UCSC CROWN 102 | OVERTURE MAPS FOUNDATION

AI Agent for Automated Release Validation

Detecting Anomalies in Geospatial Data

Sonia Sarkar

Fall 2025

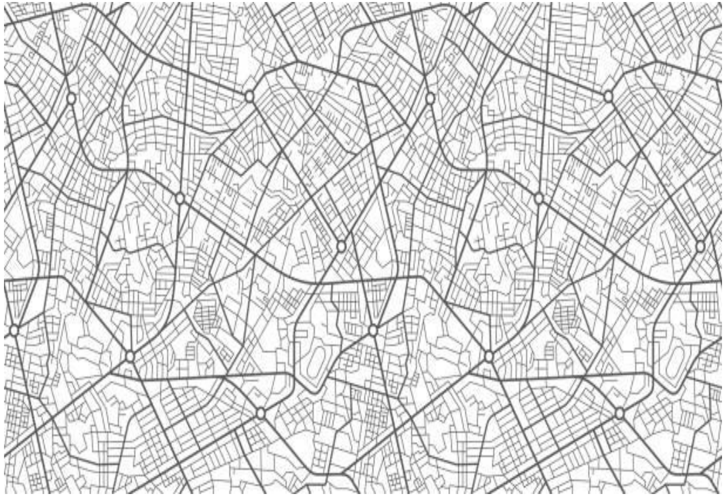
Project Overview

Problem

Each release needs validation for data anomalies: unexpected drops, duplications, quality issues. Manual review does not scale.

Solution

Build an AI agent that automatically detects anomalies, identifies root causes, and provides actionable recommendations.



OKR Journey: What Changed

REMOVED

QA Time Reduction (30%)

No access to review team workflows

False Positive Tracking

Requires multiple release cycles

5 Runs at 98% Consistency

Too strict for project scope

ADDED / REFINED

Reasoning Requirement (90%)

Measurable and deliverable

Confidence Scores (100%)

Clear success criteria

3 Runs at 95% Consistency

Realistic target

KEY LESSON

Initial OKRs were aspirational but unmeasurable. After feedback, I focused on what I could actually deliver and verify within the project timeline.

---

3 Objectives → 2 Objectives

Focused on core value

Approach:

What Problems Does This Solve?

Data Quality Issues

- Feature count drops (data disappearing)
- Feature count spikes (duplicate data)
- Missing required fields (incomplete records)
- Confidence score degradation

Data Stability Issues

- High ID churn (too many adds/removes)
- Category distribution shifts
- New categories appearing unexpectedly
- Existing categories disappearing

Geographic Issues

- Changes concentrated in specific countries
- Regional data source failures
- Geometry length changes (roads, boundaries)

