tools & data

Stewart C Russell - @scruss

#maptime TO
2015-11-18

happy GIS Day!

“be here now”

structure

1. Care & Feeding of Your Data
2. The Right Tools
3. A Tiny Little Example

unless you collect it or commission it,
data is in the wrong format

be kind to your
metadata

the undead Shapefile

```
$ ls -1
Danube_GSGS_Grid.*
Danube_GSGS_Grid.cpg
Danube_GSGS_Grid.dbf
Danube_GSGS_Grid.prj
Danube_GSGS_Grid.qpj
Danube_GSGS_Grid.shp
Danube_GSGS_Grid.shx
```

- **it's a file made up of lots of files**
- **based on a mid-80s database format**
- **too many limitations to name**
- ***will! not! die!***

the colossal GeoTIFF

- **extends the already baroque TIFF image format**
- **sometimes has aux files**
- **easily broken with graphics tools**
- **likely the biggest files you'll ever see**

check your licences

can I even publish this?

will it blend?

- **not all open data has compatible licences**
- **even Ontario municipalities aren't compatible**
- **spend all of your project budget on IP lawyers**

heid nip 1: map datum

Are we at

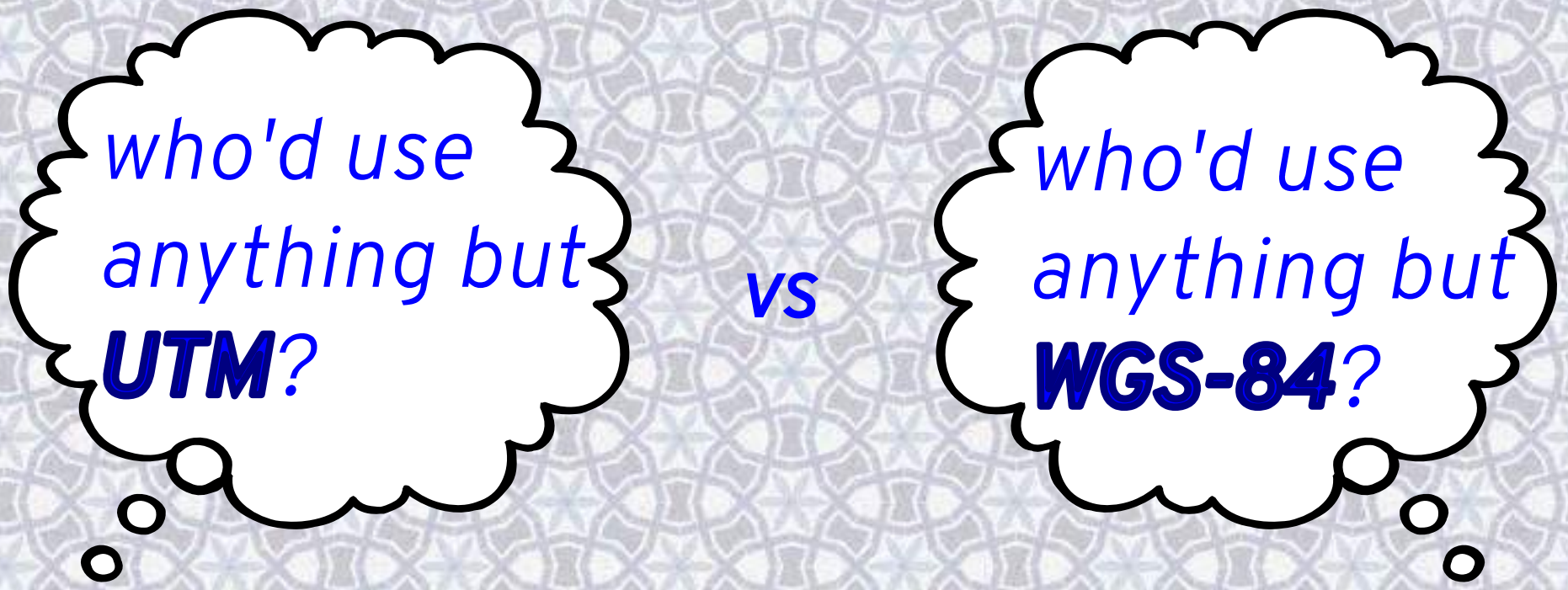
UTM 17T 629490 4833953

or

UTM 17T 629477 4833731 ?

