Identifying High-Threat Locations from Police Radio

Anna Haas
James Lovejoy
Matt Clifford
Will Holder

April 26, 2018

The Problem

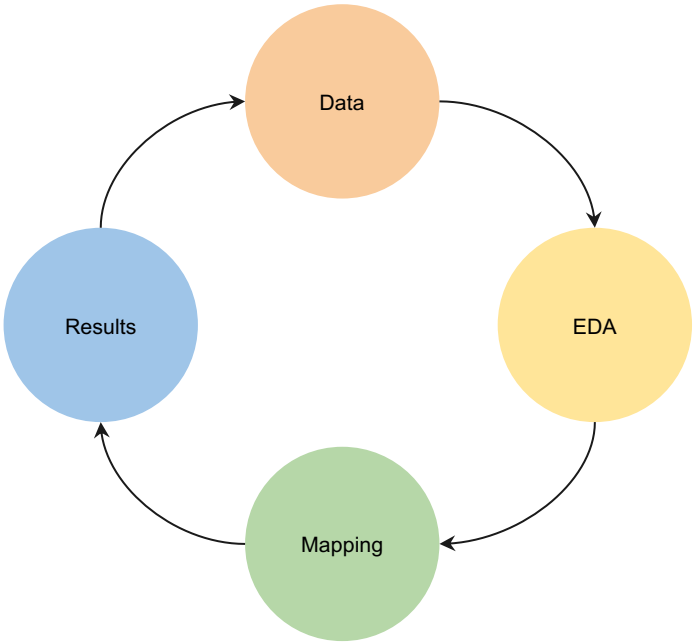
During a time of disaster, can audio from police radios be used to identify affected locations and understand severity to help FEMA more effectively allocate emergency resources?

- Map locations and frequency of mentioned locations
- Identify threat terms and frequency to determine severity

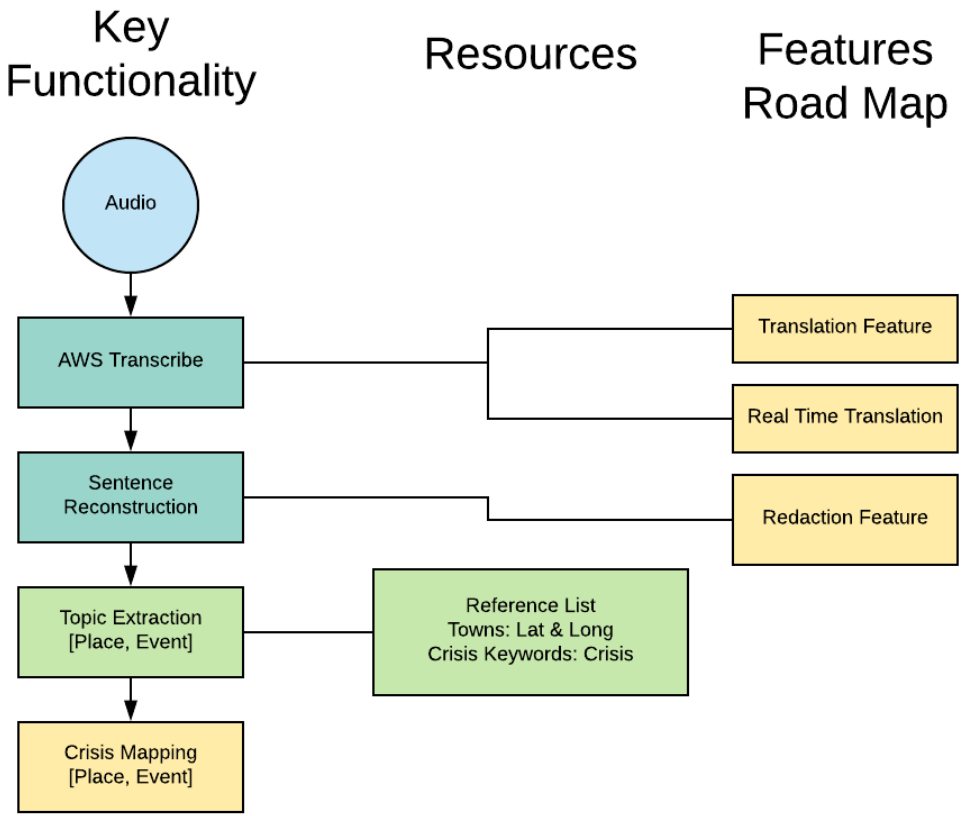
Fema Mission Statement: *“to support the citizens and first responders to promote that as a nation we work together to build sustain, and improve our capability to prepare for, protect against, respond to, recover from, and mitigate all hazards.”*

