Optimizing Evacuation Routes

By Identifying Blocked Routes in Real Time

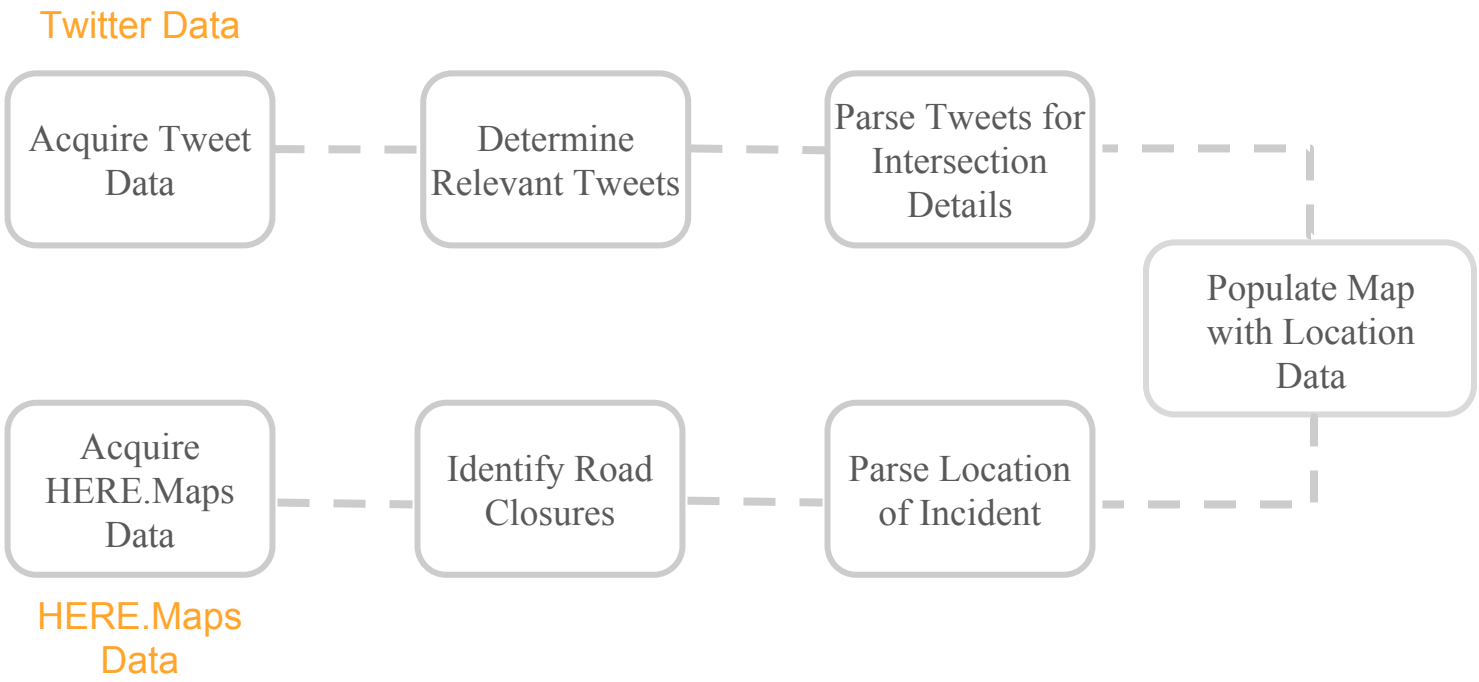
Brian Brakefield
Veronica Giannotta
Bala Krishnamoorthy
Amy Taylor