⇒ **“yes …”**

heid nip 2: projection



heid nip 3: charset

I â™¥ UTF-8

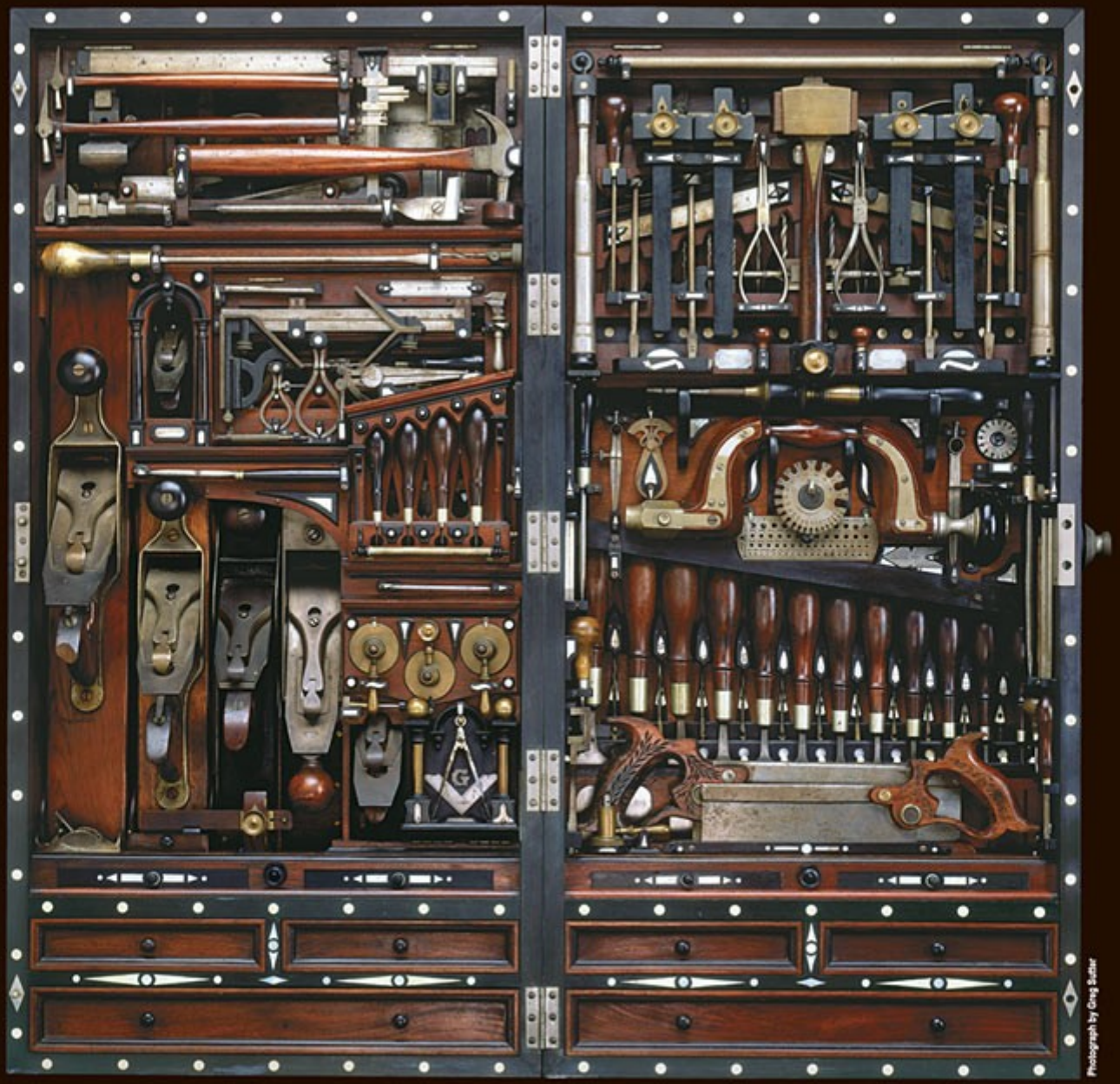
vs

I ♥ UTF-8

be kind to everyone

(not just metadata)

If you have wrangled a difficult public data set and won, *please* write up how you did it

