





Stephen Epps

GIS / Data Analyst

stephenepps@fastmail.com 

github.com/stepps00 

linkedin.com/in/stephenepps 

stepps00.io 

PROFESSIONAL EXPERIENCE

Overture Maps Foundation

June 2024 - Present

Geospatial Data Operations Analyst: 2024 - 2026; Senior Data Engineer: 2026 - Present

Oakland, CA

- Manage Overture's monthly data release process; orchestrate theme concatenation into public releases using DAGs in Apache Airflow, LakeFS, and AWS Batch
- Investigate data irregularities, new data sources, and issue patterns across Overture's six data themes using QGIS, DuckDB, S3 and Athena, and various open source mapping libraries
- Build DuckDB / Athena / Superset pipelines to audit data releases (schema conformance, null rates, geometry validity, coverage deltas) and track month-over-month changes
- Author SQL-driven validation suites to flag taxonomy violations, broken joins, and other data regressions
- Build and maintain usage documentation, data guides, and tilesets for public consumption
- Manage data quality and QA efforts across administrative, places, and address datasets; collaborate with multiple large Overture Maps Foundation member companies

Snap Inc.

March 2018 - February 2024

Data Analyst IV: 2021-2024; Data Analyst II: 2019-2021; Data Engineer: 2018-2019

San Francisco, CA

- Utilized QGIS, Python, and GitHub to validate, modify, and create location data used across Snapchat
- Produced SQLite files for Snap's internal location database and reverse geocoder, which saved \$50M+ annually
- Worked with legal, trust & safety, and product teams to develop moderation policy and training materials
- Managed and trained a team of location moderators and co-managed an international, 140+ member moderation team tasked with evaluating user feedback and editing location data in production environments
- Collaborated with monetization and sales teams to evaluate new features using data-driven insights and statistics
- Used BigQuery, SQL, and Looker to identify root causes of data quality issues and ETL failures, then offered feedback and tooling suggestions to engineering teams to enhance data pipelines and quality
- Responded promptly to data quality issues and policy escalations in GitHub, JIRA, and Zendesk
- Conducted targeted venue data edits and quarterly map tile evaluation

Mapzen

April 2016 - December 2017

Data Gardener

San Francisco, CA

- Procured, verified, and standardized open source data for use in the Who's On First gazetteer
- Used QGIS, Python, and open source tools to manage and update millions of administrative and venue records
- Reported and fixed issues, responded to user feedback and reports, and documented data management techniques, updates, and general information to engage the public about the gazetteer

Self-Employed

September 2014 - October 2017

GIS Consultant

San Francisco, CA

- TripAdvisor: Created country-specific administrative data to reverse-geocode TripAdvisor's location datasets
- National Geographic: Localized place names and analyzed geodata used in atlas and map products
- Mobility Planners: Designed location datasets, produced maps for use in municipal plans, and manipulated census-sourced economic, demographic, and population data for transportation plan overlays and analysis

Apple

March 2015 - April 2016

GIS Specialist (contractor)

Cupertino, CA

- Digitized public transit feature geometries and edited attributes using QGIS and Python
- Helped craft policy standards and design specifications for launches of new Apple Maps transit markets
- Reviewed and validated newly digitized features prior to their release in the Apple Maps iOS application

CONTRIBUTIONS

Who's On First

April 2016 - Present

Core, Active Contributor

<https://whosonfirst.org>

- Use QGIS, Python, and GitHub to maintain and update records and metadata in a global, openly licensed gazetteer of place data sponsored by the Linux Foundation and used in mapping platforms like Uber and Snap
- Source and document new data sources, transform various geospatial data formats, develop standardized schemas, conduct general project maintenance, and track issues

Protected Areas Data

September 2023 - Present

Core Maintainer

Various

- Create and maintain a series of openly licensed location datasets of national and international protected areas, parks, and areas with similar designations
- Work includes end-to-end development of location datasets, project management, and issue tracking

EDUCATION

General Assembly, August 2016

Python Programming Certification

San Jose State University, August 2013

Master of Urban Planning

San Francisco State University, December 2009

Bachelor of Arts, Urban Studies

TECHNICAL SKILLS

Geospatial: *QGIS, GDAL/OGR, Mapshaper, ArcGIS, Shapely/GEOS library, Geopandas library*

Data: *Python, SQL, BigQuery, DuckDB, BBEedit, Looker, Apache Airflow, LakeFS, AWS, Athena, Google Docs Editors*

Project Management: *Git/GitHub, Jira, Zendesk*

References available upon request