



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.

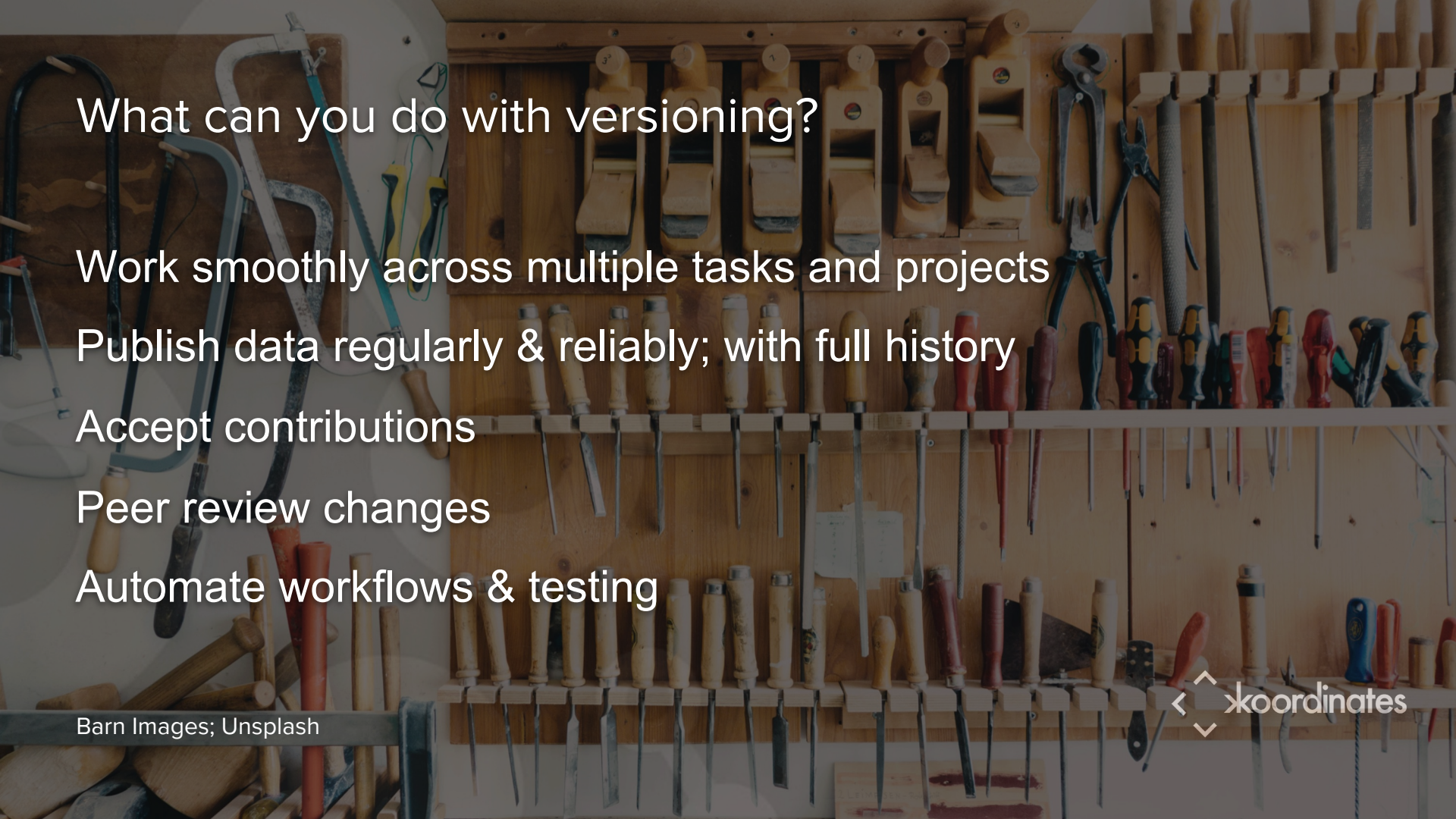


koordinates.com

@koordinates

A screenshot of the Koordinates web application interface. The top navigation bar includes the Koordinates logo, 'Pricing', 'Request a demo', a search bar for 'Data', a 'Map' button, and links for 'Get the App', 'Help', and 'Log in'. The main content area is divided into a left sidebar and a main map area. The sidebar has sections for 'Popular', 'Recent', 'Publishers', and 'Browse'. The 'Browse' section is active, showing a list of datasets with filters for 'Europe', 'Data Type', 'Publisher', 'Date', and 'License'. Four dataset cards are visible: 'World Urban Areas (1:10 million)' by Natural Earth, 'US National Waterway Network' by US Bureau of Transportation Statistics (BTS), 'Land Uptake per Person (inhabitants and jobs) (LUP) 2009 (LEAC Grid), Nov. 2016' by Vector Project, and 'High Nature Value (HNV) farmland 2012 (100 m) accounting version, Nov. 2017'. Each card includes a thumbnail map, a layer type (e.g., Polygon Layer, Line Layer, Raster Layer), and a 'Map +' button. The main map area shows a map of Europe with various layers overlaid, including a 'Protected sites habitats, birds and other sp...' layer. The map has a search bar, 'Contents', and 'Export' buttons. A scale bar at the bottom left of the map indicates 500 km and 500 mi.



A workshop wall covered in various tools. The top row features several hand planes with wooden bodies and metal blades. Below them are various chisels with wooden handles. To the left, there are several hand saws with curved blades and wooden handles. The wall is made of light-colored wood, and the tools are organized in a way that suggests a well-used workspace.