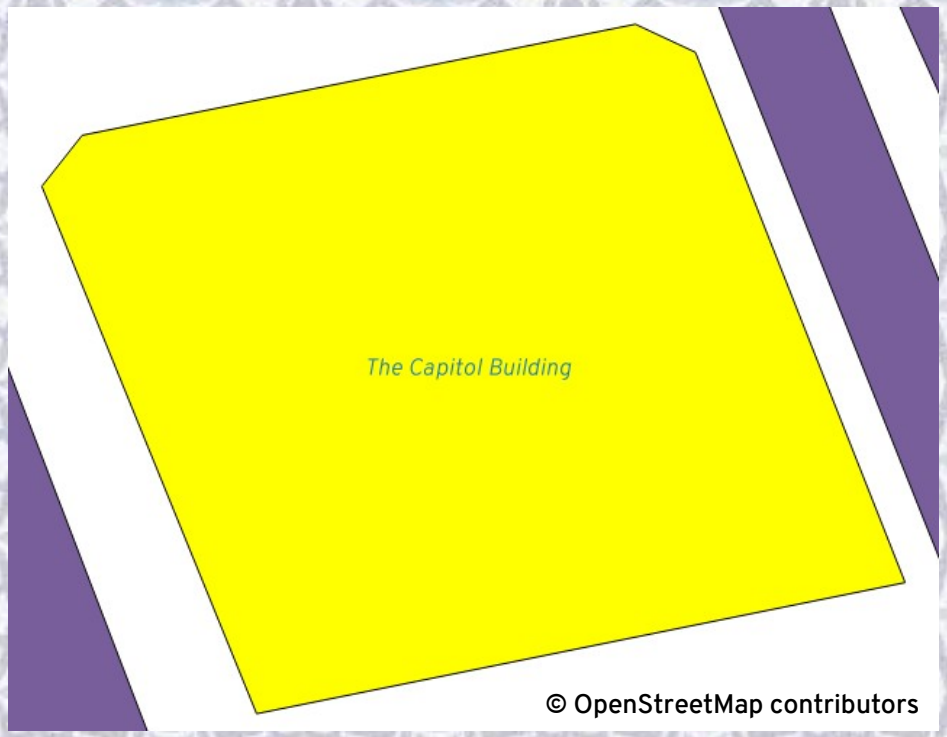


Photograph by Greg Sadler

The Studley Tool Chest

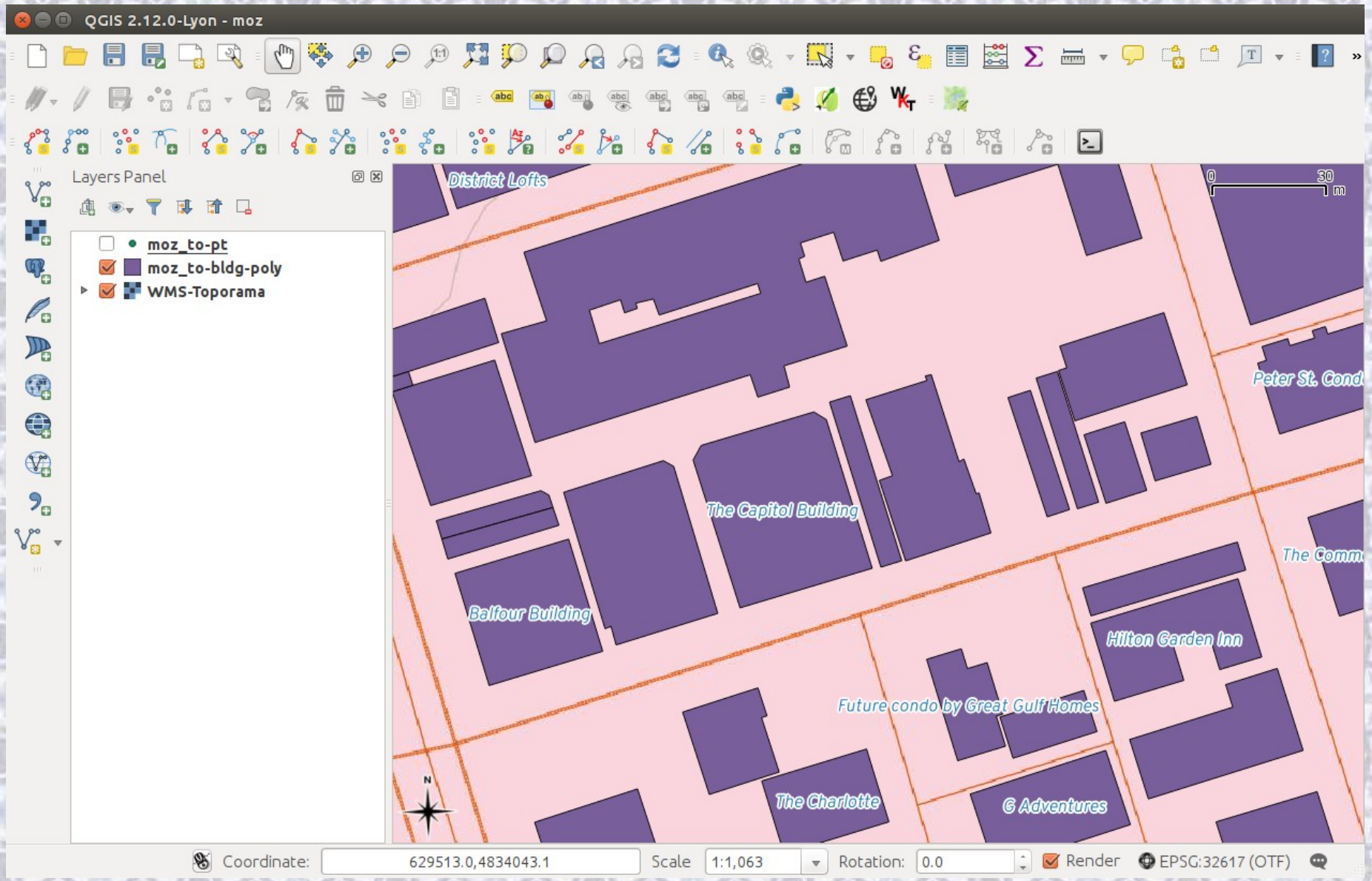
Fine WoodWorking.com

visual or batch?



{ "crs" : { "type" : "name", "properties" : { "name" : "urn:ogc:def:crs:OGC:1.3:CRS84" } }, "type" : "FeatureCollection", "features" : [{ "geometry" : { "coordinates" : [[[-79.39423, 43.64751], [-79.39409, 43.64715], [-79.39452, 43.64707], [-79.39466, 43.64742], [-79.39463, 43.64745], [-79.39427, 43.64753], [-79.39423, 43.64751]]], "type" : "Polygon" }, "type" : "Feature", "properties" : { "name" : "The Capitol Building" } }] }

qgis


QGIS 2.12.0-Lyon - moz
Layers Panel
moz_to-pt
moz_to-bldg-poly
WMS-Toporama
District Lofts
The Capitol Building
Balfour Building
Hilton Garden Inn
Future condo by Great Gulf Homes
The Charlotte
G Adventures
Peter St. Cond
The Comm
0
30
m
N
Coordinate:
629513.0,4834043.1
Scale
1:1,063
Rotation:
0.0
Render
EPSG:32617 (OTF)

gdal and ogr

GDAL reads *raster* data and:

- **scales, warps, re-projects, clips, tiles, sharpens, merges, convolutes and munges†**
- **saves in pretty much any *raster* format you need**

OGR reads *vector* data and:

- **scales, translates, warps, queries, simplifies, re-projects, clips, and hits with a big stick†**
- **saves in pretty much any *vector* format you need**

†: not a term officially recognized by the OGC.