Overview

- Data Collection
- Data Cleaning
- EDA
- Mapping
- Results
- Future Considerations



Process


Key Functionality
Resources
Features Road Map
Audio
AWS Transcribe
Sentence Reconstruction
Topic Extraction [Place, Event]
Crisis Mapping [Place, Event]
Translation Feature
Real Time Translation
Redaction Feature
Reference List
Towns: Lat & Long
Crisis Keywords: Crisis

Scope and Limitations

- Availability of data
- Potentially most relevant feed (Search & Rescue) was unavailable
- Poor audio quality
- Sentiment analysis was not ideal due to unintelligible transcription
    - Could be implemented into severity categorization
- Police forces moving towards encrypted streams
- NLP considerations (“aboutness”; discourse communities)

The Event

- **What:** The Camp Fire
- **Where:** Butte County, CA
- **When:** November 8, 2018
    - 6:44 AM - 9:12 AM

Data Collection: Audio to Text Translation

- Broadcastify Archives
  - Archival (not live) approach for focused case study
  - Sole resource of this type
- Amazon Transcribe
  - Up and running quickly with web interface
  - Key features: custom vocabulary, speaker labeling (up to 10) per job, per-word confidence
- Process
  - Iterative approach to optimizing transcription parameters
  - Tested transcription with cleaned audio, but did not yield significant improvements

Constructing Dataframes: Word Structure

	content	confidence	start_time	end_time	type	feed	speaker_start	speaker_end	sentence	speaker
0	FIRE	0.7845	8.24	8.52	pronunciation	201811080929-467022-25641_	8.24	9.88	0.0	spk_4
1	years	0.5646	8.60	8.84	pronunciation	201811080929-467022-25641_	8.24	9.88	0.0	spk_4
2	advising	0.9512	8.84	9.32	pronunciation	201811080929-467022-25641_	8.24	9.88	0.0	spk_4
3	ninety	0.9866	9.33	9.59	pronunciation	201811080929-467022-25641_	8.24	9.88	0.0	spk_4
4	seven	0.9827	9.59	9.88	pronunciation	201811080929-467022-25641_	8.24	9.88	0.0	spk_4

Constructing Dataframes: Sentence Structure

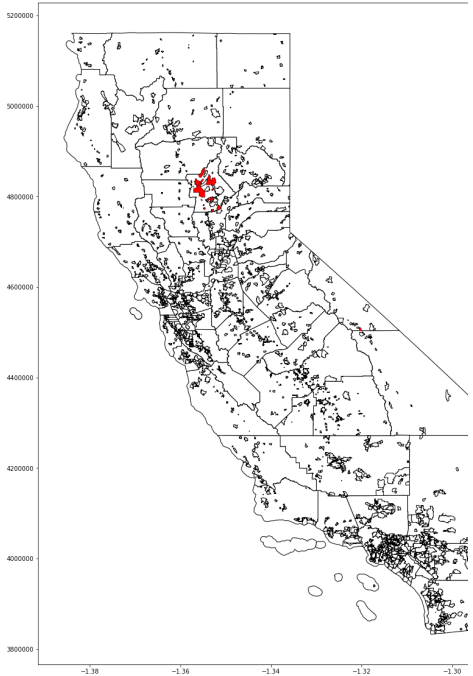
	start_time	end_time	text	avg_confidence	location	feed_name
26	2018-11-08 08:11:57-08:00	2018-11-08 08:11:57-08:00	Affirmative response . Teo CONCOW . That one ,...	0.767456	[OROVILLE, CONCOW, CHEROKEE]	Chico_Paradise_Fire__CalFire
199	2018-11-08 08:12:22-08:00	2018-11-08 08:12:24-08:00	building . There's that . Wait ,	0.661750	[]	Butte_Sheriff_Fire__Paradise_Police
305	2018-11-08 08:12:30-08:00	2018-11-08 08:12:30-08:00	DISPLAYAS fourteen . Thirteen . Can you give m...	0.802148	[GRIDLEY, BERRY-CREEK, CONCOW, NORD]	Butte_Sheriff_Fire__Paradise_Police

