Team FE
Final Presentation

Virginia Tech, Blacksburg, VA, 24060

●●●

CS 5604
Fall 2017
12/12/2017

Haitao Wang
Yali Bian
Shuo Niu
Jieun Chon

Table of contents

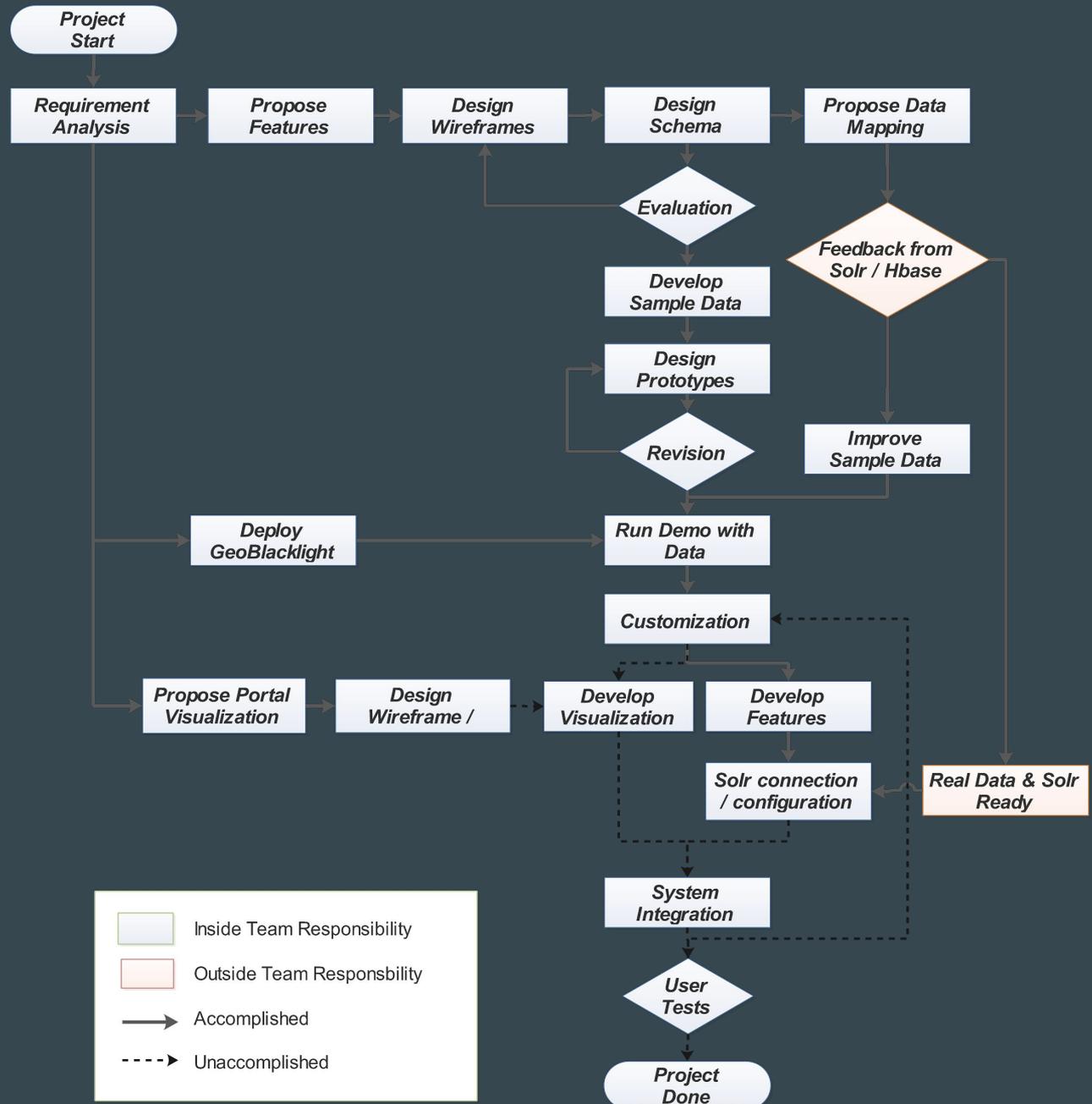1. GETAR Portal Front-end
   - Project Workflow
   - Frontend Modules
   - Core Metadata
   - Schema
   - Functionality
   - Future Work
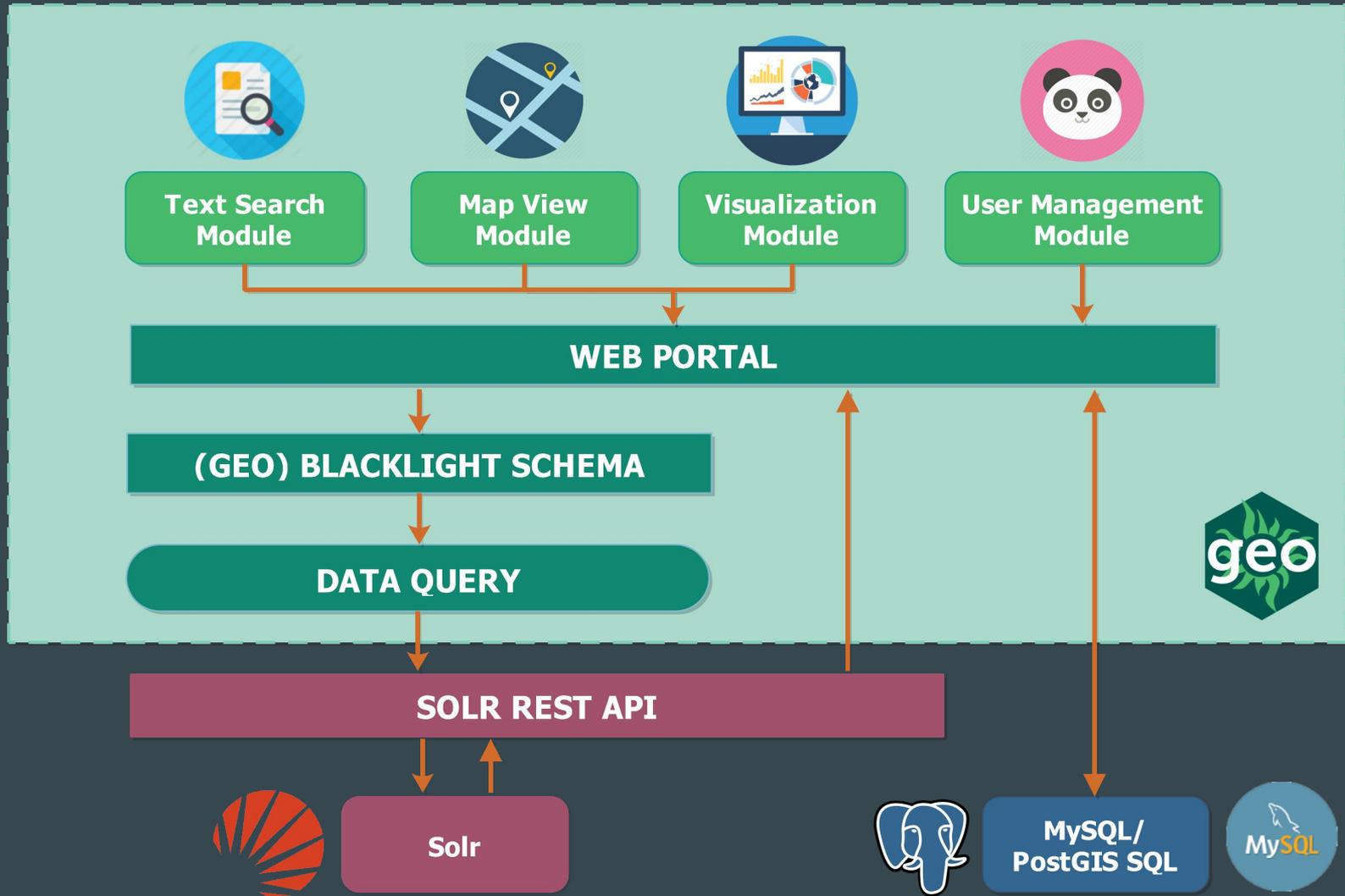
2. GETAR Visualization
   - Functions
   - Framework
   - Search
   - TweetList
   - SNS
   - User Information
   - Timeline
   - Views