even lower level tools

- **proj.4 – the reprojection standard**
- **fiona – python vector data import library**
- **shapely – python geoprocessing library**
- **openrefine – unruly data tamer**
- **unfolding – beautiful interactive maps for Processing 2.*n***

example:

war diaries

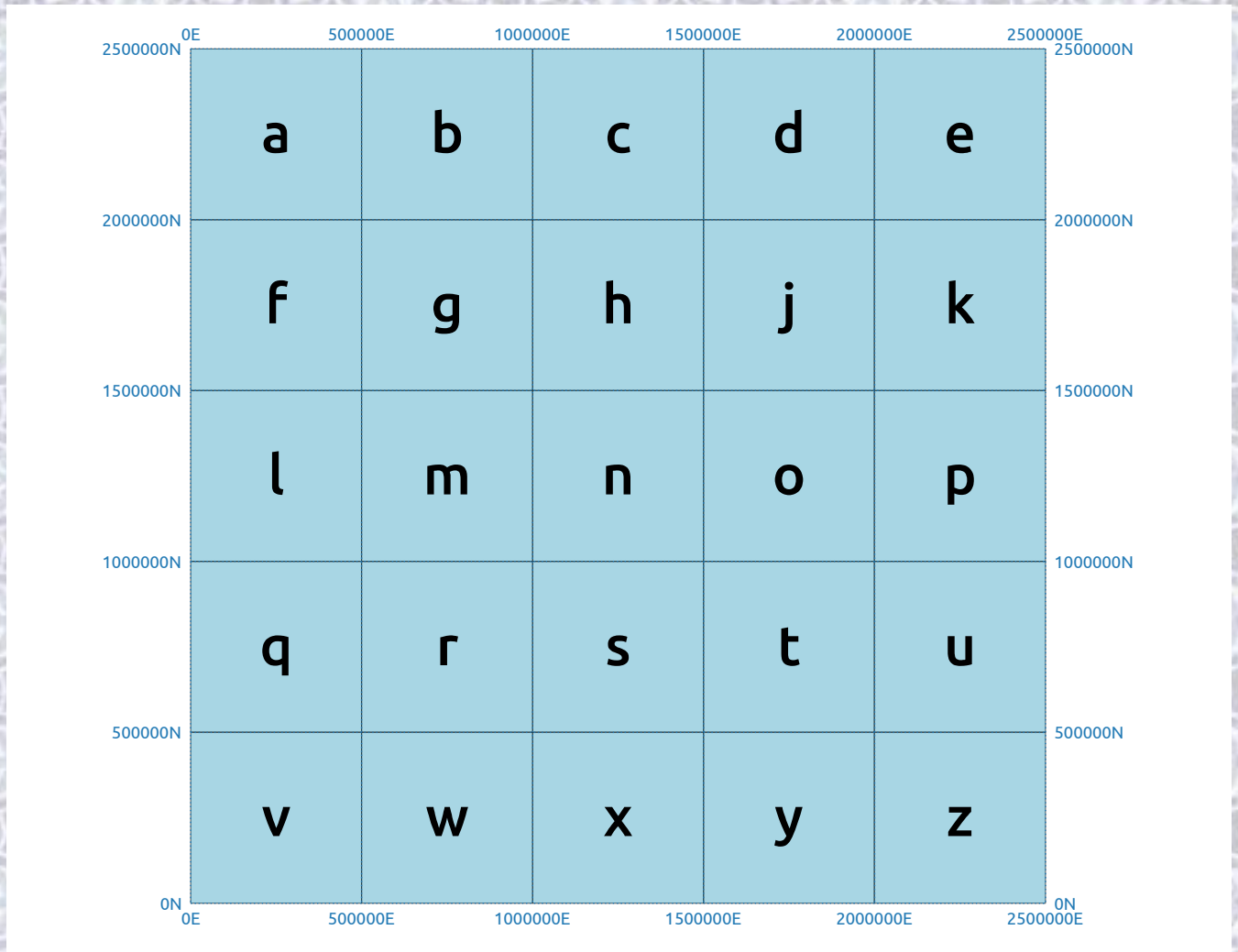
the data

20 Dec 43

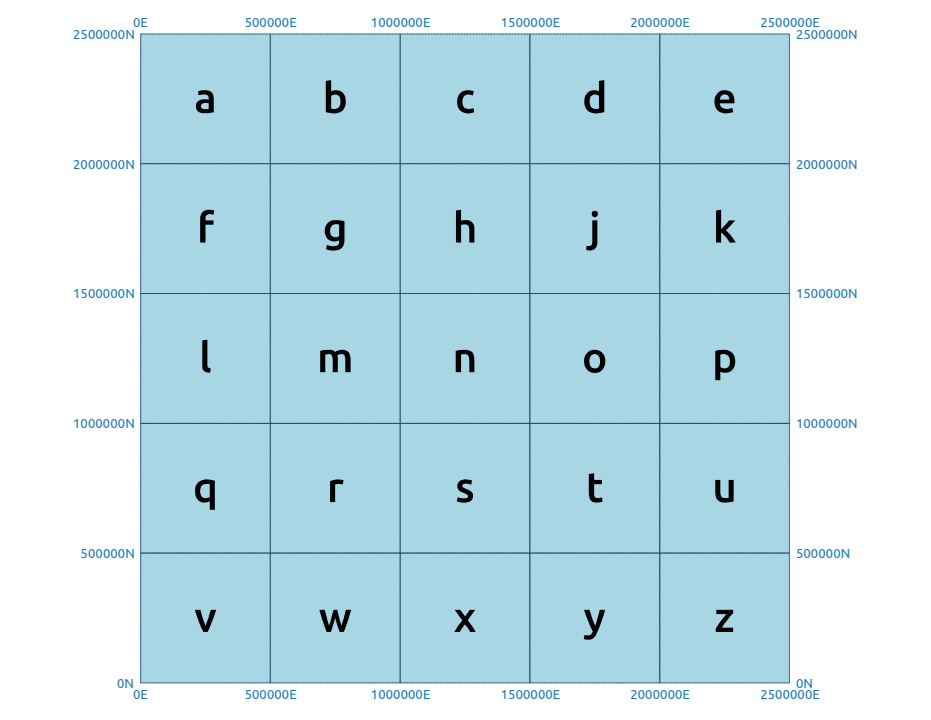
The regiment, supported by 'C' Squadron, The Three Rivers Tank Regt, advanced under a barrage on a two Coy, two Troop front at 1200 hrs today. The start line was astride the ORTONA road 100 yards forward of the crossroads at MR 322142. After very heavy fighting the regiment reorganized as follows: 'B' Coy the general area MR 333164, where liaison was established with 'C' Coy of the Seaforth H of C who had fought their way up the coast road; 'D' Coy to the northwest of 'B' Coy astride the main road; 'C' Coy at MR 328158; 'A' Coy and Bn Hq at MR 332160. ...