Approach: Hybrid Architecture

LAYER 1

Rule-Based Detector

Threshold-based detection

1004

Raw Anomalies

→

LAYER 2

AI Agent

Pattern synthesis + root cause

5

Actionable Patterns

→

OUTPUT

Dashboard

Interactive visualization

97%

Noise Reduction

Speed vs Intelligence

Rule: 2 sec | AI: 10-20 sec

Cost vs Value

Rule: Free | AI: ~$0.05/run

Reliability

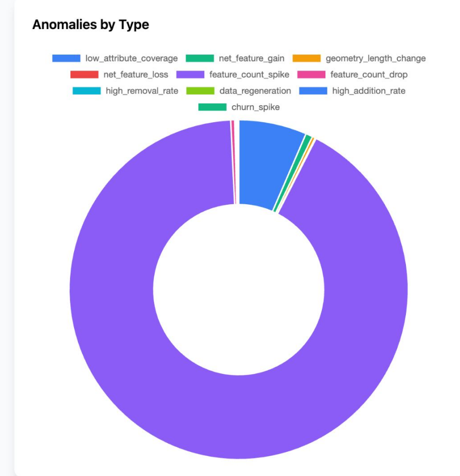
Rule-based fallback when API down

Dashboards

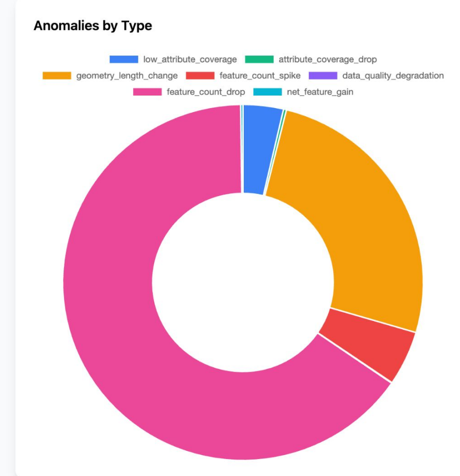
[Anomaly Report (Rule Based)]

[Comparison Dashboard]