GETAR Portal Front-end Project Workflow

(w/ GeoBlacklight)

GETAR Portal - Frontend Modules

GETAR Portal - Core Metadata

Metadata	Field	Description	Example
Document ID	id	Unique document identifier	899664353946996738
Author	dc_creator_sm	Document author	jhdee
Year	solr_year_i	Year when publishing the document	2017
Spatial coverage	dct_spatial_sm	Interpreted location as place name	Radford, virginia
Institution	dct_provenance_s	Institution of data collection holder	Virginia Tech
Spatial bounds	dct_bounds_s	Bounding box for S,W,N,E of interpreted location	ENVELOPE (34.0, -118.2, 36.5, -117.4)
URL	dct_url_s	URL of orginal documents	http://xxxx.yy.zz
Full text	dc_description_s	Full text of tweet / webpage	Full text
Topics	dc_topics_s	Related topic name	"midflight experience"
Collections	dct_isPartOf_sm	Related collection name	"Solar eclipse"
Geometry type	dc_geotype_s	Point or polygon	"point"
Document type	dc_format_s	Tweet or webpage	"webpage"
Hashtag	dct_hashtag_s	Hashtags of tweets	#solareclipse

GETAR Portal - Schema

for text search
- dc_creator_sm
- dc_description_s
- dc_format_s
- dc_identifier_s
- dc_publisher_s
- dc_rights_s
- dc_subject_sm
- dc_title_s
- dct_isPartOf_sm
- dct_provenance_s
- dct_spatial_sm
- dct_temporal_sm
- layer_geom_type_s
- layer_slug_s

for sorting text fields
- dc_publisher_s
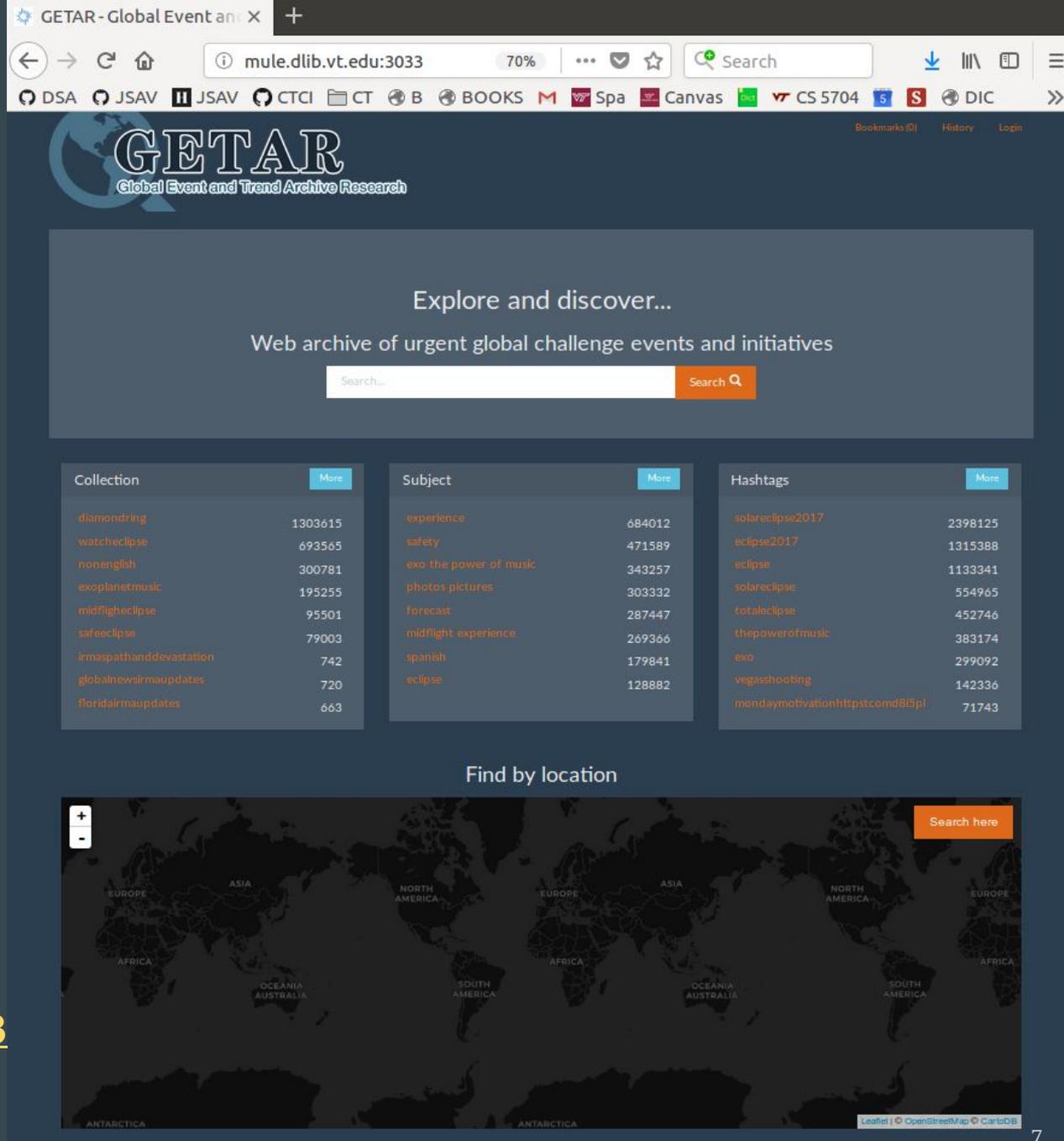- dc_title_s
- dct_provenance_s

for spell checking
- dc_creator_sm
- dc_publisher_s
- dc_subject_sm
- dc_title_s
- dct_provenance_s
- dct_spatial_sm

for suggestions
- dc_creator_sm
- dc_publisher_s
- dc_subject_sm
- dc_title_s
- dct_provenance_s
- dct_spatial_sm