— War Diary, The Loyal Edmonton Regiment, vol. 50 (Dec 1943)

modified british grid

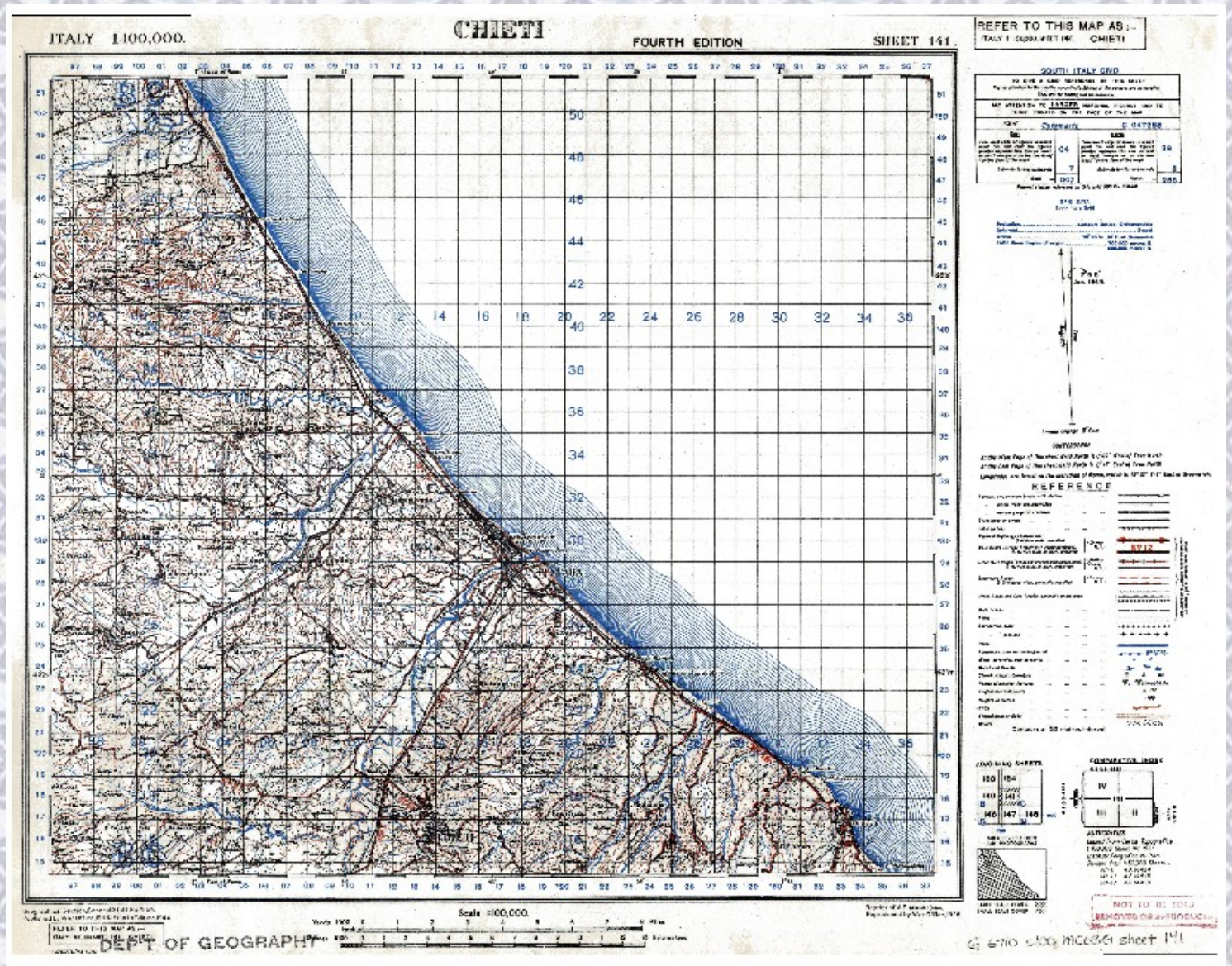


decoding the grid




```
result = SEARCH(letter,
"VWXYZQRSTULMNOPFGHJKABCDE")-
1
easting = result MOD 5
northing = INT(result / 5)
```

the map



Map CC-BY-NC, McMaster University Library

the map details

SOUTH ITALY GRID

TO GIVE A GRID REFERENCE ON THIS SHEET

Pay no attention to the smaller co-ordinate figures at the corners and in margins. They are for finding full co-ordinates.

PAY ATTENTION TO LARGER MARGINAL FIGURES AND TO THOSE PRINTED ON THE FACE OF THE MAP.

