Maps vs. the ground truth in Phoenix, AZ: ready for autonomous driving?

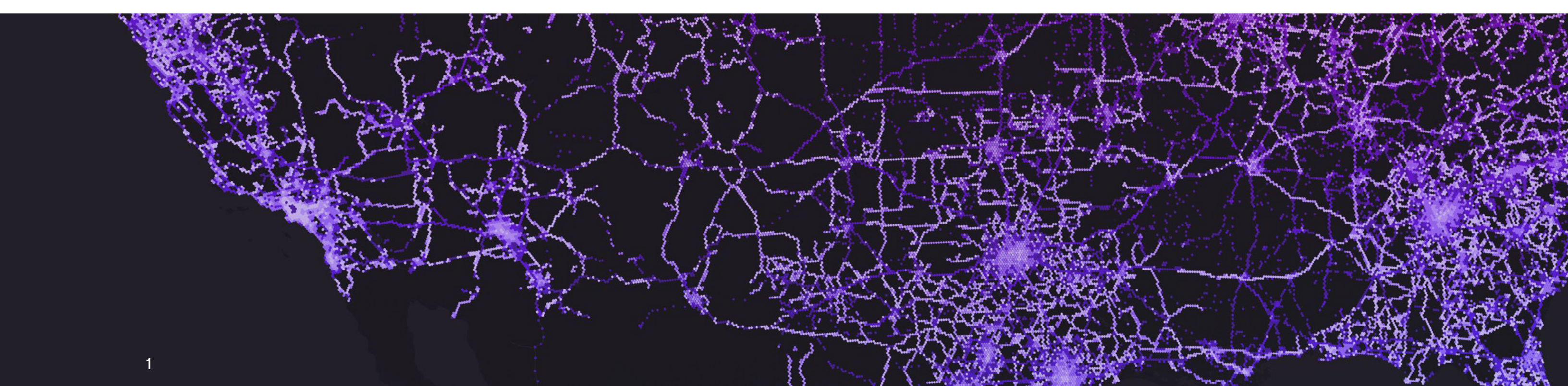
There's a difference between what maps show and what actually exists in the real world. While humans can adjust to missing data, new (software-) map users, from AVs to driver assist functions (such as Intelligent Speed Assist), cannot. While some map layers, such as the base map, seldom change, additional map layers, such as road furniture (e.g. traffic signs), or lane information change more often. To ensure safety for AV applications, the data in all relevant map layers should be highly accurate, with any change quickly detected and updated on the map.

In this ground truth study we compared fresh crowd-sourced vision data about speed limit signs in Phoenix, AZ, with pre-existing speed limit data from the City of Phoenix and OSM. We wanted to check the accuracy of speed limit data vs the ground truth (which is required for AV applications), and infer the rate of change (or error) for road speed limits, which would help define how often map data should be updated or verified.

Creating maps and keeping them updated with the ground truth is an arduous and non-trivial process; yet this is a key requirement for AVs. Current practices by map providers cannot track this level of change nor represent it in a timely enough manner.

Summary and Findings

- Comparing our Phoenix, AZ ground truth data with both City of Phoenix and OSM speed limit data shows an error rate of 3-6%, meaning that speed limits posted in streets are not what appear in map data. This may be due to changes that happened after the data was inserted into the map, or errors in map creation. This correlates with our findings in Las Vegas, NV, based on Nexar's vision network and change detection capabilities, that the annual rate of street sign change is 2.5%.
- Additionally, in the case of OSM, 47% of speed limits didn't appear at all, pointing to a lack of coverage that may also exist with traditional map players.
- Crowdsourcing vision data to "see" ground truth speed limits - as an additional layer on a map, made for driver assist and AVs - is feasible, scalable and useful to ensure map accuracy and change detection.
- Assuming the speed limit layer for driver assist and AVs needs to be 99.5% or higher, as with many automotive applications, AV maps should be updated more often. Using the Las Vegas Nexar numbers shows that updating a map just once a year would result in 97.5% accuracy.