GETAR Portal - Functionality

Home Page

http://mule.dlib.vt.edu:3033

GETAR Portal- Functionality

Facet Search

GETAR Portal- Functionality

Spatial Search
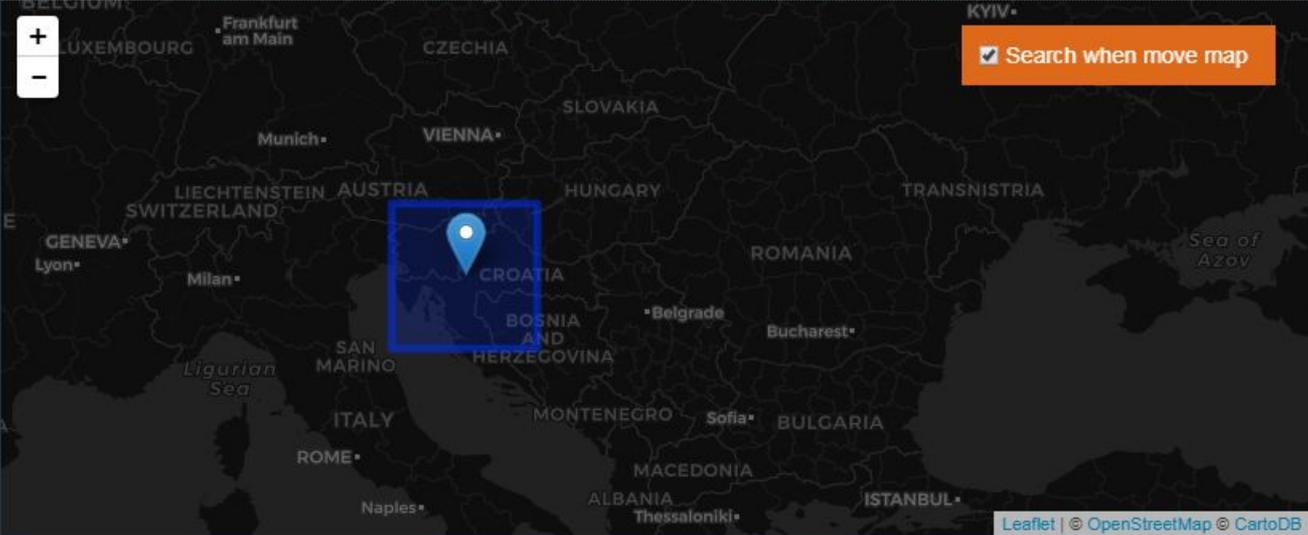
GETAR Portal
- Functionality

Document Details

UnitAPI 🐦

Source Type:	tweet
Full Text:	america solar eclipse watch history solareclipse2017 .
Subject(s):	experience
Collection:	WatchEclipse
Author(s):	UnitAPI
Publisher:	https://twitter.com
Place(s):	America
Data & Time:	2017-08-21T23:47:03Z
Hashtags:	#SolarEclipse2017
Document URL:	https://twitter.com/UnitAPI/status/899779917046583297

GETAR Portal
- Functionality

Search History

GETAR Portal - Functionality

User login

GETAR Portal - Future Work

- Spatial and temporal representation visualization of data
- Integration of GETAR Portal with GETAR Visualization
- More Customization
- Performance Analysis
- Usability Testing
- Documentation for the future users/developers

GETAR Visualization

Functions:

- Search
- View Tweets
- Social Network
- User Information
- Timeline
- Keywords
- Geo-location

GETAR Visualization

Functions:

- Search
- View Tweets
- Social Network
- User Information
- Timeline
- Keywords
- Geo-location

http://mule.dlib.vt.edu/cs5604f17_fe/TweetBank/src/

GETAR Visualization - Project Framework

User input → Solr Query → PHP → Solr/Jena → JSON → Decoded Values → D3
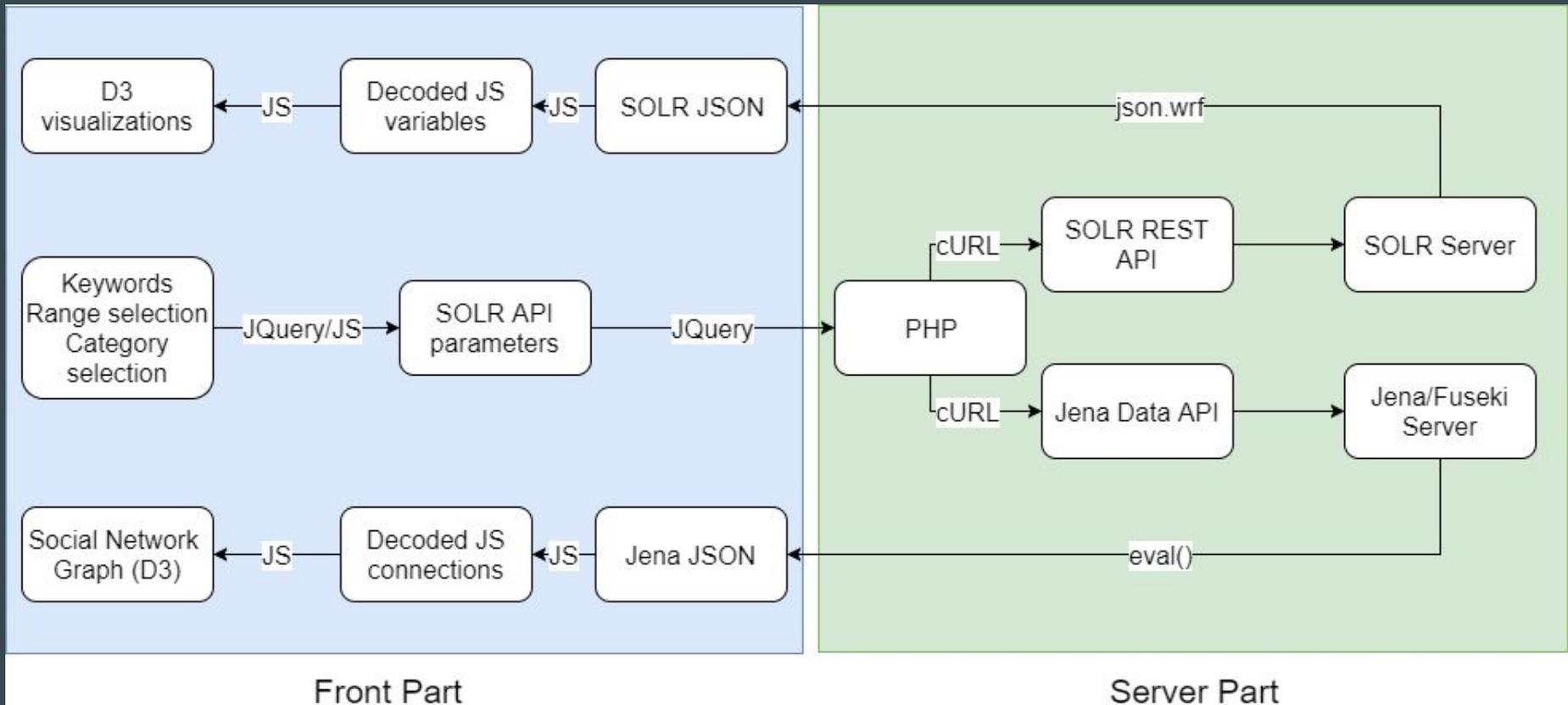
GETAR Visualization - Framework