POINT *Casamarte* C 047288

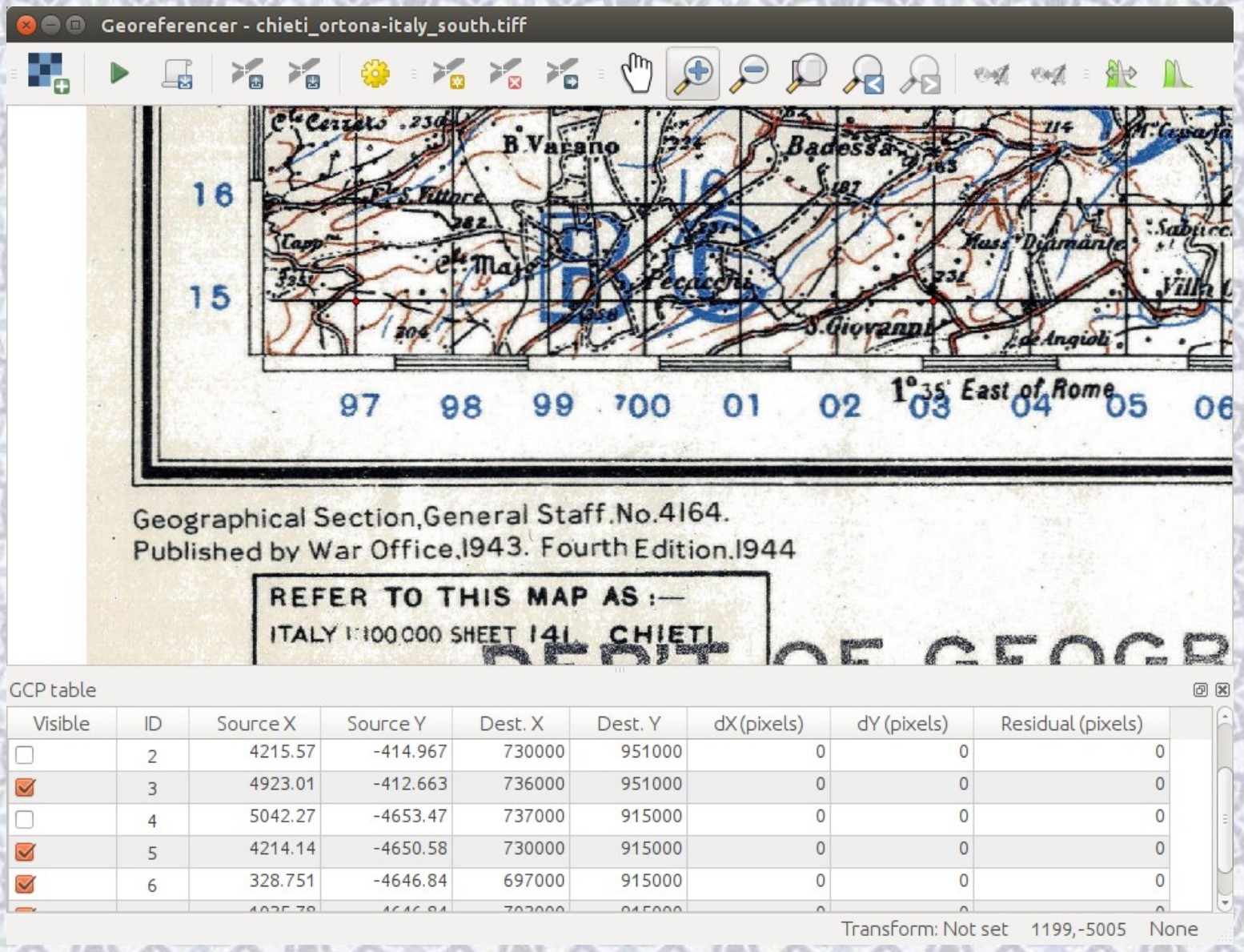
East		North	
Take west edge of square in which point lies and read the figures printed opposite this line on north or south margin or on the line itself (on the face of the map)	04	*Take south edge of square in which point lies and read the figures printed opposite this line on east or west margin or on the line itself (on the face of the map)*	28
Estimate tenths eastwards	7	*Estimate tenths northwards*	8
East ...	047	North ...	288

Nearest similar reference on this grid 500 Km. distant.