Introduction

As the real world around us changes, maps are predestined to become stale. AVs need to understand the road, and need to become "aware" of any change on the road. Maps will need to reflect change so they can stay accurate, showing everything on the road from work zones to lane markings and road signs, such as speed limits. While others have checked OSM's applicability for rideshare, we wanted to look at its applicability to speed limit data, which is of special importance to the future of self-driving. Speed limits aren't very dynamic, certainly less so than work zones which can change daily or hourly - but they do change. By collecting crowd-sourced data from our vision network and comparing it to OSM data and to City of Phoenix known road speed limits, we wanted to check the accuracy of speed limit data and its rate of change. Understanding the accuracy of both city-sourced data and OSM against the ground truth represents the rate of real-world change (or map error data) and how it impacts the accuracy of maps and their usability for autonomous driving. It also presents a roadmap for the future use of crowd sourced data to create better maps.

Printed-on-paper maps are destined to become stale at the exact moment of their publication, since the ground truth always changes. Mapping giants invest endless resources to update maps in an expensive and laborious process. Continuous coverage is not a viable option for traditional mapping since they either use expensive vehicles for mapping or rely on reports of change; if the data changes (a speed limit was decreased/increased) or is transient (a work zone) it likely won't get picked up automatically. Since map elements don't contain a time stamp it's difficult to know whether they are accurate or not.

Historically, humans used maps mainly to navigate; they wanted them to tell them how to get somewhere, not to understand road rules. Yet, increasingly, maps are not only used by humans, but also by software. AVs and other automated driving scenarios (Intelligent Speed Assist, for instance) need a different kind of map - a map that is detailed, at lane level or better, with a high degree of accuracy, and that reflects transient changes on the road. Autonomous Vehicles need maps that are precise and contain fresh data, at scale. Humans do not refer to posted speed limits on a map while driving, but software does. This means that the question of how much change happens in maps, and who uses it (software) is about to change how we think about maps and their freshness and accuracy.

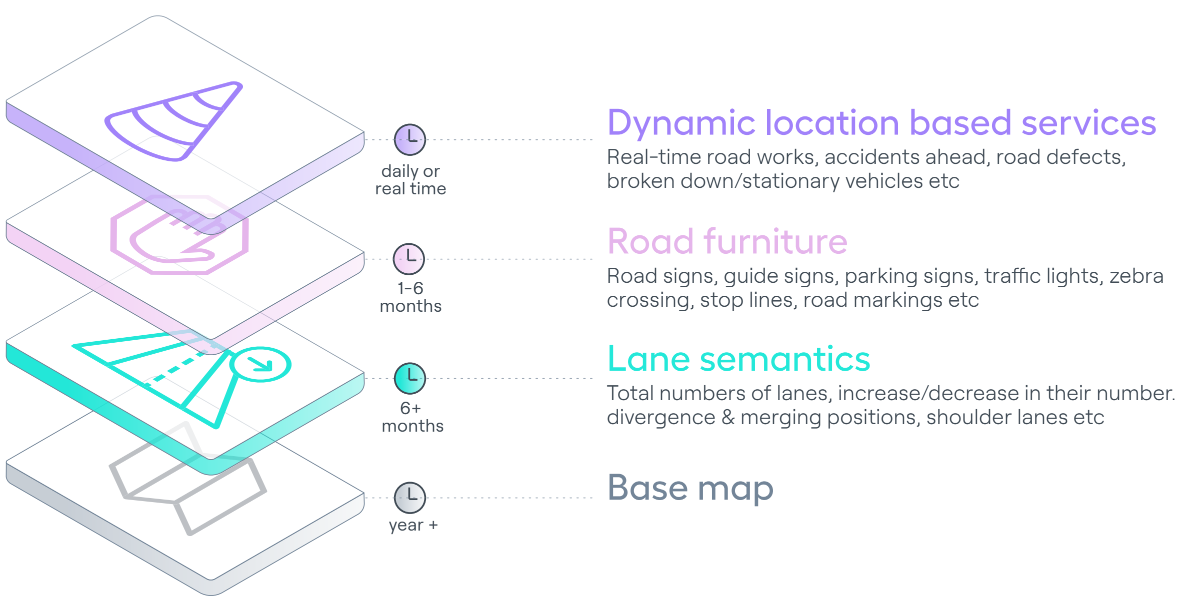
In general, one can identify 4 layers in maps that AVs and driver assist apps will require:

→ Basemap

→ Lane semantics

→ Road furniture

→ Dynamic location-based services



Collecting ground truth fresh data from crowd-sourced vision

Nexar utilizes a network of smart dash cams to collect vision at scale, using low-cost cameras. Using AI (on-device and on the cloud), detections of road elements are made, and this detection data is aggregated and localized, providing a real time (map-agnostic) layer, of transient (such as work zones) or static road elements, on top of an existing map.

Nexar dash cams pay frequent visits to a given area, creating continuous coverage and fresh detections of ground truth changes to the road - while keeping economics feasible. When maps are updated using crowd sourced data, freshness of detections is what determines the accuracy of maps, since this approach continuously detects change and maps it, applying a virtual timestamp on any element on a map, noting when was it last monitored for change.

Here's a diagram explaining crowd-sourced vision at a very high level:

Dash Cams Collect Images

Users contribute anonymous images and videos of the world

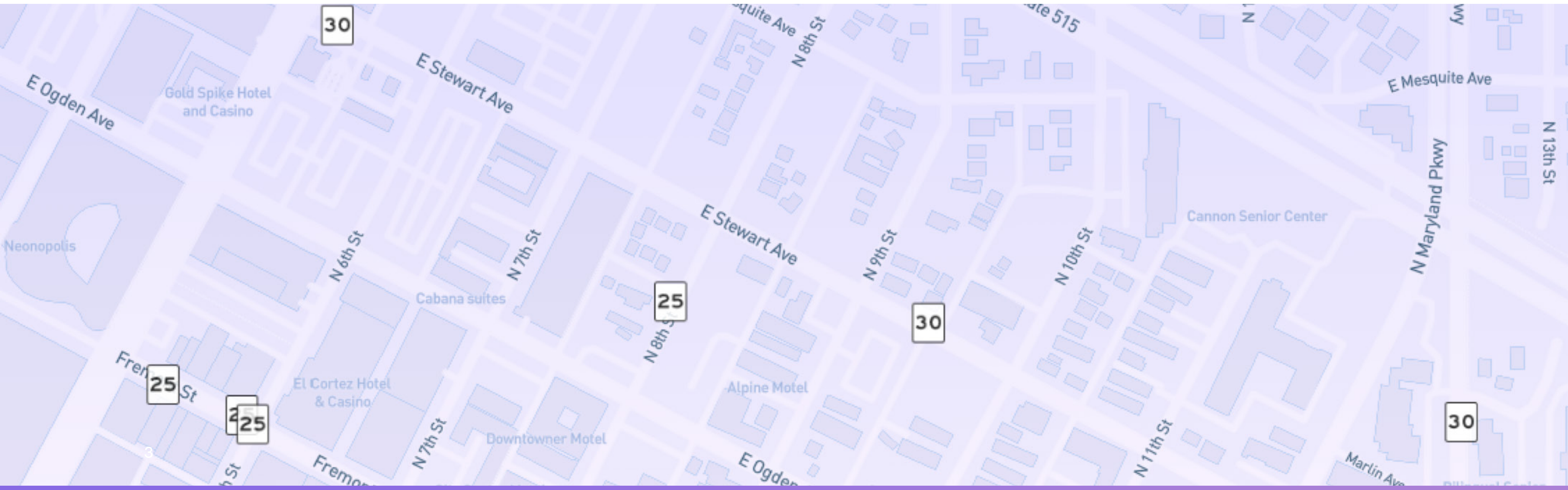
Fresh Images and Detections

Unique collection of imagery and AI to make sense of the data

In Context and Localized

Images and detections are accurately located, for context and time-based change detection

To show crowd-sourced vision in action and the changes it detects, let's take a look at Phoenix. We will compare Nexar's approach to road inventory detection (focusing on street signs) with existing data about Phoenix.