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

Opportunities

Data integrity & verifiability

Bi-directional data supply chains

Cross eco-system data handling

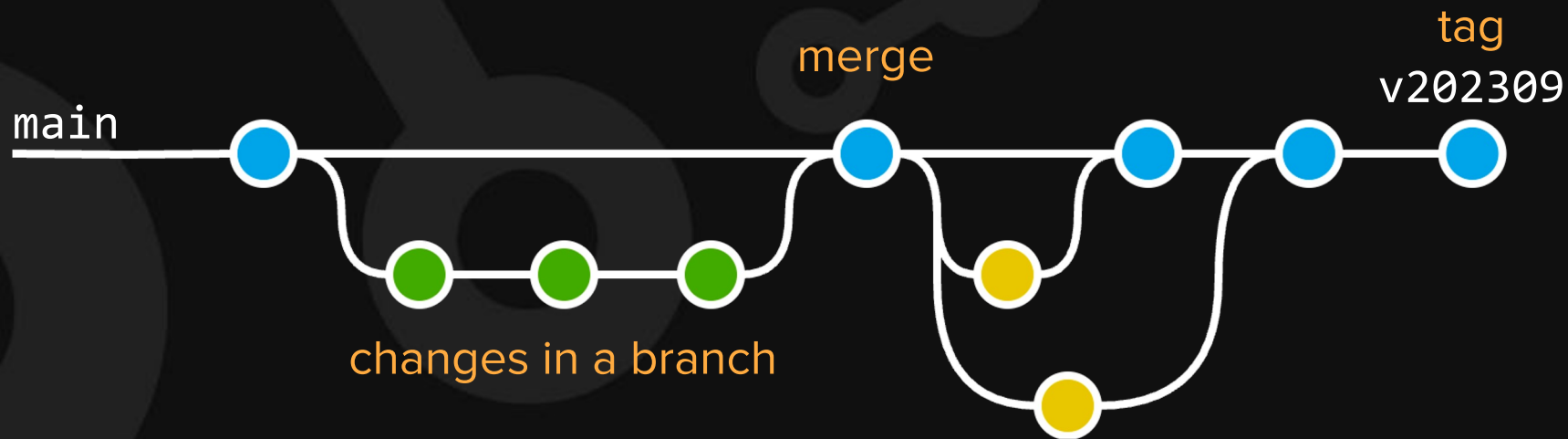
Reliably use the latest data

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

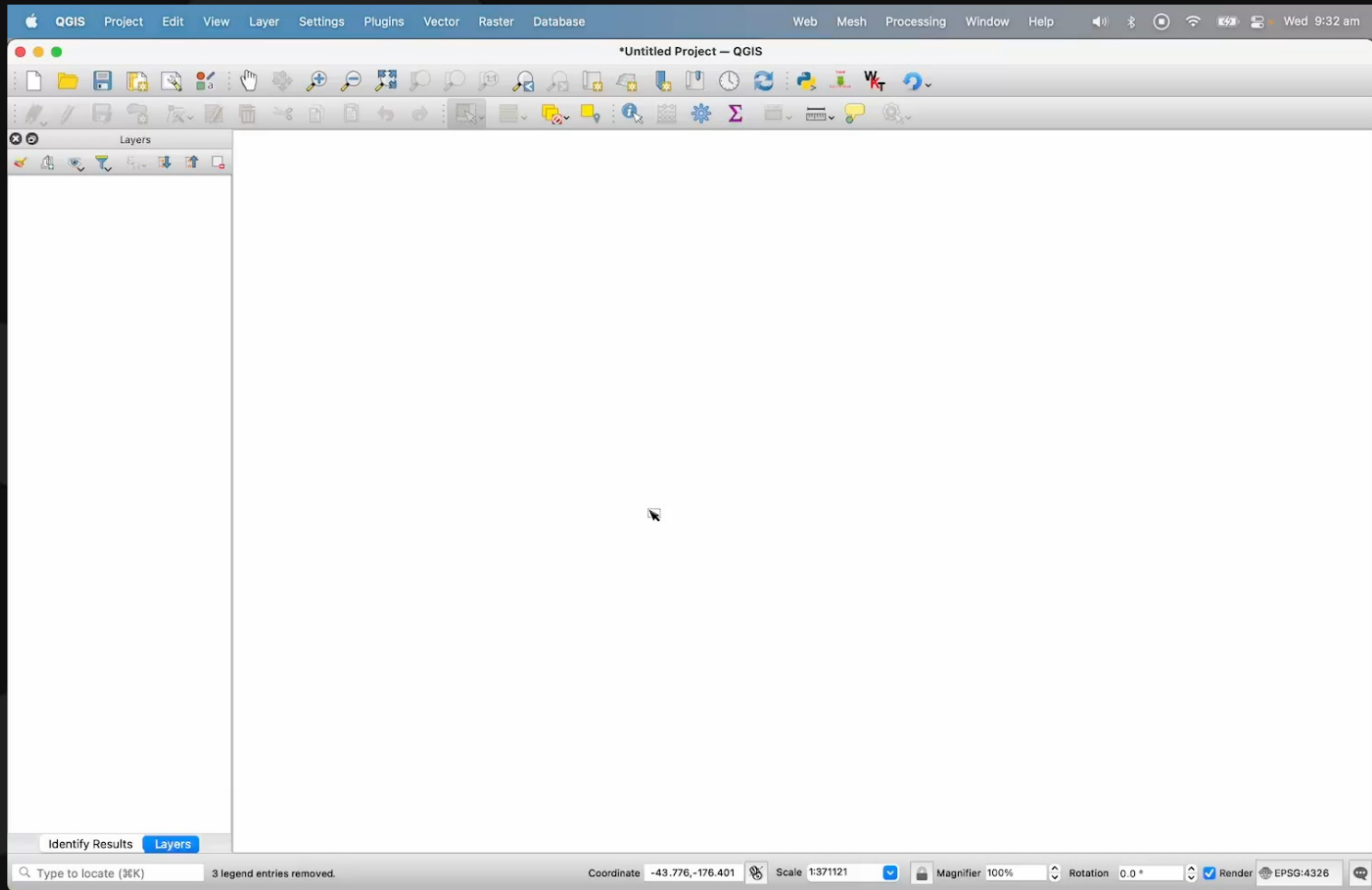
