



Practical Geospatial Data Versioning

kartproject.org

Robert Coup

robert.coup@koordinates.com



Koordinates

A geospatial data management platform inspired by cracking GIS data out of vendor silos. You can host, manage, share, publish, access, and build, alongside thousands of others worldwide.



koordinates.com



@koordinates

The screenshot shows the Koordinates web application. At the top, there's a navigation bar with links for 'koordinates', 'Pricing', 'Request a demo', 'Data', 'Map', 'Get the App', 'Help', and 'Log in'. Below the navigation is a search bar with the placeholder 'Search & browse thousands of geospatial datasets'. To the left, a sidebar has sections for 'Popular', 'Recent', 'Publishers', 'Browse' (with a dropdown set to 'Europe'), 'Data Type', 'Publisher', 'Date', and 'License'. The main area displays a grid of 12 dataset cards. Some visible cards include 'World Urban Areas (1:10 million)' from Natural Earth, 'US National Waterway Network' from US Bureau of Transportation Statistics (BTS), 'Land Uptake per Person (inhabitants and jobs) (LUP) 2009 (LEAC Grid)' from Nov. 2016 Vector Project, and 'High Nature Value (HNV) farmland 2012 (100 m) accounting version, Nov. 2017'. On the right side of the interface is a large map of Europe and surrounding regions, showing various green and blue shaded areas representing different datasets. A legend at the bottom left of the map indicates scales of 500 km and 500 mi.



What can you do with versioning?

Work smoothly across multiple tasks and projects

Publish data regularly & reliably; with full history

Accept contributions

Peer review changes

Automate workflows & testing

Barn Images; Unsplash



Opportunities

Data integrity & verifiability

Bi-directional data supply chains

Cross eco-system data handling

Reliably use the latest data

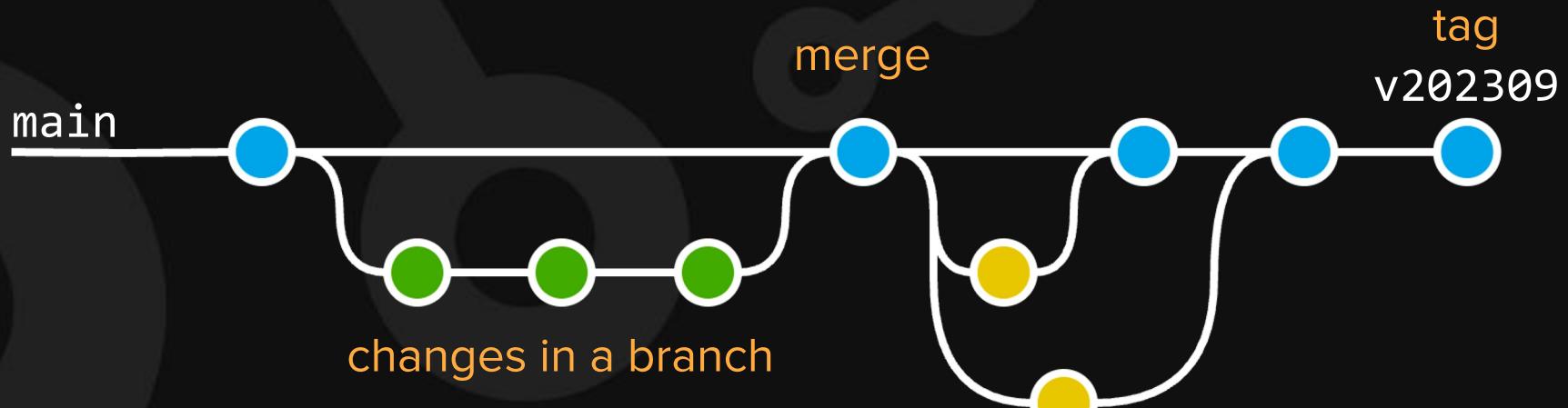
J.Zamora; Unsplash

All data is versioned



what changed?
who changed it?
when did it change?
why did it change?

