Kart

Practical Geospatial Data Versioning

kartproject.org

Hamish Campbell

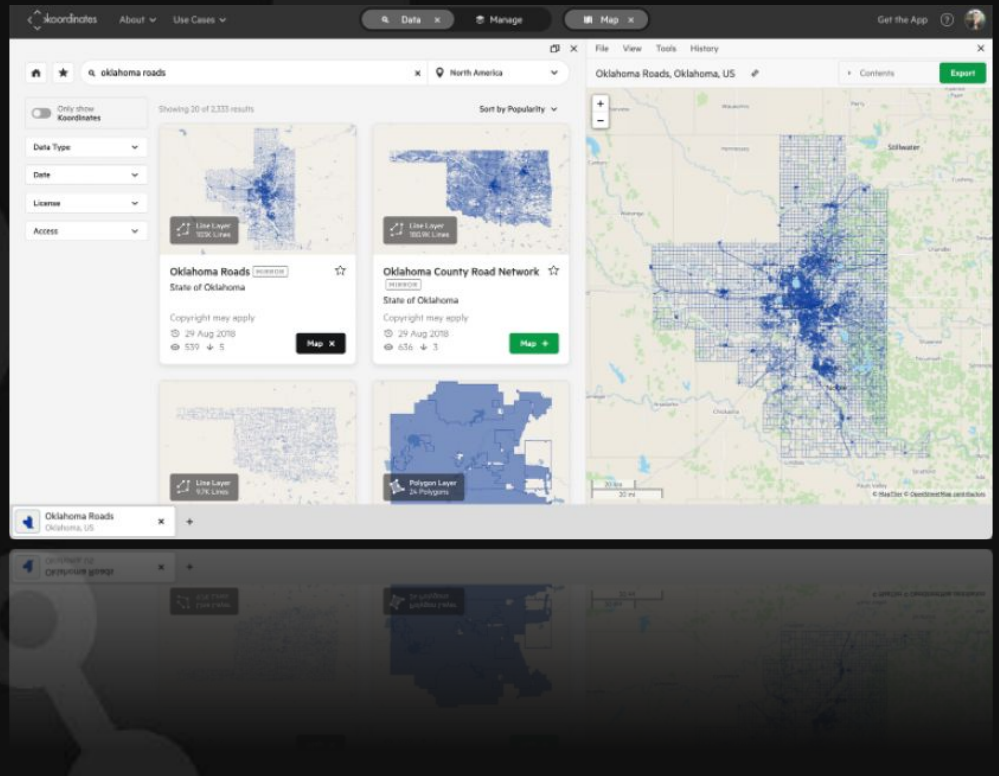
hamish.campbell@koordinates.com

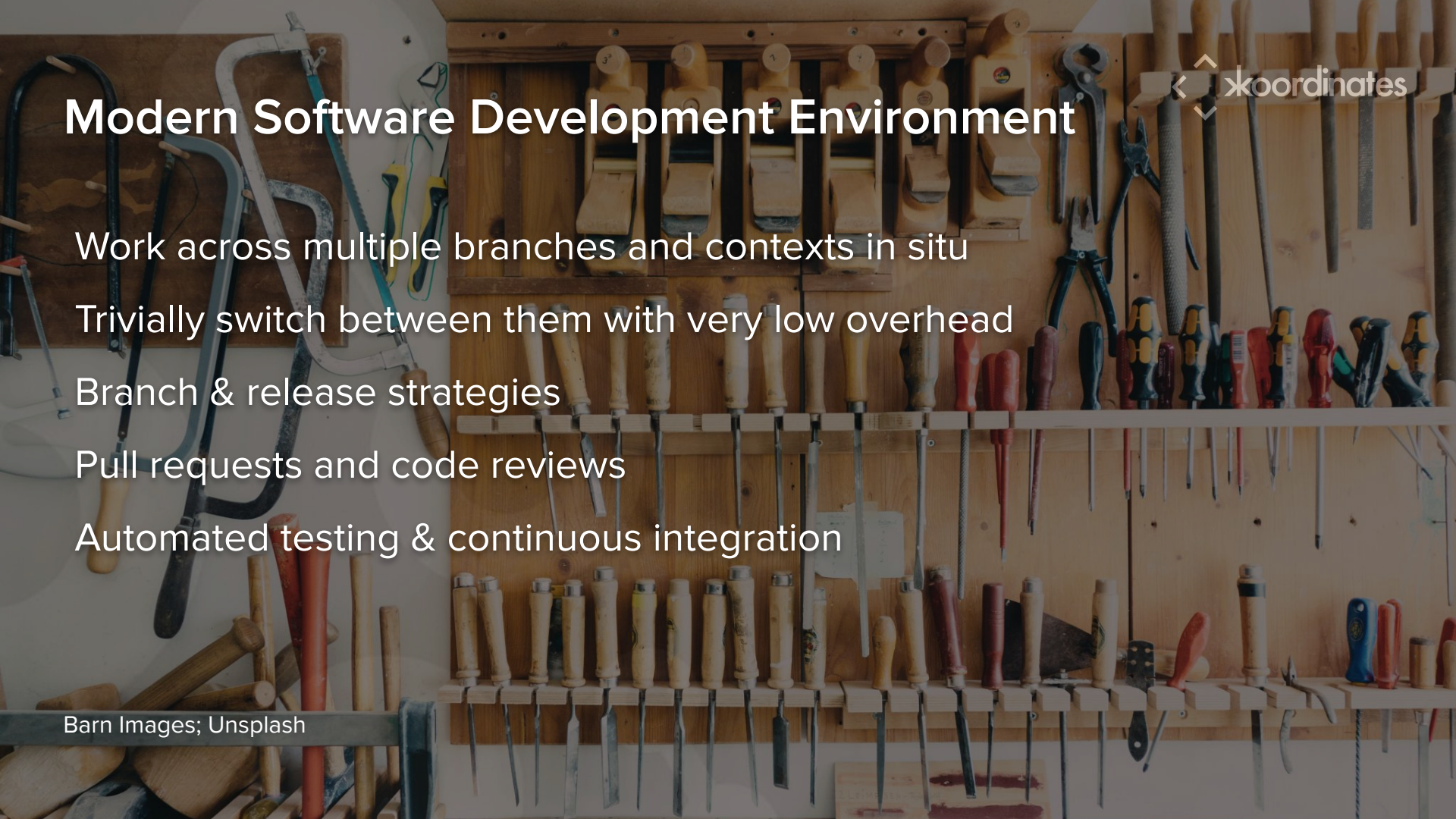
Koordinates

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