Feature 1 *- Flexible*

    a.   Easy to add new data types

    b.   Easy to extend new views

Feature 2 *- Interactive*

    a.   overview and detail

    b.   brushing and linking

    c.   exploratory searching
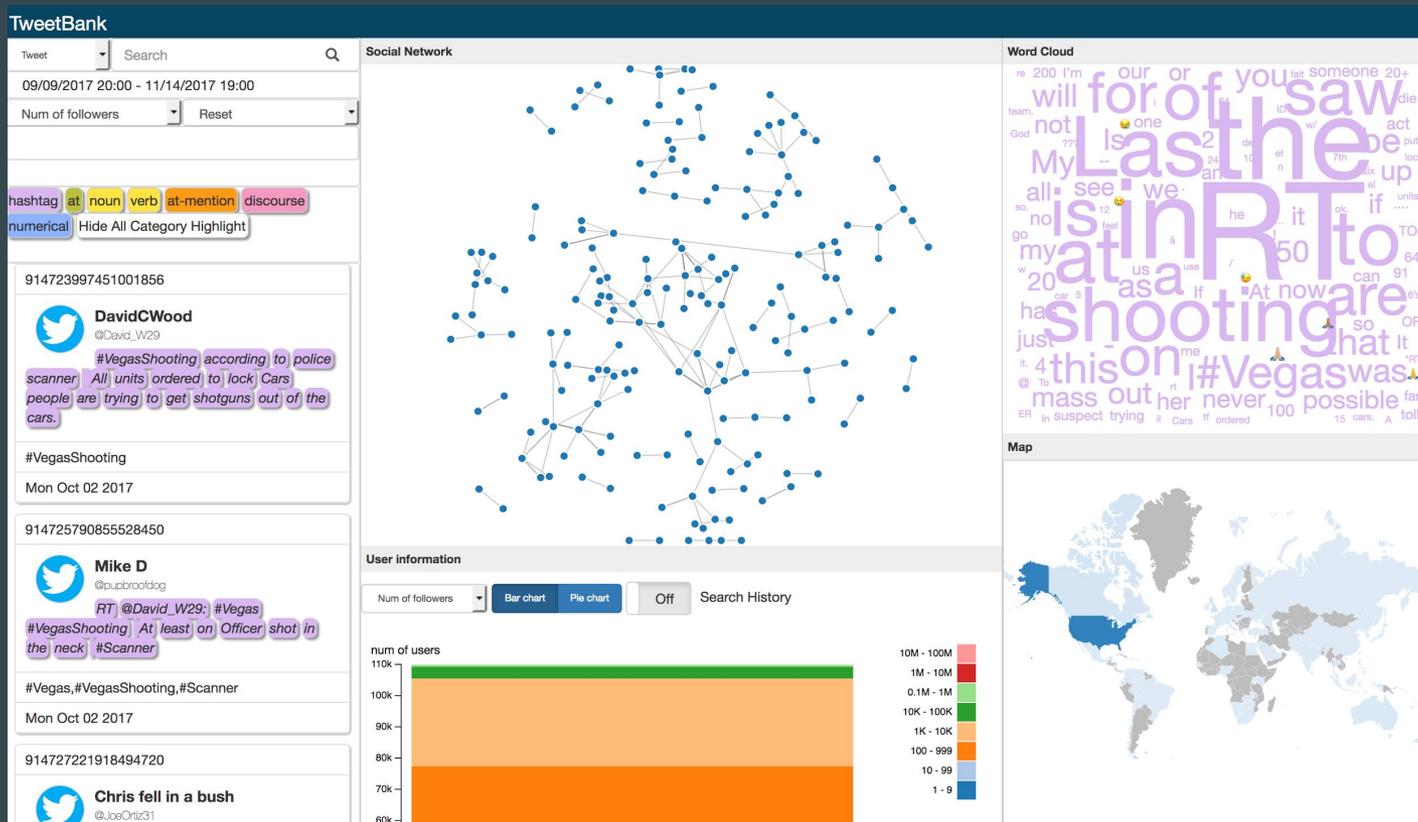
Search

Keyword:
- Single word
- Multiple words

Search fields:
- Tweet
- User name
- Hashtag

Search   Q   Tweet ▼

DiamondRing ▼

12/31/2016 19:00 - 12/31/2017 19:00

Num of friends ▼   100 TO 999 ▼

Cluster:
- DiamondRing
- WatchEclipse
- NoEnglish
- ...

Time range

User info
- Follower #
- Friends #
- Tweet #
- Favorite #
- User lang

User info range:
- 0-10
- 10-100
- 1000-10000
- ...