Built on Git

Photo by James Sullivan on Unsplash

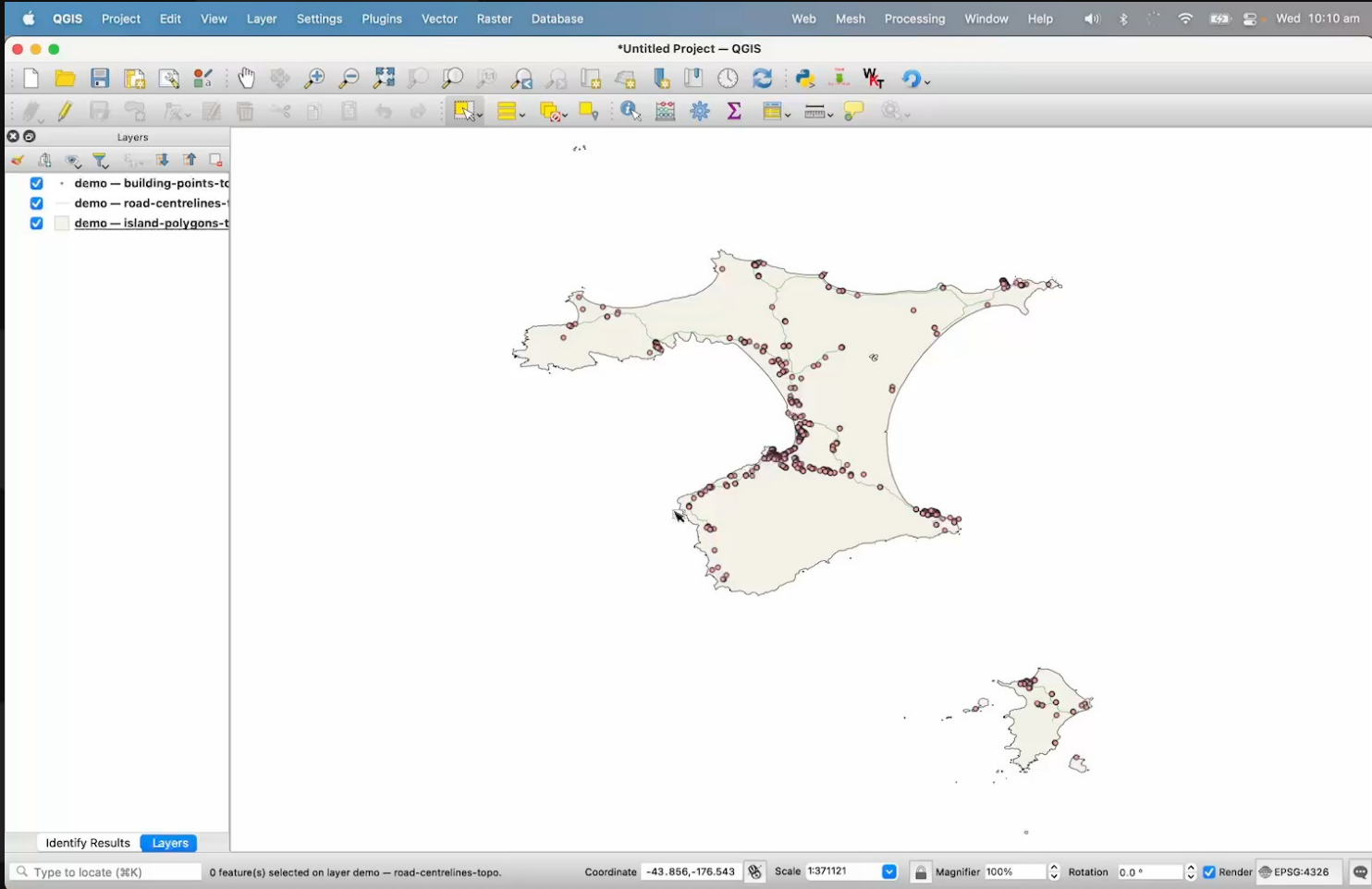


\$

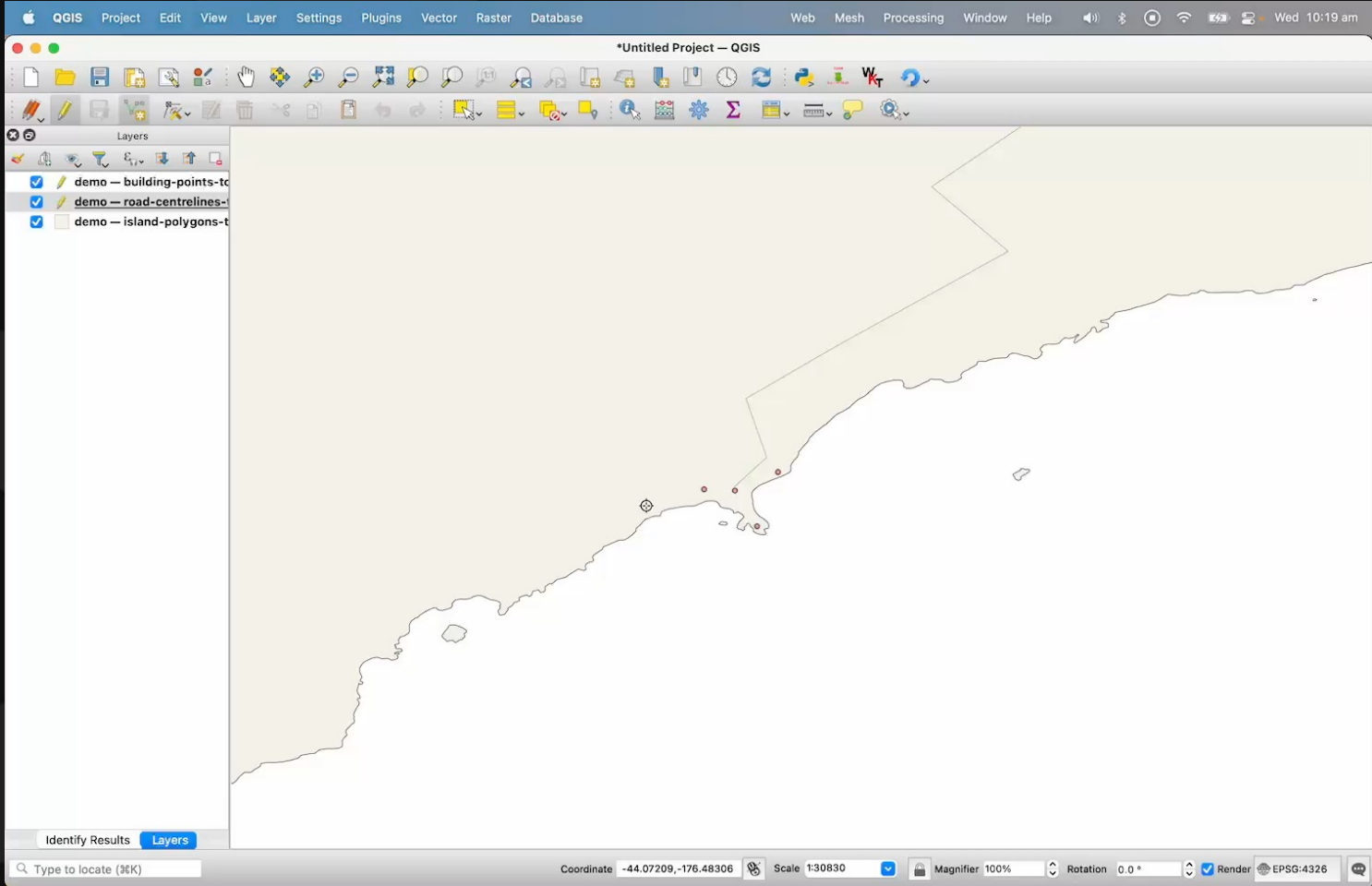




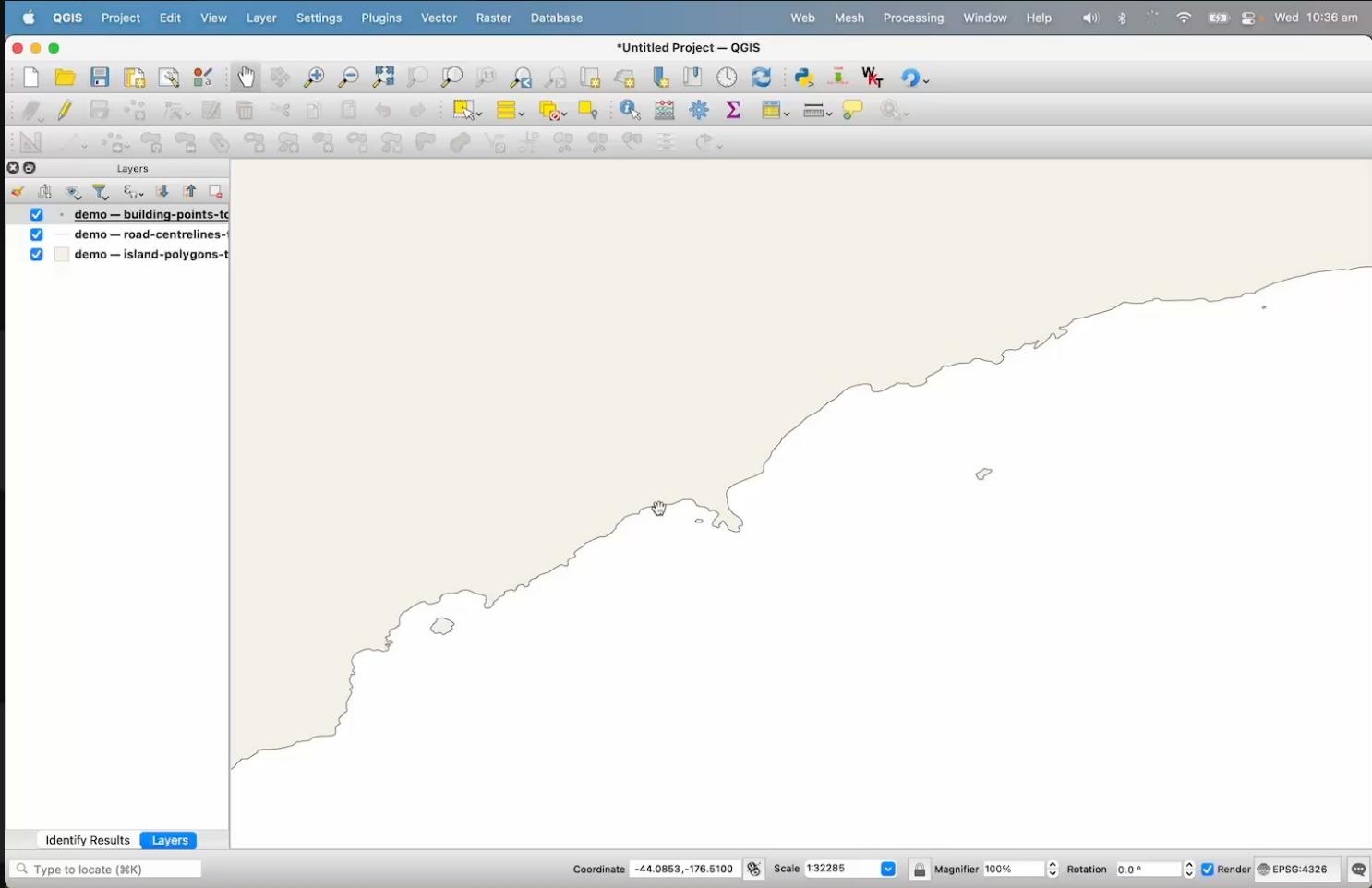
