

Dispatch Radio Mapping

Grant Wilson, J. Hall, Gabriel Perez Prieto





Agenda

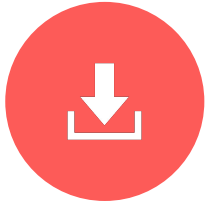
- Problem Statement
- Data Collection and Processing
- Audio Analysis
- Mapping
- Conclusion
- Next Steps



Problem Statement

- Currently, FEMA identifies areas that require immediate attention (for search and rescue efforts) either by responding to reports and requests put directly by the public or, recently, using social media posts.
- This tool will utilize live police radio reports to identify:
 - Hot spots representing locations of people who need immediate attention
 - Flag neighborhoods or specific streets where the police and first-respondents were called to provide assistance related to the event

Data Collection and Processing



DOWNLOAD
AUDIO



PARSE AUDIO



SPEECH TO TEXT



EXTRACT STREET
NAMES

Download Audio Files



● BroadCastify - San Francisco City Police Dispatch Feed

- "The world's largest source of Public Safety, Aircraft, Rail and Marine Radio Live Audio Streams"
- BART - BroadCastify Archive Tool
 - Selenium Library for Scraping
 - Feed id
 - Dates
- Save MP3 audio files





Audio

- Converting the audio to text for NLP was a challenge, given the quality of the source audio.
- Police are typically active when radioing dispatch, at varying distance from the microphone, with significant noise from
 - movement,
 - Environment,
 - static,
 - Interference
- Many of the transmissions were difficult for a human to interpret without domain expertise in police coding and terminology.
- Posed a significant challenge for the speech to text



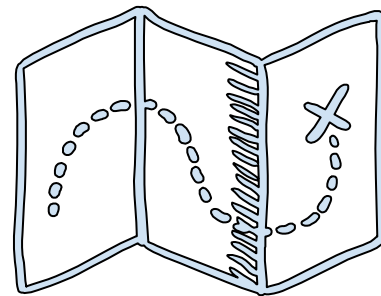
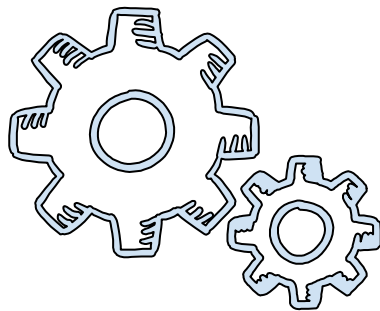
Audio

- Noise
 - fix audio recorded in noisy environments or with poor equipment to create consistency across your recordings.
- Sibilance
 - This is a characteristic of harsh consonant sounds like "s", "sh", "x", "ch", "t", and "th".
- Tone
 - Police radio transmits audio on a very thin range of sound, sounding tinny
 - Help with equalization to shape the audio from your recorded files to match a listener profile and audio that the text-to-speech model was trained on.



Audio

- Live Audio
 - Disaster mapping in Real Time
- Input can be microphone or direct line in from a radio



Parse and Segment Saved Audio Files

- Break Audio into Chunks
 - Pydub Library
 - Detect Audio - Average File Loudness
 - Break if Silence > 4s
 - Convert to WAV
 - 6 s < File < 60 s



Speech To Text

- Google Speech to Text API
 - Transcribe Audio Chunks
 - All streets of San Francisco as a Parameter
 - Give context to Google's API to search for
- Average Confidence on Transcripts
 - 79%



Extract Street Names

- Spacy Library

- Get Street Names from Transcript
- All streets of San Francisco as a Parameter
 - Give context to search for

- USAddress Library

- Get Address Numbers

- Google Maps Geocoding API

- Join Address Number and Street Name
- Get Latitude and Longitude



Example 1

- ‘David once again for the night **1910** we **1404 15th Street** **treat** everyone sits out of can I just want to play for yeah but go back to the temporary place and the Seven Roberts Roberts zebra **211** should come back to **black** Infinity first like I was like a regular **California** place it's a paper plate but I have several Robert Roberts it was fabulous party Wright place we try it was he that does go back to a **2004** is that the entity that correct yeah okay’
- Transcript Confidence = 0.8180
- Model Results:
 - 1404 California Ave, **1404 15th Ave**, 1404 Treat Ave, 1404 Black Pl

Example 2

- 'case manager Michelle is **97 121 Leland** a client is experiencing events induced psychosis getting increasingly activated and aggressive and as much customers are impossibly self-induced you doing it for back to the your doesn't get out of here with that license plates some of that about which car are you referring to the **1030** at **1141** of our courts I copy'
- Transcript Confidence = 0.8699
- Model Results:
 - **1141 Lealand Ave**

