our data for the long term and use the sexy online services wisely and within their licenses.

Licenses are really important; you can't just grab online maps and use them however you like, if you want preservation-quality maps it's a whole different issue than just throwing up a Google map. More on that in the future.

I wanted to link to an example of some great mashing up of data and maps from the Bidwern project at ANU, but today [the map page](#) there is returning an error about keys for Google maps. That helps make my point that we need to think about storing important data independently of any particular service that might go away at any time. I'm sure that in the case of Bidwern the data are safe, but researchers, don't just go and 'build a service where the data you need are all Google dependent and stored only on Google's servers, OK?

Over a [series of posts](#) I will look at some stuff I'm learning about how to use documents and pictures with geographical information embedded in them, and think about how we should use services like Google Maps without locking ourselves in. On one level this is about something I did on the weekend, but it's also relevant to the kinds of things our researchers are wanting to and will be wanting to do. Just be thankful I didn't plot all my veggie plants on a map and bore you with that.

This time, I report on a rather exciting expedition I went on with the [Toowoomba Bicycle Users Group](#) Hazard Inspection Team – aka the TBUG HIT Squad. The squad last Sunday was Hugh, David and me. Not sure what Hugh was doing there though 'cos he rides a [trike](#); I'll see what I can do about getting him expelled or at least compelled to turn up on a proper conveyance. Really, those [recumbent riders](#) should have got the message [back in 1934](#).

We took a slow ride from the Southern edge of town to the Oxygen cafe, looking at Ruthven street with new eyes. I have ridden parts of that hundreds of times in *just deal with it* mode, but it's an eye-opener to look in detail at all the hazards.

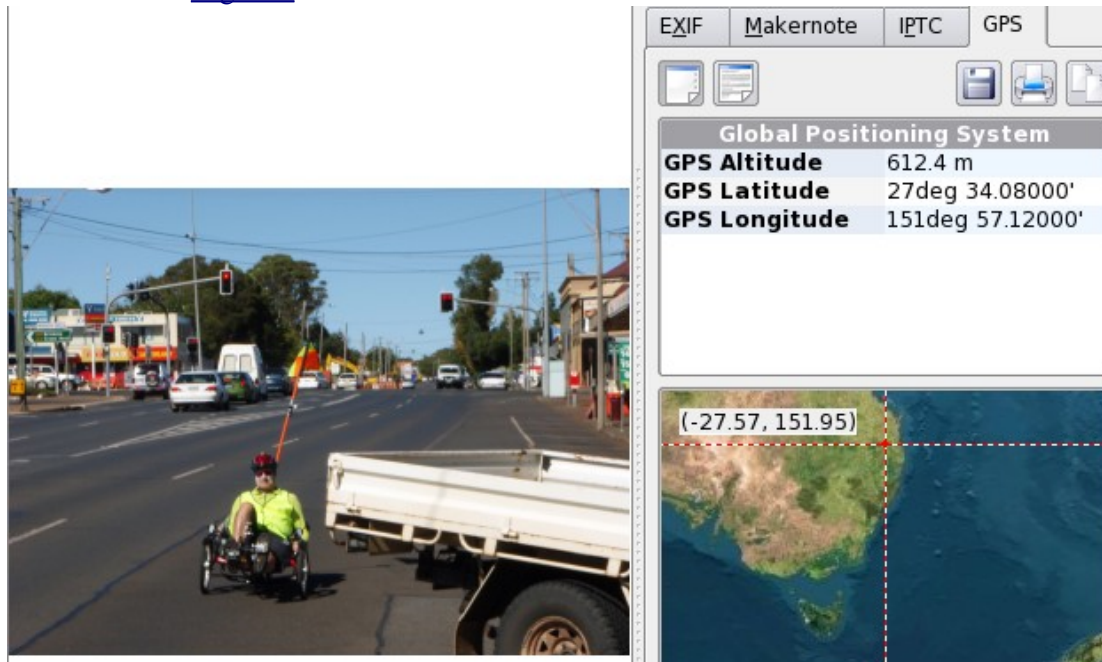
I strapped-on a GPS (a three year old antique Garmin eTrex) and a compact digital camera.

Hugh kept notes and both David and I took pictures.

I have been able to synchronize the tracklog from my GPS with the pictures, by getting the clock in both cameras set to within a minute or so of the spot-on GPS. We'll see more and more software and hardware that can make this a smooth process but for now I had to use a couple of commandline tools to get the job done. I will leave boring you with the how-to stuff for another post.

Anyway – the result is some very arty pictures complete with embedded metadata.

The view from [Digikam](#) is this:



There's Hugh at 612.5 metres above sea level. We reckoned he could have gone under that tray-back but he took his life in his hands and rode around.

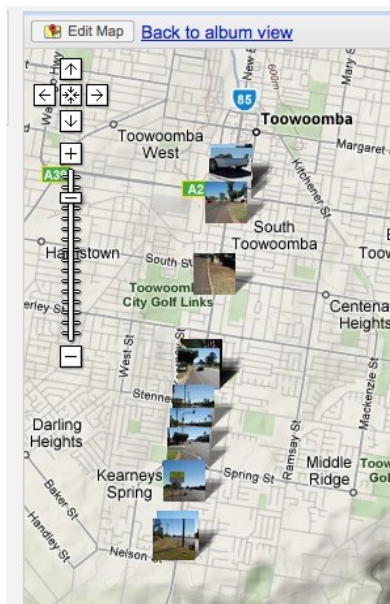
Now that I had my geotagged piccies I uploaded them to Flickr. but guess what? While Flickr recognized the tags I had embedded. (What/ToowoombaBUG/HazardInspectionTeam) it didn't like the geo-metadata. It wants the metadata expressed as tags in its own Flickr way.

Google's Picasa site, on the other hand recognizes the geographical data nicely as long as you don't let its uploader software scale down the images. But guess what? While it groks my mistyped captions Google's service doesn't recognize the tags.

It's the usual story. You have to figure out what works and what doesn't. Mistakes with valuable data could be very upsetting so we need people helping our research communities with this stuff.

By the way, this is an interesting kind of narrative genre. I can type parts of a story into the metadata of the pictures and the computer can sort them for me using the time stamp, to assemble the story but they could be served up other ways so the captions should try to be clear enough to understand on their own. I know I need practice at that.

Here's a screenshot of [the result](#):



The pictures have metadata in them that should stick with them forever. My insightful comments and skillful photography will no doubt delight road geeks of the future. Google's Picasa service might go away or I might take them down, but I'm reasonably happy that I have future-proof-ish data that will work with other services in the future. At the very least I can get a paper map or some old fashioned navigational instruments and a clock that knows what time it is in Greenwich and use the latitude and longitude.

When we get ICE version 2 out the door at the end of June we can consider how we might make ICE more aware of geotagged images, and potentially other data, but before that I will post on another quick experiment I did that finds points in web pages and shows them on a map.