\$ █



\$ █



\$ █



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
```

\$



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

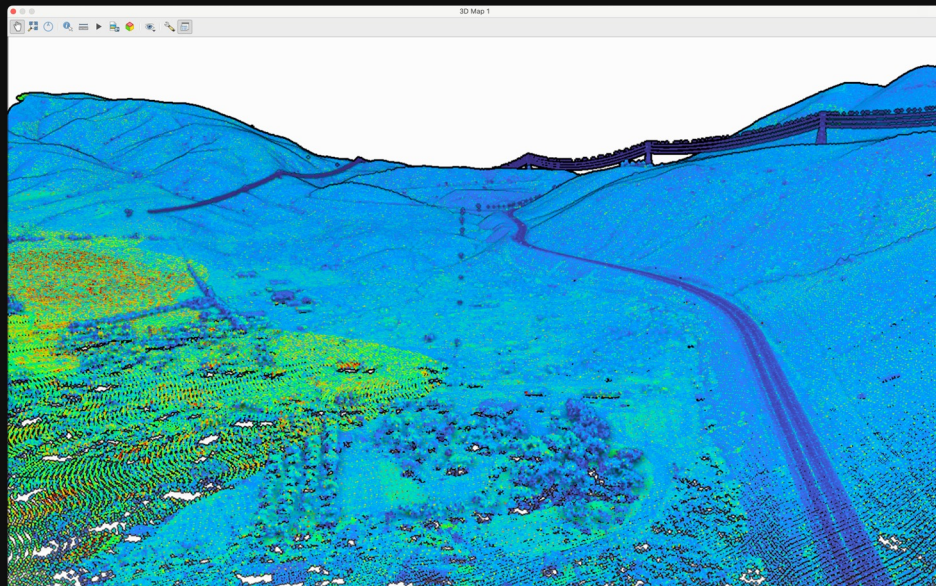
Built on Cloud Optimised Point Clouds (COPC) — see copc.io

Supports LAS/LAZ

0→TB sized datasets

S3/object storage support