GETAR Visualization
- Views: TweetList

1. Embed all tweet metadata (Query from SOLR Team)
    a. UTF-8
    b. Links
    c. Emoji
2. Highlighted with keywords (NER, POS from Tweet Team)
    a. HASHTAG
    b. Place
    c. Organization
    d. Time
    e. Number
3. Interactions:
    a. Search and highlight, rerank
    b. Search and update tag cloud
    c. highlight/unhighlight

914732206718947328

Scrambull
@Scrambu11

RT @David_W29: #Vegas #VegasShooting Officer Transporting 5 critical injured in his patrol car to ER

#Vegas,#VegasShooting

Mon Oct 02 2017

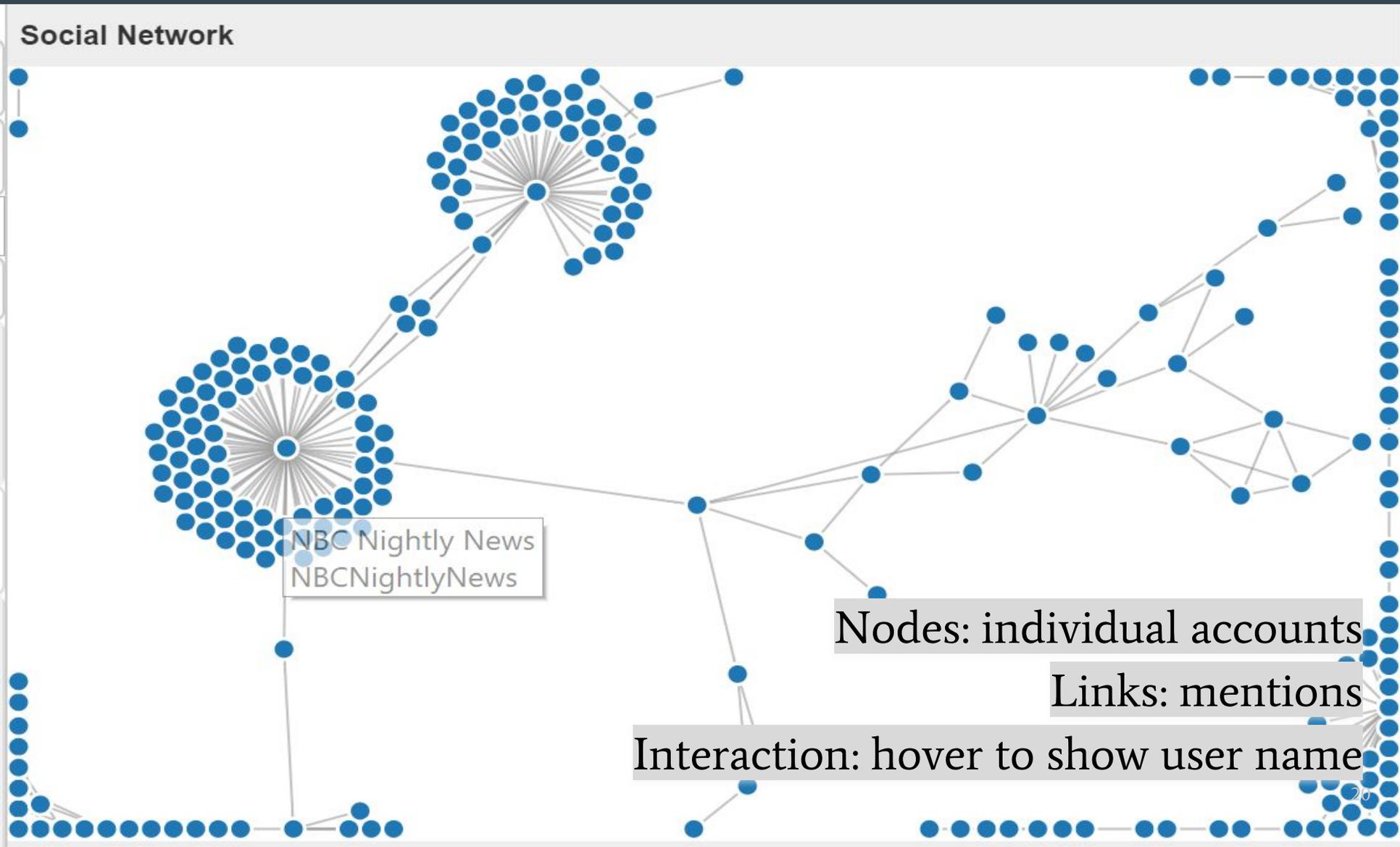
914732720957227009

Ganjdre
@andrenotlaw

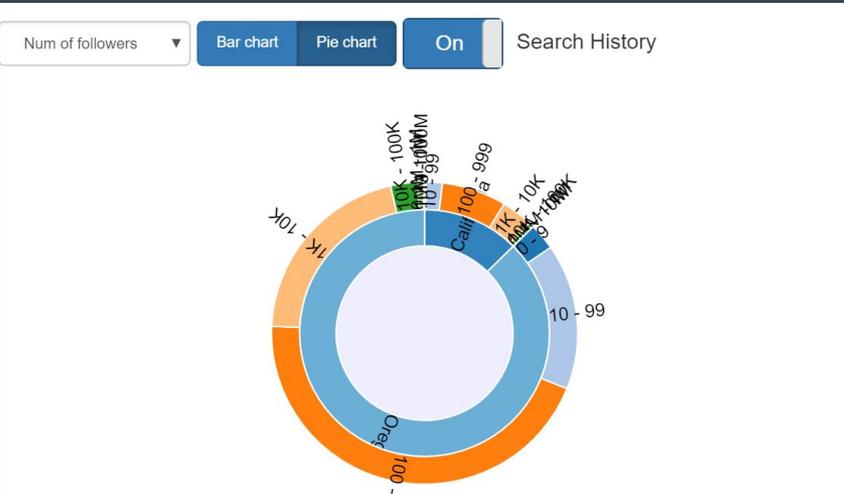RT @David_W29: #VegasShooting according to police scanner 2 shooters fully automatic weapons Mandalay Bay
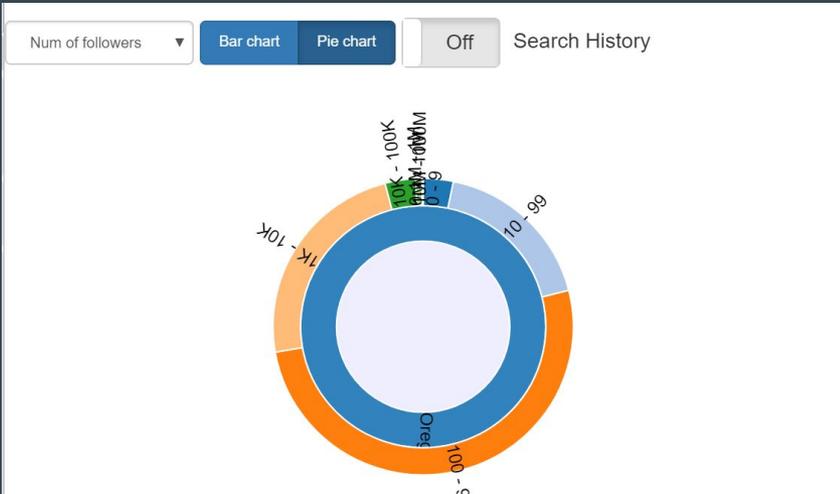
#VegasShooting