Methodology: How many speed limit signs can Nexar “see” in Phoenix, AZ?

While there have been reports on the accuracy of the OpenStreetMap, its freshness with regards to various road elements, stop signs and speed limit signs wasn’t checked (to the best of our knowledge). In late 2021, we released our analysis of crowd-sourced stop sign detections in Seattle, comparing OSM to Nexar data. This report will present our analysis of road speed limits vs ground truth in Phoenix, AZ, the fifth largest city in the US and a hotbed of AV trials. Speed limit detection is interesting for a variety of reasons. It is required for Intelligent Speed Assist applications (which also require additional capabilities of determining permitted speed where there is no posted speed limit sign, a subject we won’t touch on in this report) and for AVs to better understand the road.

Specifically, the question we were asked was how many speed limit signs can Nexar detect in a given area in Phoenix, AZ. We compared the Nexar associated speed limit information with the road data, with two disparate datasets: one was OSM and the other, which we received after completing the initial OSM analysis, was against the City of Phoenix’s known mapped speed limit data. This post will describe the changes we found vs OSM and vs the City of Phoenix speed limit data. It’s worth noting that all such detections are time-stamped and therefore provide both a validation or audit of the data (against Phoenix data) as well as a method of determining ground truth change.

How well does OSM account for speed limit signs in Phoenix, AZ?

First, we collected Nexar speed limit sign detections made during October 2021, in the area appearing below. What’s interesting is that we collected detections just during one month, using our existing dash cam network.

Using Nexar’s network of connected AI powered dash cams and dedicated data pipeline, we curated an MUTCD_ R2_1 compliant data set of 3,562 max speed limits in Phoenix. We focused on the five speed limit signs (30, 35, 40, 45 and 50 Mph).

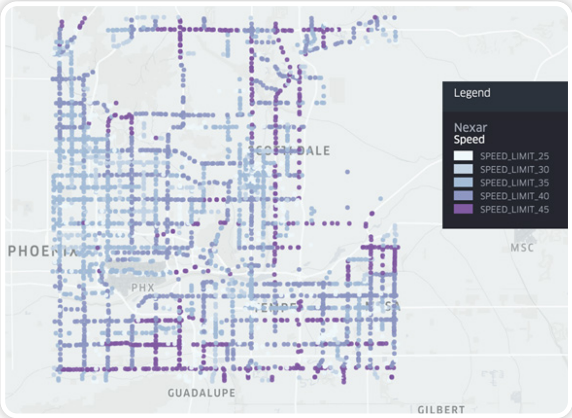
To validate our dataset, we ran our aggregation pipeline including human-in-the-loop editorial feedback, in order to ensure maximum available recall of signs with 100% provided detection precision.

To verify that we did not miss any signs, we annotated 5,450 random frames where no speed limits signs were detected from an evenly distributed sample across different road types, also comparing the frames to Google StreetView images of the relevant areas. In this test, we found that of the 5,450 frames, we had missed just one speed limit sign, giving us a 99.98% true negative rate.

Since Nexar detects signs from frames, and to associate frames with road segments, we snapped each frame from a ride to a specific road segment using a map-matching utility that uses the trace or trajectory of the ride to understand which road segment the car drove on. As part of our aggregation pipeline, we snapped each frame to a specific OSM segment. This allowed us to compare our detections with the metadata information on OSM. In this way we could validate our detected road signs with OSM speed limits.

First, we tested how well frames were snapped to OSM road segments using the gps trace of the ride up to the detection. For this purpose we annotated how many detections out of the same randomly chosen 5,450 frames were snapped to the correct road by comparing the snapped frame location with Google Street View of the same reported location. 94% of frames were snapped correctly.