1/18/2019

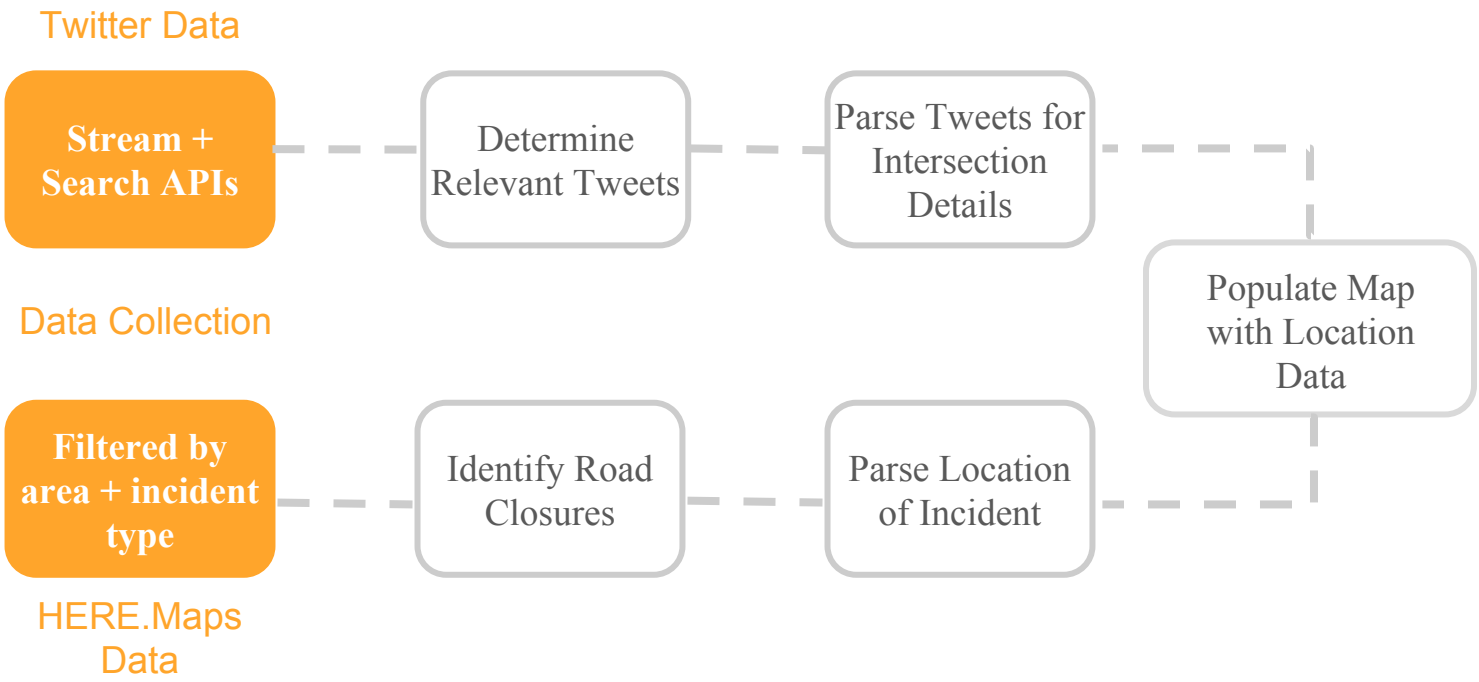
During emergency situations, every second matters.

The latest GIS and navigation systems can calculate optimal routes to a given destination. However, few current platforms rely on real-time data to identify blocked routes and damaged roads. We sought to solve this problem.