Mon Oct 02 2017
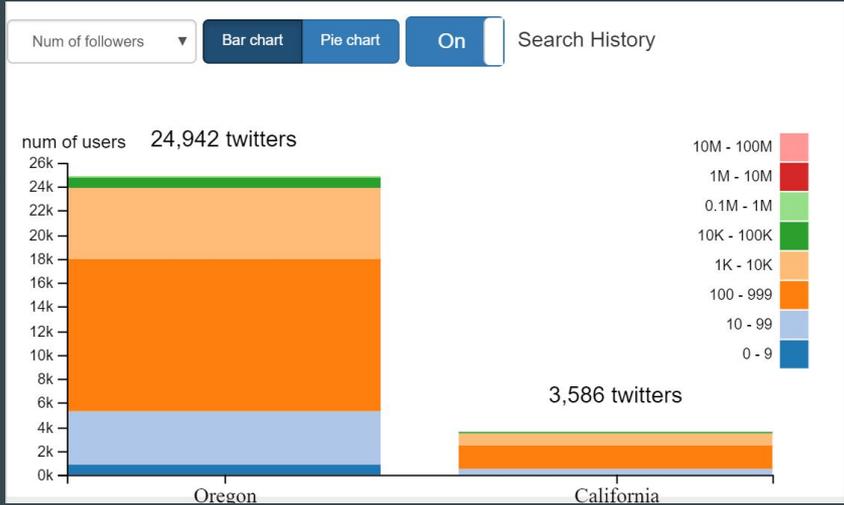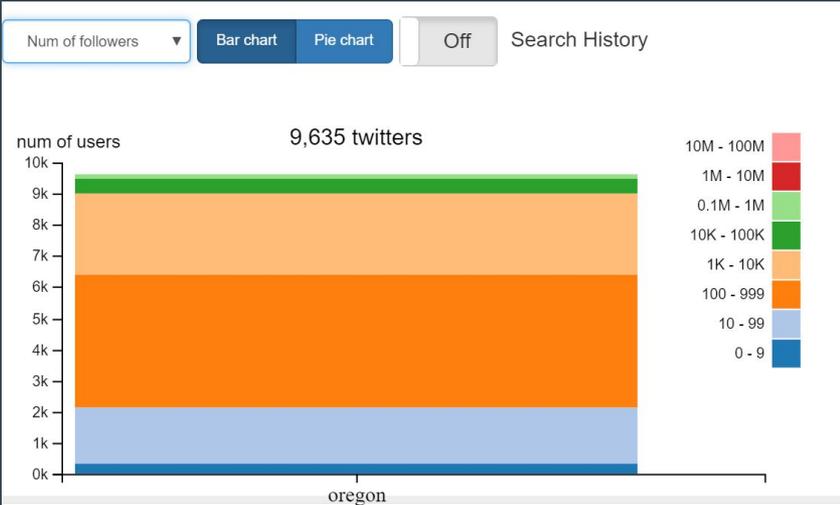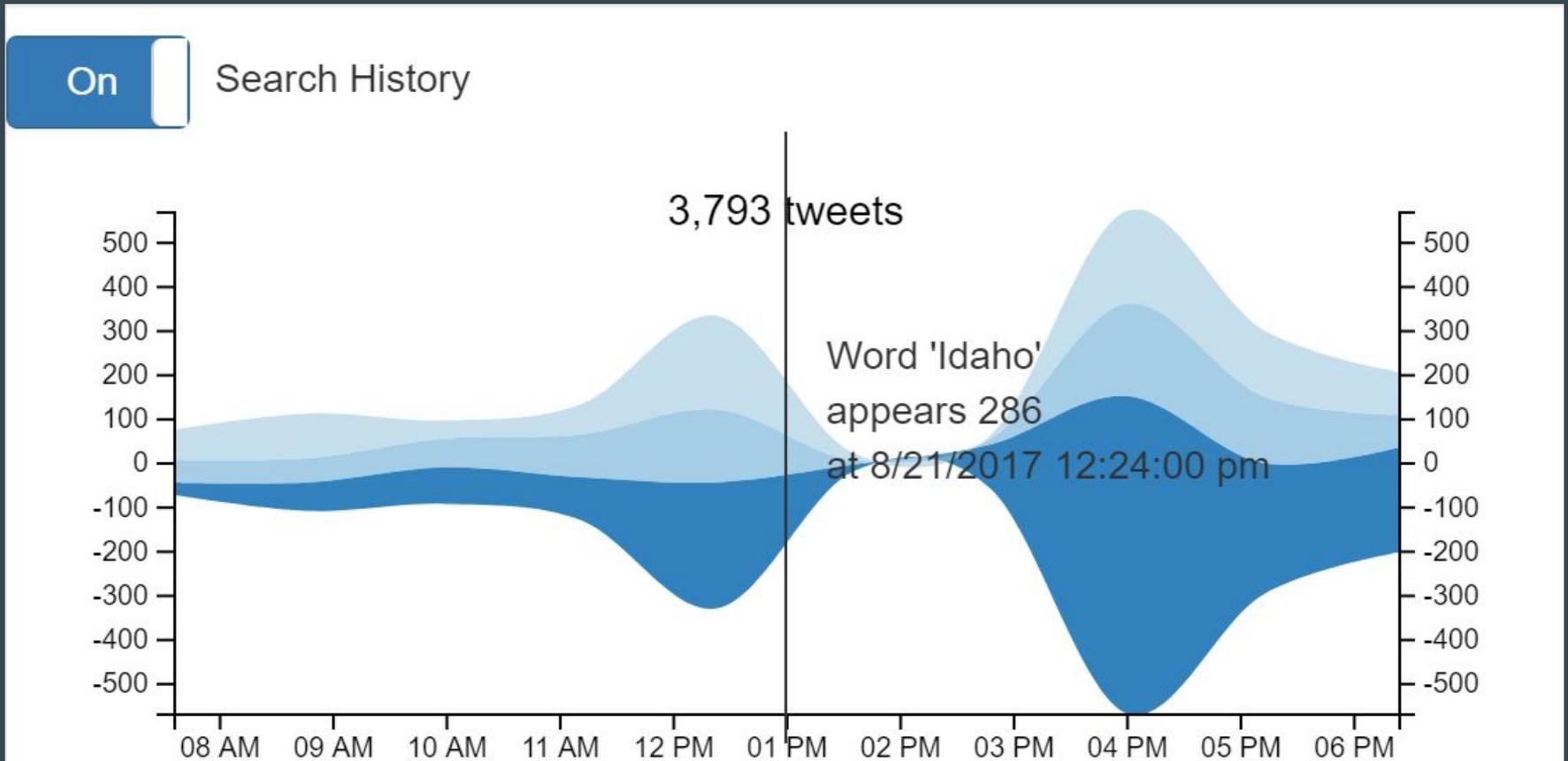
Social Network