Automatic Virtual Point Cloud (VPC)
creation for QGIS



Raster Datasets

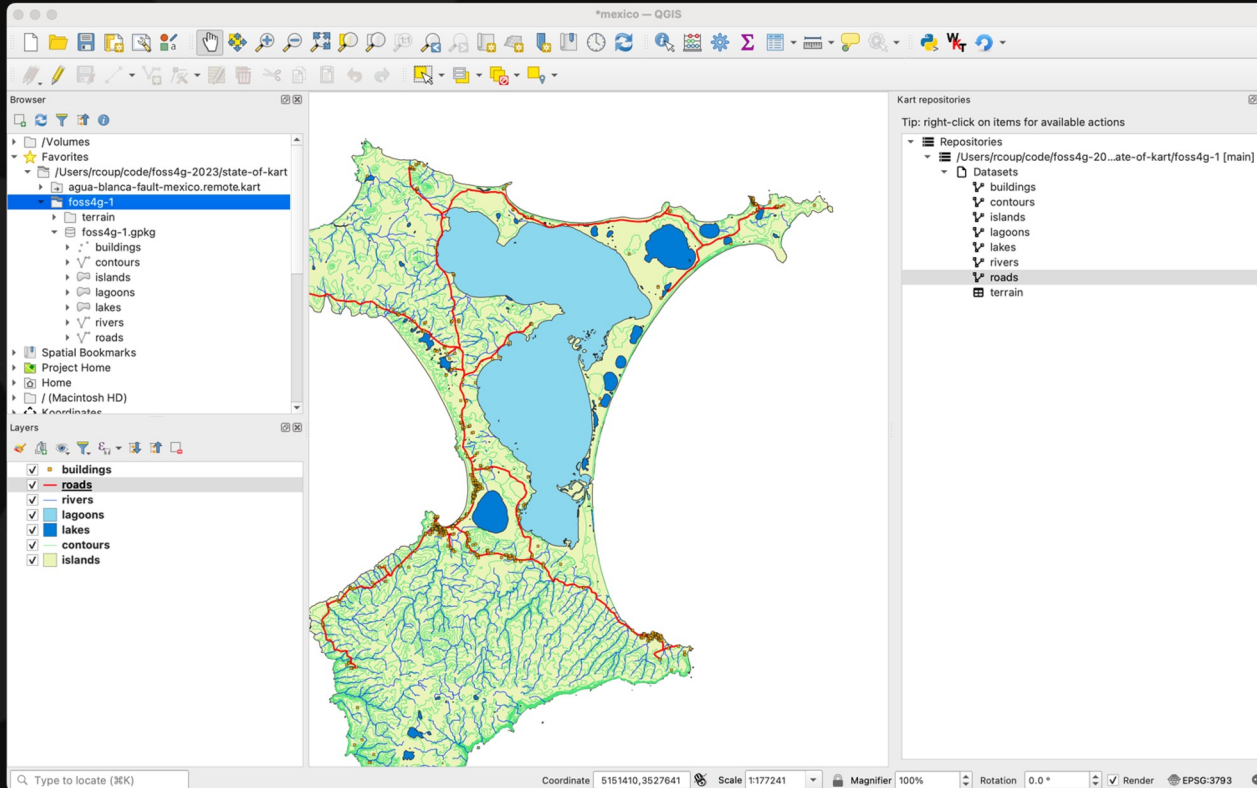
Built on Cloud Optimised GeoTIFFs

0→TB sized datasets