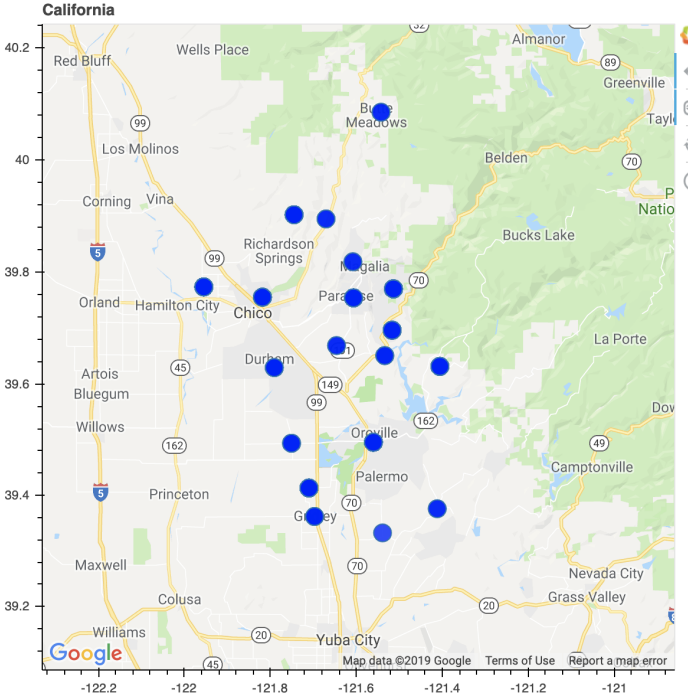
Constructing Dataframes: Mapping

	id_places	latitude	longitude	start_time	end_time	text	confidence	feed	department
0	PARADISE	39.7542	-121.606	2018-11-08 06:59:15-08:00	2018-11-08 06:59:17-08:00	justin maguire is clear and counting down from...	0.806101	Oroville_Police_Fire	BOTH
1	PARADISE	37.4829	-118.602	2018-11-08 06:59:15-08:00	2018-11-08 06:59:17-08:00	justin maguire is clear and counting down from...	0.806101	Oroville_Police_Fire	BOTH
2	OROVILLE	39.4955	-121.56	2018-11-08 07:04:34-08:00	2018-11-08 07:04:34-08:00	left thirty eleven the one you re with and thi...	0.784975	Chico_Paradise_Fire__CalFire	FIRE
3	BERRY CREEK	39.6315	-121.405	2018-11-08 07:04:34-08:00	2018-11-08 07:04:34-08:00	left thirty eleven the one you re with and thi...	0.784975	Chico_Paradise_Fire__CalFire	FIRE
4	CONCOW	39.7703	-121.513	2018-11-08 07:04:34-08:00	2018-11-08 07:04:34-08:00	left thirty eleven the one you re with and thi...	0.784975	Chico_Paradise_Fire__CalFire	FIRE

Mapping: Web Platform Integration

- Tools used:
  - GeoPandas
  - Bokeh
  - Google Maps API
- State, County, Town geographic data
  - California Open Data Portal
- Conclusion: creating a web-based user interface is entirely possible