8.20.1 → 9.24.0



7.23.0 → 8.20.1



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	nomaly_typ	severity	theme	type	subtype	class	country	column	description	metric	evious_valu	urrent_valu	rcent_chan	ease_curre	ease_previ	pected_ran	itional_con	_duplicatio	ear_total_l	abs_change	maly_category
2	low_attribu	critical	addresses	address				postal_city	99.2% of ac	postal_city	0	0.799704	0	2025-09-2	2025-08-20.1		{'field': 'pos	FALSE	FALSE	0	coverage
3	low_attribu	critical	base	infrastructure				wikidata	99.8% of in	wikidata_c	0	0.212146	0	2025-09-2	2025-08-20.1		{'field': 'wik	FALSE	FALSE	0	coverage
4	low_attribu	critical	base	infrastructure				surface	97.2% of in	surface_cov	0	2.821108	0	2025-09-2	2025-08-20.1		{'field': 'surf	FALSE	FALSE	0	coverage
5	low_attribu	critical	base	infrastructure				height	99.1% of in	height_cov	0	0.914094	0	2025-09-2	2025-08-20.1		{'field': 'heig	FALSE	FALSE	0	coverage
6	net_feature	critical	base	infrastructure					Net gain of	net_change	1.37E+08	1.39E+08	1	2025-09-2	2025-08-20.1		{'added': 15	FALSE	FALSE	1	churn
7	low_attribu	critical	base	land				wikidata	99.7% of la	wikidata_c	0	0.312303	0	2025-09-2	2025-08-20.1		{'field': 'wik	FALSE	FALSE	0	coverage
8	low_attribu	critical	base	land				names	97.7% of la	names_cov	0	2.264823	0	2025-09-2	2025-08-20.1		{'field': 'nan	FALSE	FALSE	0	coverage
9	low_attribu	critical	base	land				elevation	98.7% of la	elevation_c	0	1.312958	0	2025-09-2	2025-08-20.1		{'field': 'elev	FALSE	FALSE	0	coverage
10	low_attribu	critical	base	land				surface	99.8% of la	surface_cov	0	0.249971	0	2025-09-2	2025-08-20.1		{'field': 'surf	FALSE	FALSE	0	coverage
11	low_attribu	critical	base	land				level	100.0% of l	level_cover	0	0.022228	0	2025-09-2	2025-08-20.1		{'field': 'leve	FALSE	FALSE	0	coverage
12	low_attribu	critical	base	land_use				wikidata	99.4% of la	wikidata_c	0	0.598012	0	2025-09-2	2025-08-20.1		{'field': 'wik	FALSE	FALSE	0	coverage
13	low_attribu	critical	base	land_use				surface	95.4% of la	surface_cov	0	4.614979	0	2025-09-2	2025-08-20.1		{'field': 'surf	FALSE	FALSE	0	coverage
14	low_attribu	critical	base	land_use				level	99.9% of la	level_cover	0	0.079215	0	2025-09-2	2025-08-20.1		{'field': 'leve	FALSE	FALSE	0	coverage
15	geometry_l	critical	base	water	lake	lake			Total geom	total_geom	3331.769	2602.096	-21.9	2025-09-2	2025-08-20.1		{}	FALSE	FALSE	21.9	geometry
16	low_attribu	critical	base	water				wikidata	99.5% of w	wikidata_c	0	0.488219	0	2025-09-2	2025-08-20.1		{'field': 'wik	FALSE	FALSE	0	coverage
17	low_attribu	critical	base	water				is_salt	99.7% of w	is_salt_cov	0	0.270038	0	2025-09-2	2025-08-20.1		{'field': 'is_s	FALSE	FALSE	0	coverage
18	low_attribu	critical	buildings	building				names	99.7% of bu	names_cov	0	0.314732	0	2025-09-2	2025-08-20.1		{'field': 'nan	FALSE	FALSE	0	coverage
19	low_attribu	critical	buildings	building				level	99.9% of bu	level_cover	0	0.072704	0	2025-09-2	2025-08-20.1		{'field': 'leve	FALSE	FALSE	0	coverage
20	low_attribu	critical	buildings	building				num_floors	98.2% of bu	num_floors	0	1.790495	0	2025-09-2	2025-08-20.1		{'field': 'nur	FALSE	FALSE	0	coverage
21	low_attribu	critical	buildings	building				facade_colo	100.0% of l	facade_colo	0	0.032618	0	2025-09-2	2025-08-20.1		{'field': 'faca	FALSE	FALSE	0	coverage
22	low_attribu	critical	buildings	building				facade_mat	99.9% of bu	facade_mat	0	0.078051	0	2025-09-2	2025-08-20.1		{'field': 'faca	FALSE	FALSE	0	coverage
23	low_attribu	critical	buildings	building				roof_mater	99.9% of bu	roof_mater	0	0.069808	0	2025-09-2	2025-08-20.1		{'field': 'roo	FALSE	FALSE	0	coverage