Example 3

- 'this verse **29** that teacher Elementary is cold report of a **311** occurred last Friday on the **Hudson Street** side they have pictures not online with these are the secretary okay I'll **98** had a **71** for the **600**'
- Transcript Confidence = 0.8534
- Model Results:
 - **98** **Hudson Ct**



Mapping

- Folium vs. Google Maps API
 - More creativity in mapping
 - Better represented the data as it relates to key interests of disaster response agencies
 - Surrounding businesses, parks, and bus stops
- Historic San Francisco Crime Data
 - Similar to disaster reporting
 - Time/Date/Location
 - Type of dispatch
 - EMS
 - Police
 - Fire



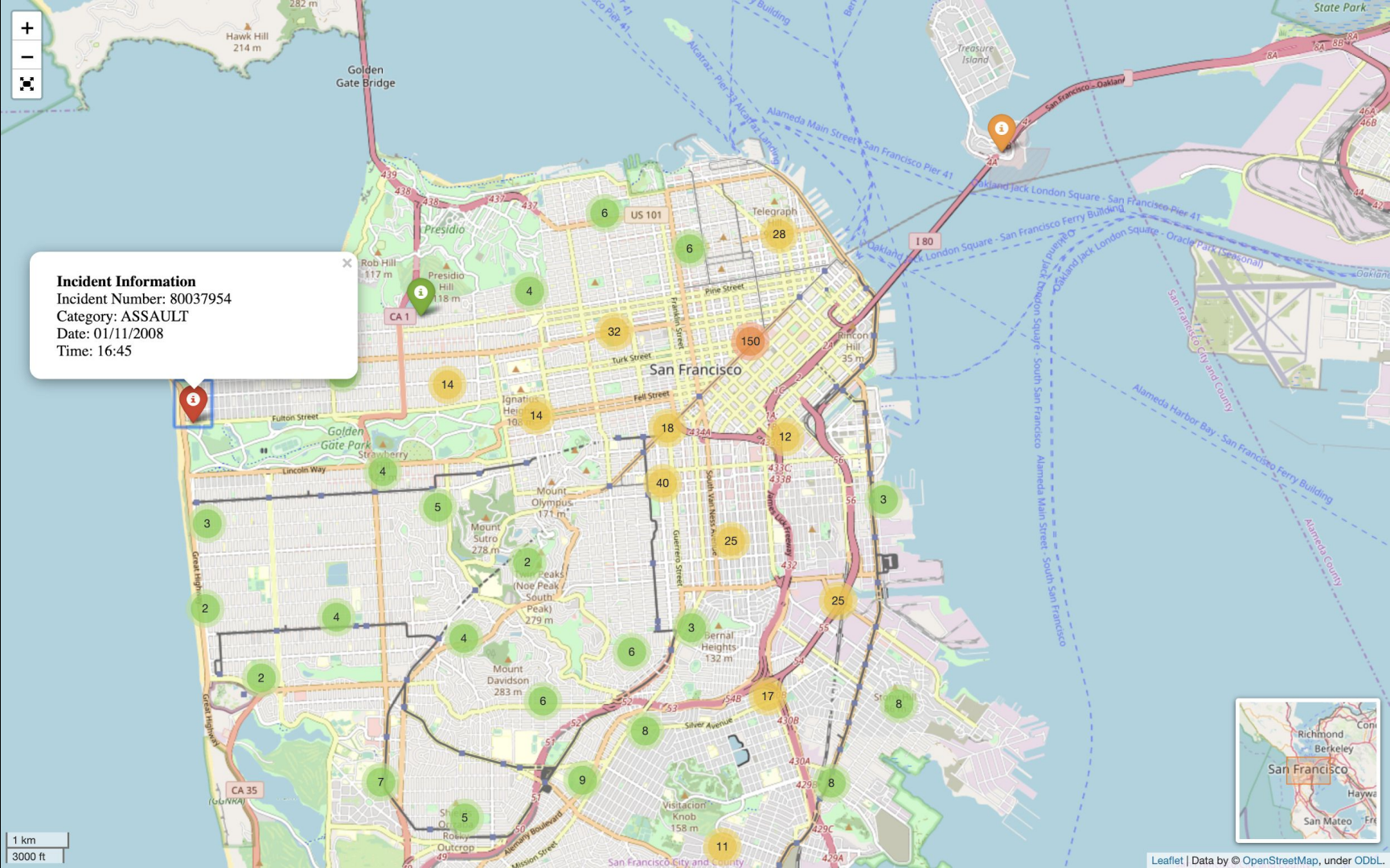
Mapping

- Two maps
 - Individual Detail
 - Point Mapping
 - Neighborhood-area reporting
 - Frequency Mapping