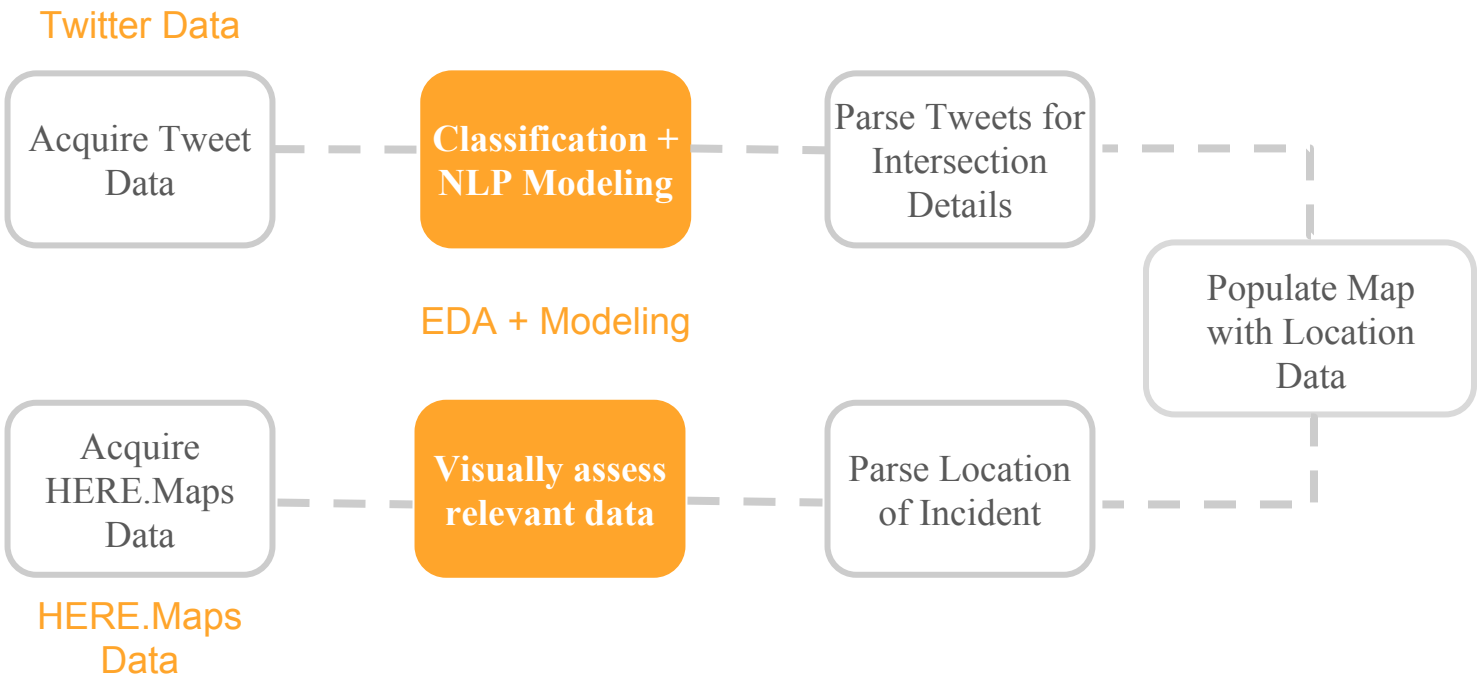
Data Pipeline



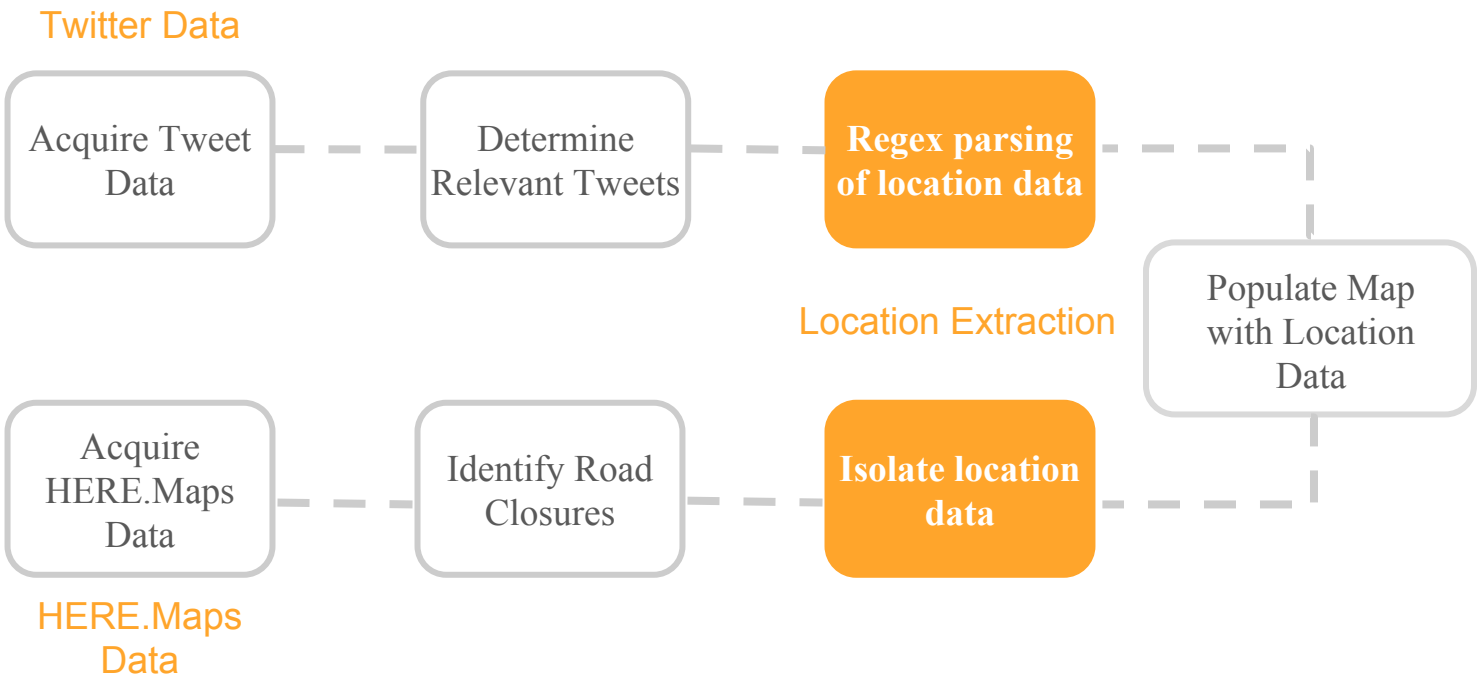
Data Pipeline



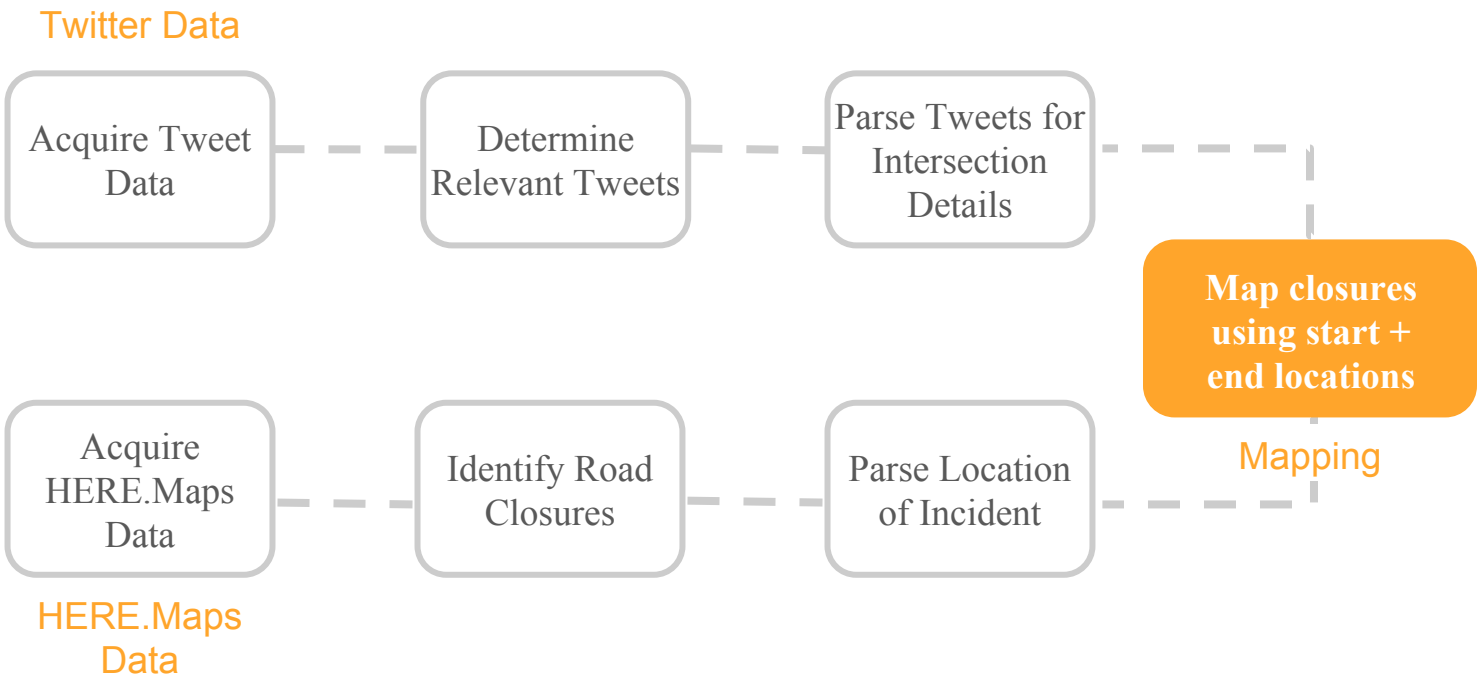
Data Pipeline



Data Pipeline



Data Pipeline



Twitter Data Collection

Search API (“historical tweets) vs. **Stream API** (live feed)