What does real versioning look like?



Kart Principles

Open & free; ecosystem agnostic

Easy to install and **batteries included**



For practical day to day use

Built on Git

Photo by James Sullivan on Unsplash

\$



QGIS Project Edit View Layer Settings Plugins Vector Raster Database Web Mesh Processing Window Help Wed 9:32 am

*Untitled Project — QGIS



Layers



Identify Results Layers

Identify Results Layers

Type to locate (36K) 3 legend entries removed.

Coordinate -43.776,-176.401

Scale 1:371121

Magnifier 100%

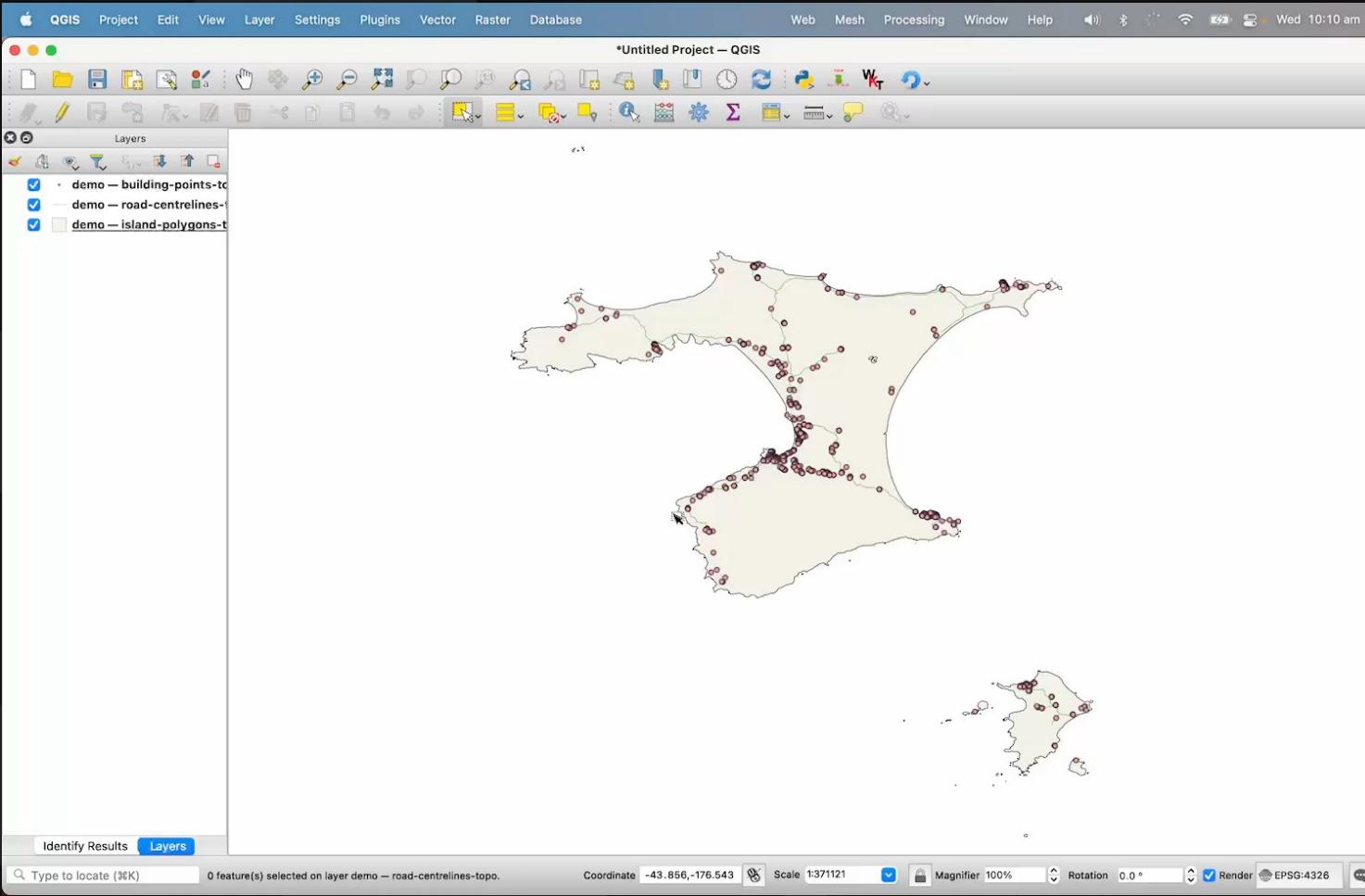
Rotation 0.0 °

Render EPSG:4326



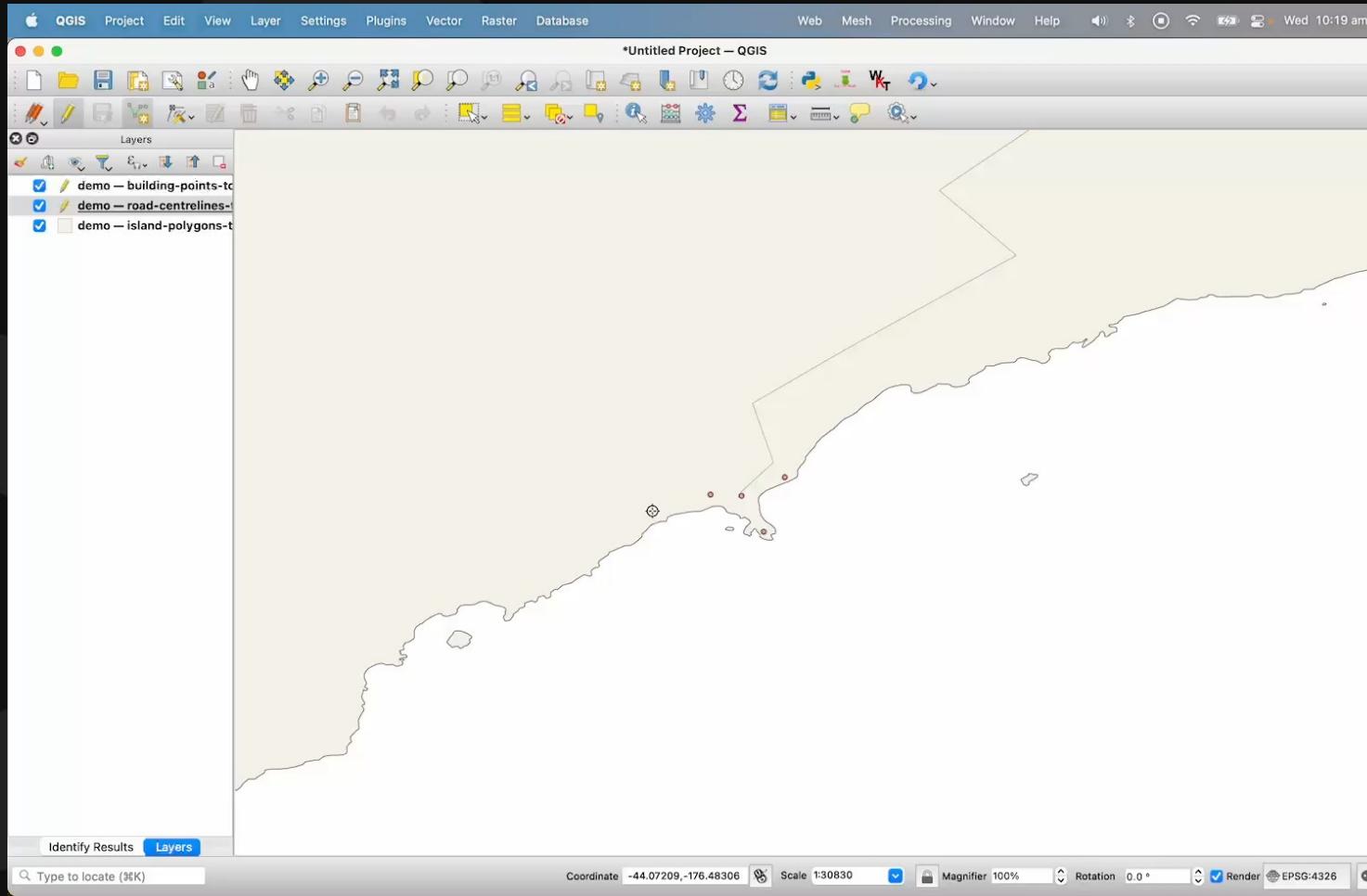
\$





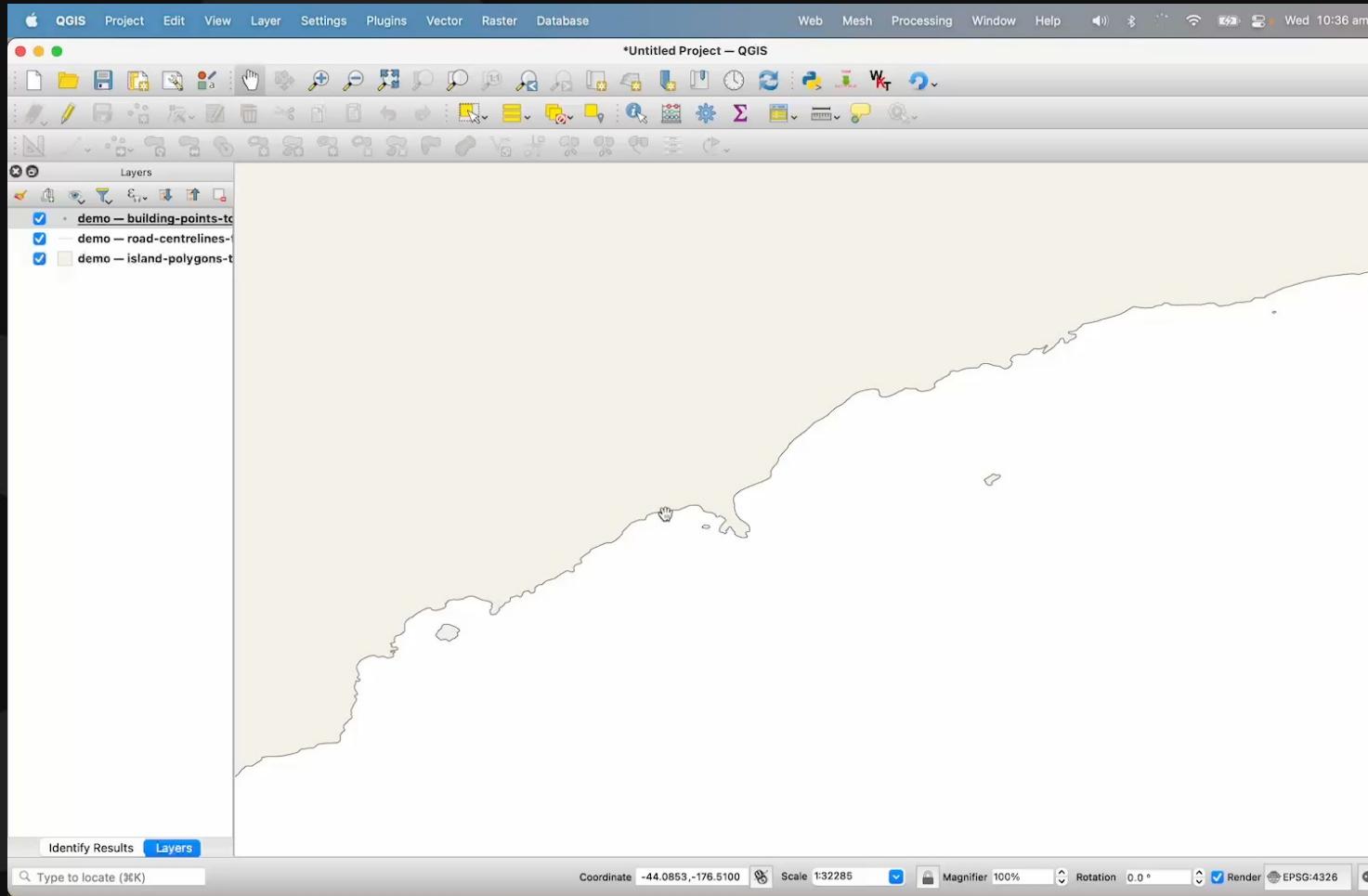
\$





\$





Working Copies

Where you work with and edit your data

Different repository users can use different working copies

- Vector & Tables: GeoPackage, PostGIS, MSSQL & MySQL
Roadmap: ESRI File GeoDatabases
- Point Clouds: LAZ & LAS