koordinates.com

@koordinates





Modern Software Development Environment

koordinates

Work across multiple branches and contexts in situ

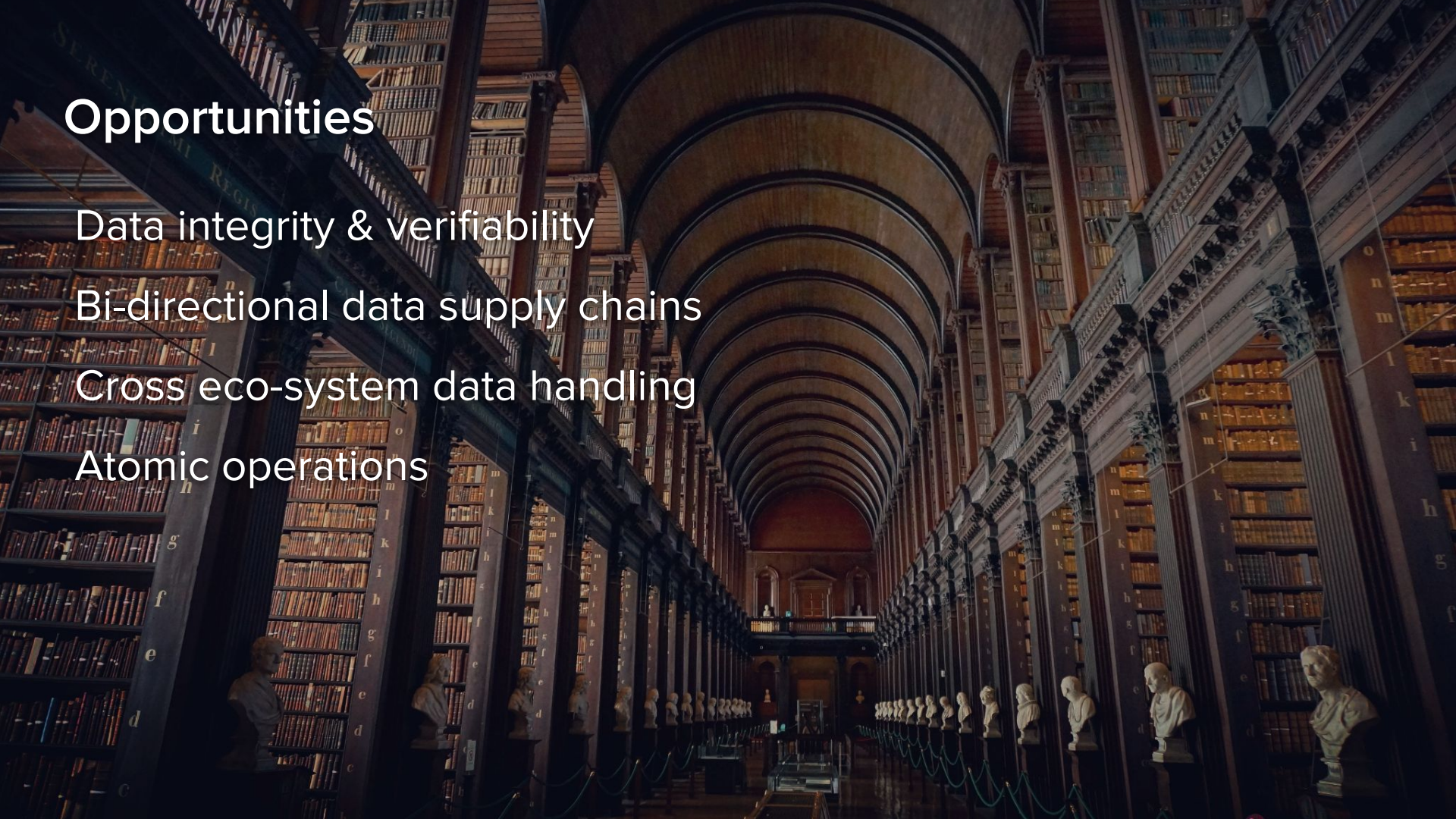
Trivially switch between them with very low overhead