NBC Nightly News
NBCNightlyNews

Nodes: individual accounts

Links: mentions

Interaction: hover to show user name

GETAR Visualization - User information

Timeline

Visualization Views - Tag Cloud

D3-cloud: tag cloud generator for D3.js
1. Dynamic Update through Interactions
   a. Summarization of current search results.
   b. Help users narrow down tweet list
   c. Exploratory search
2. Generalization and Emotional Expression
   a. Key Words from POS/NER
   b. Emoji
3. Interaction Reaction
   a. Searching: update based on tweet-list
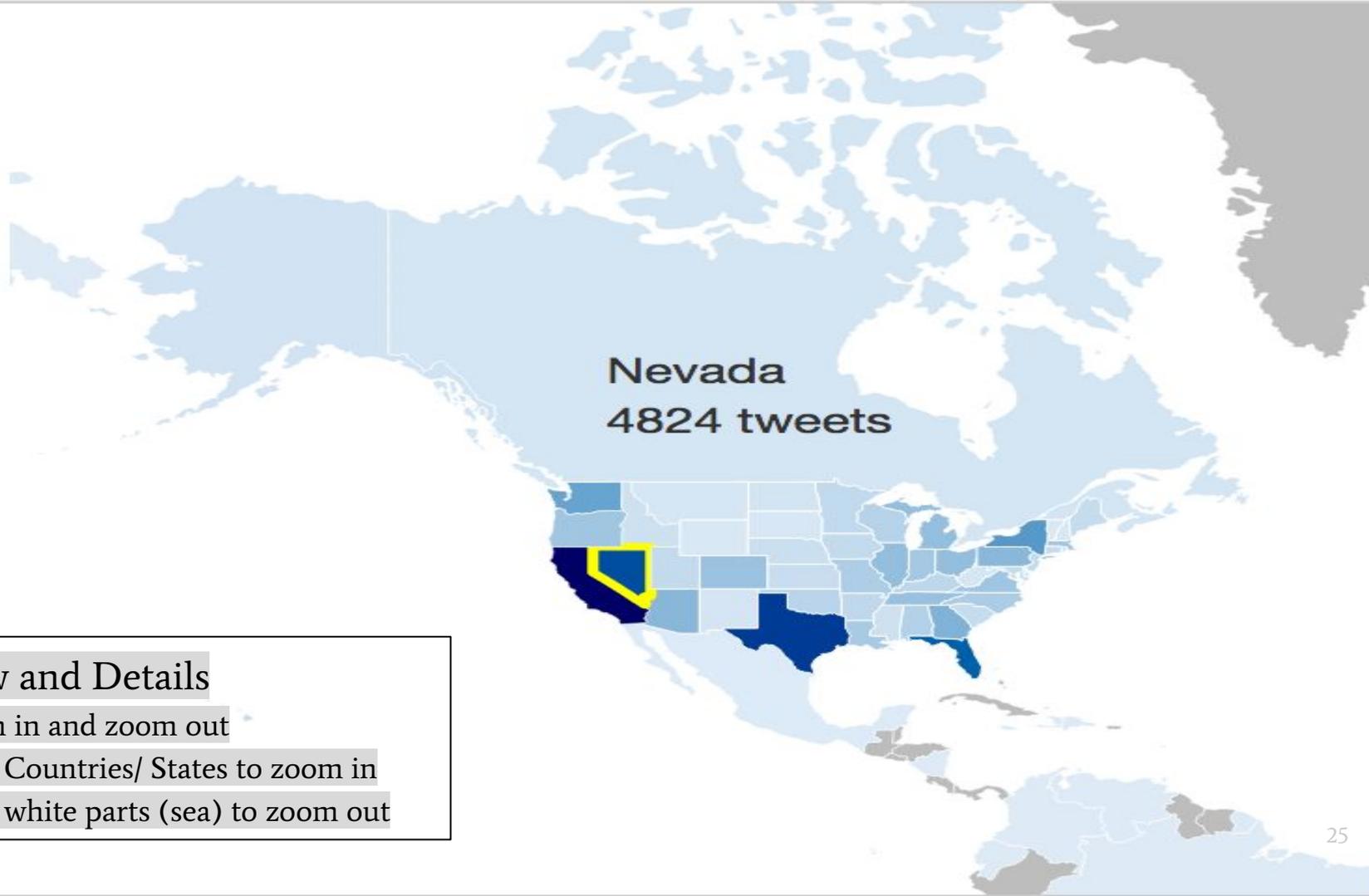   b. Highlight: update based on filtered tweet-list