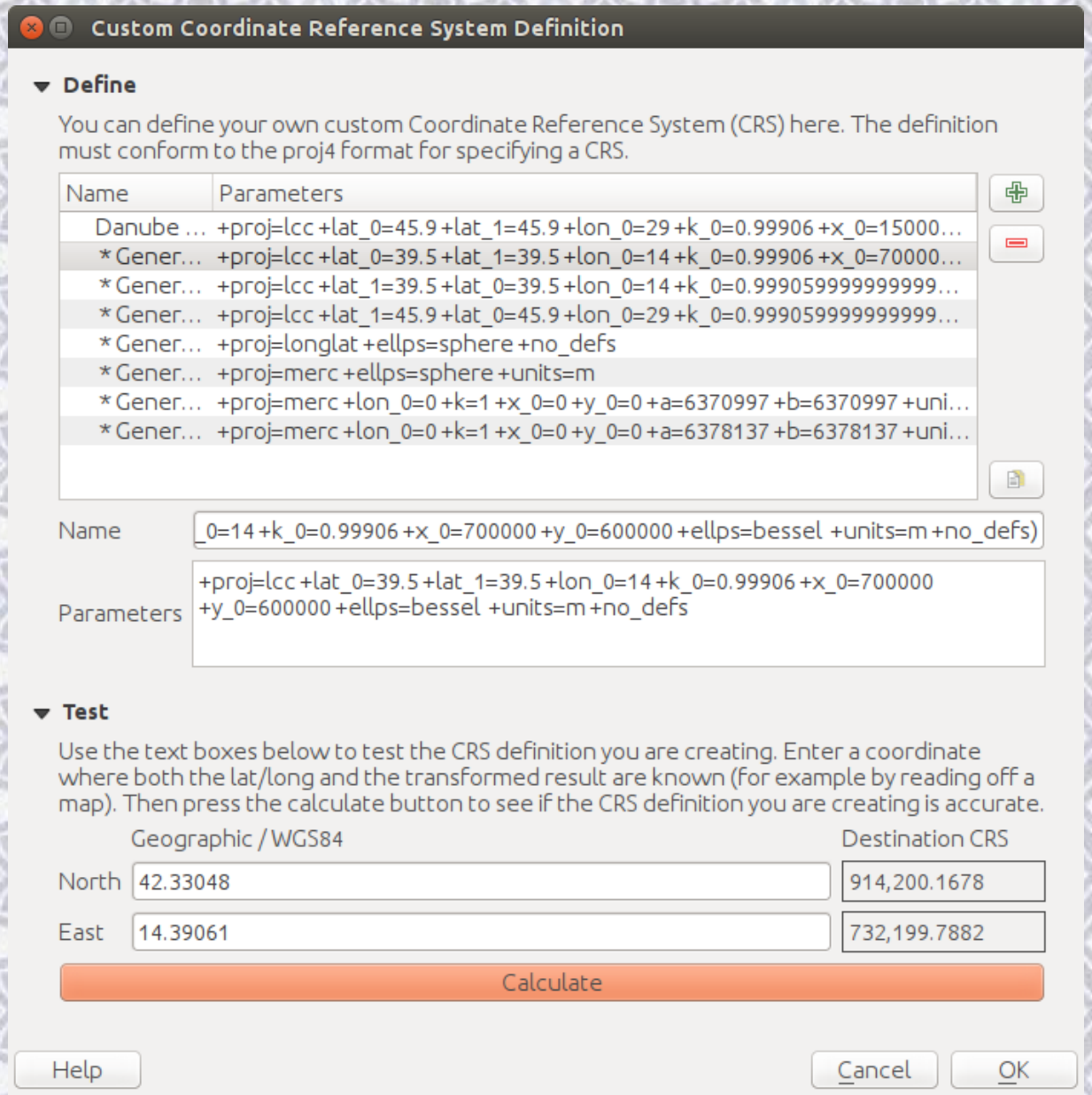
GRID DATA
South Italy Grid

Projection	Lambert Conical Orthomorphic
Spheroid	Bessel
Origin	39° 30' N., 14° E. of Greenwich
False Co-ordinates of origin	700,000 metres E. 600,000 metres N.

georeferencing ...



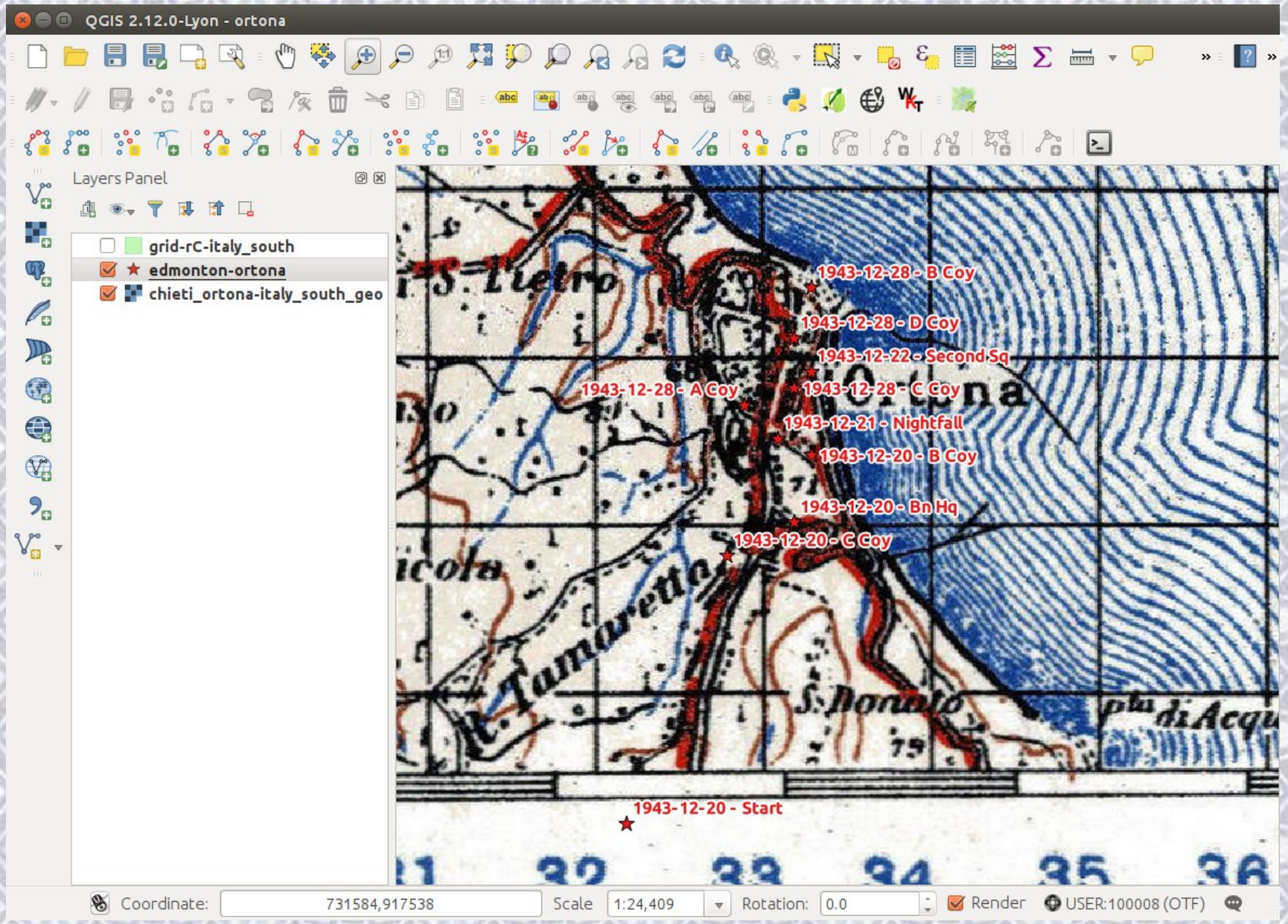
the magic of proj.4 ...