S3/object storage support

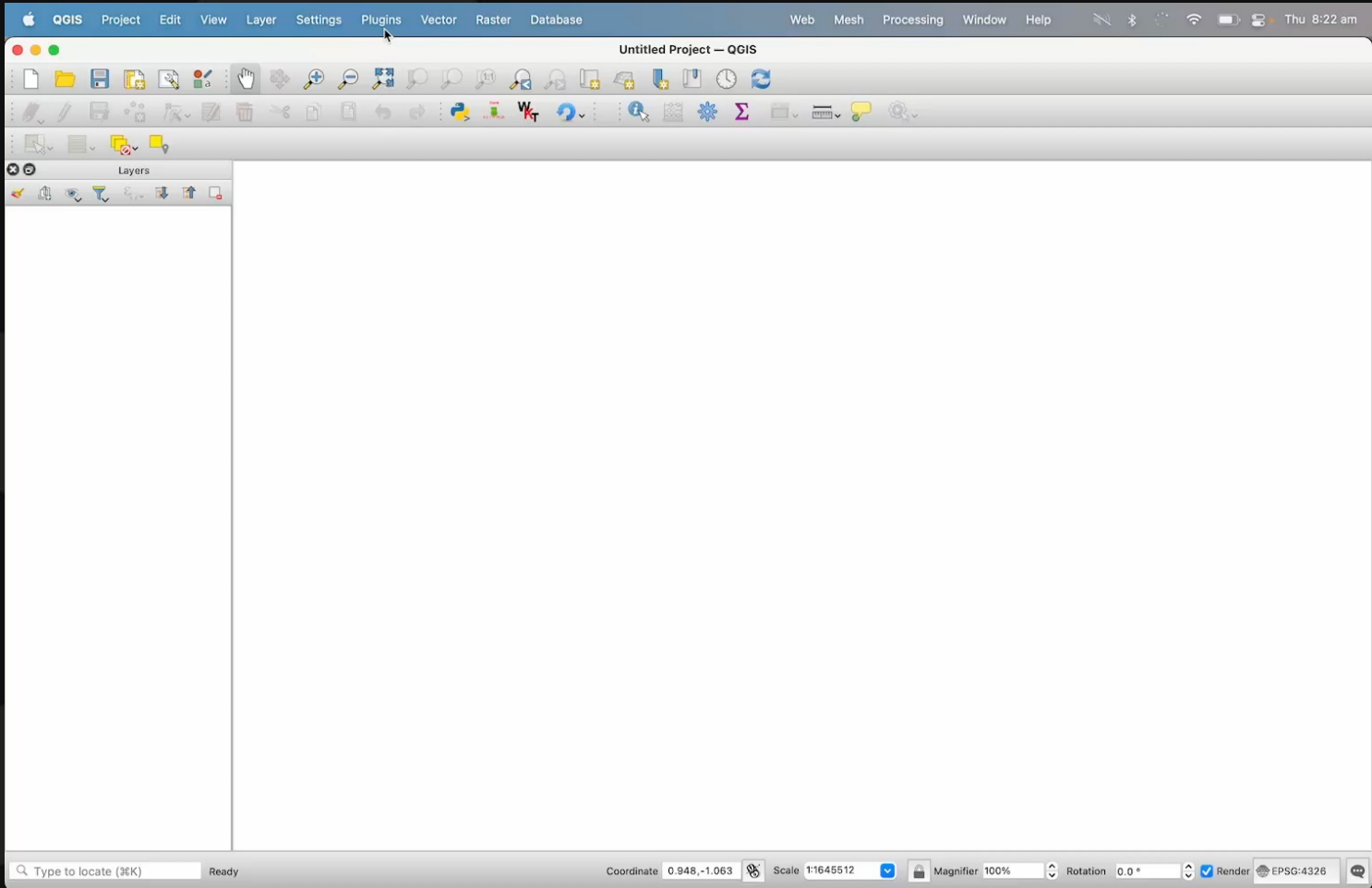
Automatic Virtual Raster (VRT) creation

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