Word Cloud

Visualization Views - Geo Map 1

Map

USA
3226 tweets

Different Levels
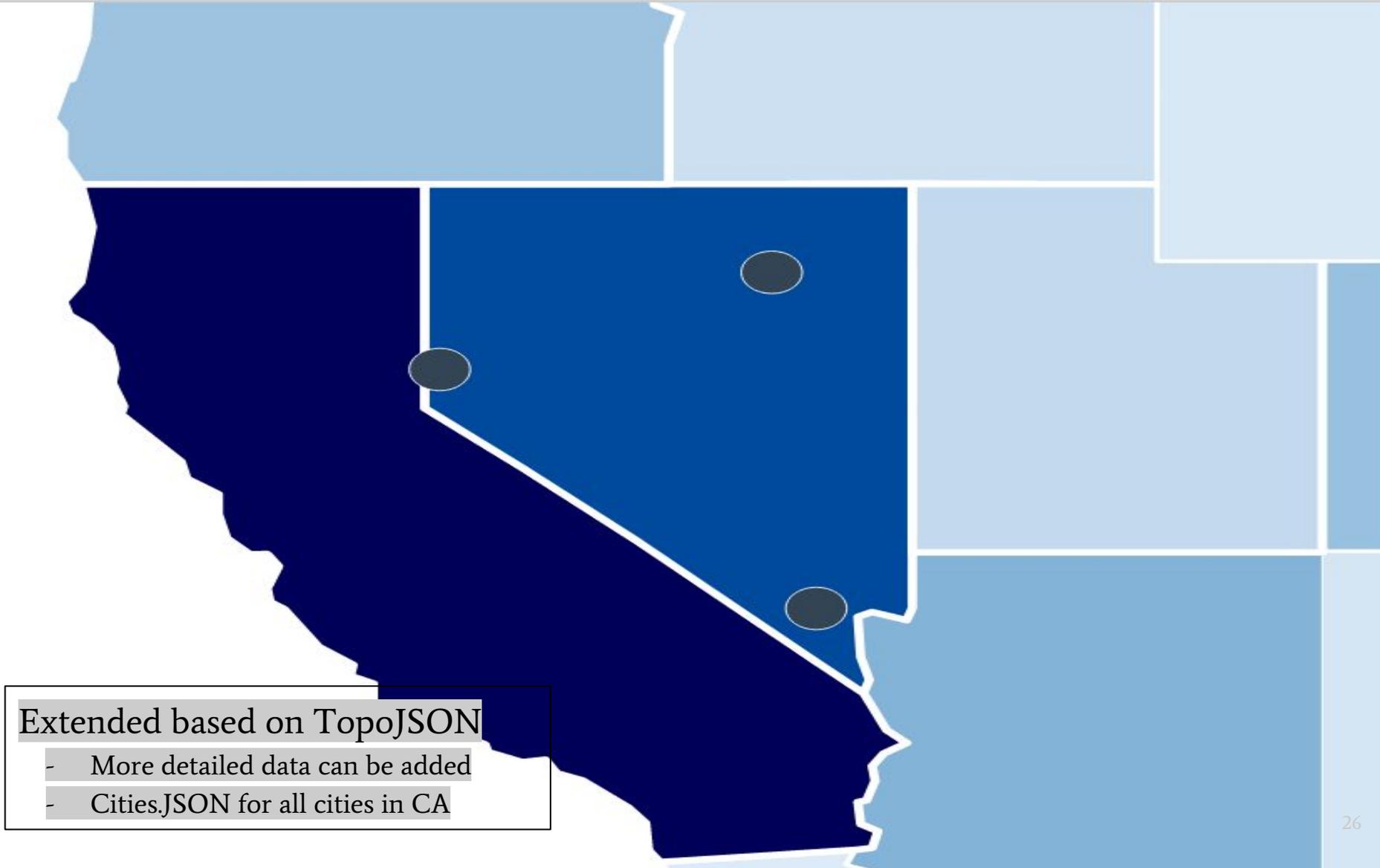- World map, US map, State map
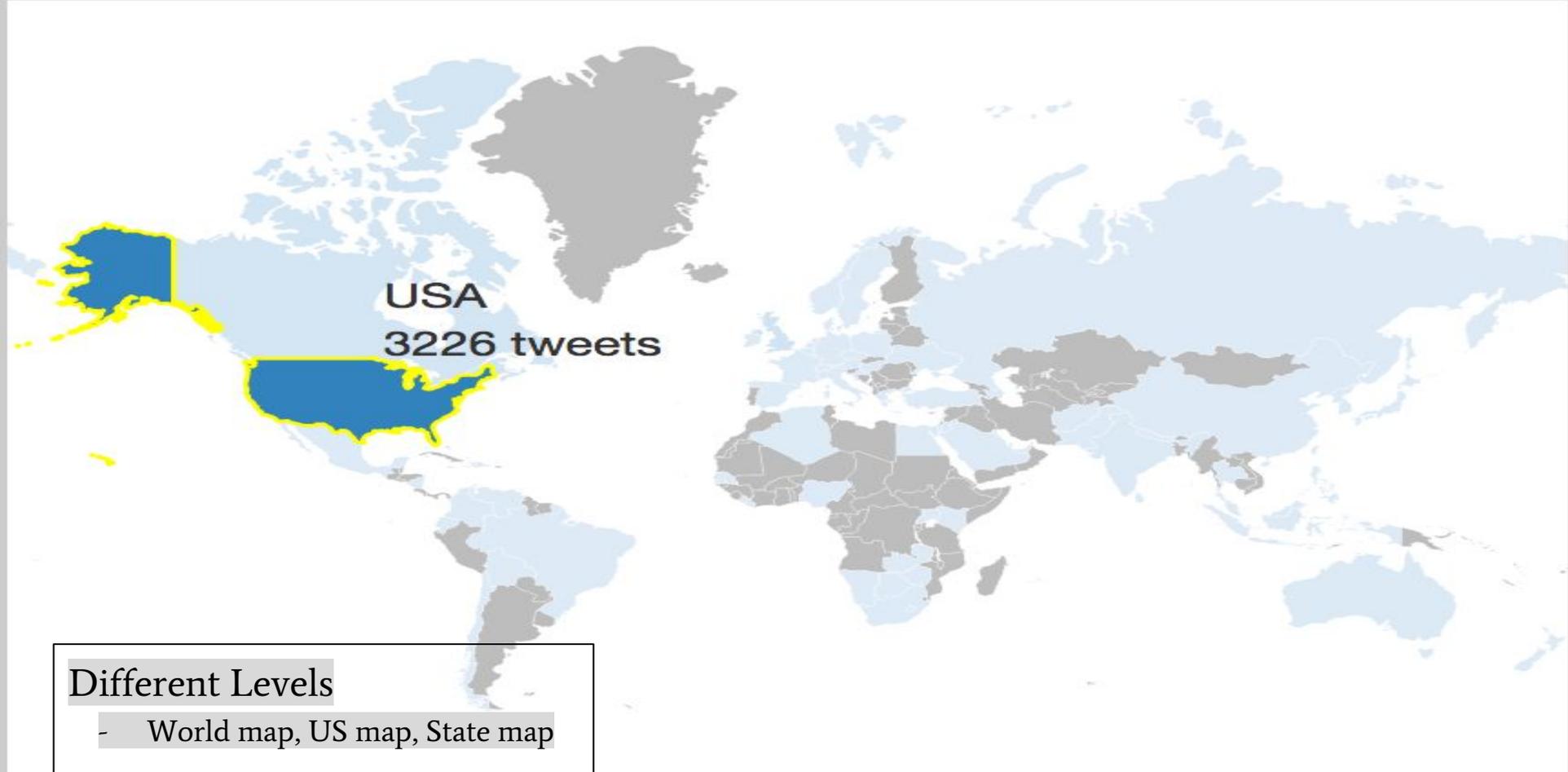
Visualization Views - Geo Map 2

Map

Nevada
4824 tweets

Overview and Details
- Zoom in and zoom out
- Click Countries/ States to zoom in
- Click white parts (sea) to zoom out

Extended based on TopoJSON
- More detailed data can be added
- Cities.JSON for all cities in CA

QUESTIONS?

Thanks go to the US National Science Foundation for supporting Global Event and Trend Archive Research (GETAR) through IIS-1619028.

Thank you!