24	low_attribu	critical	buildings	building				roof_shape	99.7% of bu	roof_shape	0	0.311569	0	2025-09-2	2025-08-20.1		{'field': 'roo	FALSE	FALSE	0	coverage
25	low_attribu	critical	buildings	building				roof_direct	100.0% of l	roof_direct	0	0.00204	0	2025-09-2	2025-08-20.1		{'field': 'roo	FALSE	FALSE	0	coverage
26	low_attribu	critical	buildings	building				roof_orient	100.0% of l	roof_orient	0	0.037932	0	2025-09-2	2025-08-20.1		{'field': 'roo	FALSE	FALSE	0	coverage
27	low_attribu	critical	buildings	building				roof_color	99.9% of bu	roof_color_	0	0.067358	0	2025-09-2	2025-08-20.1		{'field': 'roo	FALSE	FALSE	0	coverage
28	low_attribu	critical	buildings	building				min_height	100.0% of l	min_height	0	0.000715	0	2025-09-2	2025-08-20.1		{'field': 'mir	FALSE	FALSE	0	coverage
29	low_attribu	critical	buildings	building				min_floor	100.0% of l	min_floor_	0	0.000742	0	2025-09-2	2025-08-20.1		{'field': 'mir	FALSE	FALSE	0	coverage
30	low_attribu	critical	buildings	building				roof_heigh	100.0% of l	roof_heigh	0	0.012639	0	2025-09-2	2025-08-20.1		{'field': 'roo	FALSE	FALSE	0	coverage
31	low_attribu	critical	buildings	building				num_floors	100.0% of l	num_floors	0	0.007132	0	2025-09-2	2025-08-20.1		{'field': 'nur	FALSE	FALSE	0	coverage
32	net_feature	critical	buildings	building					Net loss of	net_change	2.54E+09	2.53E+09	-0.17	2025-09-2	2025-08-20.1		{'added': 44	FALSE	FALSE	0.17	churn
33	low_attribu	critical	buildings	building_part				level	99.1% of bu	level_cover	0	0.918943	0	2025-09-2	2025-08-20.1		{'field': 'leve	FALSE	FALSE	0	coverage
34	low_attribu	critical	buildings	building_part				min_floor	97.3% of bu	min_floor_	0	2.747158	0	2025-09-2	2025-08-20.1		{'field': 'mir	FALSE	FALSE	0	coverage
35	low_attribu	critical	buildings	building_part				roof_direct	97.7% of bu	roof_direct	0	2.321402	0	2025-09-2	2025-08-20.1		{'field': 'roo	FALSE	FALSE	0	coverage
36	low_attribu	critical	buildings	building_part				roof_orient	96.6% of bu	roof_orient	0	3.435883	0	2025-09-2	2025-08-20.1		{'field': 'roo	FALSE	FALSE	0	coverage
37	low_attribu	critical	buildings	building_part				roof_heigh	95.3% of bu	roof_heigh	0	4.664072	0	2025-09-2	2025-08-20.1		{'field': 'roo	FALSE	FALSE	0	coverage
38	feature_co	critical	divisions	division	locality	city	BR		Feature co	id_count	355	712	100.56	2025-09-2	2025-08-20.1		{}	TRUE	FALSE	100.56	feature_count
39	feature_co	critical	divisions	division	locality	city	CA		Feature co	id_count	184	366	98.91	2025-09-2	2025-08-20.1		{}	TRUE	FALSE	98.91	feature_count
40	feature_co	critical	divisions	division	locality	city	CN		Feature co	id_count	3115	6228	99.94	2025-09-2	2025-08-20.1		{}	TRUE	FALSE	99.94	feature_count
41	feature_co	critical	divisions	division	locality	city	EG		Feature co	id_count	101	204	101.98	2025-09-2	2025-08-20.1		{}	TRUE	FALSE	101.98	feature_count
42	feature_co	critical	divisions	division	locality	city	IN		Feature co	id_count	518	1048	102.32	2025-09-2	2025-08-20.1		{}	TRUE	FALSE	102.32	feature_count
43	feature_co	critical	divisions	division	locality	city	IR		Feature co	id_count	129	256	98.45	2025-09-2	2025-08-20.1		{}	TRUE	FALSE	98.45	feature_count
44	feature_co	critical	divisions	division	locality	city	JP		Feature co	id_count	819	1638	100	2025-09-2	2025-08-20.1		{}	TRUE	FALSE	100	feature_count
45	feature_co	critical	divisions	division	locality	city	MX		Feature co	id_count	158	314	98.73	2025-09-2	2025-08-20.1		{}	TRUE	FALSE	98.73	feature_count
46	feature_co	critical	divisions	division	locality	city	PH		Feature co	id_count	150	302	101.33	2025-09-2	2025-08-20.1		{}	TRUE	FALSE	101.33	feature_count