- Use Tweepy **Cursor** method
- Max history = **1 week**
- **Rate limit** = 45,000 tweets/15min
- For broad searches, sample is **random**

- Use Tweepy **Stream Listener** class methods
- **No limit** (live stream) BUT can’t build a queue

Twitter Data Collection

Search API (“historical tweets) vs. **Stream API** (live feed)

- Can **filter on location**

- **Cannot filter** by location

Twitter Data Collection

Search API (“historical tweets) vs. **Stream API** (live feed)

- can **filter on location**

- **cannot filter** by location

q = “road closed”

- Returns all tweets containing both “road” and “closed” anywhere in the text
- “Same” result for Search or Stream

USA place id

q = “place:1c69a67 road closed”

Returns all past tweets, located in the USA that contain “road” and “closed”

USA long/lat coordinates

q = “road closed”, locations = [-125,24, -66, 48]

Returns all incoming tweets located in the USA OR tweets that contain “road” and “closed”

NLP and Classification Modeling | Determine Relevant Tweets

- **143 tweets labeled**

Classification	Example
0 = unrelated or useless	MVHS will remain closed tomorrow due to concerns about road conditions.
1 = Related, but road not FULLY blocked	Road construction. right lanes closed in #Pima on I-10 EB at Ruthrauff Rd
2 = Relevant, road is fully blocked AND streets provided	Closed due to road construction in #FortWorth on 35W SB at Pharr, stop and go traffic back to Northside Dr

- **Train test split**
- **CountVectorize**

Baseline Accuracy = 40%

- 40%(2), 38%(1), 22%(0)

- **Classify predictions using Logisitic Regression**

Results

Training data = 100% accurate
Test data = **90% accurate**

Model is overfit, but performing okay for now.

NLP and Classification Modeling | Analyzing Incorrect Predictions

	Classifications	
Tweet	**Actual**	**Predicted**
Cougar fans traveling to Lakeland for the Girls NECC championship game. State Road 9 is closed in Wolcottville due to a structure fire, plan accordingly and find an alternative route	2	1 = related (not blocked)
Manchester road off Wayne avenue is closed off by police. Giving traffic delivery updated as I see them	2	0 = unrelated
Southbound 101 freeway still closed at lost hills. Agoura Road is jam southbound. Heading to the Civic arts Plaza f...	1	0 = unrelated

x3 other tweets

- Model minimizes false positives
- No related or useless tweets are being incorrectly labeled as relevant

Twitter Data | Location Extraction from Tweets

Step 1 — For tweets describing closed roads, visually assess patterns in text

Step 2 — Use regex to search patterns, and extract relevant street names and intersections

Step 3 — Combine parsed data into "start" and "end" intersections, and populate a dataframe

Step 4 — Feed dataframe of "start" and "end" intersections into the mapping tool

Closed due to road construction in #Southside on S Sam Houston Tollway EB between Hwy 288 and Cullen #traffic