For the same bounding box, we pulled OSM ways (a total of 13,579) that contain information about the maximum speed limit. We then compared them to Nexar detections, with the results appearing below.



Area bounding box: SW: [-112.07, 33.37]; NE: [-111.82, 33.59]

Comparing Nexar data to Phoenix's speed limit database

As we were finalizing our comparison to OSM data, we also received City of Phoenix's known mapped road speed limits and decided to compare the Nexar data against the Phoenix data. Our next step was to check our detections against data provided to us by the City of Phoenix, essentially auditing their data with real world visual detections made by Nexar. The Phoenix data contained speed limit information for the entire city polygon. We compared Nexar's data to a subset of the data set as follows:

- → The Phoenix dataset provided a total of 9,573 road segments with the speed limit information for those segments
- → Since we'd run the analysis for OSM, we compared the data that was previously analyzed to the relevant part in the Phoenix dataset

How fresh and accurate are OSM and City of Phoenix data?

Compared to OSM road segment data with speed limit data, 47% of Nexar's detections of speed limits did not exist in OSM, with 47% of Nexar detections reflecting data that was different than the data in OSM and 6% of Nexar detections showing a different road segment speed limit than the one stated in OSM (possibly due to a change, or error).

OSM vs Crowd-sourced vision

47% OSM data verified

47% No OSM data

6% OSM data incorrect

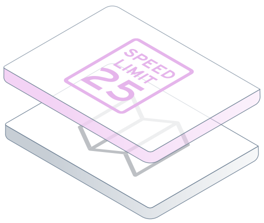
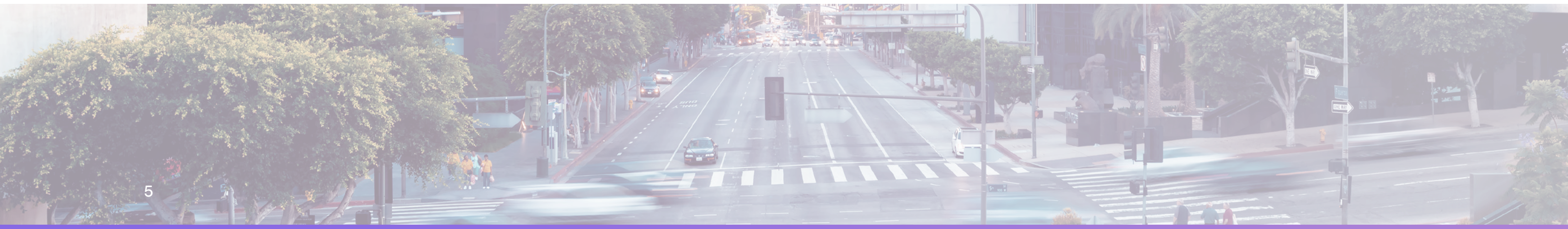


Fig.1: Total signs detected by Nexar with different OSM speed data; Nexar's signs (columns); OSM Speed (rows)

OSM / Nexar	25 mph	30 mph	35 mph	40 mph	45 mph
15 mph	0	22	22	22	6
20 mph	2	0	27	0	0
25 mph	0	8	0	16	4
30 mph	4	3	40	0	12
35 mph	0	0	7	11	0
40 mph	0	0	1	0	6
45 mph	0	0	1	0	3
50 mph	0	0	0	0	0
55 mph	0	0	0	0	0



Compared to the Phoenix dataset, 6% of Nexar's road speed limit data is different from the Phoenix known data, meaning drivers see a different speed limit posted to the area than what appears in the City of Phoenix data. A total of 98 discrepancies were found by Nexar:

- In 57 cases Phoenix data was wrong. We have determined that 3 cases are as a result of temporary work zones.
- In 41 cases Phoenix has no speed limit information although visual data indicates that in reality a speed limit sign is posted in the area. Most of the signs were 40 Mph signs.