Point Mapping

- Drops “pins” for each reported incident
 - Incident Number
 - Incident Time/Date
 - Response Category
 - Police
 - Fire
 - EMS
 - Incident Category
 - Assault
 - Downed Power Line
 - Fire



Incident Information

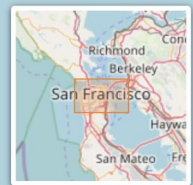
Incident Number: 80037954

Category: ASSAULT

Date: 01/11/2008

Time: 16:45

1 km
3000 ft



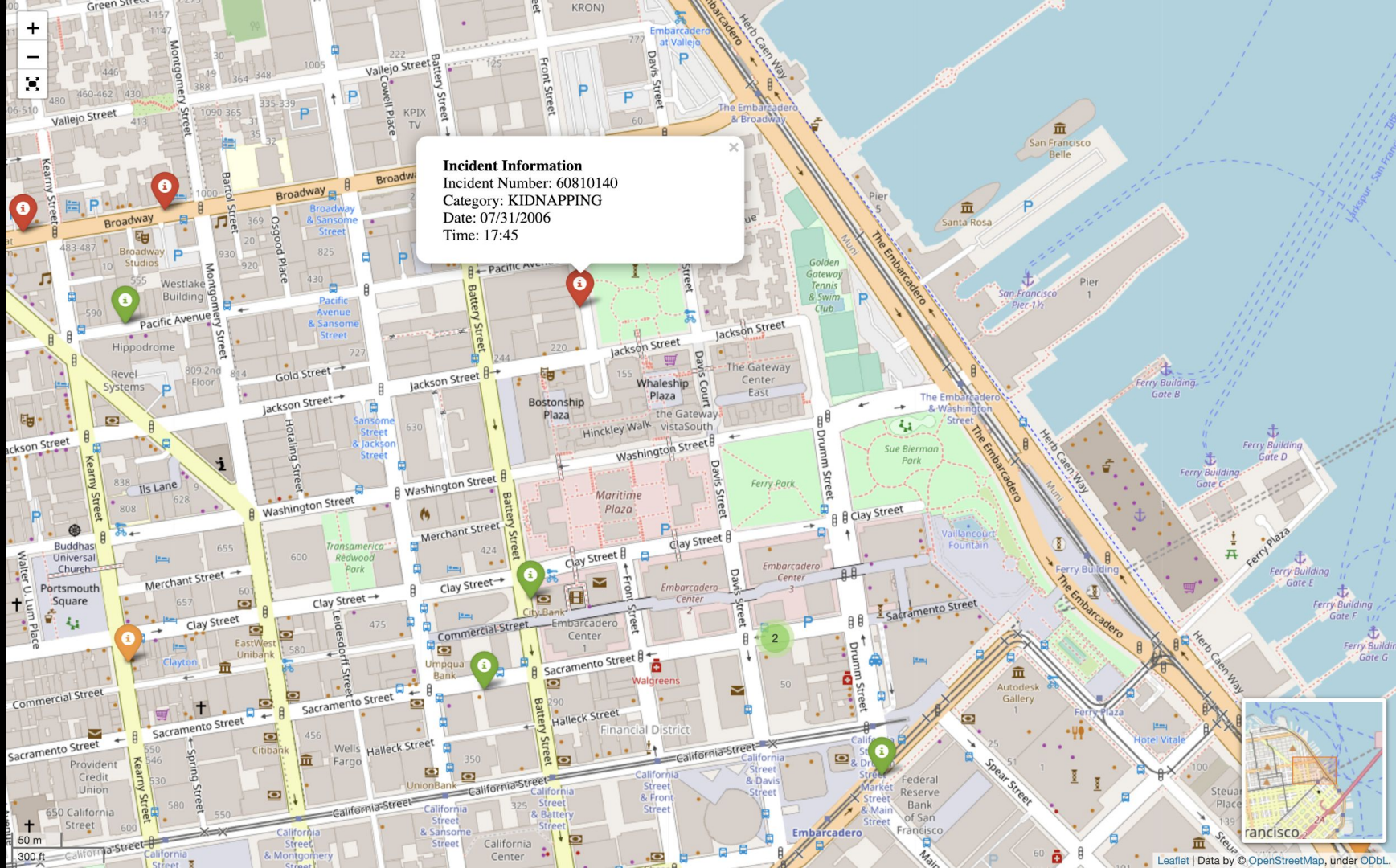
Incident Information