Branch & release strategies

Pull requests and code reviews

Automated testing & continuous integration

Barn Images; Unsplash



Opportunities

Data integrity & verifiability

Bi-directional data supply chains

Cross eco-system data handling

Atomic operations

Kart History

We’re not new to linear version control

“Git for data” is a fertile area

Variations on a theme:

- “Data in Git”
- “Inspired by Git”
- Specialist systems
- Machine learning and other data science concerns

Alexandr Popadin on Unsplash

Kart Principles

Built on Git

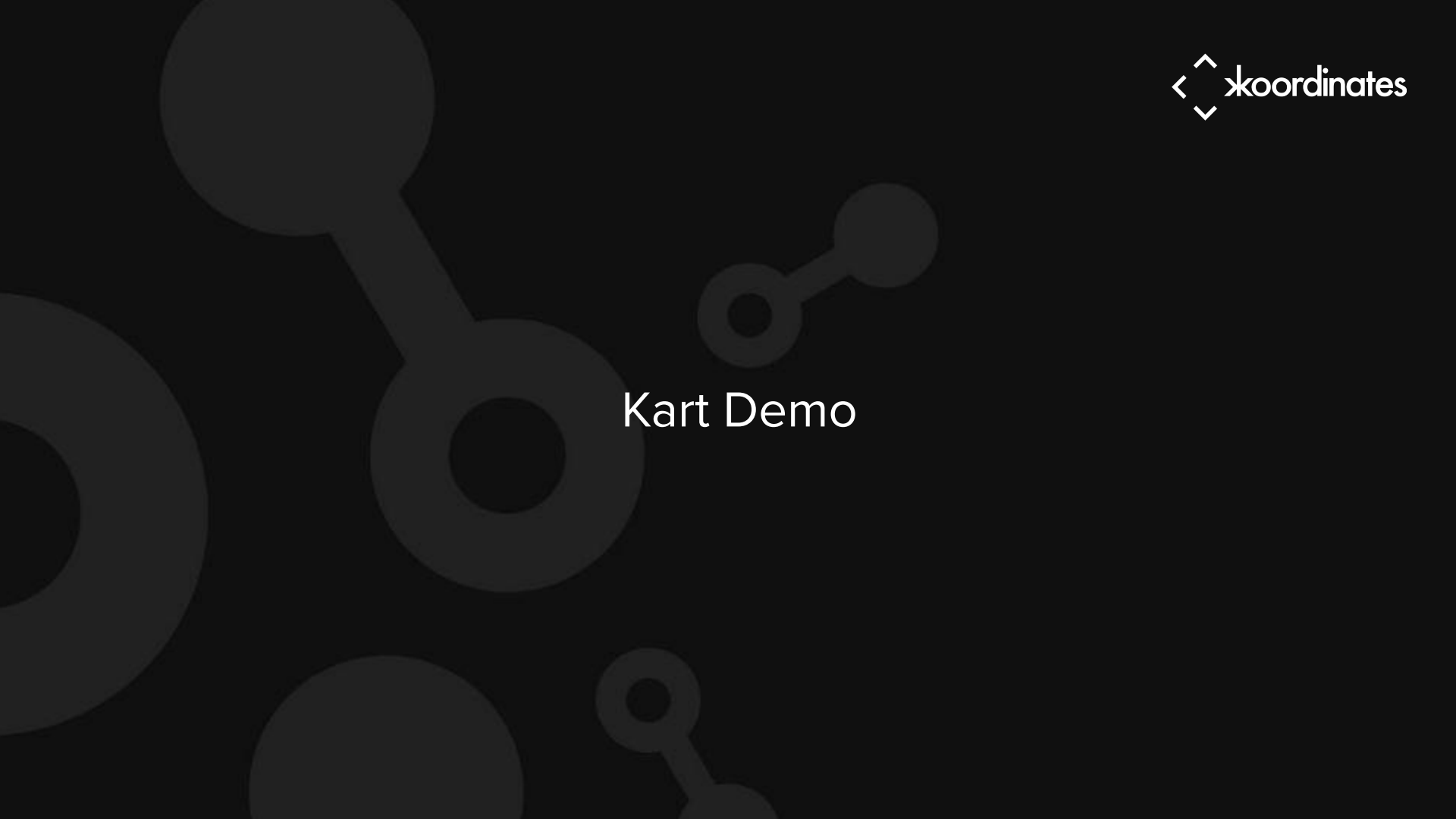
Easy to install and **batteries included** 