- Rasters & Grids: native & GeoTIFFs
- Cloud Optimised

Spatially Filtered Clones

Work with only your area of interest

Smaller working copies for better performance in your tools

Reduces the network transfer for clones & fetches



Fetch from & push relevant updates to the full dataset

Spatial Filtering

```
$ kart clone example.org:mydata  
  --spatial-filter="EPSG:4326;POLYGON((-4 55.7,  
  -4 56, -4.5 56, -4.5 55.7, -4 55.7))"
```

```
$ kart clone example.org:mydata  
  --spatial-filter=@myextent.txt
```

```
$ kart fetch
```

\$

▼ **coordinates**

Vector & Table Datasets

0→100GB sized datasets

Data types follow a SQL model

✓ CRS support ✓ Schema changes ✓ Conflict resolution

Import from many OGR formats

- 💡 Can "re-import" from a snapshot dataset into a new commit, and Kart will figure out the change.

Point Cloud Datasets

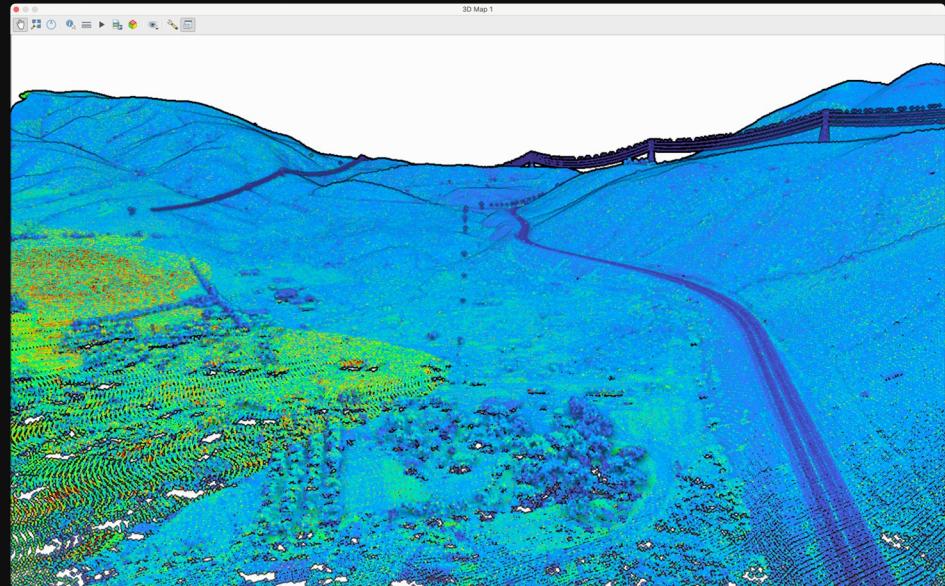
Built on Cloud Optimised Point Clouds (COPC) — see copc.io