City data vs Crowd-sourced vision

3% No data

3% Wrong data

94% Verified

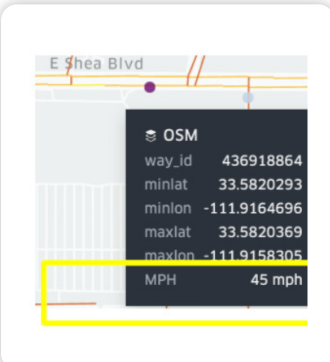
Fig.2: Total signs detected by Nexar with different OSM speed data; Nexar's signs (columns); Phoenix Speed (rows)

		Nexar			
	Speed (mph)	**30**	**35**	**40**	**45**
Phoenix	30	0	9	0	0
	35	10	0	0	3
	40	1	8	0	3
	45	0	2	17	0
	50	0	0	0	1

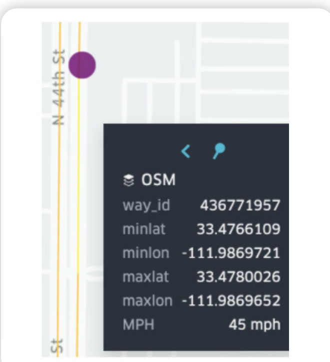
Some examples

Here are two examples of OSM data discrepancy, where the OSM road speed limit data was different than the actual sign as detected by Nexar:

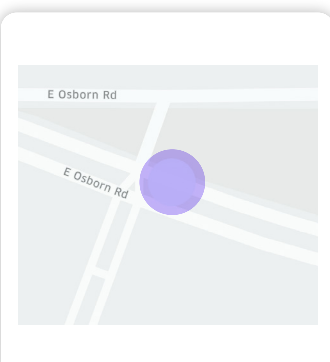
Example 1: OSM has outdated information



Example 2: OSM has outdated information



Example 3: Work Zone and fresh speed limit signs detected by Nexar; not detected by OSM



Example 4: OSM has no speed limit information

Example 5: OSM has no speed limit information

A short note on temporary detections

Some of the differences between OSM and Nexar stem from the fact that in addition to OSM's data being very outdated, Nexar's data is fresh enough to capture temporary road signs, even signs held up by a highway employee to signal road closure. The freshness of our data coupled with our change in detections capabilities is allowing us to see the changes in the road that result in new and temporary road signs, like these examples of road work zones:

Example 6: Work Zone and fresh speed limit signs detected by Nexar; (OSM 45 mph)

Conclusion

For the longest time, our requirements from maps were relatively lax. For lack of a better method to adjust to changes in ground truth, we settled for maps that are stale by definition, as long as the base layer was relatively correct and the traffic data up to date. This fits with how humans use maps. In the era we are entering, the maps of yesteryear just won't cut it anymore. As maps move from being a convenience to becoming the cornerstone of ADAS and AVs, the levels of precision, freshness and completeness that are required grow exponentially, and as a result, the core business process of map creation needs to be overhauled to better reflect the ground truth.

The most interesting part of this analysis is that the speed limit data from both City of Phoenix and OSM shows an error rate of 3-6%, meaning that speed limits posted in streets are not what appear in map data. This may be due to changes that happened after the data was inserted into the map, or errors. Additionally, in the case of OSM, 47% of speed limits didn't exist at all, pointing to a problem of coverage (on top of the accuracy issues) that may also exist with traditional map players, where there is a known problem of residential coverage. Nexar's change detection shows that in Las Vegas, road signs change at a rate of 2.5% annually.

This report used data from just one month, October 2021. This means that using crowd sourced vision data to map speed limits is feasible as a way of effectively auditing map data and changing it to reflect the ground truth.

Assuming the speed limit layer for driver assist and AVs needs to be 99.5% or higher as with many automotive applications, maps should be updated more often. Even using the Las Vegas Nexar numbers shows that a once a year map update would only result in 97.5% accuracy.

In short, to build the maps of the future - the digital twin of the physical world- we need to embrace crowd-sourced vision and change detection. At a change level of 2.5 to 6% questions of how to ensure maps are fresh should be taken seriously by the industry as a whole.