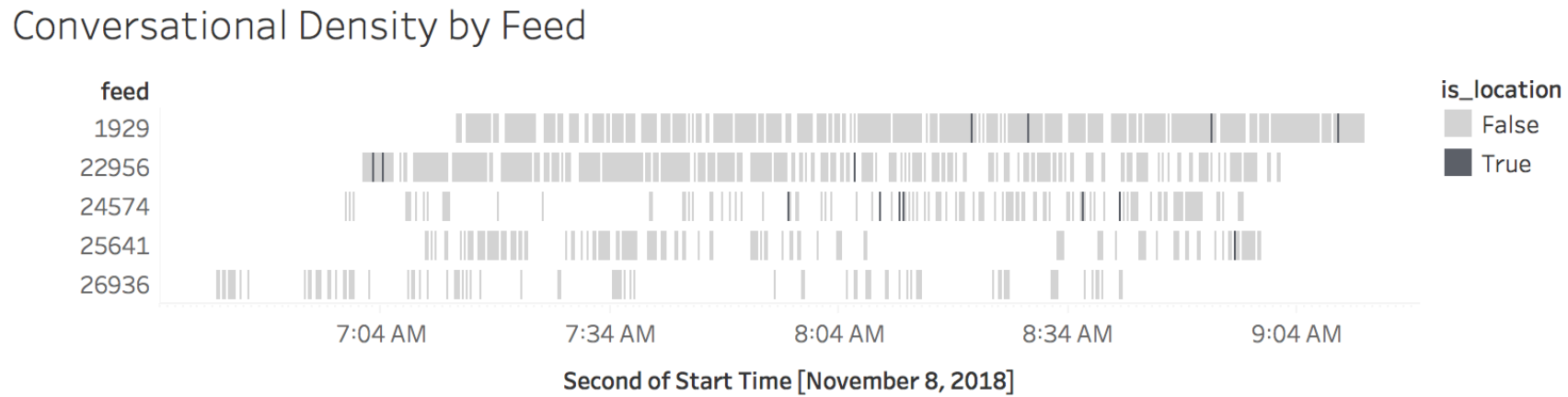




Mapping

- Running total of location mentions double-encoded as ring size and color
- Future features like Makeline may enable clearer indication of information source/flow
- Mention-based indicators are limited.
- Granularity: street-level challenges.
- Meaning-based processing could add significant value.