Supports LAS/LAZ

0→TB sized datasets

S3/object storage support

Automatic Virtual Point Cloud (VPC)
creation for QGIS



Raster Datasets

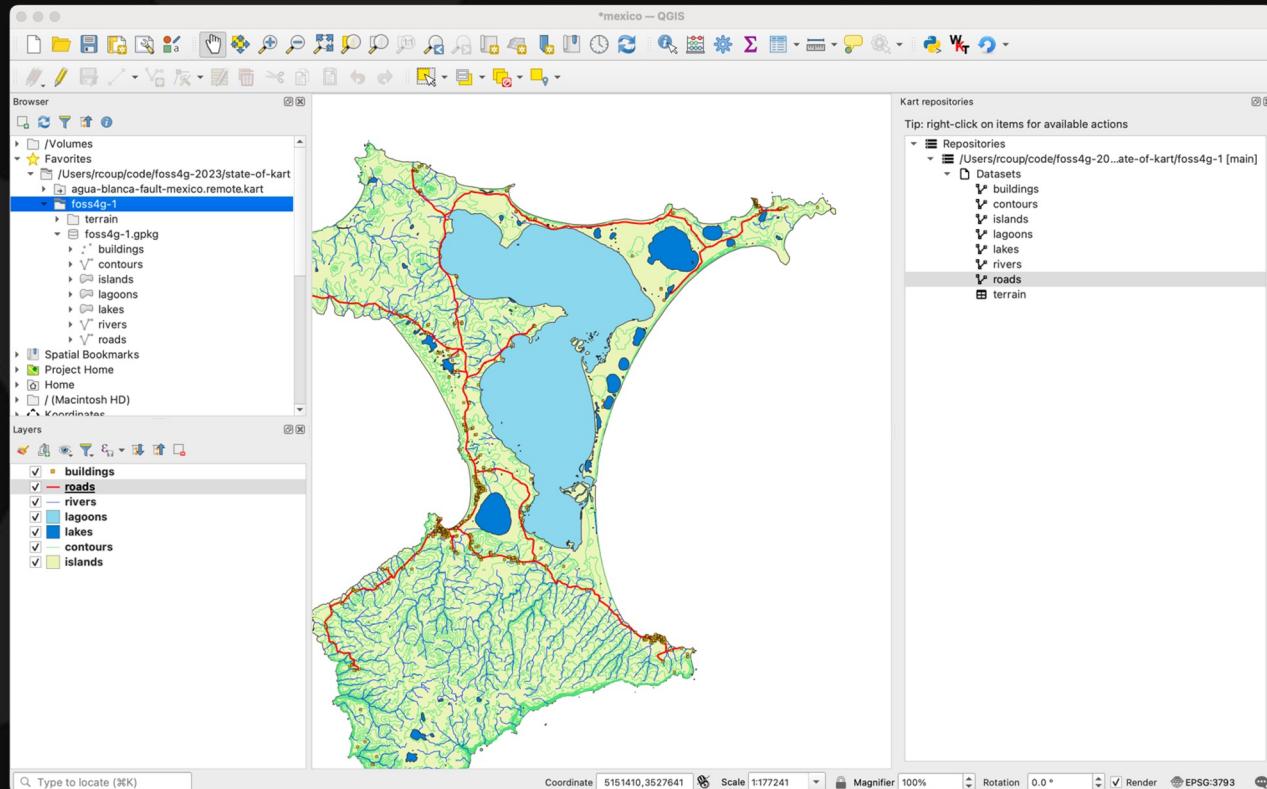
Built on Cloud Optimised GeoTIFFs

0→TB sized datasets

S3/object storage support

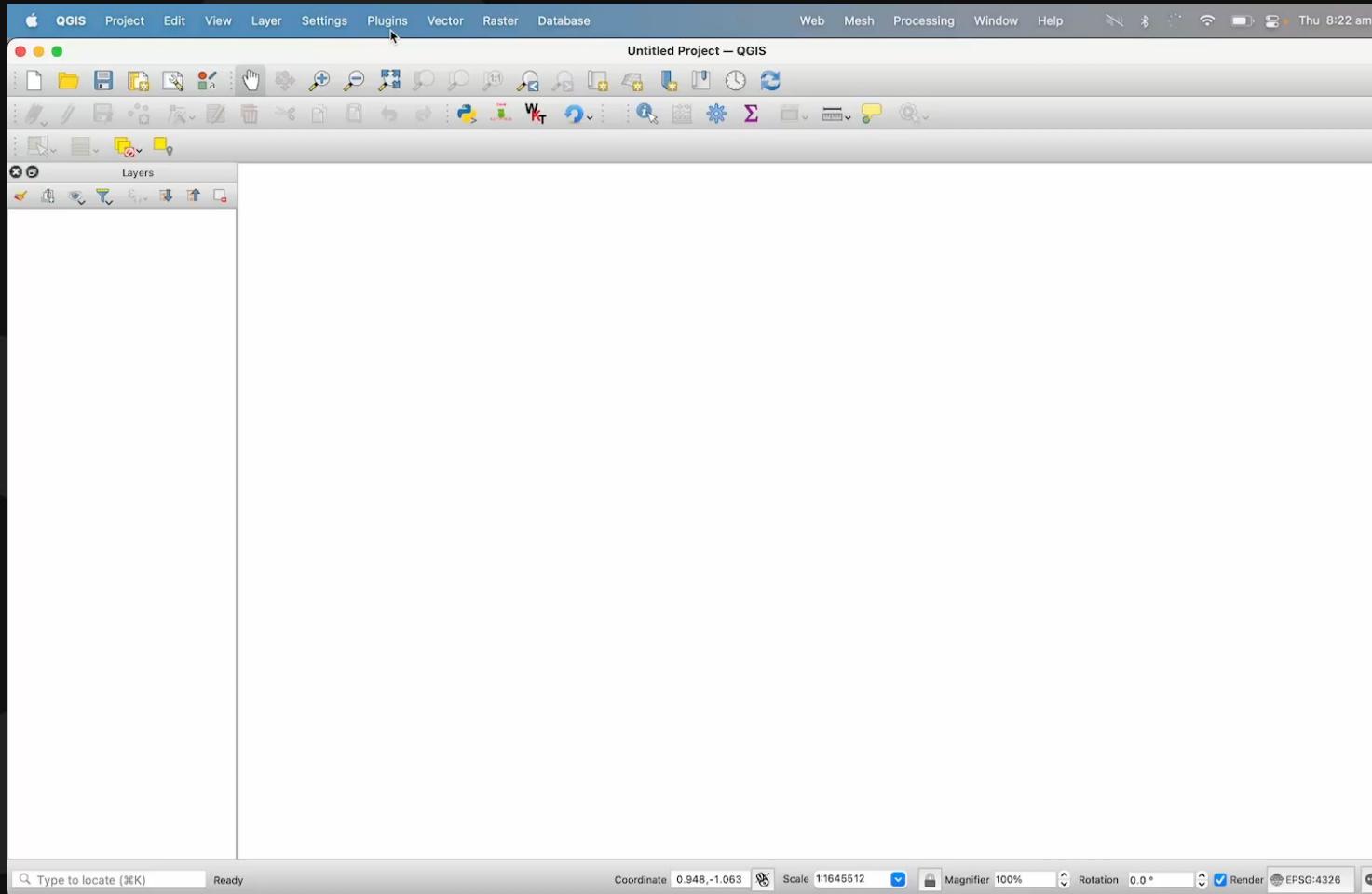
Automatic Virtual Raster (VRT) creation

QGIS Plugin



Install via the QGIS
plugin manager





Roadmap

- 💡 Referencing data from *existing* S3 buckets without copying
- 💡 File GeoDatabase working copy
- 💡 Blend local & remote Raster & Point Cloud datasets
- 💡 Inter-linking datasets for projects
- 💡 Serving tiles & APIs (like STAC) directly from repos
- 💡 Extended CRS support



kartproject.org

 @KartForData

Robert Coup

robert.coup@koordinates.com