S Sam Houston Tollway EB | Hwy 288 | Cullen

Start: S Sam Houston Tollway EB & Hwy 288

End: S Sam Houston Tollway EB & Cullen

HERE.Maps Data

Filters applied:
1) Geographic area
2) Incident type

Dig through nested dictionaries in JSON output to find relevant columns

Create a dataframe of closed roads, with location coordinates

Populate map using location coordinates

Sample dataframe output

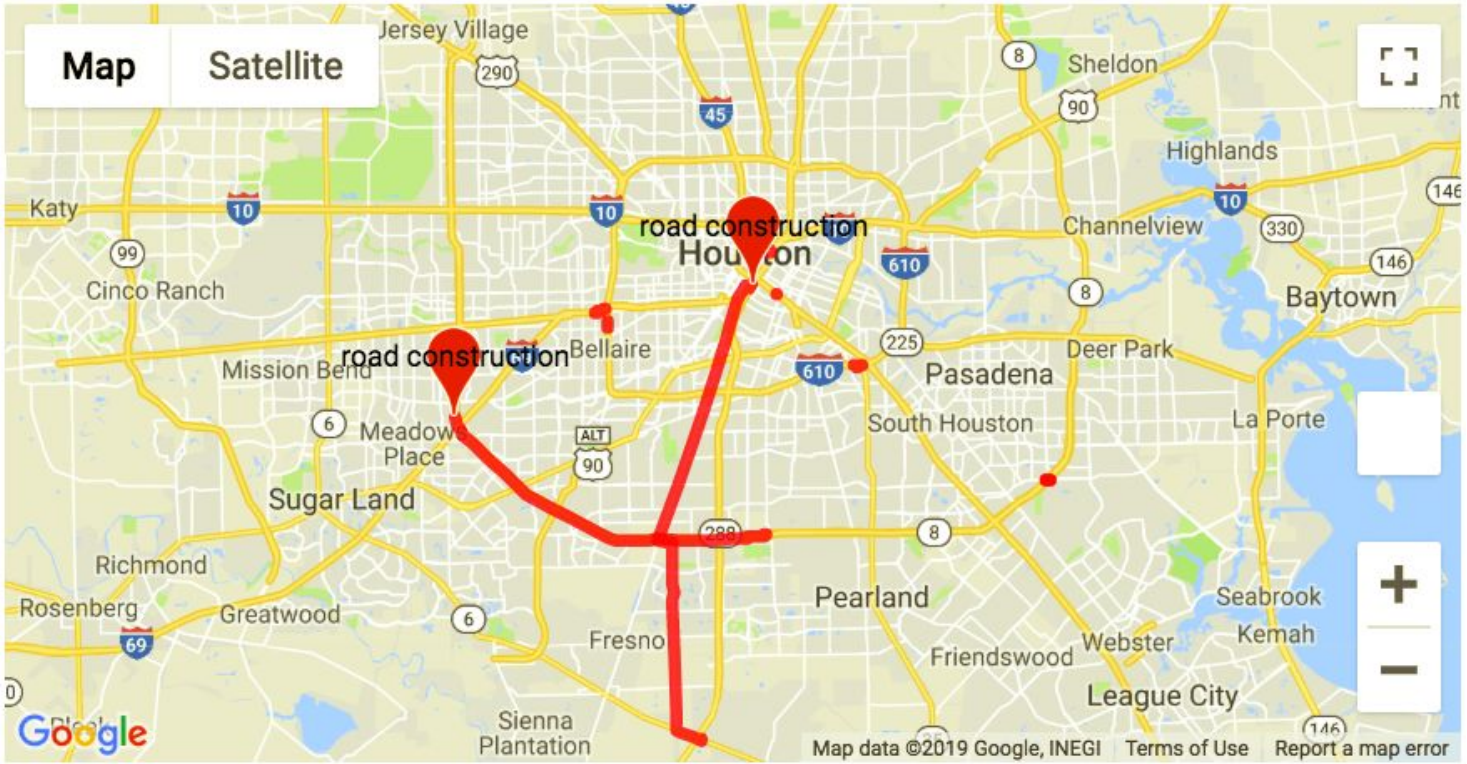
	city	source	criticality	intersection	start_lat_long	end_lat_long	start_time	end_time	entry_time
0	Houston, TX	Here Maps API	critical	{}	(29.6307, -95.16875)	(29.630614, -95.171866)	01/17/2019 19:44:58	01/17/2019 20:12:10	01/17/2019 19:44:58
1	Houston, TX	Here Maps API	critical	{}	(29.7369, -95.34869)	(29.73713, -95.34964)	01/17/2019 19:03:41	01/18/2019 00:02:45	01/17/2019 19:03:41
2	Houston, TX	Here Maps API	critical	{'ORIGIN': {'ID': '', 'STREET1': {'ADDRESS1': ...	(29.92321, -95.62916)	(29.92392, -95.62903)	01/17/2019 19:20:14	01/18/2019 00:19:40	01/17/2019 19:20:14

Mapping | Rendering Road Closure Data

Step 1: Geocoding Twitter intersections

Mapping | Rendering Road Closure Data

Step 2: Rendering with Jupyter GMaps



Future Iterations

Enable real-time tweet processing.

A trained model can filter through tweets in real-time to identify and map emerging road closures.

Fully integrate the application.

Connect scripts to allow all processes to run independently.

Collect more tweets.

Number of tweets collected should be in the 100s or 1,000s to have sufficient training data.

Improve NLP and Regex performance.

Train the NLP model to feed “useful” tweets to the Regex algorithm, and optimize the model to extract street names, intersections, and other location data without the aid of Regex.

Improve mapping feature.

Improve precision, increase maximum number of closures displayed, and deploy aesthetic improvements. Optimize alternate routes based on total travel time.


ROAD
CLOSED

Q + A


Thank You.