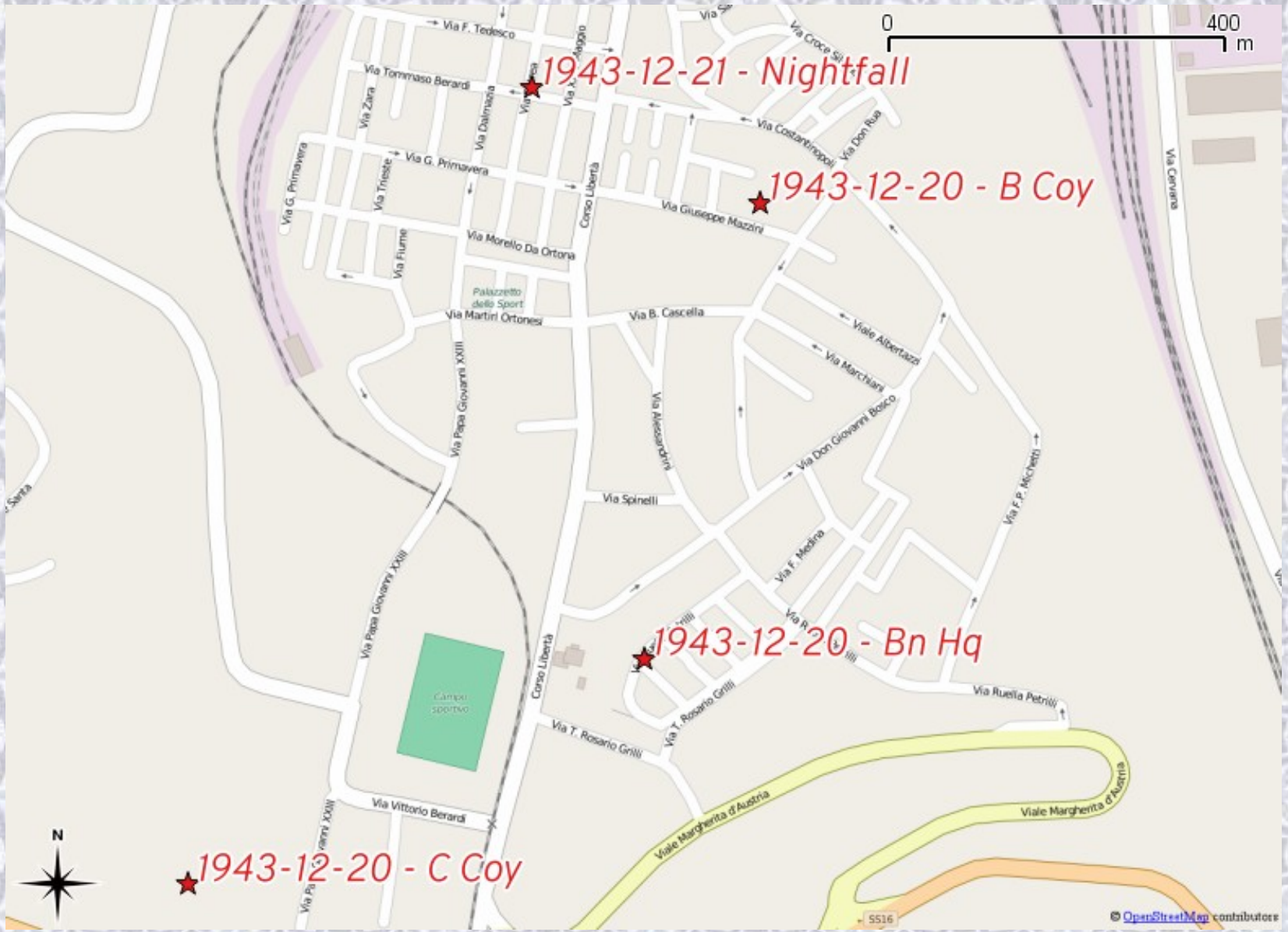
the data, transformed

MR	X	Y	Label
322142	732200	914200	1943-12-20 - Start
333164	733300	916400	1943-12-20 - B Coy
328158	732800	915800	1943-12-20 - C Coy
332160	733200	916000	1943-12-20 - Bn Hq

in historical context



in modern context



credits

- *YOU* — for showing up
- OpenStreetMap
- McMaster University Library
- The Taunton Press, Inc / finewoodworking.com
- The Loyal Edmonton Regiment Museum
- Thierry Arsicaud, echodelta.net
- GIS StackExchange
- Malcolm Hamilton
- All the GIS developers who can do hard sums so we don't have to