Incident Number: 60810140

Category: KIDNAPPING

Date: 07/31/2006

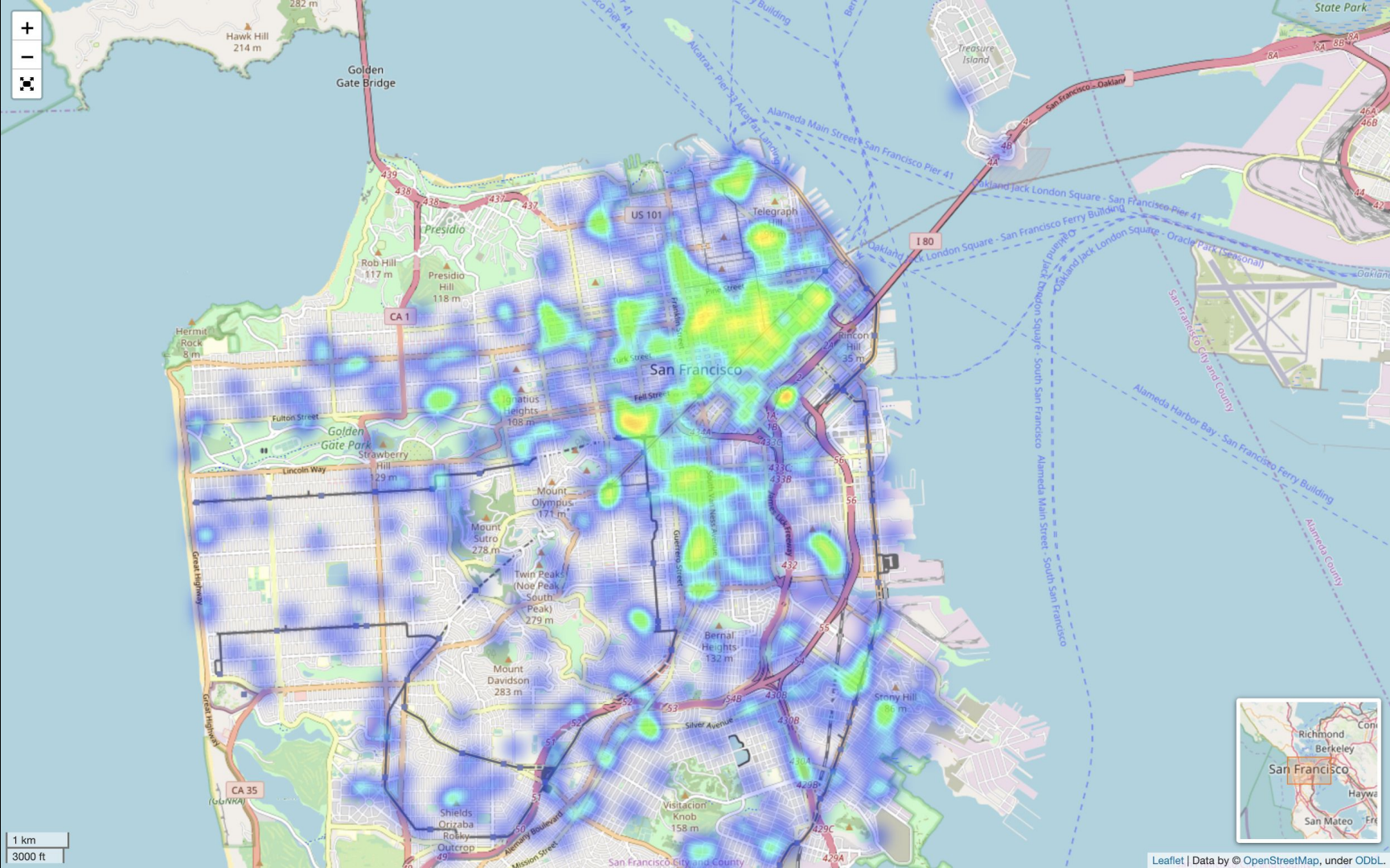
Time: 17:45

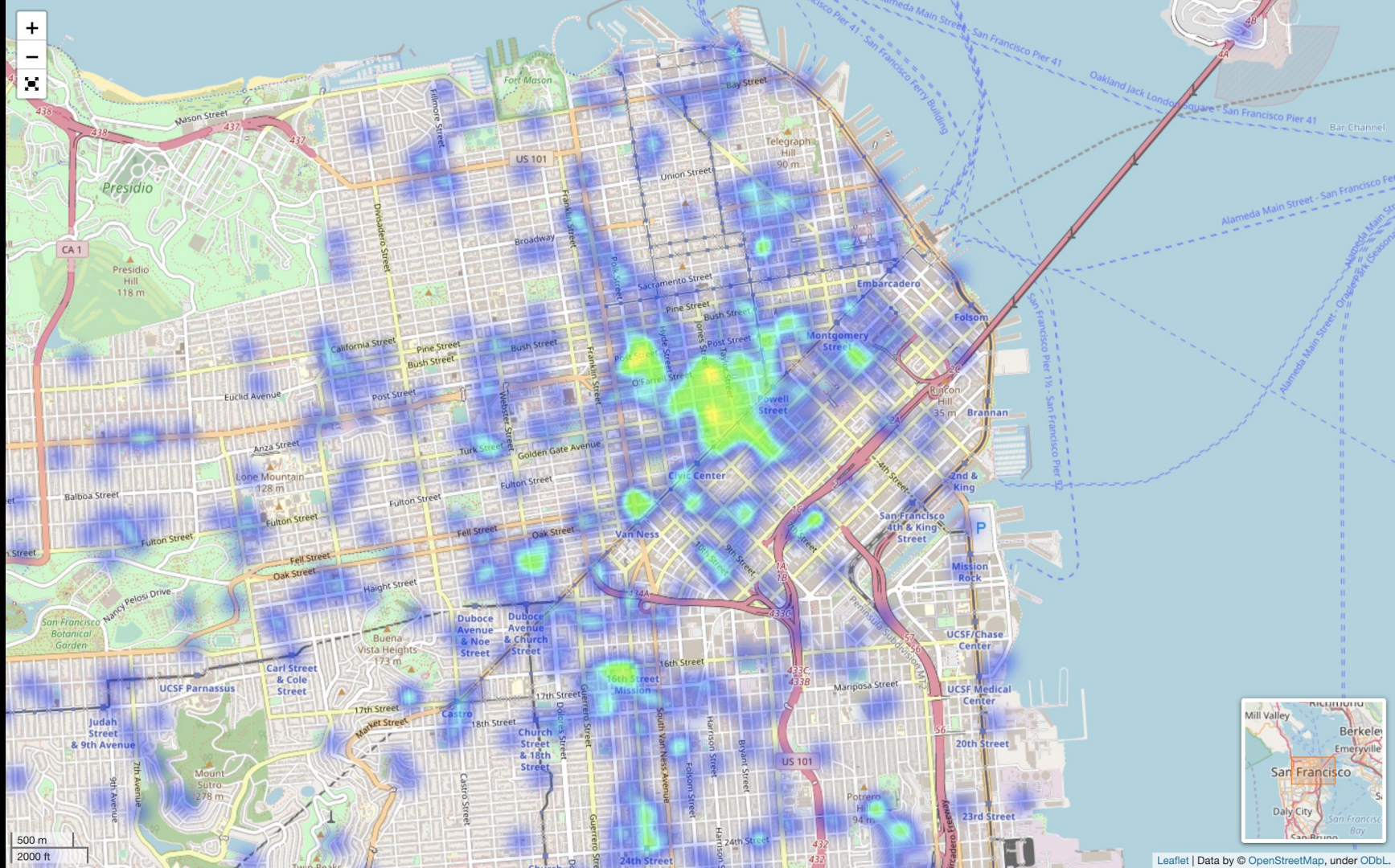




Frequency Mapping

- Neighborhood-wide incidents
- Quick read on areas requiring assistance
- Balances with detail-oriented Point Mapping
 - Demonstrates call-concentration
 - Neighborhoods needing disaster-assistance







Conclusion

- Improvements
 - Audio Analysis
 - Live Audio
 - Point Mapping
 - Frequency Mapping
- Utility
 - Identify key neighborhoods
 - Facilitate native integrating with dispatch radio
 - Cleaner address analysis
 - Isolate Critical Infrastructure weaknesses
 - Neighborhood analysis of incidents



Next Steps

- Clean audio before processing
 - Dolby
- Improve NLP
 - Sentiment/Urgency Analysis
- Address Extraction
- New avenues for speech to text
- Include damage assessment