Key Finding: Systematic Bug Detected

BUG IDENTIFIED

Data Duplication in Release 2025-09-24.0

Feature counts doubled (~100% increase) across divisions theme in 189+ countries. Pattern indicates systematic duplication bug in data pipeline.

AFFECTED COUNTRIES (SAMPLE)

US: 1,125 → 2,238

CN: 3,115 → 6,228

BR: 355 → 712

IN: 518 → 1,048

JP: 819 → 1,638

+ more

AI RECOMMENDATION

Investigate data conflation pipeline. Consider rollback of affected release. Verify GERS ID deduplication logic.

Anomalies with

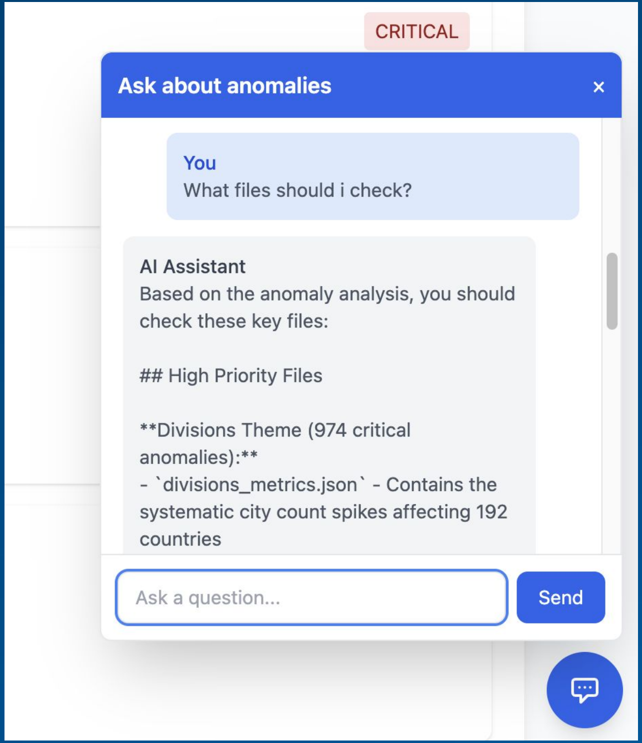
~100%

increase pattern

898

of 1004 total anomalies

Chat


CRITICAL
Ask about anomalies
×
You
What files should i check?
AI Assistant
Based on the anomaly analysis, you should check these key files:
High Priority Files
Divisions Theme (974 critical anomalies):
- `divisions_metrics.json` - Contains the systematic city count spikes affecting 192 countries
Ask a question...
Send

Outcomes: OKR Results

OBJECTIVE 1: Beat Baseline Rule Checker

KR1: RECALL ≥ 85%
~95.7%
All major patterns caught

KR2: PRECISION ≥ 95%
~95.3%
Almost all flags were real

KR3: F1 IMPROVEMENT ≥ 15%
F1 score: 95.5%
See below

KR4: REASONING ≥ 90%
100%
All patterns have root cause

OBJECTIVE 2: Establish Reliability

KR1: CONSISTENCY ≥ 95%
99%
5 near identical runs

KR2: CONFIDENCE SCORES
100%
All findings have 0-1 scores

KR3: FALLBACK ≥ 80%
100%
Rule-based works independently

Why F1 Improvement Is Only Partially Completed

The 15% target assumed the rule-based baseline would have many false positives. In practice, the baseline achieved ~99% precision because the detected anomalies (systematic duplication bug) were real issues. With both systems near 100% F1, large improvement was mathematically impossible. However, the AI agent delivered 99.6% noise reduction (981 → 4), which better captures practical value.

Impact: The Real Value

97%

Noise Reduction

1004 alerts → 5 patterns

Before: Rule-Based Only

Engineer sees 1004 alerts. Spends hours manually grouping. Might miss that they are all the same bug.

After: AI Agent

Engineer sees 5 patterns with root causes. Immediately knows: duplication bug in divisions theme affecting 189 countries. Takes action in minutes.

100%

Root Cause ID

5

Actionable Items

Reflection

What I Would Do Differently

- Begin with a solid understanding of how to read the data
- Run baseline before setting improvement targets. The 15% F1 goal assumed weak baseline - it was actually strong.
- Understand the strength of a rule-based program and focus more on using the agent to synthesize (instead of detect) problems sooner

Future Work

Geospatial Validation

Detect POIs in water, invalid geometries

Pipeline Integration

Auto-run on each release

Spam Detection

Flag spammy names and nonsense attributes

Key Lesson

The right metric matters. Noise reduction (981→4) captured value better than F1.

Be open to restructuring the project to result in better outcomes.

Thank You

sosarkar@ucsc.edu

GitHub:

https://github.com/project-terraforma/Sonia-Anomaly-Detection.git

Sponsors:

Overture Maps Foundation