For practical day to day use

Open & free; ecosystem agnostic

Photo by James Sullivan on Unsplash


Victor Rodriguez on Unsplash



Kart Demo

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

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

Vector & Table Datasets

koordinates

0➔100GB sized datasets

Data types follow a SQL model

CRS support

Schema changes are supported

Import from PostGIS, MSSQL, MySQL, GPKG, SHP, and other OGR formats

Conflict resolution

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

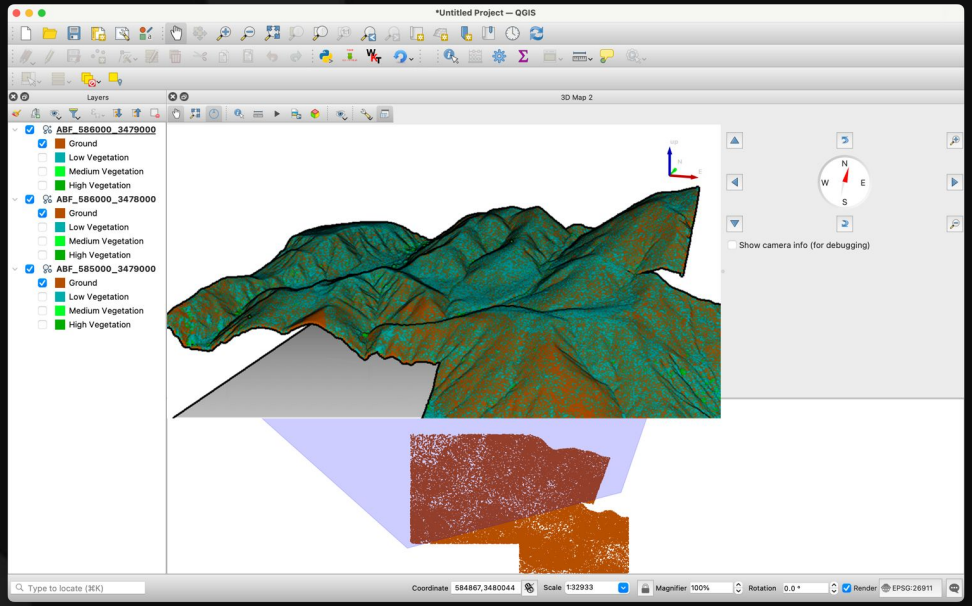
Point Cloud Datasets

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

0➔TB sized datasets

S3/object storage support

Supports LAS/LAZ



Spatially Filtered Point Cloud Demo

Roadmap I

ArcGIS Pro & FME plugins

File GeoDatabase working copy

Raster & Grid datasets

Exports

QGIS plugin improvements

Guide to self-hosting repositories

Cleaner build system

Roadmap II

Improving documentation

Work with Point Cloud/Raster/Grid data both remotely & locally

Files, Documents, Licenses, Metadata

Inter-linking datasets

Extended CRS support

koordinates

kartproject.org

@KartForData

Hamish Campbell

hamish.campbell@koordinates.com