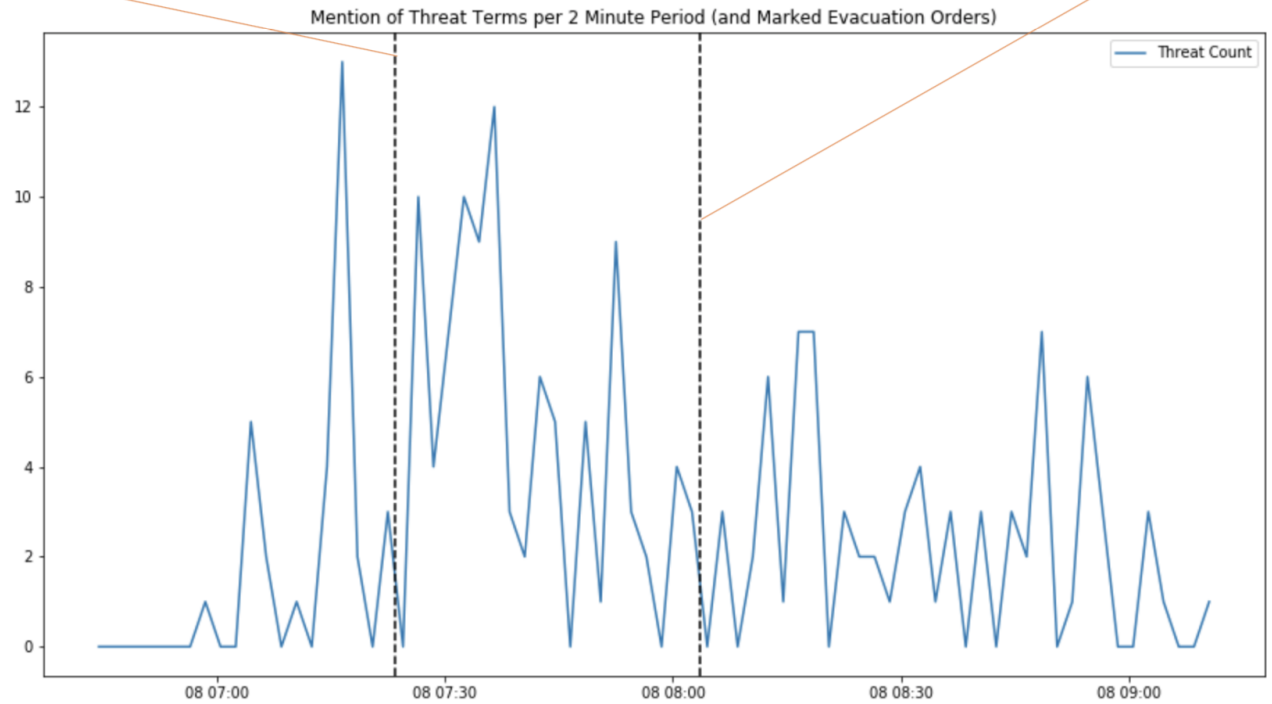
Time Distribution of Data



Start Time Second for each feed. Color shows details about is_location. Size shows details about statement_time.

Butte County Sheriff @ButteSheriff · 8 Nov 2018
EVACUATION ORDER: Due to a fire in the area, an evacuation order has been issued for the town of Pulga. If assistance is needed in evacuating, call 9 1 1.
#ButteSheriff

Butte County Sheriff @ButteSheriff · 8 Nov 2018
EVACUATION ORDER: Due to a fire in the area, an evacuation order has been issued for all of Pentz road in Paradise East to Highway 70.
#ButteSheriff #CampFire



The Conclusion

- The Problem: During a time of disaster, can audio from police radios be used to identify affected locations and understand severity to help FEMA more effectively allocate emergency resources?
  - Yes, there is value in this information, but obstacles remain to a useful solution
- What value does this add?
  - An aggregation of multiple feeds that is unavailable elsewhere
  - Interactive and interpretable information that can be accessed by all
  - Could help evaluate how resources/solutions are implemented and elucidate more effective solutions

Future Considerations

- Consider investment in high-quality training data
- Provide access to audio clips for individual incident points
- Many obstacles to solve for scaled, live solution
- Next Steps:
    - Create web-based mapping solution
    - Add level of human approval - flag things that might be an issue and add the human option to attend to or ignore the issue (see on a map the quote, location and maybe an estimated severity level)
    - Update custom vocabulary from twitter (live - applicable to all types of events)

Sources

- Data
  - Broadcastify Archives (required premium subscription)
  - CA Geographic Data: https://data.ca.gov/dataset/ca-geographic-boundaries
  - Butte selected locations: https://en.wikipedia.org/wiki/Butte_County,_California
- Images:
  - https://www.washingtonpost.com/resizer/I2T-yNy6jndvNUG-dMTUeJ658pc=/1484x0/arc-anglerfish-washpost-prod-washpost.s3.amazonaws.com/public/VVZRSOHO4EI6RC2HXUEXL7LBTE.jpg
- Background Information:
  - https://www.sfchronicle.com/california-wildfires/article/What-we-know-about-the-deadly-Camp-Fire-13401383.php
  - https://www.nytimes.com/interactive/2018/11/18/us/california-camp-fire-paradise.html?login=google

Thank You!

Anna Haas
James Lovejoy
Matt Clifford
Will Holder

April 26, 2018